From Conventional Methodologies to Error-Corrected Sequencing: High-Resolution Profiling of Dose-, Time-, and Tissue-Specific Benzo[*b*]fluoranthene-Induced Mutagenesis

David Schuster

Thesis submitted to the University of Ottawa in partial fulfillment of the requirements
for the Doctorate in Philosophy degree in Biology with
“Specialization in Chemical and Environmental Toxicology“

Department of Biology
Faculty of Science
University of Ottawa

Abstract

Accurate assessment of *in vivo* somatic mutagenesis is essential for human health risk assessment. However, current regulatory testing guidelines rely heavily on short-term exposures and single-gene phenotypic reporter assays. This thesis characterizes the *in vivo* mutagenic dynamics of the priority polycyclic aromatic hydrocarbon (PAH) benzo[*b*]fluoranthene (BbF) across multiple tissues and extended exposure durations using Duplex Sequencing (DS). This approach simultaneously addresses existing knowledge gaps and validates DS as a promising tool to modernize *in vivo* mutagenicity assessments based on quantitative and mechanistic data.

BbF exposure induced dose-dependent increases in micronucleus frequency in peripheral blood and mutation frequency (MF) in the liver and bone marrow (BM) of MutaMouse males after 28 days. Benchmark dose (BMD) modeling revealed comparable clastogenic and mutagenic potencies. Mutations were preferentially induced in intergenic loci, suggesting a protective activity of transcription-coupled nucleotide excision repair in transcribed regions. Furthermore, mutagenesis was characterized predominantly by C:G>A:T and secondary C:G>T:A mutations. Investigation of the trinucleotide mutation pattern revealed similarities towards human tobacco-associated cancer signatures, aligning with a canonical route of PAH exposure.

Evaluating prolonged exposures up to 90 and 180 days demonstrated that mutation accumulation mechanisms are tissue-specific. The liver exhibited a higher maximum mutational burden than BM, reflecting its primary role in xenobiotic metabolic activation. In the highly proliferative BM, mutagenesis was heavily driven by the clonal expansion of mutant cells, whereas predominantly unique *de novo* mutations were detected in liver. This indicates that the physiological propagation of mutant cells contributes differently to the overall mutation burden depending on tissue characteristics. Despite these distinct tissue dynamics, the underlying mutation spectrum remained consistent over time, indicating a stable mutagenic mechanism. Furthermore, BMD modeling highlighted that prolonged exposures increase mutagenic potency and yield lower points of departure compared to standard 28-day designs.

Direct comparisons revealed strong quantitative concordance, highlighted by robust Pearson correlations between DS-derived MF on the endogenous mouse mutagenesis panel or the exogenous *lacZ* transgene versus the conventional transgenic rodent (TGR) assay. These endpoints yielded overlapping potency estimates, validating sequence-resolved analysis as robust quantitative alternatives to the phenotypic TGR assay. Mutation spectrum analysis revealed comparable mutation subtypes between the MMP and *lacZ* loci, hinting towards endpoint independent mutagenic mechanisms. Conversely, targeted sequencing of the endogenous *Pig-a* locus was less informative and did not detect dose-related responses. Mutation spectrum analysis was predominantly driven by insertions and deletions in the *Pig-a* locus due to sequencing artifacts within homopolymeric sequences. This underscores the importance of distributed loci selection over single-gene sequencing. To address existing barriers to regulatory adoption, power analyses demonstrated that sample sizes of four animals per dose group provide sufficient statistical power, representing a reduction in animal use that supports 3Rs principles. Further collaborative efforts demonstrated the high inter-laboratory reproducibility of DS across eight laboratories and developed an open-source bioinformatic approach to harmonize mutation frequency calculations, statistical analysis and quantitative modeling.

Collectively, this research defines the dose-, time-, and tissue-dependent determinants of BbF-induced mutagenesis and shows that these factors directly influence potency estimates. Sequence-resolved analyses confirm stable mutational mechanisms across endpoints and exposure durations. Concordance with OECD assays, together with reproducibility and standardized workflows, establishes Duplex Sequencing as a sensitive and quantitatively robust approach. It reduces animal use, integrates into existing toxicology study designs, and enables high-resolution mechanistic data generation. These features support its readiness for broader adoption in *in vivo* mutagenicity assessment.

Abstract:

L'évaluation précise de la mutagenèse somatique *in vivo* est essentielle pour l'évaluation des risques pour la santé humaine. Cependant, les lignes directrices réglementaires actuelles en matière d'essais reposent largement sur des expositions à court terme et des tests phénotypiques sur un gène rapporteur unique. Cette thèse caractérise la dynamique mutagène *in vivo* d'un hydrocarbure aromatique polycyclique (PAH) prioritaire, le benzo[*b*]fluoranthène (BbF), dans de multiples tissus et sur des durées d'exposition prolongées en utilisant le séquençage en duplex (DS - *Duplex Sequencing*). Cette approche permet à la fois de combler les lacunes actuelles dans les connaissances et de valider le DS comme un outil prometteur pour moderniser les évaluations de la mutagénicité *in vivo* en se basant sur des données quantitatives et mécanistiques.

L'exposition au BbF a induit des augmentations dose-dépendantes de la fréquence des micronoyaux dans le sang périphérique et de la fréquence de mutation (FM) dans le foie et la moelle osseuse (MO) de mâles MutaMouse après 28 jours. La modélisation de la dose de référence (BMD - *Benchmark Dose*) a révélé des puissances clastogènes et mutagènes comparables. Les mutations ont été préférentiellement induites dans les locus intergéniques, ce qui suggère une activité protectrice de la réparation par excision de nucléotides couplée à la transcription dans les régions transcrites. De plus, la mutagenèse a été caractérisée principalement par des mutations C:G>A:T et, secondairement, par des mutations C:G>T:A. L'étude du profil de mutation trinucléotidique a révélé des similitudes avec les signatures des cancers humains associés au tabac, ce qui s'aligne avec une voie d'exposition canonique aux HAP.

L'évaluation d'expositions prolongées allant jusqu'à 90 et 180 jours a démontré que les mécanismes d'accumulation des mutations sont spécifiques aux tissus. Le foie présentait une charge mutationnelle maximale plus élevée que la MO, reflétant son rôle principal dans l'activation métabolique des xénobiotiques. Dans la MO, hautement prolifération, la mutagenèse était fortement stimulée par l'expansion clonale de cellules mutantes, tandis que des mutations *de novo* principalement uniques ont été détectées dans le foie. Cela indique que la propagation physiologique des cellules mutantes contribue différemment à la charge mutationnelle globale selon les caractéristiques du

tissu. Malgré ces dynamiques tissulaires distinctes, le spectre de mutation sous-jacent est resté constant au fil du temps, indiquant un mécanisme mutagène stable. De plus, la modélisation de la BMD a mis en évidence que les expositions prolongées augmentent la puissance mutagène et produisent des points de départ inférieurs par rapport aux protocoles standards de 28 jours.

Des comparaisons directes ont révélé une forte concordance quantitative, soulignée par de robustes corrélations de Pearson entre la FM dérivée du DS sur le panel de mutagenèse endogène de la souris (MMP) ou le transgène *lacZ* exogène et le test conventionnel sur rongeurs transgéniques (TGR). Ces paramètres d'évaluation ont produit des estimations de puissance qui se chevauchent, validant l'analyse résolue par séquençage comme une alternative quantitative robuste au test phénotypique TGR. L'analyse du spectre de mutation a révélé des sous-types de mutation comparables entre les locus MMP et *lacZ*, suggérant des mécanismes mutagènes indépendants du paramètre d'évaluation. À l'inverse, le séquençage ciblé du locus endogène *Pig-a* s'est avéré moins informatif et n'a pas permis de détecter de réponses liées à la dose. L'analyse du spectre de mutation a été principalement influencée par des insertions et des délétions dans le locus *Pig-a* en raison d'artefacts de séquençage au sein de séquences homopolymères. Cela souligne l'importance de la sélection de locus distribués par rapport au séquençage d'un seul gène. Pour surmonter les obstacles existants à l'adoption réglementaire, des analyses de puissance ont démontré que des échantillons de quatre animaux par groupe de dose fournissent une puissance statistique suffisante, ce qui représente une réduction de l'utilisation des animaux conforme aux principes des 3R. D'autres efforts de collaboration ont démontré la grande reproductibilité inter-laboratoires du DS à travers huit laboratoires et ont permis de développer une approche bioinformatique open-source pour harmoniser les calculs de fréquence de mutation, l'analyse statistique et la modélisation quantitative.

Collectivement, cette recherche définit les déterminants dépendants de la dose, du temps et du tissu de la mutagenèse induite par le BbF et montre que ces facteurs influencent directement les estimations de puissance. Les analyses résolues par séquençage confirment l'existence de mécanismes mutationnels stables quels que soient les paramètres d'évaluation et les durées d'exposition. La concordance avec les essais de l'OCDE, associée à la reproductibilité et à des flux de travail standardisés,

établit le séquençage en duplex comme une approche sensible et quantitativement robuste. Il permet de réduire l'utilisation d'animaux, s'intègre aux plans d'études toxicologiques existants et permet la génération de données mécanistiques à haute résolution. Ces caractéristiques confirment qu'il est prêt pour une adoption plus large dans l'évaluation de la mutagénicité *in vivo*.

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisors, **Dr. Carole Yauk** and **Dr. Francesco Marchetti**. Thank you for your mentorship, guidance, and patience throughout this PhD. You have shaped me both scientifically and personally, and I can only hope to become as strong a researcher and person as you both are. Your continued encouragement, especially during the inevitable ups and downs, and your confidence in my abilities have made this journey possible. By connecting me with international researchers and encouraging leadership within the field, you helped me build an invaluable network. I could not have wished for better supervisors, and I am grateful to have you in my life long beyond this project.

I also extend a special thank you to my mentors by heart, **Dr. Paul White** and **Dr. Naveed Honarvar.** Paul, you first supported me in Canada and had already guided me during my Master's at BASF in Germany. Without you, this PhD would not have happened. Thank you for your continued guidance, your support in our publications, and your role on my TAC committee. Naveed, thank you for introducing me to genetic toxicology and for teaching me the skills that made this work possible. Your mentorship and character laid the foundation for the scientist I am today.

I am highly appreciative of my Thesis Advisory Committee, **Dr. Kin Chan**, **Dr. Bruce McKay**, and **Dr. Paul White**. Thank you for your valuable feedback and guidance throughout this thesis. I also thank my external evaluator, **Dr. Silvia Balbo**, for serving on my committee and for the many engaging scientific discussions over the years.

This research was made possible through the generous support of the Burroughs Wellcome Fund and the Chemical Management Plan. I also thank the broader genetic toxicology community. Scientific progress depends on collaboration, and I am grateful to the many colleagues and collaborators who contributed along the way. Thank you for your trust, your openness to my ideas, and for the shared scientific and personal moments.

To my colleagues in the GReAT Lab at the University of Ottawa and Health Canada, thank you for your support and collaboration. You welcomed me to Canada with open arms, and the past four years have been both productive and enjoyable. I

especially want to highlight **David Eickmeyer**. Thank you for a friendship that will last a lifetime; the “Double D” squad will not be forgotten. **Dr. Geronimo Parodi-Matteo**, thank you for your energy, your support, and for always being there to listen. I am also grateful to all the colleagues I had the privilege to work with, including **Dr. Matthew Meier**, **Dr. Alper James Alcaraz**, **Dr. Eunnara Cho**, **Annette Dodge**, **Danielle LeBlanc**, **Gu Zhou**, **Andrew Williams**, **Elena Esina,** and many others across the GReAT Lab and Health Canada.

To my friends in Germany**, Dr. Sebastian Graute**, **Dr. Christoph Zink**, and **Markus Wolfram**, and to my close friends in Canada, **Dr. Paul Agnani**, **Nicolas Bonin**, **Alex Hicks**, **Wyeth Blumberg**, and many more: thank you for your unwavering friendship across time zones. You kept me grounded throughout this journey.

A very special thank you goes to **Emmanuelle Monniez**. We met as colleagues in the GReAT Lab and became partners along this shared journey. I would not have been able to do this without you by my side. I am incredibly grateful to have you as my partner, and I look forward to the bright future ahead of us. I love you, and together we will be unstoppable. An honorary thank you also goes to Calvin for his unwavering feline support and companionship through many long working hours.

Finally, to my family, whom I deeply love. Starting this PhD far from home was a challenge, but it was made much easier knowing that you always believed in me and remain my safe haven. Thank you to my parents, **Thomas** and **Gabrielle Schuster**, for raising me into the person I am today. You are my rock and my foundation; no matter what happens, I know I can always rely on you. Thank you to my sister, **Dr. Pia Haug**, and my brother, **Simon Schuster**, and your families. As older siblings, you have always been my role models, providing guidance and perspective. I am incredibly proud to be your brother and can only hope to follow your example.

Thank you all for your support throughout this journey. Collectively, you have shaped my path and prepared me for what comes next. As I begin the next chapter with my colleagues at Toxys in the Netherlands, I know that distance will never weaken these connections. I look forward to staying in touch with all of you.

Table of Contents

Abbreviations

AhR:	Aryl hydrocarbon receptor
AIC:	Akaike Information Criterion
AKR:	Aldo-keto reductase
BaP:	Benzo[*a*]pyrene
BbF:	Benzo[*b*]fluoranthene
BER:	Base excision repair
BM:	Bone marrow
BMD:	Benchmark dose
BMDL / BMDU:	Benchmark dose lower / upper confidence limit
BMR:	Benchmark response
COSMIC:	Catalogue of Somatic Mutations in Cancer
CYP:	Cytochrome P450
CII:	Transcriptional activator II
DS:	Duplex Sequencing
DCS:	Duplex / Double strand consensus sequence
ECS:	Error-corrected sequencing
ENU:	N-ethyl-N-nitrosourea
GLM / GLMM:	Generalized linear model / Generalized linear mixed model
GPI:	Glycosylphosphatidylinositol
IARC:	International Agency for Research on Cancer
Indels:	Insertions and deletions
lacZ:	β-Galactosidase
MF:	Mutation frequency
MFmax:	Mutation frequency maximum
MFmin:	Mutation frequency minimum
MMP:	Mouse Mutagenesis Panel
MN:	Micronucleus
MNV:	Multi-nucleotide variant
NER:	Nucleotide excision repair
NGS:	Next-generation sequencing
OECD:	Organisation for Economic Co-operation and Development
PAH:	Polycyclic aromatic hydrocarbon
pfu:	Plaque forming unit
Pig-a:	Phosphatidylinositol glycan-class A
PTFS:	Peak tag family size
PoD:	Point of departue
RBC:	Red blood cells
RET:	Reticulocytes
ROS:	Reactive oxygen species
SBS:	Single base substitution / mutational signature
SNV:	Single nucleotide variant
SSCS:	Single strand consensus sequence
TC-NER:	Transcription-coupled nucleotide excision repair
TG:	Test guideline
TGR-assay:	Transgenic rodent mutation assay
UMI:	Unique molecular identifier
VAF:	Variant allele fraction

List of Figures

Figure 8: Body weight per animal was measured at the beginning and end of the BbF exposure and used to calculate the average weight gain ± SEM (green) over time for each dose group after 28, 90 and 180 days. The liver weight (blue) per animal (absolute liver weight) was measured and used to calculate the relative liver weight compared to the final body weight per dose ± SEM over time for each dose group after 28, 90 and 180 days. Significant differences compared to respective vehicle control at each exposure duration were identified using a Dunnett's test and are indicated as * $p < 0.05$ and ** $p < 0.01$

Figure 9: Estimated vehicle control and BbF-induced mutation frequency minimum (MFmin ± SEM) in bone marrow (red) and liver (blue) of MutaMouse males after BbF exposure for 28, 90, and 180 days. Statistical testing was done using a generalized linear mixed model of each dose against their respective vehicle control; *$p < 0.05$ and **$p < 0.01$; $n = 4$ per dose group.

Figure 10: Bubble plots indicating the alt depth (size) of unique mutations (bubble) at the top dose after 28-, 90- and 180-days of exposure to BbF in bone marrow and liver. Colours are indicative of the normalized subtype of mutations according to the legend on the right.

Figure 11: Mutation frequency minimum (MFmin) with SEM per subtype in bone marrow (A; red) and liver (B; blue) after 180 days of exposure to olive oil or increasing doses of BbF. Statistical testing was done using generalized linear models for each dose and subtype. Significant differences between each dose and the vehicle controls are indicated through * $p < 0.05$ and ** $p < 0.01$.

Figure 12: Mean mutation frequency minimum (MFmin) ($x10^{-8}$) per dose of each 20 loci captured within the Mouse Mutagenesis panel using Duplex Sequencing after 28, 90 and 180 days in bone marrow and liver. The loci are arranged from the highest to lowest MFmin at the top dose for each exposure duration and tissue, with target font colour indicating intergenic (red) and genic (black) loci according to Mouse GenCode v.M25. The colour indication is consistent throughout all radar plots per dose group, and the scale is identical within each tissue

Figure 13: Overview of the C>A (blue), C>G (black), C>T (red), T>A (grey), T>C (green) and T>G (pink) mutation profiles of the mouse single base substitutions (SBS) SBS5, SBS4, SBS18 and SBS40a (top) and the trinucleotide mutation pattern (middle) of the bone marrow and liver top dose after 180 days of exposure to 25 mg/kg/day of BbF. The bottom panels display the results of the signature reconstruction of each bone marrow (left) and liver (right) dose, with indication of the total mutation count and the contribution of SBS1 (yellow), SBS4 (red), SBS5 (green), SBS17b (blue), SBS18 (orange), and SBS40a (purple). Cosine similarities are highlighted as † = greater than 0.8 and ‡ = greater than 0.9

Figure 14: Log_{10} confidence intervals (BMDL to BMDU) of benchmark dose (BMD) modeling using a 50% benchmark response on the mutation frequency minimum (circle) and mutation frequency maximum (triangle). BMD modeling on bone marrow is indicated in red, and liver is indicated in grey.

Figure 15: Mutant frequency (mut/pfu) ± SEM measured using the TGR assay in BM (left) and liver (right) from MutaMouse males exposed to increasing doses of BbF or

olive oil for 28+3 days via oral gavage (n = 8 per dose group). Statistically significant increases relative to the vehicle control were determined by one-way ANOVA with post-hoc Dunnett test and are indicated as * $p < 0.05$ and ** $p < 0.01$.

Figure 16: Mutant frequency per reticulocyte (RET; left) and per red blood cell (RBC; right) ± SEM measured using the Pig-a assay on male MutaMouse blood extracted after 28-days of exposure to increasing doses of BbF or olive oil via oral gavage. Statistically significant increases compared to the respective vehicle controls were determined after Box-Cox transformation of the mutant frequencies followed by ANOVA with post-hoc Dunnett test and are indicated as * $p < 0.05$ and ** $p < 0.01$.

Figure 17: Mutation frequency minimum (MFmin; filled bar) and maximum (MFmax; striped bar) ± SEM estimates measured by duplex sequencing in MutaMouse males exposed to increasing doses of BbF or olive oil for 28+3 days via oral gavage. Results are shown for the mouse mutagenesis panel (MMP) (left) and lacZ locus (right), in bone marrow (BM: top) and Liver (bottom) with n = 4 per group. Statistically significant increases compared to respective vehicle controls were determined using generalized linear mixed models and are indicated as * $p < 0.05$ and ** $p < 0.01$.

Figure 18: Mutation frequency minimum (MFmin; left) and maximum (MFmax; right) ± SEM measured by duplex sequencing in MutaMouse males exposed to increasing doses of BbF or olive oil for 28+3 days via oral gavage. Results are shown for the Pig-a locus with n = 4 per group. No dose showed a statistically significant increase relative to respective vehicle controls, as determined using generalized linear mixed models.

Figure 19: Mutation frequency minimum (MFmin) ± SEM estimates per mutation subtype measured by duplex sequencing in MutaMouse males exposed to increasing doses of BbF or olive oil for 28+3 days via oral gavage. Results are shown for the mouse mutagenesis panel (MMP) (left) and lacZ locus (right), in bone marrow (BM: top) and Liver (bottom) with n = 4 per group. Statistically significant increases compared to respective vehicle controls were determined using generalized linear mixed models and are indicated as * $p < 0.05$ and ** $p < 0.01$.

Figure 20: Mutation frequency minimum (MFmin) ± SEM estimates per mutation subtype measured by duplex sequencing in MutaMouse males exposed to increasing doses of BbF or olive oil for 28+3 days via oral gavage. Results are shown for the Pig-a locus in bone marrow (BM: top) with n = 4 per group. Statistically significant increases compared to respective vehicle controls were determined using generalized linear mixed models and are indicated as * $p < 0.05$ and ** $p < 0.01$.

Figure 21: Pearson correlation analysis between mutation frequency minimum of duplex sequencing using the MMP (left) or the lacZ-transgene (right) versus the mutant frequencies obtained through the phenotypic TGR assay. The dose groups are indicated by colour for 0 (olive), 6.25 (blue), 12.5 (green), 25 (yellow), 50 (orange) and 100 (red) mg/kg/day. The 95% confidence level is visualized in grey.

Figure 22: Log_{10} confidence intervals (BMDL to BMDU) of benchmark dose (BMD) modeling using a 50% benchmark response on the mutation frequency minimum for the MMP (top) and lacZ loci (middle) and mutant frequency of the TGR assay (bottom) measured in MutaMouse males exposed to increasing doses of BbF or olive oil for

28+3 days via oral gavage. BMD modeling on bone marrow is indicated in red, and liver is indicated in grey.

Statement of Contributions

Chapter 2: Dose-Related Mutagenic and Clastogenic Effects of Benzo[*b*]fluoranthene in Mouse Somatic Tissues Detected by Duplex Sequencing and the Micronucleus Assay

Published in:

Environmental Science & Technology, 58(49): 21450–21463, November 27, 2024

DOI:

10.1021/acs.est.4c07236

Authors:

David M. Schuster[1], Danielle P. M. LeBlanc[2], Gu Zhou[2], Matthew J. Meier[2], Annette E. Dodge[1], Paul A. White[2], Alexandra S. Long[3], Andrew Williams[2], Cheryl Hobbs[4], Alex Diesing[4], Stephanie L. Smith-Roe[5], Jesse J. Salk[6], Francesco Marchetti[2,7] and Carole L. Yauk[1]

Affiliations:

[1]Department of Biology, University of Ottawa, Ottawa, Canada

[2]Environmental Health Science and Research Bureau, Health Canada, Ottawa, Canada

[3]Existing Substances Risk Assessment Bureau, Health Canada, Ottawa, Canada

[4]Integrated Laboratory Systems, LLC, an Inotiv Company, Research Triangle Park, NC, USA

[5]Division of Translational Toxicology, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA

[6]University of Washington School of Medicine, Department of Medicine, Division of Hematology and Oncology

[7]Department of Biology, Carleton University, Ottawa, Canada

Experimental design:

Francesco Marchetti, Paul A. White and Carole L. Yauk

Manuscript writing lead:

David M. Schuster

***In vivo* exposure and tissue collection**:

David M. Schuster, Danielle P. M. LeBlanc and Gu Zhou

Micronucleus assay scoring:

Cheryl Hobbs and Alex Diesing

Duplex Sequencing library building:

David M. Schuster

Bioinformatic data processing:

David M. Schuster, Matthew J. Meier, Annette E. Dodge and Jesse J. Salk

Statistical analysis:

David M. Schuster and Andrew Williams

Data analysis and interpretation:

David M. Schuster, Paul A. White, Francesco Marchetti, Carole L. Yauk

Manuscript review and approval:

All authors

Chapter 3: Prolonged Benzo[*b*]fluoranthene Exposure Increases Tissue-Specific Mutational Burden, Expansion of Mutant Cells, and Mutagenic Potency

Submitted and in review at:

Environment International

Authors:

David M. Schuster[1], Danielle P. M. LeBlanc[2], Gu Zhou[2], Matthew J. Meier[2], Annette E. Dodge[1], Paul A. White[2], Andreas Zeller[3], Andrew Williams[2], David Eickmeyer[1], Sabrina Kehm[3], Juergen Funk[3], Francesco Marchetti[2,4] and Carole L. Yauk[1]

Affiliations:

[1]Department of Biology, University of Ottawa, Ottawa, Canada

[2]Environmental Health Science and Research Bureau, Health Canada, Ottawa, Canada

[3]Pharmaceutical Sciences, pRED Innovation Center Basel, F. Hoffmann-La Roche Ltd, Basel, Switzerland

[4]Department of Biology, Carleton University, Ottawa, Canada

Experimental design:

Francesco Marchetti, Paul A. White, Andreas Zeller and Carole L. Yauk

Manuscript writing lead:

David M. Schuster

***In vivo* exposure and tissue collection:**

David M. Schuster, Danielle P. M. LeBlanc and Gu Zhou

Histopathological analysis:

Andreas Zeller, Sabrina Kehm and Juergen Funk

Duplex Sequencing library building:

David M. Schuster

Bioinformatic data processing:

David M. Schuster, Matthew J. Meier and Annette E. Dodge

Statistical analysis:

David M. Schuster and Andrew Williams

Data analysis and interpretation:

David M. Schuster, Francesco Marchetti, Paul A. White and Carole L. Yauk

Manuscript review and approval:

All authors

Chapter 4: Comparative *In Vivo* Mutagenicity Assessment of Benzo[*b*]fluoranthene: Concordance of Error-Corrected Sequencing and OECD Mutation Assays

Target journal:

Archives of Toxicology

Authors:

David M. Schuster[1], Gu Zhou[2], Danielle LeBlanc[2], Matthew J. Meier[2], Paul A. White[2], Andrew Williams[2], Jesse J. Salk[3], Jake Higgins[4], Ann-Karin Olsen[5], Steven Dertinger[6], Jeffrey Bemis[6], Connie L. Chen[7], Carole L. Yauk[1] and Francesco Marchetti[2,8]

Affiliations:

[1]Department of Biology, University of Ottawa, Ottawa, Canada

[2]Environmental Health Science and Research Bureau, Health Canada, Ottawa, Canada

[3]University of Washington School of Medicine, Department of Medicine, Division of Hematology and Oncology

[4]Independent Researcher (formerly TwinStrand Biosciences), Seattle, WA, USA.

[5]Division of Climate and Environmental Health, Department of Chemical Toxicology, Norwegian Institute of Public Health, Oslo, Norway.

[6]Litron Laboratories, Rochester, New York, USA.

[7]Health and Environmental Sciences Institute, Washington, DC, USA.

[8]Department of Biology, Carleton University, Ottawa, Canada

Experimental design:

Francesco Marchetti, Paul A. White, Connie L. Chen, Ann-Karin Olsen and Carole L. Yauk

Manuscript writing lead:

David M. Schuster

***In vivo* exposure and tissue collection:**

David M. Schuster, Gu Zhou, Danielle LeBlanc

Phenotypic assays:

David M. Schuster, Steven Dertinger, Jeffrey Bemis and Gu Zhou

Duplex Sequencing library building:

David M. Schuster

Bioinformatic data processing:

David M. Schuster, Matthew J. Meier, Jesse J. Salk and Jake Higgins

Statistical analysis:

David M. Schuster, Steven Dertinger, Jeffrey Bemis and Andrew Williams

Data analysis and interpretation:

David M. Schuster, Francesco Marchetti, Paul A. White and Carole L. Yauk

Manuscript review and approval:

All authors

Appendix A: Power Analyses to Inform Duplex Sequencing Study Designs for MutaMouse Liver and Bone Marrow

Published in:

Environmental and Molecular Mutagenesis, 65(8): 234–242, 2024

DOI:

10.1002/em.22619

Authors:

Elena Esina[1], Annette E. Dodge[1,2], Andrew Williams[2], David M. Schuster[1], Danielle P. M. LeBlanc[2], Francesco Marchetti[2,3] and Carole L. Yauk[1]

Affiliations:

[1]Department of Biology, University of Ottawa, Ottawa, Ontario, Canada.

[2]Environmental Health Science and Research Bureau, Health Canada, Ottawa, Ontario, Canada.

[3]Department of Biology, Carleton University, Ottawa, Ontario, Canada.

Experimental design:

Francesco Marchetti and Carole L. Yauk

Manuscript writing lead:

Elena Esina

Duplex Sequencing library building:

David M. Schuster and Danielle P. M. LeBlanc

Bioinformatic data processing

David M. Schuster, Elena Esina and Danielle P. M. LeBlanc

Statistical and power analysis:

David M. Schuster, Danielle P. M. LeBlanc, Elena Esina and Andrew Williams

Data analysis and interpretation:

Elena Esina, Andrew Williams, Francesco Marchetti and Carole L. Yauk

Manuscript review and approval:

All authors

Appendix B: Transferability, Reproducibility and Sensitivity of Mutation Quantification by Duplex Sequencing

Published in:

Environmental and Molecular Mutagenesis, 66(6-7): 311-326, 2025 **DOI:** 10.1002/em.70020

Authors:

Shaofei Zhang[1], Barbara L. Parsons[2], Devon Fitzgerald[3], Anne Ashford[4], James Todd Auman[5], Tao Chen[2], Annette Dodge[6,10], Azeddine Elhajouji[7], Lena Pfaller[7], Shawn Harris[8], Jake Higgins[3], Cheryl A. Hobbs[5], Francesco Marchetti[6], Matthew J. Meier[6], Meagan B. Myers[2], Jesse Salk[3], Rebecca Sahroui[9], David M. Schuster[10], Raja Settivari[9], Stephanie L. Smith-Roe[11], Carole L. Yauk[10], Jian Yan[2], Andrew Williams[6] and Connie L. Chen[12]

Affiliations:

[1]Pfizer Worldwide Research, Development, and Medical, Groton, Connecticut, USA.

[2]US Food and Drug Administration/National Center for Toxicological Research, Jefferson, Missouri, USA.

[3]TwinStrand Biosciences, Seattle, Washington, USA.

[4]Safety Innovation, Safety Sciences, Clinical Pharmacology & Safety Sciences, R&D, Cambridge, UK.

[5]Inotiv, Durham, North Carolina, USA.

[6]Environmental Health Science and Research Bureau, Health Canada, Ottawa, Canada.

[7]Novartis AG, Biomedical Research, Preclinical Safety, Basel, Switzerland.

[8]DLH Corporation, Durham, North Carolina, USA.

[9]Corteva Agriscience, Newark, New Jersey, USA.

[10]University of Ottawa, Ottawa, Canada.

[11]Division of Translational Toxicology, National Institute of Environmental Health, Research Triangle Park, North Carolina, USA.

[12]Health and Environmental Sciences Institute, Washington, DC, USA.

Experimental design:

Shaofei Zhang, Connie L. Chen, Barbara L. Parsons, Francesco Marchetti, Carole L. Yauk and Jesse Salk

Manuscript writing lead:

Shaofei Zhang

Independent laboratory library generation:

David M. Schuster, Anne Ashford, Tao Chen, Azeddine Elhajouji, Lena Pfaller, Rebecca Sahroui, Raja Settivari, Stephanie L. Smith-Roe and Jian Yan

Bioinformatic data processing:

Shaofei Zhang, Barbara L. Parsons, Devon Fitzgerald, Annette Dodge, Matthew J. Meier and Jake Higgins

Statistical analysis:

Andrew Williams, Shawn Harris and Meagan B. Myers

Data analysis and interpretation:

Shaofei Zhang, Barbara L. Parsons and Connie L. Chen

Manuscript review and approval:

All authors

Appendix C: MutSeqR: An Open-Source R Package for Standardized Analysis of Error-Corrected Next-Generation Sequencing Data in Genetic Toxicology

Published in:

Bioinformatics Advances, 5(1): vbaf265, October 23, 2025

DOI:

10.1093/bioadv/vbaf265

Authors:

Annette E. Dodge[1,2], Andrew Williams[1], Danielle P.M. LeBlanc[1], David M. Schuster[2], Elena Esina[2], Clint C. Valentine[3], Jesse J. Salk[4], Alex Maslov[5], Chris Bradley[5], Carole L. Yauk[2], Francesco Marchetti[1,8] and Matthew J. Meier[1]

Affiliations:

[1]Environmental Health Science and Research Bureau, Health Canada, Ottawa, ON K1A 0K9, Canada.

[2]Department of Biology, University of Ottawa, Ottawa, ON K1N 6N5, Canada.

[3]Fulcrum Genomics LLC, Somerville, MA 02144, United States.

[4]Division of Hematology and Oncology, University of Washington School of Medicine, Seattle, WA 98195, United States.

[5]Department of Genetics, Albert Einstein College of Medicine, Bronx, NY 10461, United States.

[6]Matter Bioworks, Brooklyn, NY 11237, United States.

[7]MutagenTech Corp, New York, NY 10014, United States.

[8]Department of Biology, Carleton University, Ottawa, ON K1S 5B6, Canada.

Experimental design:

Matthew Meier, Carole L. Yauk and Francesco Marchetti

Manuscript writing lead:

Annette E. Dodge, Matthew Meier

Software development:

David M. Schuster, Annette E. Dodge, Andrew Williams, Matthew Meier, Elena Esina

Duplex Sequencing library building:

David M. Schuster, Annette E. Dodge and Danielle P.M. LeBlanc

Data visualization and software testing:

David M. Schuster, Annette E. Dodge, Danielle P.M. LeBlanc, Elena Esina and Matthew J. Meier

Manuscript review and approval:

All authors

Chapter 1: Introduction

1.1 Genetic Toxicology and Relevance for Human Health

Genetic toxicology evaluates whether exogenous agents damage genetic material and increase the probability of genetic alterations relevant to adverse health outcomes (Heflich et al., 2020; Kirkland et al., 2019; Ren et al., 2017). One of these adverse outcomes includes different types of mutations. These mutations are broadly classified into gene mutations (mutagenicity), clastogenicity and aneugenicity (Menz et al., 2023; Sun et al., 2025). Mutagenicity refers to permanent DNA sequence changes, including single nucleotide variants and small insertions or deletions (Menz et al., 2023; Ren et al., 2017). Clastogenicity refers to structural chromosome damage, and aneugenicity refers to changes in chromosome number (Menz et al., 2023). These classifications reflect distinct biological events, but share a core feature: genetic damage persists once it is fixed during DNA replication and propagated through cell division (Chatterjee & Walker, 2017).

Mutagenicity can occur in somatic tissues or germ cells. While germ cell mutation can be passed to the offspring, somatic mutations are relevant directly to human health because they can accumulate within tissues and alter cellular function. This accumulation can lead to accelerated aging or diseases such as cancer (Hanahan & Weinberg, 2011; Martincorena, 2019; Vogelstein et al., 2013). Environmental and occupational exposures can generate DNA damage, including bulky adducts and oxidative damage. If DNA damage is unrepaired prior to replication, it can manifest into mutations (Błaszczyk & Mielżyńska-Švach, 2017; Chatterjee & Walker, 2017).

Regulatory agencies require information on a chemical's mutagenic potential to make decisions on chemical safety and risk (Heflich et al., 2020; Müller et al., 1999; White & Johnson, 2016). *In vivo* data are crucial for mutagenicity assessment. Whole-organism studies capture absorption, distribution, metabolism, and excretion alongside tissue-specific biology while current *in vitro* test systems cannot fully represent these dynamics (Bell et al., 2018; Bessems & Geraets, 2013; Dixit et al., 2003). Established *in vivo* assays are therefore widely used to support chemical evaluation (Gollapudi et al., 2013; Heflich et al., 2020; White & Johnson, 2016; Wills et al., 2016). Therefore, robust chemical evaluation requires an informed understanding of the biological mechanisms that govern mutation formation and fixation.

1.2 Biological Basis of Mutagenesis

Mutagenesis is the biological process through which DNA lesions are converted into stable, heritable sequence alterations within a cell lineage (Chatterjee & Walker, 2017; Mao & Wyrick, 2019). DNA damage refers to chemical or physical alterations of the DNA structure, such as bulky adducts, oxidized bases, abasic sites, or strand breaks. A mutation is a permanent change in the nucleotide sequence (Mao & Wyrick, 2019; Ren et al., 2017). Distinct DNA repair pathways recognise and process different types of DNA damage. Small base modifications and abasic sites are primarily resolved by base excision repair (BER) (Mao & Wyrick, 2019). Bulky helix-distorting lesions are removed by nucleotide excision repair (NER). NER occurs through global genome NER and transcription-coupled NER (TC-NER). The NER pathway includes lesion recognition, dual incision, and resynthesis (Shuck et al., 2008; Spivak, 2016; Staresincic et al., 2009). Mismatch repair corrects replication mispairs and small insertion or deletion loops that escape polymerase proofreading (Manhart & Alani, 2017; Marra & Schär, 1999). This process reduces the probability of replication errors becoming fixed in the genome. Double-strand breaks are repaired mainly by non-homologous end joining or homologous recombination (Chiruvella et al., 2013; Johnson & Jasin, 2000). Non-homologous end joining has a higher error rate, whereas homologous recombination uses sequence information to enable accurate restoration. When replication forks encounter unrepaired lesions, translesion synthesis polymerases can bypass the damage. These polymerases often exhibit lower fidelity, which increases mutagenesis and links lesion tolerance directly to mutation formation (Kokoska et al., 2002; Yang & Woodgate, 2007). Mutation fixation is the key step that makes genetic change permanent. It converts transient molecular damage into a persistent sequence variant that propagates through subsequent cell divisions (Liu et al., 2016; Mao & Wyrick, 2019; Ren et al., 2017; Stinson et al., 2020).

Mutation frequency (MF) is quantified as the number of mutations per nucleotide and induced mutations are evaluated relative to concurrent controls (Gingerich et al., 2014; Valentine et al., 2020; White et al., 2019). MF is challenging to quantify because spontaneous somatic mutations are rare in healthy tissues. This rarity reflects the combined action of high-fidelity DNA replication and genome maintenance pathways (Kane & Shcherbakova, 2014; Ueda et al., 2022). Replicative polymerases achieve high fidelity through intrinsic nucleotide selectivity and proofreading while mismatch

repair removes errors that escape proofreading. These processes result in very low mutation rates per cell division (Manhart & Alani, 2017; Sale, 2013). Consequently, the absolute frequency of new mutations remains low even when a genotoxicant increases mutation formation. Robust measurement therefore depends on biological replication and adequate sampling of cells or DNA molecules.

Mutagenesis varies across tissues because they differ in exposure, metabolism, proliferation kinetics, and repair capacity (Blokzijl et al., 2016; Hoch, 2023; Sun et al., 2019). The tissue-specific internal dose of a genotoxicant is determined by toxicokinetics. Furthermore, tissue-dependent metabolism creates strong organ-to-organ differences in the type and abundance of DNA lesions caused by reactive metabolites (Dixit et al., 2003; Sun et al., 2019). Proliferation rate influences how quickly lesions are converted into mutations. Replication provides the opportunity for misincorporation or error-prone bypass at damaged sites (Liu et al., 2016; Mao & Wyrick, 2019; Yang & Woodgate, 2007). Repair pathway activity and chromatin context modulate lesion removal. They also affect the probability that unrepaired damage is converted into mutations. These factors contribute to tissue-specific mutation burdens even under identical external exposures (Hoch, 2023; Mao & Wyrick, 2019).

The impact of mutations is magnified in somatic tissues when they clonally expand. Clonal expansion occurs when a single mutational event in a self-renewing progenitor propagates to many descendant cells. This propagation leads to the repeated detection of the same variant (Heddle, 1999; Martincorena, 2019). This expansion can occur randomly due to genetic drift. Alternatively, expansion occurs via positive selection when a specific mutation confers a growth or survival advantage to the host cell. In rapidly proliferating tissues, clonal expansion contributes significantly to the total number of detectable mutations. This process alters how MF relates to initial lesion formation (Heddle, 1999). Clonality dictates how mutagenesis must be quantified. Counting all mutant observations equally conflates the number of independent mutation events with the size of the mutant clones. Distinguishing independent mutations from clonally derived mutations is essential for accurate interpretation. This distinction is critical when comparing tissues with different cell turnover dynamics (Heddle, 1999; Nishino et al., 1996).

Exposure duration and sampling time influence measured mutagenesis. These parameters determine the opportunity for damage induction, the time available for mutation fixation and clonal outgrowth (Marchetti et al., 2021). With increasing exposure duration, mutation burden rises through sustained lesion formation and repeated opportunities for replication-dependent fixation. In highly proliferative tissues, mutation burden also increases through the clonal expansion of mutated progenitors (Heddle, 1999). While the expansion of neutral mutations occurs steadily over time, clonality also increases during regenerative proliferation induced by tissue toxicity. Furthermore, extended exposures increase the probability of mutations occurring in cancer driver genes that actively drive cell proliferation. The balance between mutation fixation and clonal expansion does not scale identically across tissues (Blokzijl et al., 2016; Cosentino & Heddle, 1999; Xie et al., 2014). Sampling time relative to exposure is critical. Mutations are detectable only after DNA lesions are fixed during replication. This fixation is delayed in slowly proliferating tissues. Consequently, capturing the full mutational response in these tissues requires longer post-exposure intervals (Douglas et al., 2022; Marchetti et al., 2021; OECD, 2025d). Tissue biology, clonal dynamics, and time dictate how an exposure translates into an observed mutation signal. These factors must be evaluated to accurately interpret the dose-, tissue-, and duration-dependent effects of environmental genotoxicants.

1.3 Benzo[*b*]fluoranthene (BbF) as an Exemplar Environmental Mutagen

Polycyclic aromatic hydrocarbons (PAHs) are a group of environmental genotoxicants of great concern. These organic compounds consist of two or more fused aromatic rings. They are generated largely by the incomplete combustion of organic material, such as coal, petroleum fuels, biomass, and tobacco (McMullin et al., 2025; Srogi, 2007). In ambient air, PAHs exist in both gas and particle phases. Higher-molecular-weight PAHs are predominantly particle-bound. Consequently, they are enriched in particulate matter and indoor dust reservoirs (Maertens et al., 2008; Srogi, 2007; Yang et al., 2021). Human exposure occurs through inhalation, dermal contact, and ingestion, particularly of contaminated food and dust. Dietary ingestion typically dominates background human exposure. Inhalation becomes more prominent in high-emission microenvironments, such as traffic corridors and combustion-intensive spaces (Bernard & Dudler, 2022; X. Zhang et al., 2021; Y. Zhang et al., 2014). Humans are typically exposed to PAHs as complex mixtures. However, evaluating the toxicity

of complex mixtures is challenging for regulatory agencies. Therefore, risk assessment frameworks rely on representative index chemicals with well-described sources, toxicokinetics, and genotoxic mechanisms to model PAH toxicity (Boström et al., 2002; Kelly et al., 2021).

Benzo[*b*]fluoranthene (BbF; C20H12, 252.3 g/mol, CAS 205-99-2) is a five-ring PAH. It is classified as one of the 16 priority PAHs by the U.S. Environmental Protection Agency due to its environmental prevalence and toxicological concern (Keith, 2015; Thurston et al., 2000; Tomei et al., 2006; Y. Zhang et al., 2020). It has been detected in mixtures from traffic emissions, residential heating, industrial combustion, and tobacco smoke. Consistent with its high molecular weight and low volatility, BbF associates with particulate matrices rather than remaining in the gas phase (Maertens et al., 2008; McMullin et al., 2025; Srogi, 2007). Environmental BbF concentrations in air are typically in the sub-ng/m³ range with environmental BbF levels at 0.61–0.65 ng/m³ in outdoor and 0.20–0.30 ng/m³ in indoor air (Gustafson et al., 2008; Jung et al., 2010). BbF in settled house dust, which acts as a reservoir for higher-molecular-weight PAHs, ranges from 0.160 to 54.0 µg/g, with an arithmetic mean of 4.87 µg/g (Maertens et al., 2008; Whitehead et al., 2013). Ingestion, as the primary pathway for PAH exposure in non-smokers, accounts for more than 70% of total PAH exposure. The estimated mean dietary intake of BbF varies around 0.09 µg/day, depending on the individual lifestyle (Martorell et al., 2010). Thus, BbF is considered a ubiquitous environmental pollutant.

BbF disposition in the body is determined by its physicochemical profile. BbF is a highly hydrophobic, nonpolar hydrocarbon with a low aqueous solubility of ~0.001 mg/L, vapour pressure ~5 × 10^{-7} mm Hg, Henry's constant ~5 × 10^{-7} atm·m³/mol, and log Koc ~6.3–7.3 (Patel et al., 2020). This profile supports extensive partitioning into lipid-rich compartments. It also causes slower biological mobilization compared to more polar PAHs (Patel et al., 2020; Srogi, 2007). Following absorption, the liver serves as the major site of BbF biotransformation (Amin et al., 1982; Franchin et al., 2026).

BbF does not directly damage DNA. It must be metabolically activated to generate genotoxic intermediates. PAH bioactivation initiates through the aryl hydrocarbon receptor (AhR). The AhR is a ligand-activated transcription factor. Upon activation, it translocates to the nucleus, dimerizes with the AhR nuclear translocator,

and drives the transcriptional induction of xenobiotic-metabolizing enzymes (Moorthy et al., 2015; Nebert et al., 2004; Sundberg & Hankinson, 2019). These AhR-dependent pathways induce Phase I Cytochrome P450 (CYP) enzymes, notably CYP1A1, CYP1A2, and CYP1B1 (Moorthy et al., 2015; Vorrink et al., 2014).

Reactive metabolites are created via two primary pathways. The first is the diol-epoxide activation route. This process begins with CYP-mediated monooxygenation of the parent PAH to epoxide intermediates. Epoxide hydrolase then converts these epoxides into trans-dihydrodiols. A second CYP-dependent oxidation produces reactive dihydrodiol epoxides capable of forming bulky covalent DNA adducts (Boström et al., 2002; Shiizaki et al., 2017). Rat-liver metabolic studies have identified multiple hydroxylated metabolites and dihydrodiol intermediates. These include trans-11,12-dihydro-11,12-dihydroxybenzo[*b*]fluoranthene, confirming the formation of reactive precursors consistent with this activation pathway (Amin et al., 1982). In parallel, a second metabolic pathway relies on aldo-keto reductases (AKRs) to produce redox-active PAH *o*-quinones (Moorthy et al., 2015; L. Zhang et al., 2012).

Ultimately, the balance between bioactivation and clearance dictates the DNA damaging potential of a chemical. Clearance of BbF is driven by Phase II conjugation, which includes glucuronidation, sulfation, and glutathione conjugation (Moorthy et al., 2015; Shiizaki et al., 2017; L. Zhang et al., 2012).

Reactive dihydrodiol-epoxides covalently bind DNA, predominantly at the N2 position of guanine, to form bulky PAH-DNA adducts. These helix-distorting lesions block replicative DNA polymerases and stall replication forks. If these lesions persist into the S phase, error-prone translesion synthesis polymerases bypass the damage, characteristically producing C:G to A:T transversions and frameshift deletions (Gowda et al., 2017; Hess et al., 1997; Paules et al., 1988). The NER pathway primarily mediates the removal of these bulky adducts. Excision efficiency varies according to adduct stereochemistry and local DNA conformation, establishing a mechanistic basis for site- and context-dependent mutagenesis (Hess et al., 1997; Mocquet et al., 2007). In parallel, *o*-quinones generated via the AKR pathway undergo redox cycling to produce ROS that induces oxidative DNA damage, notably 8-oxoguanine (8-oxoG). Unrepaired 8-oxoG readily mispairs with adenine during replication, driving additional C:G>A:T transversions. The BER pathway maintains genomic integrity by removing

these oxidative base lesions before mutation fixation occurs (David et al., 2007; Wang et al., 2025; Zhang et al., 2012).

Consistent with these mechanisms, the International Agency for Research on Cancer (IARC) classifies BbF as a Group 2B compound, or a possible human carcinogen (IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, 2010). Within the IARC framework, the carcinogenicity of higher-molecular-weight combustion PAHs is mechanistically defined by the metabolic activation of parent compounds into electrophilic intermediates, followed by covalent interaction with DNA and the subsequent conversion of persistent lesions into mutations within critical genes. Empirical *in vivo* data support this sequence of events for BbF. Following a 28-day oral exposure in MutaMouse males, BbF produced DNA adducts across multiple somatic tissues (Long et al., 2016). Collectively, these findings establish BbF as a robust, mechanistically interpretable PAH, providing an ideal exemplar for evaluating dose- and tissue-dependent *in vivo* mutagenesis using mutation-resolved methodologies

1.4 Current Framework for *In Vivo* Mutagenicity Assessment

Regulatory assessment of mutagenic hazard operates through a tiered framework. Initial hazard identification relies on *in vitro* tests. Subsequently, *in vivo* assays confirm biological relevance under conditions that incorporate absorption, distribution, metabolism, excretion, and tissue-specific biology (Müller et al., 1999; OECD, 2016a). The Organisation for Economic Co-operation and Development (OECD) reviews and endorses internationally applicable test guidelines (TGs) (OECD, 1995, 2016b, 2017, 2020a, 2025d, 2025c). These TGs define standardized study designs, sampling regimens, acceptability criteria, and interpretation principles that are optimized for routine evaluation and hazard identification. Recent OECD revisions also explicitly incorporated efficiency and animal welfare considerations, in alignment with the 3R principle: replacement, reduction and refinement (Kroeger, 2006; OECD, 2017, 2025d, 2025c; Thybaud et al., 2017). When executed in compliance with Good Laboratory Practice, data generated under these TGs ensure international portability across regulatory programs via the OECD Mutual Acceptance of Data framework (OECD, 1995, 2017). Furthermore, the OECD strongly encourages combining

endpoints within repeat-dose study designs to reduce unnecessary animal use (OECD, 1995).

Initial *in vitro* screening evaluates both gene mutation and chromosomal damage. The bacterial reverse mutation test (OECD TG 471) serves as the primary screen for point mutation activity (OECD, 2020a). Subsequent mammalian cell follow-up includes gene mutation assays (e.g., OECD TG 476) and chromosomal damage assays (e.g., OECD TG 487) to detect clastogenic and aneugenic events (OECD, 2016b, 2023). While these *in vitro* assays provide broad endpoint coverage, they inherently lack organism-level kinetics and tissue-specific metabolism. Consequently, international guidance mandates that positive *in vitro* results trigger *in vivo* testing to establish true biological relevance (OECD, 2017). This tiered logic ensures that hazard assessments for chemicals and pharmaceuticals accurately account for complex exposure contexts and metabolic activation.

In vivo testing is conventionally organized into distinct endpoint classes, primarily chromosomal damage and gene mutation. To assess chromosomal damage, the mammalian erythrocyte micronucleus test (OECD TG 474) evaluates chromosome breakage and chromosomal loss during erythropoiesis by enumerating micronucleated immature erythrocytes sampled from bone marrow (BM) or peripheral blood (OECD, 2016a). The transgenic rodent (TGR) somatic and germ cell gene mutation assay (OECD TG 488) is the current gold standard to assess *in vivo* gene mutagenicity (OECD, 2025d). The TGR assay provides a practical approach to quantify mutations in reporter transgenes recovered from germ cells or somatic tissues (Gingerich et al., 2014; OECD, 2025d). The TGR can be used to link tissue choice to the route of exposure and tissue-specific metabolism, enabling mutation assessment in critical site-of-contact and target organs (Douglas et al., 2022; Long et al., 2016; OECD, 2025d). Alternatively, the Phosphatidylinositol glycan class A (*Pig-a*) gene mutation assay (OECD TG 470) measures *in vivo* mutagenesis using a blood-based phenotypic readout (OECD, 2025c). The assay detects mutant-phenotype blood cells that lack glycosylphosphatidylinositol-anchored surface proteins resulting from mutations in the endogenous *Pig-a* gene (Dobrovolsky et al., 2010; OECD, 2025c; Peruzzi et al., 2010). Because this assay utilizes peripheral blood, it permits non-lethal serial sampling within the same animal, facilitating time-course evaluations and seamless integration into

repeat-dose toxicology study designs (Bryce et al., 2008; Dertinger et al., 2021; Gollapudi et al., 2015).

Historically, OECD genetic toxicology guidelines focused primarily on binary hazard identification (OECD, 2016a, 2017, 2025d, 2025c). However, the field is shifting toward using multi-dose datasets for quantitative human health risk assessment. This modern approach relies specifically on benchmark dose modeling and point-of-departure derivation (Gollapudi et al., 2015; MacGregor et al., 2015; White et al., 2025; White & Johnson, 2016). Quantitative analysis demands assay comparability and distinction between transient DNA damage indicators and fixed mutation endpoints (Gollapudi et al., 2013; MacGregor et al., 2015; White & Johnson, 2016). Accordingly, studies intended for quantitative dose-response characterization require specific protocol modifications beyond routine hazard identification frameworks.

1.5 Limitations of Conventional *In Vivo* Assays for Modern Genotoxicity

Conventional *in vivo* genotoxicity and mutagenicity assays provide a robust foundation for regulatory hazard identification. However, modern genetic toxicology increasingly requires data beyond binary hazard calls (White & Johnson, 2016). Modern risk assessment demands quantitative dose-response modeling and broad tissue coverage to accurately evaluate a chemical's genotoxic properties. Furthermore, modern approaches require precise mechanistic interpretations that link specific DNA lesion classes to mutational outcomes (Gollapudi et al., 2013; MacGregor et al., 2015; Salk & Kennedy, 2020). Updated OECD TGs explicitly note that studies intended for quantitative risk assessment require protocol modifications, such as additional dose groups, alternative dose-selection strategies, and increased statistical power, that exceed routine hazard identification frameworks (OECD, 2017, 2025d, 2025c). Consequently, a critical gap exists between the capabilities of conventional assays and the stringent demands of modern quantitative mutagenicity assessment.

The *in vivo* mammalian erythrocyte micronucleus test exclusively evaluates structural chromosome damage in hematopoietic tissues (OECD, 2016a). Chemicals often exhibit strong tissue-specific genotoxicity patterns, particularly in site-of-contact tissues and metabolically active organs, that are not mirrored in hematopoietic compartments. Despite recent efforts, the inclusion of additional somatic target tissues lacks standardisation and validation for regulatory application (Morita et al., 2011).

Additionally, the mammalian erythrocyte micronucleus test measures clastogenic and aneugenic events rather than gene mutations (Hayashi, 1984; OECD, 2016a). Consequently, reliance on hematopoietic micronucleus endpoints alone does not inform on mutagenicity and might miss chromosome damage if a chemical does not cause an effect in hematopoietic tissues.

The *Pig-a* assay provides a non-lethal, blood-based phenotypic readout to assess gene mutations (OECD, 2025c). However, because the assay exclusively evaluates hematopoietic tissues, it cannot inform on mutagenicity at initial sites of contact or bioactivating tissues such as the liver (Chikura et al., 2019; Heflich et al., 2019; Peruzzi et al., 2010). Furthermore, the assay measures the phenotypic loss of glycosylphosphatidylinositol-anchored surface proteins rather than direct DNA sequence changes. Thus the *Pig-a* assay only captures loss-of-function mutations within the *Pig-a* gene and does not detect sequence variants that fail to alter the selectable phenotype (Dobrovolsky et al., 2010; Gollapudi et al., 2015; Heflich et al., 2019). The reliance on a single reporter gene also inherently constrains the mutational target space. The assay cannot provide genome-wide information regarding mutation density, distribution, or context-dependent effects across diverse endogenous chromatin environments (Bryce et al., 2008; Gollapudi et al., 2015; Heflich et al., 2019). Ultimately, these restrictions prevent the precise mechanistic interpretation of tissue-specific mutagenicity.

The TGR assay addresses the tissue-restriction weakness of the *Pig-a* assay by enabling mutation quantification in any tissue yielding sufficient DNA (Long et al., 2016; OECD, 2025d). However, the assay requires specialized transgenic animals carrying exogenous reporter transgenes, which increases logistical complexity compared to standard wild-type strains (Lambert et al., 2005; OECD, 2025d). Furthermore, mutant scoring relies on recovering the transgene and introducing it into a bacterial host for phenotypic quantification (Gingerich et al., 2014; Lambert et al., 2005). Consequently, this bacterial readout only captures mutations that alter the selectable phenotype, meaning sequence variants such as synonymous substitutions remain undetected (Besaratinia et al., 2018; Lambert et al., 2005). Because the exogenous reporter lacks the dynamic sequence contexts and specific transcriptional activity of the native mammalian genome, the assay cannot resolve how mutational mechanisms operate across diverse endogenous environments (Besaratinia et al.,

2018; Gingerich et al., 2014; OECD, 2025d). Additionally, TGR protocols are rigidly designed around mutation fixation kinetics. The required 28+3 or 28+28 sampling designs limit the flexibility necessary for dense time-course sampling to resolve *in vivo* mutation dynamics (Long et al., 2016; Marchetti et al., 2021; OECD, 2025d). Finally, while sequencing mutant plaques can provide mechanistic insight, it is rarely performed in routine studies and remains fundamentally constrained to the exogenous reporter locus (Beal et al., 2015; Meier et al., 2019; OECD, 2025d).

Collectively, the conventional *in vivo* test battery relies predominantly on phenotypic readouts, restricted genomic targets, and specific tissue compartments that fail to capture the full biological complexity of mutagenesis (OECD, 2016a, 2025c, 2025d). These methodological constraints prevent the detection of silent sequence variants, obscure tissue-specific metabolic responses, and fundamentally conflate independent *de novo* mutations with clonally expanded events (Beal et al., 2015; Heddle, 1999; Martincorena, 2019). While these limitations do not invalidate the use of conventional assays for hazard identification, they reveal a critical gap in the ability to inform on mechanisms of mutagenesis across the endogenous mammalian genome. Therefore, modern genetic toxicology requires complementary approaches capable of directly quantifying rare mutations in endogenous DNA, generating comprehensive mutational spectra, and supporting harmonized, clonality-aware comparisons across diverse tissues and study designs while adhering to the 3R principles.

1.6 Error-Corrected Sequencing (ECS) and Duplex Sequencing (DS) for *In Vivo* Mutagenicity Assessment

DNA sequencing served as a fundamental tool for validating *in vivo* phenotypic mutation assays. During the development of the OECD test guidelines targeted sequencing was required to confirm that the observed mutant phenotypes were driven by genotypic alterations. Beyond initial assay validation, conventional next-generation sequencing (NGS) is historically utilized as a follow-up tool to sequence selected reporter-gene mutants recovered from TGR assays (Beal et al., 2015). This targeted sequence analysis informs on mutational spectra, mutational hotspots, and corrects for clonal expansion of mutants while providing additional information on the genomic context of the model organism (Beal et al., 2015; Meier et al., 2019). Consequently, these sequence data strengthen the quantitative interpretation of mutant frequencies

and link chemical exposures to plausible modes of action. However, sequencing enriched reporter-gene plaques or phenotypically selected cells does not directly measure rare, endogenous somatic mutations within normal mammalian tissues.

A fundamental challenge of NGS is the technical error rate that exceeds the ultra-rare variant frequencies found within normal mammalian tissues. Standard NGS platforms exhibit an inherent error rate of approximately 1% that could result in false-positive calls during somatic mutation detection (Ma et al., 2019; Marchetti et al., 2023b; Salk & Kennedy, 2020; Schmitt et al., 2012). These technical errors are systematically introduced throughout sample handling, library preparation, PCR amplification, and the sequencing process itself (Costello et al., 2013; Ma et al., 2019). While basic computational filters reduce the error rate to approximately 10^{-4} to 10^{-5} per base, this threshold remains insufficient for resolving ultra-low frequency mutations (Ma et al., 2019; Schmitt et al., 2012). Specifically, the median spontaneous MF in somatic human and mouse cells is approximately 2.8×10^{-7} and 4.4×10^{-7} per base pair, respectively (Milholland et al., 2017). Because conventional suppression cannot reach this ultra-low frequency threshold, the precise detection of rare *de novo* somatic mutations requires highly sensitive, error-corrected sequencing (ECS) technologies.

ECS resolves the technical limitations of conventional NGS by creating molecular redundancy. To accurately quantify ultra-rare variants, the field has developed diverse ECS strategies that build consensus sequences from repeated observations of original DNA templates (Marchetti et al., 2023b; Schmitt et al., 2012; Yauk et al., 2026). These platforms utilize different molecular designs to generate this necessary redundancy. For instance, Duplex Sequencing (DS) utilizes exogenous barcodes and strict complementary strand matching to suppress errors (Salk & Kennedy, 2020; Schmitt et al., 2012; Valentine et al., 2020). CODEC physically links the forward and reverse strands of a DNA fragment to achieve single-duplex resolution (Bae et al., 2023). Other methods bypass exogenous barcodes entirely; SMM-seq employs circle amplification to generate multiple independent copies per fragment (Maslov et al., 2022), while PacBio HiFi sequencing relies on repeated-pass circular consensus sequencing of individual molecules (Wenger et al., 2019). Furthermore, platforms such as Hawk-Seq and PECC-Seq construct consensus sequences by matching complementary strands via precise shear-point mapping (Matsumura et al., 2019; Mulugeta et al., 2026; You et al., 2023). Regardless of the specific molecular

design, these ECS technologies successfully lower the background error rate to between 10^{-7} and 10^{-8} (Maslov et al., 2022; Matsumura et al., 2019; Salk & Kennedy, 2020; Valentine et al., 2020; Wenger et al., 2019; You et al., 2023). The reduced error rate enables the direct, sensitive quantification of mutations in mammalian tissues (Schmitt et al., 2012).

DS is currently the most applied ECS technology for mutagenicity assessment (Yauk et al., 2026). A primary advantage of DS is its inherent portability across diverse experimental systems because it does not require specialized transgenic models or phenotypic selection (Kennedy et al., 2014; Schmitt et al., 2012). This versatility allows the application of DS across a wide range of *in vivo* rodent tissues (Ashford et al., 2025; Bercu et al., 2023; Dodge et al., 2023; LeBlanc et al., 2022; Smith-Roe et al., 2023) and wild-type strains (Simon et al., 2025), generating results concordant with conventional transgenic rodent mutation assays (Ashford et al., 2025; Bercu et al., 2023). Furthermore, this portability seamlessly extends to *in vitro* cell culture systems (Armijo et al., 2023; Cho et al., 2023; Wang et al., 2021; Zhivagui et al., 2023)

In addition, DS utilizes customizable hybrid-capture panels to interrogate specific genomic loci enabling flexibility in target selection (Kennedy et al., 2014; Schmitt et al., 2012). This approach enables study designs to employ either genome-representative panels or targeted sequences tailored to particular biological and regulatory questions. For example, the standardized Mouse Mutagenesis Panel (MMP) targets twenty 2.4-kb regions distributed across genic and intergenic sequences (Valentine et al., 2020). This panel successfully quantified induced mutagenesis and identified chemical specific mutation spectra while informing on how specific genomic contexts influence mutation susceptibility (Dodge et al., 2023; LeBlanc et al., 2022). An equivalent rat genome-representative panel successfully quantified target-specific mutation frequencies and mutational spectra (Bercu et al., 2023; Smith-Roe et al., 2023). The DS custom hybrid-capture panels can directly bridge to established regulatory endpoints by targeting conventional reporter loci such as the *β*-Galactosidase *(lacZ)* transgene alongside endogenous targets to enable endpoint specific quantitative comparisons (Ashford et al., 2025). Finally, DS resolves sequence variants at the individual molecular level. This resolution allows repeated observations of identical variants to be appropriately classified as either independent *de novo* mutations or clonally derived events (Dodge et al., 2023; LeBlanc et al., 2022).

Ultimately, the application of DS provides a high analytical resolution that fundamentally improves the interpretation of *in vivo* mutagenesis.

Despite the advantages of ECS, regulatory acceptance requires further validation, possible incorporation into existing testing strategies and standardization of data processing. Within the context of this thesis, validation requires demonstrating that ECS, specifically DS, produces conclusions that are concordant with established *in vivo* mutagenicity assays. Furthermore, this validation must demonstrate that DS provides relevant mechanistic information. Concordance with conventional tests requires systematic evaluation at multiple levels. First, the basic directionality of the mutagenic response must align across analytical platforms (Valentine et al., 2020). Second, the response must demonstrate consistency across diverse tissue compartments (Long et al., 2016). Third, the quantitative dose-response features that drive interpretation, such as potency and reproducibility across study designs, must remain consistent (Ashford et al., 2025; Bercu et al., 2023; LeBlanc et al., 2022; Schmitt et al., 2012). Validation is hereby not limited to agreement with a single endpoint. It must also demonstrate that DS captures biologically plausible mutational patterns across different tissues and biological models (Marchetti et al., 2023a, 2023b; Yauk et al., 2026). Additionally, seamless integration of DS into conventional *in vivo* study designs is required to align with the 3R principle. ECS technologies need to be applicable on extracted DNA of diverse tissues collected during standard necropsy workflows of short-term and prolonged, repeat-dose toxicology studies. Finally, DS-based conclusions must remain robust and stable under the defensible mutation-counting assumptions (Yauk et al., 2026). Thus, standardized data processing needs to be established to ensure transparency and comparability of conclusions derived from ECS based mutagenicity assessment. Overall, ECS is a promising tool to address key knowledge gaps in the field of genetic toxicology and could advance modern mutagenicity assessment.

1.7 Thesis Rationale

While ECS provides the technical capacity to advance mutagenicity testing, significant biological and regulatory knowledge gaps currently restrict the comprehensive evaluation of chemical exposures. At the regulatory level, quantitative human health risk assessment frequently extrapolates mutagenic risk from short-term

in vivo assays to predict the consequences of chronic, lifelong exposures (Dankovic et al., 2015; Escher et al., 2020; Suciu et al., 2023). This mathematical extrapolation fails to capture the dynamic biological realities of prolonged exposure and to this date no mutagenicity studies have been conducted to inform on effects of lifelong exposure duration on mutagenesis. At the physiological level, the magnitude of somatic mutation accumulation is not determined by external exposure concentration alone. Tissue-specific characteristics, including metabolic competence and cellular proliferation kinetics, strictly modulate how exposure duration alters the total mutational burden and drives the clonal expansion of mutated cells (Blokzijl et al., 2016; García-Nieto et al., 2019; Li et al., 2021; Long et al., 2016). At the molecular level, empirical sequence-resolved data are urgently needed to define precise mutational signatures. Identifying specific patterns is essential to establish mechanisms of action directly from endogenous DNA (Phillips, 2018). Furthermore, sequence-level resolution is required to determine how localized genomic features dictate regional susceptibility to mutation formation (Evstyukhina et al., 2023; Wyrick & Roberts, 2015). At the technical level, validation requires the rigorous benchmarking of modern sequencing technologies against established phenotypic assays.

This thesis utilizes BbF as the central experimental anchor across short-term and prolonged exposure approaches to systematically address these knowledge gaps. BbF is classified as one of the 16 priority polycyclic aromatic hydrocarbons due to its widespread environmental prevalence and established carcinogenic potential (Keith, 2015). Despite classification, the specific mutagenic characteristics and DNA-damaging mechanisms of BbF on endogenous mammalian DNA remain severely undercharacterized. DS enables the quantification of BbF induced mutation frequencies while providing high-resolution mechanistic data regarding mutation spectra, similarities towards mutational signatures, mutation location, and the extent of clonal expansion within exposed tissues (Dodge et al., 2023; LeBlanc et al., 2022; Smith-Roe et al., 2023; Valentine et al., 2020).

The liver and BM were selected as the primary target tissues to empirically evaluate how tissue specific characteristics influence chemically induced mutagenesis. The liver represents a metabolically active, slowly proliferating organ that acts as the primary site for xenobiotic biotransformation (Moorthy et al., 2015; Pelkonen & Raunio, 1997; Smith et al., 2022; Wu et al., 2025). Conversely, BM represents a highly

proliferative hematopoietic compartment with comparatively lower xenobiotic metabolic capacity (Cartiser et al., 2011; Pelkonen & Raunio, 1997; Yeh et al., 2024). Furthermore, conventional *in vivo* genotoxicity testing, such as the micronucleus and *Pig-a* assay, relied primarily on hematopoietic compartments to establish regulatory safety and chemical hazard (OECD, 2016a, 2025c). Evaluating mutagenesis simultaneously across these two distinct biological environments explicitly defines how cellular proliferation and metabolic competence drive mutation fixation, while ensuring the resulting empirical data remain directly relevant to established regulatory paradigms.

Applying this single model chemical across short-term and prolonged exposure study designs enables the precise determination of duration-dependent effects on mutagenesis. This integrated approach allows for the simultaneous evaluation of exposure duration dynamics in both the liver and bone marrow. Specifically, it identifies potential shifts in the mechanistic mode of action and the progressive accumulation of clonally derived mutations over time. The exposure concentrations for the short-term, 28-day study design were selected in alignment with established frameworks evaluating BbF-induced DNA damage. However, the maximum exposure concentrations for the prolonged study designs were systematically reduced based on a preliminary dose-range finding study to prevent severe systemic toxicity over 180 days. Concurrently, additional lower exposure concentrations were incorporated into the extended studies to accurately characterize the lower bounds of the dose-response curve. This calibrated dosing strategy enabled the rigorous application of benchmark dose modeling to explicitly quantify time-dependent shifts in chemical mutagenic potency.

1.8 Thesis Objectives and Hypotheses

The three main objectives of this thesis are to: (1) apply DS to deeply characterize the mutagenic effects of short-term BbF oral exposure in MutaMouse liver and BM; (2) determine how prolonged exposure durations and tissue-specific characteristics dictate the accumulation of mutations, clonal expansion, and shifts in mutagenic potency; and (3) evaluate the qualitative and quantitative concordance of ECS against established conventional *in vivo* mutation assays and inform on endpoint-specific drivers of mutagenesis.

These overarching objectives are addressed in three data chapters, each with specific objectives, hypotheses, and importance as follows:

Chapter 2: Dose-Related Mutagenic and Clastogenic Effects of Benzo[*b*]fluoranthene in Mouse Somatic Tissues Detected by Duplex Sequencing and the Micronucleus Assay.

Specific objectives:

1. Quantify the dose-dependent effects of BbF on chromosome damage and mutation frequency.
2. Determine the mutation spectra in control and BbF-treated animals.
3. Characterize variations in the mutagenic response between somatic tissues and across 20 endogenous genomic regions.
4. Extract the BbF trinucleotide signature and identify similarities to the catalogue of somatic mutations in cancers (COSMIC).
5. Compare tissue- and endpoint-specific BbF mutagenic potency.

Hypothesis: BbF induces dose-dependent mutagenicity in the liver and BM, with mutation spectral shifts consistent with polycyclic aromatic hydrocarbon exposure and enhanced mutagenicity in non-genic regions.

Importance: This chapter establishes the baseline *in vivo* genotoxic profile of BbF. The findings demonstrate that DS captures highly informative mechanistic mutagenicity data under a standard 28-day study design, ensuring alignment with current regulatory testing practices.

Chapter 3: Prolonged Benzo[*b*]fluoranthene Exposure Increases Tissue-Specific Mutational Burden, Expansion of Mutant Cells, and Mutagenic Potency.

Specific objectives:

1. Quantify the dose- and time-dependent effects of BbF exposure on mutation frequency using Duplex Sequencing.
2. Assess the contribution of clonally derived mutations to the mutational burden over time.
3. Evaluate changes in mutation types and localization within 20 endogenous loci.

4. Characterize the BbF trinucleotide mutation spectrum and identify similarities to known human cancer signatures from COSMIC.
5. Apply benchmark dose modeling to evaluate time-dependent changes in chemical potency.

Hypotheses: First, prolonged exposure will increase mutational burden and chemical potency due to increased mutation induction in parallel with clonal expansion over time. Second, mutation accumulation with exposure duration will be driven by tissue-specific characteristics, including cellular proliferation rate and metabolic activity.

Importance: This chapter demonstrates that MF is a composite endpoint fundamentally shaped by tissue biology, clonal dynamics, and exposure duration. It establishes that quantitative interpretation for risk assessment is influenced by study design.

Chapter 4: Comparative *In Vivo* Mutagenicity Assessment of Benzo[*b*]fluoranthene: Concordance of Error-Corrected Sequencing and OECD Mutation Assays

Specific objectives:

1. Directly evaluate concordance between conventional *in vivo* gene mutation assays and DS approaches for BbF-induced somatic mutagenicity.
2. Determine whether these methods yield comparable quantitative potency estimates and points of departure using benchmark dose modeling.
3. Inform how sequence context affects mutagenesis.
4. Identify predominant mutation types that drive endpoint-specific mutagenicity.

Hypothesis: ECS produces qualitative and quantitative mutagenicity conclusions that are concordant with established phenotypic *in vivo* assays, while simultaneously providing mechanistic resolution including information on mutation spectra and genomic context.

Importance: This chapter validates DS against established regulatory assays. It proves that sequencing-based endpoints support similar hazard and potency

conclusions as established tests while expanding interpretive value without relying on exogenous reporter transgenes.

Collectively, this thesis addresses fundamental biological and regulatory knowledge gaps of somatic mutagenesis. The empirical data characterize the *in vivo* mutational dynamics of a priority environmental pollutant while rigorously validating DS against established OECD assays. Ultimately, this project provides the mechanistic and quantitative foundation required to advance the field of genetic toxicology and supports the adoption of ECS into modern regulatory risk assessment frameworks.

Chapter 2: Dose-Related Mutagenic and Clastogenic Effects of Benzo[*b*]fluoranthene in Mouse Somatic Tissues Detected by Duplex Sequencing and the Micronucleus Assay

Modified from: David M. Schuster, Danielle P. M. LeBlanc, Gu Zhou, Matthew J. Meier, Annette E. Dodge, Paul A. White, Alexandra S. Long, Andrew Williams, Cheryl Hobbs, Alex Diesing, Stephanie L. Smith-Roe, Jesse J. Salk, Francesco Marchetti and Carole L. Yauk. Dose-Related Mutagenic and Clastogenic Effects of Benzo[*b*]fluoranthene in Mouse Somatic Tissues Detected by Duplex Sequencing and the Micronucleus Assay. Environmental Science & Technology, 2024 Nov 27; 58, 49, 21450–21463. doi.org/10.1021/acs.est.4c07236

2.1 Abstract

Polycyclic aromatic hydrocarbons (PAHs) are common environmental pollutants that originate from the incomplete combustion of organic materials. We investigated the clastogenicity and mutagenicity of benzo[*b*]fluoranthene (BbF), one of 16 priority PAHs, in MutaMouse males after a 28-day oral exposure. BbF causes robust dose-dependent increases in micronucleus frequency in peripheral blood, indicative of chromosome damage. Duplex sequencing (DS), an error-corrected sequencing technology, reveals that BbF induces dose-dependent increases in mutation frequencies in bone marrow (BM) and liver. Mutagenicity is increased in intergenic relative to genic regions, suggesting a role for transcription-coupled repair of BbF-induced DNA damage. At higher doses, the maximum mutagenic response to BbF is higher in liver, which has a lower mitotic index but higher metabolic capacity than BM; however, mutagenic potency is comparable between the two tissues. BbF induces primarily C:G>A:T mutations, followed by C:G>T:A and C:G>G:C, indicating that BbF metabolites mainly target guanines and cytosines. The mutation spectrum of BbF correlates with cancer mutational signatures associated with tobacco exposure, supporting its contribution to the carcinogenicity of combustion-derived PAHs in humans. Overall, BbF's mutagenic effects are similar to benzo[*a*]pyrene, a well-studied mutagenic PAH. Our work showcases the utility of DS for effective mutagenicity assessment of environmental pollutants.

2.2 Introduction

Identifying and characterizing the mutagenic properties of environmental chemicals is an essential component of human health risk assessment. Chemicals can cause gene and/or chromosomal mutations that can drive genetic diseases, such as cancer (Jackson et al., 2018; Martínez-Jiménez et al., 2020). Presently, genetic toxicology studies are used for hazard identification; however, there is growing acceptance that mutations are endpoints of regulatory concern (Heflich et al., 2020). Consequently, emphasis is now being placed on quantifying mutations with higher accuracy and precision, characterizing mechanisms of action, and conducting quantitative dose–response analyses of mutagenic endpoints.

Polycyclic aromatic hydrocarbons (PAHs) are ubiquitous environmental pollutants that pose health risks because of their reactivity with DNA (Boström et al., 2002). PAHs are formed through incomplete combustion of organic materials (International Agency for Research on Cancer, 1983, 1984). Humans are exposed to PAHs through inhalation of diesel exhaust, asphalt fumes, combustion emissions such as tobacco smoke, ingestion of barbecued and smoked foods, or dermal contact. PAH exposure is linked to reproductive (Mackenzie & Murray Angevine, 1981), cardiovascular (Paigen et al., 1985), and other tissue toxicities (Cartiser et al., 2011; Long et al., 2016). Most PAHs are classified by the International Agency for Research on Cancer as probable (Group 2A) or possible (Group 2B) human carcinogens, with benzo[*a*]pyrene (BaP) being classified as carcinogenic to humans (Group 1) (International Agency for Research on Cancer, 1985, 2008, 2023). Sixteen PAHs have been prioritized by the US Environmental Protection Agency due to their prevalence in the environment and known mutagenic and carcinogenic properties (Keith, 2015).

Benzo[*b*]fluoranthene (BbF) is a priority PAH that is associated with genotoxicity, carcinogenicity, adverse hormonal changes, kidney damage, and reproductive diseases (Mass et al., 1996; Thurston et al., 2000; Tomei et al., 2006; Zhang et al., 2020). Humans are exposed to environmental BbF at concentrations ranging from 0.61 to 0.65 ng/m^3 in outdoor air, 0.20 to 0.30 ng/m^3 in indoor air, (Gustafson et al., 2008; Jung et al., 2010) and 1.66 µg/g in dust (Maertens et al., 2008). Exposure levels can vary based on location and individual lifestyle factors, such as smoking and diet. BbF is metabolized into reactive epoxides by Cytochrome P450 (CYP) enzymes (Amin et

al., 1982; Harvey, 1985; Yang et al., 2021). Previous studies have shown that BbF is mutagenic in multiple tissues (LaVole et al., 1982; Long et al., 2016); however, these studies analyzed mutagenesis in transgenic reporter genes. The effects of BbF on endogenous DNA and its mutational mechanisms are not well characterized.

Modern next-generation sequencing (NGS) technologies paired with error-correcting processes provide opportunities to comprehensively characterize mutagenesis (Marchetti et al., 2023b). Duplex Sequencing (DS) is a targeted error-corrected NGS technology that allows the detection of rare mutations (Dodge et al., 2023; LeBlanc et al., 2022; Smith-Roe et al., 2023; Valentine et al., 2020). It reduces the technical error rate of conventional NGS by uniquely tagging both strands of the original DNA fragments during library building, enabling the development of consensus sequences for each strand to eliminate polymerase chain reaction and sequencing errors (Kennedy et al., 2014; Marchetti et al., 2023b; Salk & Kennedy, 2020). Importantly, DS quantifies mutation frequencies while in parallel providing detailed information on types of mutations, location, and clonal expansion. Information on target nucleotides can be used to link chemically induced trinucleotide mutation patterns to human cancer signatures to inform potential disease relevance (Alexandrov et al., 2020; Salk & Kennedy, 2020; Valentine et al., 2020). DS has recently been used to investigate the mutagenic properties of BaP (LeBlanc et al., 2022) and other genotoxicants (Dodge et al., 2023; Smith-Roe et al., 2023), demonstrating its ability to precisely quantify chemical potency and identify mutational signatures of cancer.

Here, we used DS to study the mutagenicity of BbF in mouse liver, a metabolically active tissue with a moderate mitotic index, and bone marrow (BM), a tissue with low metabolic activity that is highly proliferative. We paired DS with the micronucleus assay to provide complementary information on the clastogenic effects of BbF in the same mice. The objectives were to (1) quantify the dose-dependent effects of BbF on chromosome damage and mutation frequency; (2) determine the mutation spectra in control and BbF-treated animals; (3) characterize variations in the mutagenic response between somatic tissues and across 20 genomic regions; (4) extract the BbF trinucleotide signature and identify similarities to the catalogue of somatic mutations in cancers (COSMIC); and (5) compare tissue- and endpoint-specific BbF mutagenic potency.

2.3 Methods

2.3.1 Animal Husbandry and Chemical Exposure

MutaMouse, Mus musculus, animals with a CD2F1 (BALB/c × DBA/2) strain background were maintained and exposed at Health Canada. The study was approved by the Health Canada Ottawa Animal Care Committee (No: 2023-011) and followed the Canadian Council on Animal Care guidelines. Eight- to 10-week-old males were exposed to doses of 0, 6.25, 12.5, 25, 50, or 100 mg/kg body weight/day (hereafter mg/kg/day) BbF (CAS 205-99-2; Sigma-Aldrich, Oakville, ON Canada) dissolved in olive oil for 28 consecutive days via oral gavage. Each dose group contained eight animals, and the three top doses (25, 50, and 100 mg/kg/day) were selected based on a previous study that investigated the genetic toxicity of nine PAHs, including BbF (Long et al., 2016). We added two doses between the vehicle control and the lowest dose from the previous study to better identify the point of departure for BbF to induce genotoxicity (Long et al., 2016).

2.3.2 Micronucleus Assay

Two days following completion of chemical administration, we collected blood (n = 8 per group) from the saphenous vein using a Microvette capillary tube coated with ethylenediaminetetraacetic acid. For the micronucleus (MN) assay, 200 µL of peripheral blood was processed according to the MicroFlowPLUS-M Kit (Litron Laboratories, Rochester, NY, USA). Briefly, blood was diluted in heparin and fixed in ultracold methanol (Sigma-Aldrich, St. Louis, MO, USA). Samples were stored at −80 °C for a minimum of 3 days before performing a buffer wash and stabilizing the fixed samples in the long-term Storage Solution of the MicroFlowPLUS-M Kit. The peripheral blood was then shipped to Integrated Laboratory Systems (ILS), an Inotiv Company (North Carolina, USA), on cold packs. Eight samples per dose group were labeled with antimouse CD71-FITC and antirodent CD61-PE antibodies and run on a BD FACSCalibur dual-laser benchtop flow cytometer.

The assay detects micronucleated reticulocytes (MN-RET) and red blood cells (MN-RBC) as per Organisation for Economic Co-operation and Development (OECD) Test Guideline 474 (OECD, 2016a). At least 20,000 ± 2000 RET and approximately 1,000,000 RBC were enumerated per sample. The frequencies of MN-RET and MN-RBC were expressed per 1000 cells, and the %RET was used as an indicator of BM

toxicity. We performed a one-way ANOVA to compare the frequency of micronuclei across multiple treatment groups in RBC and RET. Normality of the residuals was assessed using a Q–Q plot and confirmed by the Shapiro–Wilk test ($p > 0.05$), while homogeneity of variance was verified with a residuals vs predicted values plot and Levene's test for equality of variances ($p > 0.05$). Both analyses were conducted using GraphPad Prism. Afterward, we applied a Bonferroni multiple comparison test to assess pairwise differences against the vehicle control with a p value of less than 0.05 being considered statistically significant.

2.3.3 Tissue Collection and DNA Extraction

BM and liver tissues were harvested from euthanized animals 3 days after the end of exposure as per the OECD Test Guideline 488 (OECD, 2022). The liver was separated into its right posterior, right anterior, median, caudate, and left lobe. We collected BM by centrifugation of the femurs after removing the proximal and distal ends, and the cell pellet was then resuspended in phosphate-buffered saline (Corning cellgro, Manassas, VA, USA) (Amend et al., 2016). Liver and BM samples were flash frozen and stored at −80 °C until analysis.

DNA was extracted from the right anterior liver lobe and BM of four randomly selected animals per dose group. The extractions were performed using the Qiagen DNeasy blood and tissue kit protocol (Cat. #69504, Qiagen, Hilden, Germany) with modifications. Specifically, tissue disruption by bead beating the samples twice at 4000 rpm for 30 s was added, and proteinase K digestion was conducted with a decreased temperature of 37 °C and an increased incubation duration of 180 min. The extracted DNA was analyzed using the Qubit 1× dsDNA Broad Range Assay Kit (Thermo Fischer, Waltham, MA, USA) to determine DNA concentration and the Agilent TapeStation Genomic ScreenTape assay (Agilent Technologies, Santa Clara, CA, USA) to evaluate its integrity. All samples had DNA integrity numbers >7.0 and concentrations >25 ng/μL as required for DS library preparation (Valentine et al., 2020).

2.3.4 DS Library Preparation and Sequencing

Library preparations for targeted DS were performed using the TwinStrand Duplex Sequencing Mutagenesis Panel (Mouse-50) v1.0 according to the DS manual provided by TwinStrand Biosciences. Briefly, 500 ng of DNA was enzymatically

fragmented to a final size of approximately 300–400 base pairs (bp). Fragments were then end-polished, poly A-tailed, and ligated to DS adapters with unique molecular identifiers (UMIs) (TwinStrand Biosciences, Seattle, WA, USA). Quantification of the library concentrations was performed using Qubit 1× dsDNA High Sensitivity Assay Kits (Thermo Fisher, Waltham, MA, USA). The Agilent TapeStation High Sensitivity D1000 Screen Tape assay (Agilent Technologies, Santa Clara, CA, USA) was used to determine the average fragment sizes. The mouse mutagenesis panel (MMP) is designed to capture 20 endogenous loci (2.4 kb each) that are distributed across all mouse autosomes, with chromosome one containing two loci. The loci were selected by TwinStrand Biosciences to broadly represent the nucleotide composition, GC content (the percentage of guanine and cytosine bases in a DNA sequence), heterochromatin state, and transcriptional properties of the mouse genome (LeBlanc et al., 2022). Repetitive regions and pseudogenes were excluded to ensure efficient hybrid capture. Cancer driver genes or genes under strong negative selective pressure were also omitted. Of the 20 loci, nine are located within genic and 11 within intergenic regions according to the Mouse GenCode gene database version M25 (Frankish et al., 2019).

Sequencing was performed at Psomagen (Rockville, MD 20850, USA) on a NovaSeq 6000 S4 flow cell, targeting a minimum of 600 million raw reads passing filters per sample with at least 10,000-fold target locus coverage. All samples surpassed the minimum sequencing threshold; all libraries produced over 700 million informative duplex-bases, mean on-target sequencing higher than 10,000×, and sufficient peak tag families.

Sequencing data are found in the Sequence Read Archive (SRA) under SRA accession number PRJNA1131313. These FASTQ files were processed using the TwinStrand DS Mutagenesis App version 3.4.2 (TwinStrand Biosciences, Seattle, WA, USA) (Valentine et al., 2020). This application extracts the 12-nucleotide barcode sequence and removes the 5 bp adapter from the read sequence. After checking for non-nucleotide characters, the barcode sequence is parsed to create a duplex tag header. After paired-end read alignment and sorting, a single strand consensus sequence (SSCS) is produced via ConsensusMaker.py. Afterward, SSCS are sorted and analyzed via DuplexMaker.py to obtain the double strand consensus sequence (DSCS). This data processing uses UMIs to identify and eliminate technical errors and

false mutation discoveries. The error-correction reduces the error rate to less than one error per 10^{-7} to 10^{-8} nucleotides, enabling the detection of rare mutagenic events (Schmitt et al., 2012).

2.3.5 DS Data Interpretation and Statistical Analysis

We further processed the DS data to calculate the tissue-specific mutation frequency (MF). First, we identified and counted each mutation by aligning the sequencing data to the reference genome of *M. musculus* (mm10, GRCm38.p6). Since cells carrying mutations can clonally expand, tissues may contain multiple identical mutations that arose from a single mutagenic event. We first calculated mutation frequencies under the conservative assumption that identical mutations arose from the same mutagenic event. Under this assumption, all identical mutations within a sample were counted as one single mutation to calculate the minimum mutation frequency (MFmin). We also counted all mutations to obtain the maximum mutation frequency (MFmax) (Dodge et al., 2023). The variant allele fraction was set to 0.01 to exclude mosaicism and germline mutations.

MFmin and MFmax estimations by dose were calculated using R software based on generalized linear models with an overdispersed binomial error distribution. The doBy R package was used to perform pairwise comparisons and the estimates were back-transformed (Højsgaard & Halekoh, 2018). We calculated the back-transformed standard error using the delta method, and a Holm-Sidak correction was applied to adjust the p values for multiple comparisons. Binomial error distribution of generalized linear mixed models (GLMM) was used to calculate the locus-specific MFs using the “lme4” package to obtain a binomial error distribution via “glmer” function and pairwise comparisons were made using the doBy R package as described above (Højsgaard & Halekoh, 2018). We obtained the single base spectrum by direct comparison of unique substitution mutations post data processing to the mm10 reference genome. All statistical analyses were performed using the MutSeqR package (github.com/EHSRB-BSRSE-Bioinformatics/MutSeqR).

Mutation counts were summed across dose group using the SigProfiler Matrix Generator (Bergstrom et al., 2019). The dose and tissue specific 96-trinucleotide signature were determined by considering the adjacent nucleotides located upstream and downstream of single base substitutions. These trinucleotide signatures were then

reconstructed using single base substitution signatures (SBS) of the Catalogue of Somatic Mutations in Cancer (COSMIC) database with the SigProfiler Assignment tool (Alexandrov et al., 2020). The accuracy metrics for this reconstruction include the calculation of cosine similarity between the chemically induced trinucleotide pattern and the reconstructed spectrum (Díaz-Gay et al., 2023). In general, the threshold for cosine similarities is set to 0.8 with cosine similarities greater than 0.9 being considered a reliable reconstruction (Islam et al., 2022).

2.3.6. Benchmark Dose Modeling and Potency Evaluation

Benchmark dose (BMD) modeling was performed to determine the mutagenic and clastogenic potency of BbF; the BMD represents the dose at which there is a predefined change relative to background levels of mutations and MN. We applied the PROAST v70.1 web tool on the MN-RET frequencies, MN-RBC frequencies, and the MFmin values for liver and BM. The value for the Akaike Information Criterion (AIC), a measure of the goodness of fit for each model, was used to identify best fit models. A difference of 2 units was used to indicate that one model had a better fit than another following the European Food Safety Authority recommendations (EFSA Scientific Committee et al., 2022). BMD calculations were fit to the 3- and 5-parameter Hill, Exponential, Inverse Exponential, and Log-Normal models. The most probable BbF dose–response curve for the various endpoints was determined by averaging 200 bootstrap runs. The AIC was calculated throughout all bootstrap runs and unfit models were excluded from the model averaging. We used a 50% increase compared to the vehicle control as the benchmark response (BMR) (Slob, 2017; White et al., 2020; Zeller et al., 2017). The respective BMD lower and upper confidence limit values (i.e., BMDL and BMDU) were derived from 90% confidence intervals.

2.4 Results

We evaluated the effect of repeat-dose, 28-day oral exposure to BbF on clastogenicity and mutagenicity in MutaMouse males. BbF-induced chromosome damage was measured using the flow-cytometric peripheral blood MN assay. MF and spectral changes caused by BbF were quantified using DS. We note that no significant changes in body and organ weights were detected in BbF exposed animals (data not shown).

2.4.1 Dose-Dependent Increase in Micronucleus Frequency

To quantify the extent to which BbF exposure causes chromosome damage, we measured MN-RET and MN-RBC frequencies in peripheral blood collected 2 days following the 28-day exposure. BbF caused a dose-dependent increase in the frequencies of MN-RET and MN-RBC (Figure 1a, Figure 1b and Appendix D1 Table 1), with significant increases ($p < 0.05$) in all dose groups relative to olive oil controls. MN-RET frequencies per 1000 cells increased from 2.34 ± 0.1 in the control group to 8.11 ± 0.4 in the high dose group (100 mg/kg/day), and MN-RBC frequencies increased from 1.63 ± 0.1 in the control group to 4.84 ± 0.2 in the high dose group. We also observed a significant increase from 1.60 ± 0.1 to 2.17 ± 0.1 in %RET (Appendix D1 Table 1), which is indicative of BbF-induced BM toxicity (Dertinger et al., 2012). These results provide evidence of BM toxicity and a robust increase in chromosome damage caused by oral exposure to BbF.

2.4.2 BbF Induces Dose-Related Increases in Mutation Frequency Detectable within Endogenous DNA

Next, we applied DS to a panel of 20 endogenous loci to investigate the mutagenic impacts of BbF exposure. Our sequencing recovered between 2.21×10^8 to 7.06×10^8 raw reads passing filter per sample in BM, and 2.10×10^8 to 6.81×10^8 in liver (Appendix D1 Table 2). On average, there were approximately 1 billion informative duplex bases per sample in both tissues and DS reads were distributed evenly across the 20 target regions. Single nucleotide variants (SNVs), insertions and deletions (indels), and multinucleotide variants (MNVs) were quantified. An average of 36, 37, 64, 66, 98, and 131 unique mutations were identified in BM and 40, 48, 79, 88, 196, and 327 in liver for 0, 6.25, 12.5, 25, 50, and 100 mg/kg/day BbF dose groups, respectively.

Our analysis reveals that oral exposure to BbF causes robust and statistically significant dose-dependent increases in MFmin in BM and liver (Figure 2). MFmin ($\times 10^{-8}$ ± SEM) in control animals was 4.13 ± 0.45 in BM and 5.33 ± 1.01 in liver, respectively. MFmin increased to 14.78 ± 0.76 in BM and 36.27 ± 2.49 in liver at the high dose (Figure 2 and Appendix D1 Table 3). The fold change relative to controls in the top dose was greater in liver (6.8-fold increase) than in BM (3.6-fold increase). Significant increases in MFmin were found for all but the lowest dose in BM and the two lowest doses in liver.

2.4.3 Mutational Burden is not Driven by Clonal Expansion

To explore the extent of clonal expansion of cells carrying mutations, we examined the DS data in each individual animal to quantify all mutagenic events (MFmax). MFmax was consistently greater than MFmin in each animal in both tissues (Appendix D1 Table 4). In BM, there were two animals with notable amounts of clonal expansion, i.e., animals 6 and 10 in the 6.25 and 12.5 mg/kg/day dose groups, respectively (Figure 3a). Clonal expansion in these animals was driven by multiple clonally expanded mutations rather than being caused by one or two mutational events. In contrast, there were no major instances of clonal expansion events in liver (Figure 3b). Moreover, animals 6 and 10 did not have the same clonal expansion patterns in liver as were observed in BM. Finally, there was higher inter animal variability in MFmax compared to MFmin (Figure 3). These results indicate that clonal expansion events in our study were tissue-specific and relatively low in frequency.

2.4.4 Benzo[*b*]fluoranthene Induced Mutations Predominantly in Intergenic Loci

Next, we investigated whether there are locus-specific differences in mutagenicity across the MMP loci in vehicle control and BbF-treated mice. In BM, there was a statistically significant dose-dependent increase in MFmin in 10 out of 20 loci relative to vehicle controls (i.e., with MFmin per locus at the top dose being significantly higher than vehicle control) (Figure 4a). The top five most BbF-responsive loci at the highest dose are on chromosomes 14, 11, 8, 17, and 16, with MFmin of 37.50 ± 3.58, 35.18 ± 6.74, 34.73 ± 10.20, 34.70 ± 3.78, and 30.38 ± 6.63 × 10^{-8}. The lowest MFmin of 3.80 ± 0.72 × 10^{-8} occurred in the locus on chromosome 3 (Figure 4a and Appendix D1 Table 5). There was a 9.9-fold difference in MFmin at the top dose between the most and least responsive loci in BM.

In liver, all loci exhibited significant dose-dependent increases in MFmin following BbF exposure (Figure 4b). At the top dose, the top 5 most responsive loci are located on chromosomes 14, 1.2, 18, 5, and 17, with MFmin of 63.35 ± 12.50, 45.28 ± 6.47, 44.88 ± 4.12, 42.63 ± 5.68, and 42.50 ± 5.09 × 10^{-8}, respectively. There was a 3.7-fold difference in MFmin between the most and least responsive loci in liver. The locus on chromosome 13 had the lowest MFmin (17.00 ± 1.72 × 10^{-8}) (Appendix D1 Table 6). As in BM, the locus located on chromosome 14 had the highest MF. Overall, the majority of most BbF-responsive loci in both tissues are intergenic. The average MFmin was 21.8 and 40.5 × 10^{-8} in intergenic loci, and 6.7 and 31.2 × 10^{-8} in genic loci, in BM

and liver, respectively. The difference between intergenic and genic loci was significant between the two and the four highest doses in BM and liver, respectively. Overall, our results indicate that the observed susceptibility to BbF-induced mutations is locus-specific.

2.4.5 DS-based Spectrum Analysis Reveals C:G>A:T and C:G>T:A as Drivers of Benzo[*b*]fluoranthene Induced Mutagenesis

DS enables the detailed characterization of chemically induced mutation subtypes to provide insight into underlying mutagenic mechanisms. To understand the mutation spectrum induced by BbF, we analyzed SNVs, categorized into single base substitution types according to the pyrimidine reference code, indels and MNVs. C:G>T:A was the most common mutation subtype in controls, with a MFmin of 2.34 ± 0.24 × 10^{-8} in BM and 4.75 ± 0.57 × 10^{-8} in liver. BbF induced dose-dependent increases in C:G>A:T and C:G>T:A mutations compared to the vehicle control ($p < 0.01$ at the top dose for BM and liver), with highly similar spectra observed in the two tissues (Figure 5 and Appendix D1 Table 7). C:G>A:T substitutions had the highest MFmin at the top dose reaching 12.78 ± 1.41 × 10^{-8} in BM (6.1-fold above control) and 35.48 ± 5.62 × 10^{-8} in liver (29.3-fold above control). C:G>T:A MFmin increased to 7.84 ± 0.81 (3.3-fold) and 16.53 ± 1.37 × 10^{-8} (3.5-fold) in BM and liver, respectively. Overall, the BbF mutation spectrum is consistent between tissues and is mainly driven by C:G>A:T and C:G>T:A mutations.

MFmin for insertions in the vehicle controls was 0.12 ± 0.08 × 10^{-8} in BM and 0.24 ± 0.07 × 10^{-8} in liver. Insertion MFmin at the top BbF dose increased to 0.21 ± 0.07 and 0.45 ± 0.03 × 10^{-8} for BM and liver, respectively. Deletions occurred in the vehicle control with frequencies of 0.95 ± 0.17 × 10^{-8} in BM and 0.81 ± 0.27 × 10^{-8} in liver. Deletion MFmin at the top BbF dose was 1.03 ± 0.26 and 1.65 ± 0.33 × 10^{-8} for BM and liver, respectively. The detected increases in indels were not statistically significant in both tissues. The frequencies of MNVs increased from 0.06 ± 0.03 to 0.41 ± 0.04 × 10^{-8} in BM and from 0.14 ± 0.14 to 1.06 ± 0.09 × 10^{-8} in liver reaching statistical significance ($p = 0.004$;Appendix D1 Table 7).

2.4.6 Induced Mutation Spectrum is Similar to Tobacco Related Cancer Signature

We characterized the trinucleotide mutation spectrum to determine the influence of adjacent nucleotides on BbF-induced mutagenicity. The control samples were

enriched in C:G>T:A mutations with the highest frequencies observed in the ACA, CCA, and TCA trinucleotide context. BbF caused a shift in the trinucleotide spectrum driven predominantly by increases in C:G>A:T mutations at CCT, CCA, ACA, and TCT trinucleotides (Appendix D1 Figure 1). We used the SigProfiler toolset to reconstruct the trinucleotide signature of vehicle controls and BbF-treated tissues. COSMIC signature SBS5 was enriched in control BM (25%) and liver (36%) (Appendix D1 Figure 2); SBS5 is associated with aging and is found in most cancers and in normal cells. Exposure to BbF shifted the trinucleotide mutation pattern in BM to be associated with SBS29 (32%), which originates from tobacco chewing, and SBS5 (19%). The reconstructed signature had a cosine similarity of 0.94 with the original signature (Figure 6a). The enriched signatures in liver at the top dose were SBS4 (75%), which is associated with tobacco smoking, and SBS5 (14%). The liver reconstructed signature had a cosine similarity of 0.96 with the original signature (Figure 6b). SBS4 showed a dose-dependent increase in liver mutation profiles; no other SBS showed dose-dependent increases. Our data demonstrate dose-dependent changes in the trinucleotide mutation pattern with enrichment of cancer signatures SBS4 and SBS29.

2.4.7 Potency Estimates of Benzo[*b*]fluoranthene Overlap Between Bone Marrow and Liver

We employed BMD modeling to compare the potency of BbF in inducing clastogenicity and mutagenicity across endpoints and tissues. The BMDs are presented as confidence intervals (BMDL–BMDU plotted in order of increasing BMDs) for comparison (Figure 7). The BMD confidence interval for MFmin in BM overlapped with all other confidence intervals, indicating comparable BbF potencies between the endpoints. However, the BMD confidence intervals for MN-RET and MN-RBC were overlapping but were both lower than the confidence interval for MFmin in liver (Figure 7 and Appendix D1 Table 8). Thus, a 50% increase in clastogenicity in RET and RBC occurs at lower doses than a 50% increase in liver MFmin. Confidence interval ranking from least sensitive to most sensitive endpoints/tissue follows the pattern: MFmin in liver (10.1 to 18.4 mg/kg/day), followed by MFmin in BM (3.6 to 12.8 mg/kg/day), then MN frequency in RBCs (2.2 to 9.3 mg/kg/day), and finally MN frequency in RET (0.4 to 4.1 mg/kg/day). Our results show that the BMD confidence intervals were generally within a comparable range, indicating that there are no major tissue- or endpoints-specific differences in mutagenic potencies of BbF.

2.5 Discussion

We investigated the effects of subchronic oral exposure of MutaMouse males to BbF on mutagenicity and clastogenicity. As expected, our results show that BbF causes dose-dependent increases in MF and MN. The magnitude of the mutagenic response was larger in liver than in BM at the highest dose tested; however, dose–response modeling to estimate the BMD50 for MFmin reveals overlapping confidence intervals (i.e., similar potencies) for the two tissues. Greater increases in MF due to BbF exposure were observed in intergenic than genic loci, consistent with transcription-coupled repair of bulky adducts. BbF induces mostly C:G>A:T mutations with a spectrum similar to other PAHs such as BaP (LeBlanc et al., 2022). Mutational signature analysis following this short-term exposure shows a correlation between BbF-induced mutations in both liver and BM and signatures observed in human cancers associated with tobacco exposure. The data support that BbF contributes to the carcinogenicity of PAH mixtures in humans and underscores the human health significance of exposure to this PAH. Our work also showcases the high value of DS in deeply characterizing the mutagenic effects of exposure to environmental genotoxicants.

We observed robust increases in chromosome damage in blood cells and mutagenic responses in liver and BM. Our MN analysis extends the findings of Long et al. (Long et al., 2016) to demonstrate significant clastogenic effects at lower BbF doses. Long et al. also showed a higher mutagenic response in liver *lacZ* reporter gene mutations than in BM, which is consistent with our findings; however, our analyses quantify the mutagenicity of BbF to endogenous loci rather than a single exogenous bacterial reporter gene, increasing the biological relevance of the findings to human health. Indeed, our data demonstrate large differences in BbF response across the 20 loci in the MMP, supporting the need to look beyond individual genes. The liver is more metabolically active, proliferates more slowly, and has a lower DNA replication rate than BM. Although more reactive metabolites are created in the liver than in BM, DNA damage has more time to be repaired prior to replication and has less opportunity to be converted into mutations during the exposure period (Dertinger et al., 2012). At the same time, despite the lower metabolic activity, the high proliferation rate makes the hematopoietic system susceptible to MN formation. Our results indicate systemic distribution and strong clastogenic and mutagenic effects after oral exposure to BbF

that reflect the metabolic capacity and proliferation rate of the two tissues (Pedone et al., 2017).

To further compare the response between the two tissues and endpoints, we used BMD modeling to identify the doses at which BbF exposure elicits a 50% increase in MN or MF. The BMD confidence intervals for MN-RETs and MN-RBCs, and mutations in BM, overlapped. The only notable difference was that MF in liver was slightly lower than that for MN-RETs. The magnitude of the mutagenic response was larger in liver than in BM at the highest dose tested likely due to more extensive xenobiotic metabolism in the liver; however, dose–response modeling to estimate the BMD_{50} for MFmin reveals overlapping confidence intervals (i.e., similar potencies) for the two tissues. Thus, we observe no major differences in potencies for these tissues/endpoints. Interestingly, the mutagenicity potency of BbF in BM is comparable to the mutagenic potency of BaP in BM assessed using the MMP (LeBlanc et al., 2022). The confidence intervals span 3.6 to 12.8 mg/kg/day and 4.5 to 7.5 mg/kg/day for BbF and BaP, respectively. Overall, our results indicate that BbF is a potent genotoxicant in both hematopoietic and liver tissues and that its mutagenic potency in BM is comparable to that of BaP, which is the frame of reference when evaluating PAHs.

Similarly to previous studies, (Dodge et al., 2023; LeBlanc et al., 2022; Valentine et al., 2020) our analyses revealed major differences in response to BbF across the 20 loci of the mouse mutagenesis panel, with MFmin being higher in intergenic than genic loci. Two of these studies applied the same MMP used in our study. The top responding loci are nearly identical across these studies, which analyzed chemicals with different mutagenic mechanisms. LeBlanc et al. (LeBlanc et al., 2022) studied the BM of MutaMouse males exposed for 28 days to BaP. BbF and BaP, especially their hydrodiol and quinone metabolites, cause bulky DNA adducts that target guanines or cytosines and are mainly repaired by nucleotide excision repair (NER); thus, consistencies across these exposures may be expected (Amin et al., 1982; Verma et al., 2012). In contrast, Dodge et al. (Dodge et al., 2023) measured mutations following exposure to procarbazine, which causes DNA alkylation that is restored by BER (Błaszczyk & Mielżyńska-Švach, 2017; Brambilla et al., 1978). All three chemicals also cause an increase in reactive oxygen species that can oxidize DNA, which is then repaired by BER (Chakraborty et al., 2021; Wilk et al., 2013). While NER and BER can both be coupled to active transcription, PAH adducts are mainly repaired by TC-NER

(Chakraborty et al., 2021; Zhong et al., 2010). Our study is consistent with the hypothesis that higher induced MFs in intergenic regions might be driven by the lack of TCR activity in these loci (Dodge et al., 2023; LeBlanc et al., 2022). In contrast, the same genic versus intergenic trend has not been observed in DS experiments with Sprague–Dawley rats. These studies administered N-nitrosodiethylamine and N-ethyl-N-nitrosourea, which are common alkylating agents (Bercu et al., 2023; Smith-Roe et al., 2023). We note that DNA repair mechanisms play a crucial role in defining the mutagenic potency of chemicals and that the local and specific DNA repair capacity for the DNA lesions created by the toxicant must be inadequate to induce detectable increases in MF (Cho et al., 2023; Long et al., 2018). More research across different model systems and chemicals with different mechanisms of action will be required to effectively determine the underlying drivers of these intriguing locus-specific differences in mutagenic responses.

One of the powerful attributes of DS is its ability to directly quantify the extent of clonal expansion of any mutation. Thus, for each locus we compared MFmin and MFmax. MFmin is highly informative for characterizing the potency and mutational signature of a mutagen. MFmax has physiological importance for the organism since mutated cells that clonally expand statistically increase the likelihood of multiple clonal events within a cluster of cells carrying the same mutation that may render a cell population more prone to tumor formation(Greaves & Maley, 2012; Martincorena, 2019). However, we note that the MMP loci were chosen in part because they are not thought to be under positive or negative selective pressure; thus, clonal expansion should be a random event. Herein, we observed no major differences between MFmin and MFmax in most of the samples. A key factor for clonal expansion of mutations is time (Heddle, 1999). The 28-day exposure duration in adult mice may be too short for notable clonal expansion to occur within the observed loci. With increasing exposure duration or sampling time, we may expect to see an increase in clonal expansion especially in rapidly proliferating tissues like BM. Overall, the low levels of clonal expansion of BbF-induced mutations does not significantly affect MF in the tissues studied, but differentiation between unique and clonal mutations is valuable for understanding the mutational burden in each tissue.

The application of DS also enabled an in-depth analysis of the mutation spectrum induced by BbF to provide insight into its mechanism of action. In line with

expectations, and similarly to BaP, BbF causes C:G>A:T and C:G>T:A mutations. Although the specific nucleophilic sites of BbF adduct formation are not known (Long et al., 2016), previous studies have identified the amino group of guanines as the most likely target for PAH-DNA adduct formation through reacting with diol-epoxides or hydroquinones (Mass et al., 1996; Schreck et al., 2009). The metabolism of BbF is expected to create these highly reactive intermediates that can engage in oxidation–reduction, nucleophilic addition, and electrophilic aromatic substitution reactions (Patel et al., 2020). The liver is the main xenobiotic metabolizing tissue and has high CYP levels, with CYP1A1 and CYP1B1 isoforms being key enzymes for PAH metabolism (Sadler et al., 2016; Shiizaki et al., 2017). However, due to the high instability of BbF metabolites, it is unlikely that they are transported from the liver to the BM unless they are stabilized within the blood serum(Ginsberg & Atherholt, 1989). Thus, we postulate that BbF requires metabolism in the BM through local CYP enzymes such as the highly abundant CYP1B1 (Galvan et al., 2005). Overall, our results support that both tissues metabolize BbF. The consistencies in mutation spectrum in both tissues and similarities to BaP strongly suggest that BbF metabolites predominantly form cytosine or guanine adducts and that BbF is systemically distributed.

The trinucleotide mutational spectrum provides deeper insights into underlying biological processes associated with mutagenesis; it can be used to explore similarities with human cancer signatures. We observed tissue-specific differences in the trinucleotide patterns in exposed samples. Both tissues showed dose-dependent increases in C:G>A:T mutations, but the upstream adjacent nucleotides differed. In BM, mutations were found predominantly in a GCN>GAN context; whereas, in liver they were predominantly in the CCN>CAN context. Signature reconstruction indicated that SBS29, associated with oral tobacco exposure, was the strongest contributing signature in BM; while in liver, SBS4, which is associated with tobacco smoking and is also associated with the DS mutational spectrum of BaP (LeBlanc et al., 2022), was the highest contributing signature. SBS29 is enriched in mutations in a GCA>GAA, GCC>GAC, or GCT>GAT context, whereas SBS4 consists of CCA>CAA, CCC>CAC, and CCT>CAT variants. The respective etiologies of these signatures align with common routes of BbF exposure for humans since chewing tobacco (Stepanov et al., 2010) and tobacco smoke (Vu et al., 2015) contain high levels of BbF. Both signatures are mainly driven through mutations in intergenic regions. The detection of these

cancer-related signatures after 28 days of BbF exposure emphasizes the relevance of BbF to human PAH-induced carcinogenesis. Interestingly, we were able to detect SBS4, associated with the inhalation of PAHs in human cancers, in our mice exposed to BbF by oral gavage. It is possible that chemically induced mutagenicity in diverse tissues varies according to the applied route of exposure based on first-pass metabolism. However, regardless of route, the main type of DNA adducts caused by PAH exposure are generally expected to be consistent with CYP metabolism to diol-epoxides as supported by this finding (Amin et al., 1982, 1985; Yang, 1988).

BbF administration over the time course of 28 days for this project aligns with OECD Test Guideline 488 (OECD, 2022), which is considered the “gold standard” for regulatory mutagenicity assessment. Our study confirms that this short-term exposure is also sufficient to produce a robust DS dose–response in endogenous mouse loci in both of the tissues studied. This study and others support that DS mutagenicity assessment can be integrated with MN frequency assessment in 28-day studies to comprehensively characterize chromosome damage and mutagenic responses *in vivo* (LeBlanc et al., 2022; OECD, 2016a; Valentine et al., 2020). Our results also suggest that DS could be integrated with other conventional toxicology studies such as the OECD Test Guideline 407 (OECD, 2025a). Such integration would enable an evaluation of mutagenicity and clastogenicity within the standard 28-day repeated dose studies in rodents to significantly reduce animal use in toxicity testing, aligning with the 3Rs principles (Kroeger, 2006; Russell & Burch, 1959).

In conclusion, our study provides detailed insight into the mutagenic and clastogenic characteristics of BbF and its underlying mechanisms of action in the mammalian genome. In addition to dose-dependent increases in MN and MF in tissues with different metabolic activity and mitotic indices, we showed that C:G>A:T mutations are the drivers for BbF’s mutational signature. We observed enrichment of mutational signatures found in human PAH-associated cancers in both tissues after a 28-day BbF exposure. This finding underscores the significant role of BbF in environmental carcinogenesis. Finally, our findings highlight the high value of DS in mutagenicity assessment of chemical exposures and its potential utility for regulatory testing.

2.6 Acknowledgements

C.L.Y., F.M., and P.A.W. hold an Innovation in Regulatory Science Award from the Burroughs Wellcome Fund (#1021737). C.L.Y. is grateful for funding from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2021-02806) with infrastructure support through the Canadian Foundation for Innovation John Evans Leaders Fund (#233109), the Ontario Research Fund (#40569), and the University of Ottawa. C.L.Y. acknowledges that this project was conducted, in part, thanks to funding provided through the Canada Research Chairs Program (CRC-2020-00060). Work at Health Canada was supported by the Chemicals Management Plan to F.M. F.M. and M.M. acknowledge funding from Health Canada's Genomics Research and Development Initiative. S.S.R. was funded by the intramural research program of the National Institutes of Environmental Health Sciences, National Institutes of Health (ES103378-01) and genetic toxicity testing conducted by Integrated Laboratory Systems, an Inotiv company, under NIEHS contract 75N96020C00001. We thank Dr. Scott S. Auerbach at the Division of Translational Toxicology, NIEHS, and Marc Beal from Health Canada for their helpful review of the manuscript.

2.7 Figures and Tables

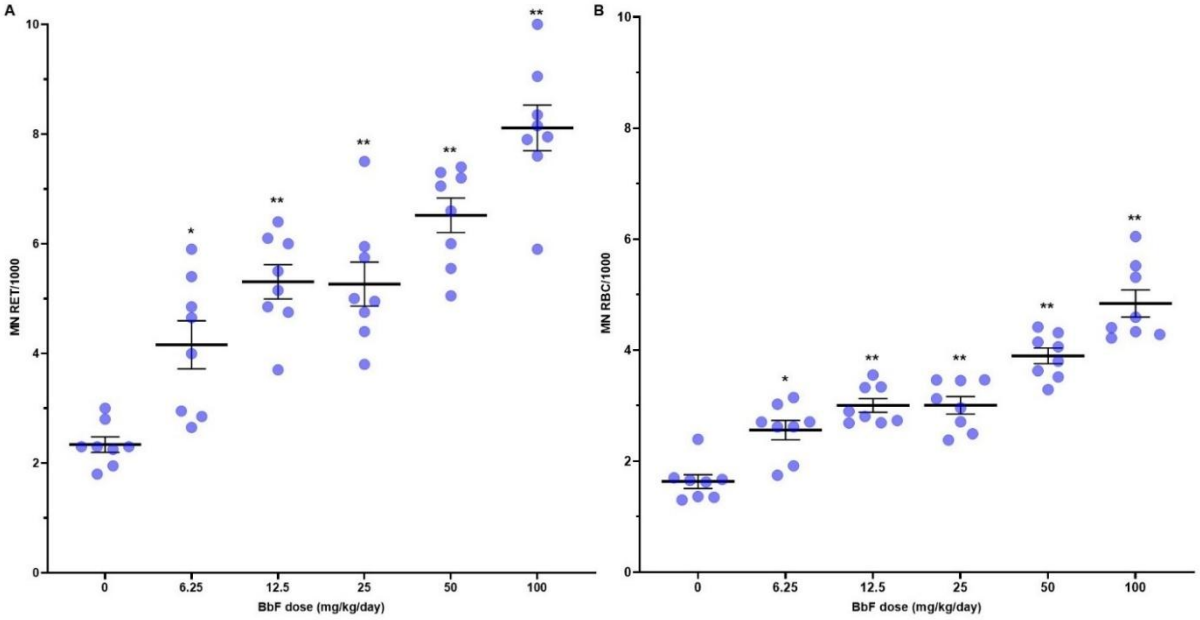


*Figure 1: Frequencies of induced micronuclei in peripheral blood of MutaMouse males after vehicle or BbF exposure for 28 days. Mean micronucleus (MN) frequencies (black bar) with indication of standard errors of the mean (SEM). MN frequencies per 1000 cells are shown for both reticulocytes (MN RET) (A) and red blood cells (MN–RBC) (B). Dots represent individual animal data. * p < 0.05 and ** p < 0.01 relative to vehicle control.*

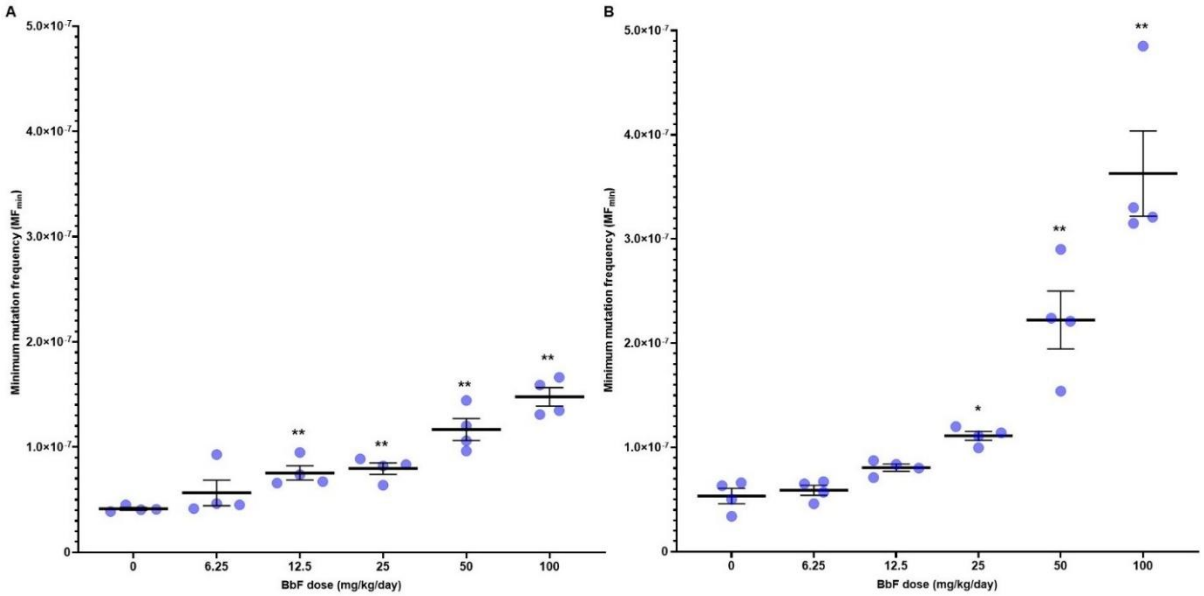


*Figure 2: Mean minimum mutation frequency (MFmin ± SEM) per bp measured by Duplex Sequencing in mouse bone marrow (A) and liver (B) after exposure to increasing doses of BbF (n = 4 per dose group). Asterisks indicate * p < 0.05 and ** p < 0.01 relative to vehicle control. Dots represent individual animal MFmin.*

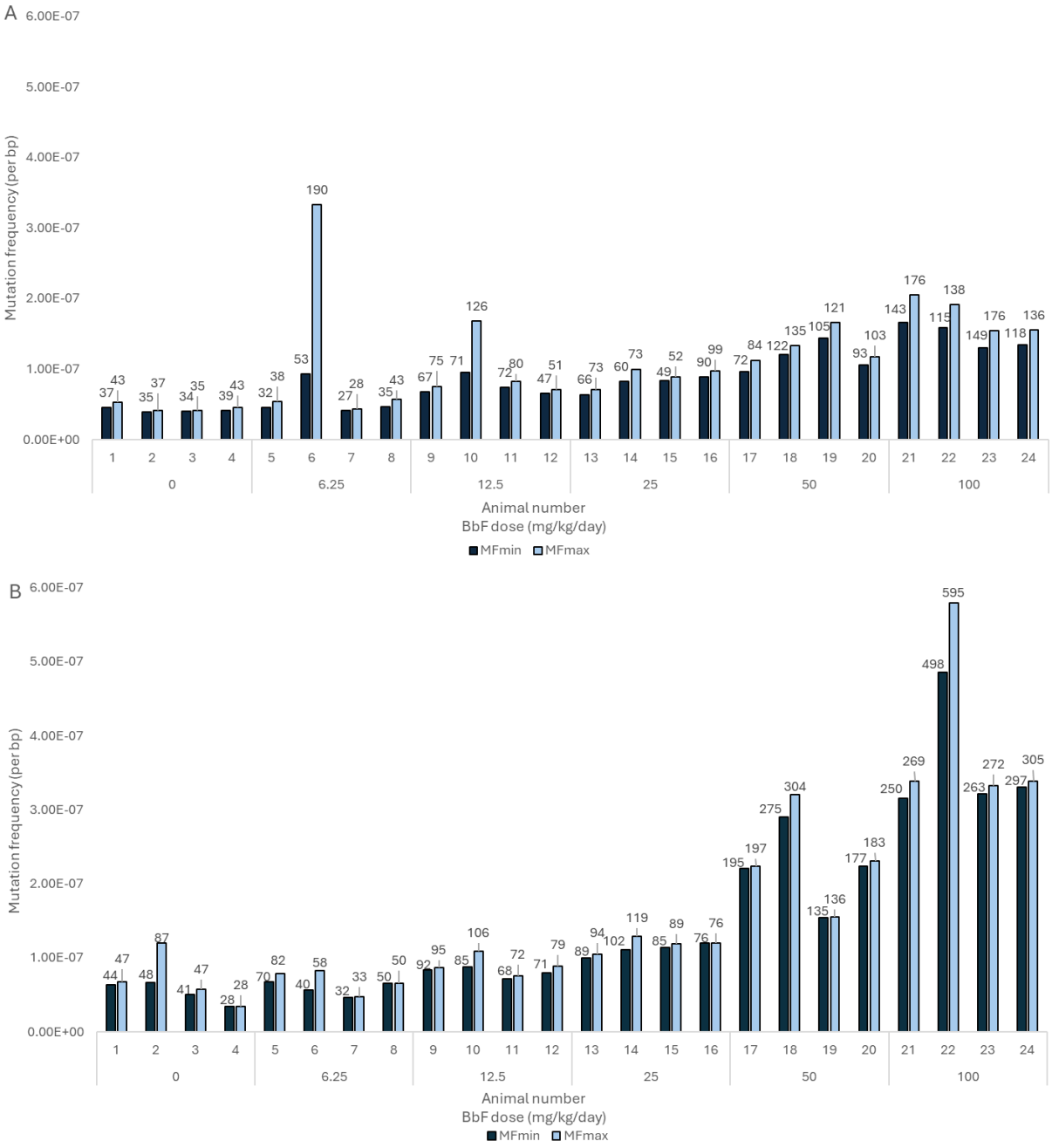


Figure 3: Mean mutation frequency per bp measured by Duplex Sequencing in bone marrow (A) and liver (B) for individual mice within each BbF dose group. The mutation frequency of unique mutations is indicated as MFmin (black). The mutation frequency including clonally expanded mutations is shown as MFmax (light blue). Numbers above the bars represent the observed number of mutations.

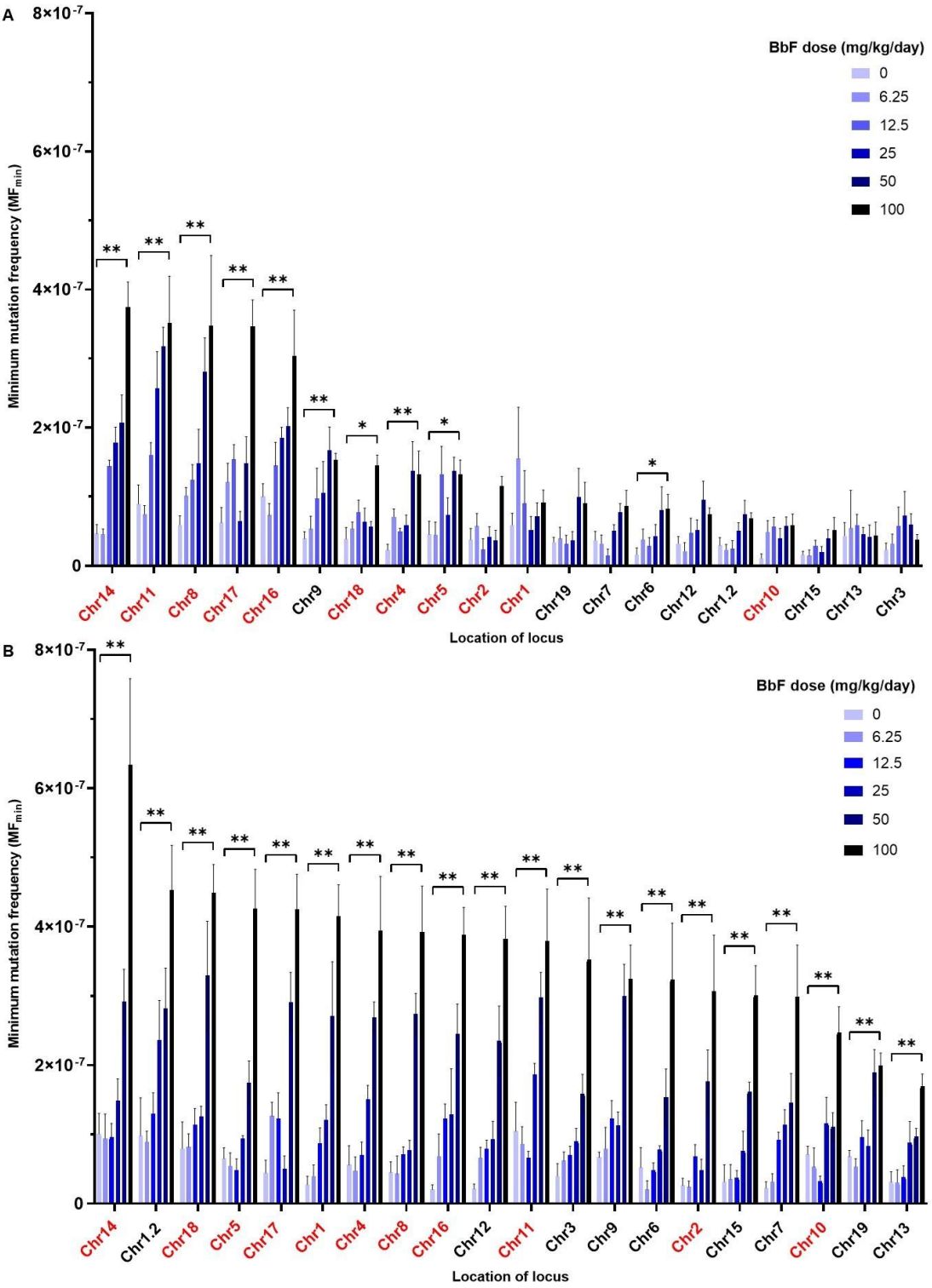


*Figure 4: Mean minimum mutation frequency (MFmin) ± SEM per locus for BbF–exposed mouse bone marrow (A) and liver (B) alongside vehicle controls. Loci are sorted from the highest top dose response to the lowest (left to right). Intergenic and genic loci are show in red and black, respectively. Statistical testing was done for the top dose with respect to the vehicle control using a generalized linear mixed model; * $p < 0.05$ and ** $p < 0.01$; $n = 4$ per dose group.*

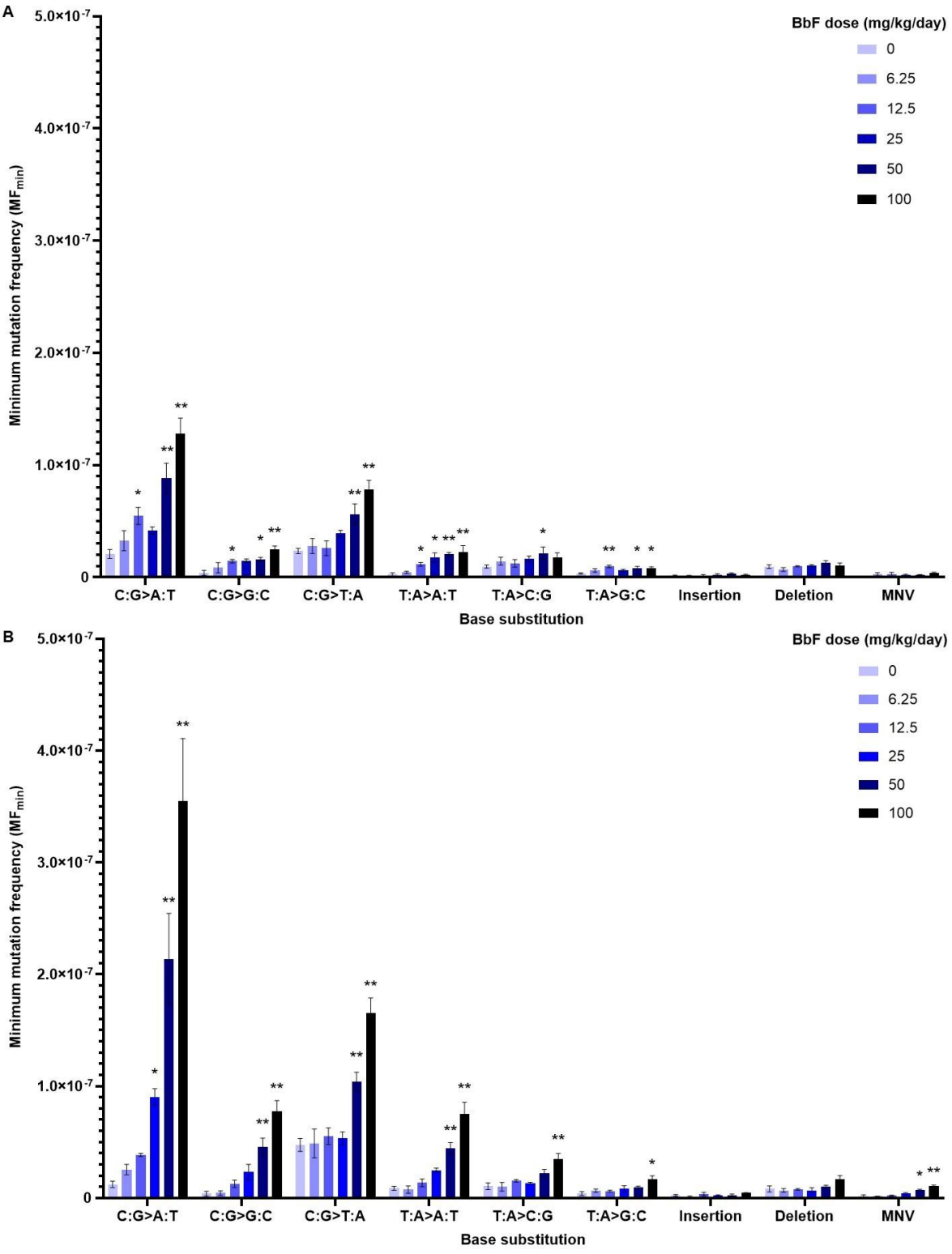


*Figure 5: Mutation spectrum for vehicle controls and BbF dose groups observed in mouse bone marrow (A) and liver (B). Mutation subtypes, insertions, deletions and multi–nucleotide variants (MNV) are presented as mean MFmin ± SEM. Mutation subtypes are based on the six single nucleotide variants using a pyrimidine reference. Statistical testing was done using a generalized linear mixed model; * $p < 0.05$ and ** $p < 0.01$; $n = 4$ per dose group.*

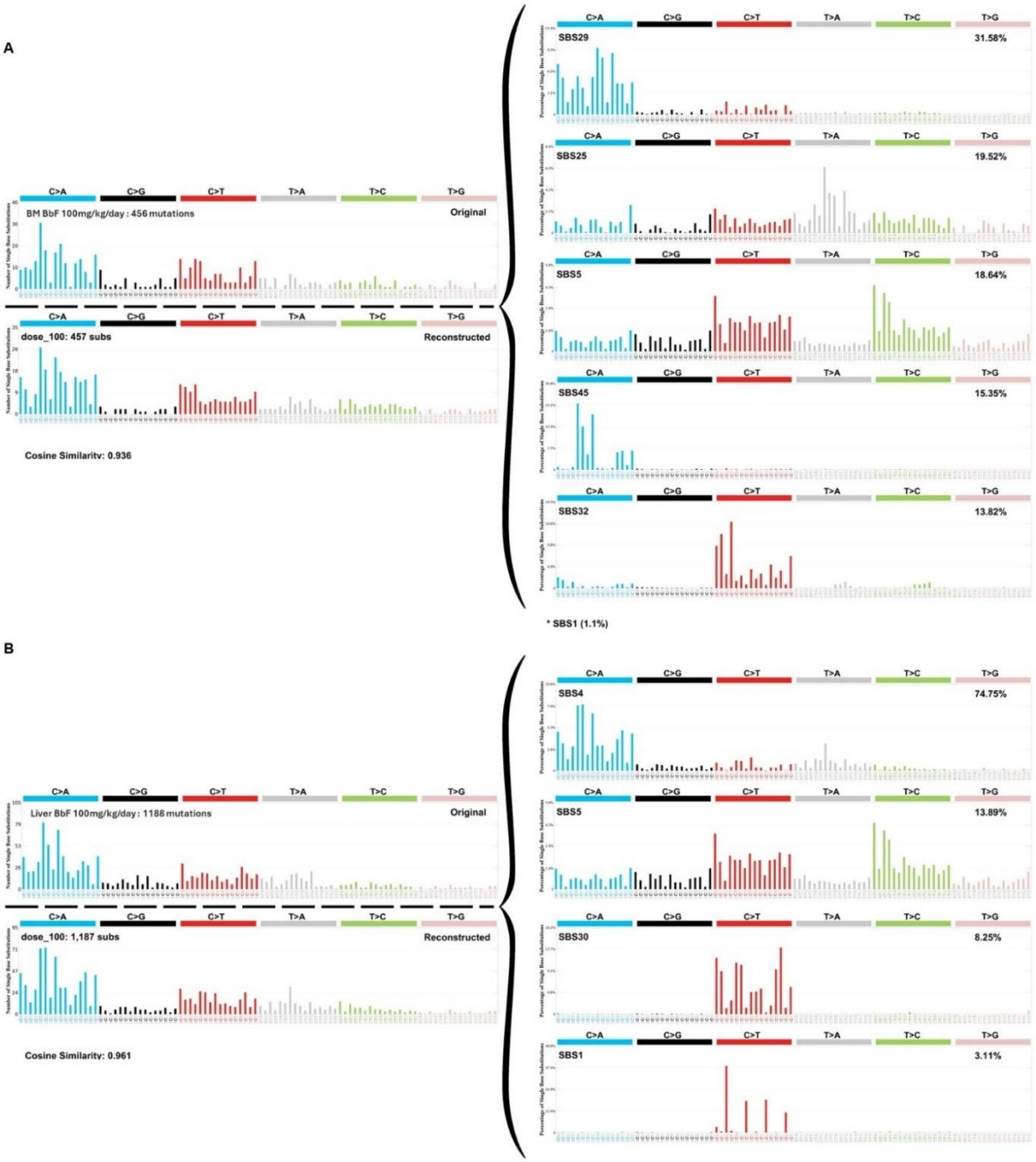


Figure 6: Mutation signature analyses for top BbF dose (100 mg/kg/day) in bone marrow (A) and liver (B). The original trinucleotide mutation profile is shown on the top left. SigProfilerAssignment used the single base substitution (SBS) signatures of the Catalogue of Somatic Mutations in Cancer (COSMIC) database to reconstruct the signatures. The SBS signatures and their relative contributions are shown on the right. The reconstructed mutation pattern, number of substitutions used for the assignment, and cosine similarity between reconstructed and observed trinucleotide mutation profile are shown on left below the original profile. The total number of mutations in the original signature are indicated on the left.

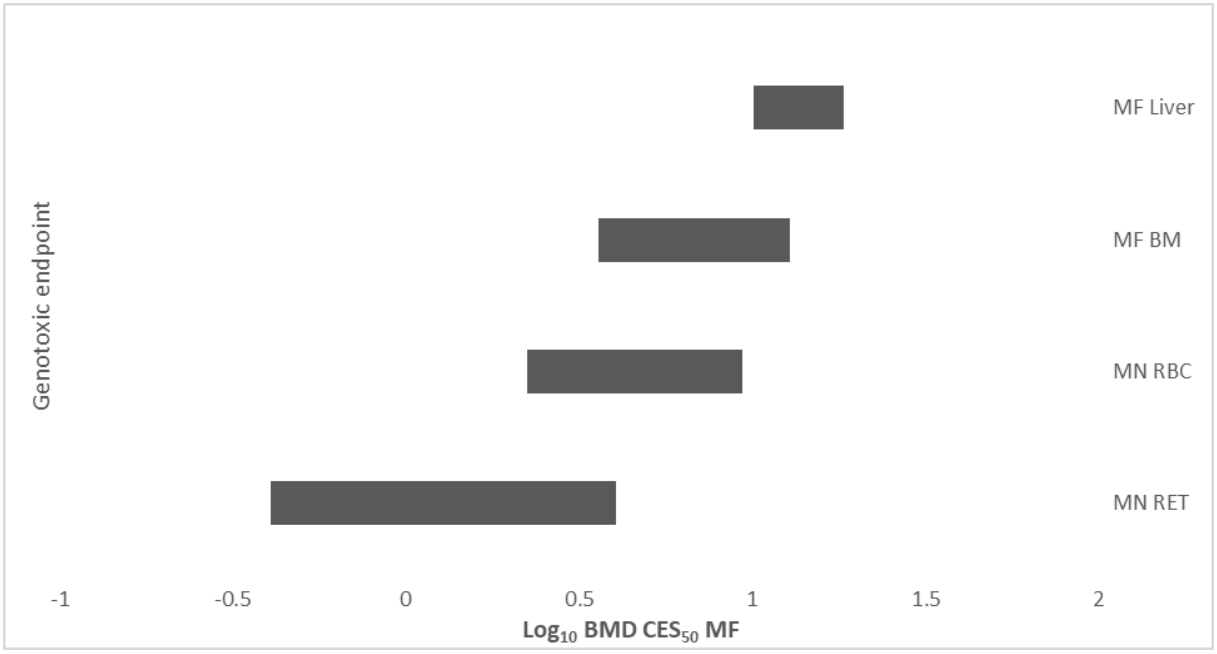


Figure 7: Log_{10} benchmark dose (BMD) for a 50% increase in micronucleus (MN) or mutation frequency (MFmin). 90% confidence intervals based on BMD model averaging analyses using PROAST are shown. Confidence intervals are organized from lowest (top) to highest (bottom) potency. Labels indicate micronuclei in reticulocytes (MN–RET) and red blood cells (MN–RBC), and mutation frequency in liver (MF liver) and bone marrow (MF BM).

Chapter 3: Prolonged Benzo[*b*]fluoranthene Exposure Increases Tissue-Specific Mutational Burden, Expansion of Mutant Cells, and Mutagenic Potency

Modified from: David M. Schuster, Danielle P. M. LeBlanc, Gu Zhou, Matthew J. Meier, Annette E. Dodge, Paul A. White, Andreas Zeller, Andrew Williams, David Eickmeyer, Sabrina Kehm, Juergen Funk, Francesco Marchetti and Carole L. Yauk. Prolonged Benzo[*b*]fluoranthene Exposure Increases Tissue-Specific Mutational Burden, Expansion of Mutant Cells, and Mutagenic Potency. In review at Environment International.

3.1 Abstract

Chronic exposure to polycyclic aromatic hydrocarbons (PAHs) is associated with increased risk of cancer through a mutagenic mode of action. We investigated how exposure duration influences mutagenicity in bone marrow (BM) and liver of MutaMouse males orally exposed to increasing doses of benzo[*b*]fluoranthene (BbF), a priority PAH, for 28, 90, or 180 days. We applied Duplex Sequencing (DS) across 20 endogenous loci to assess dose- and time-dependent effects on mutation frequency, clonal expansion of mutant cells, mutation distribution, and spectra. Mutation frequencies increased with BbF dose and exposure duration in both tissues. After 90 and 180 days of exposure to 25 mg/kg/day BbF, mutation frequency was 1.7x and 2.1x higher in BM, and 3.8x and 8.6x higher in liver, than 28-day exposures. Clonally derived mutations increased significantly with exposure duration in BM but not liver. Intergenic loci incurred more mutations than genic regions in both tissues across all exposure durations. Mutation spectra were dominated by C:G>A:T transversions and was enriched in SBS4, the mutation signature associated with tobacco-induced human lung cancer. Benchmark dose (BMD50) modeling showed increasing BbF potency with longer exposures. BMD50 decreased from 8.2 (28d) to 3.7 (90d) and 3.0 (180d) mg/kg/day in BM; in liver, BMD50 shifted from 14.3 (28d) to 4.3 (90d) and 3.9 (180d) mg/kg/day. The observed tissue-specific responses may be attributed to differences in proliferation rate, metabolic activity, and cellular lifespan. Our findings provide insights into the mutagenic impacts of prolonged PAH exposure and highlight tissue-specific differences in mutation susceptibility.

3.2 Introduction

Chronic exposure to diverse environmental pollutants raises concerns about cumulative health consequences and their associated economic impact (Hou et al., 2012; R. Singh et al., 2007; Sørensen et al., 2003). Air pollution alone is associated with over 15,300 premature deaths and $120 billion in health care costs per year in Canada (Health Canada, 2021). Despite the key role of environmental mutagenicity in human diseases such as cancer and heritable genetic disorders, our understanding of how prolonged exposure to genotoxicants influences mutational burden and landscape over time remains limited. Understanding how mutations accumulate over sustained exposures is crucial to inform risk assessment and disease prevention strategies.

Among the diverse array of environmental pollutants, polycyclic aromatic hydrocarbons (PAHs) are of particular interest due to their pervasive presence and well-documented genotoxic and carcinogenic properties (Boström et al., 2002; IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, 2010; International Agency for Research on Cancer, 1985, 2023; Long et al., 2016). Most PAHs originate from the incomplete combustion of organic materials and are found in vehicle exhaust, industrial emissions, wildfire emissions, tobacco smoke, and char-broiled foods (Boström et al., 2002; IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, 2010). Consequently, humans are chronically exposed to PAHs predominantly via ingestion and inhalation (Luo et al., 2021; Ramesh et al., 2004). Current *in vivo* genetic toxicology test guidelines have been instrumental in identifying PAH hazards; however, these tests primarily rely on acute and subchronic exposures (OECD, 2016a, 2025c, 2025d, 2025a). Thus, the dynamics of PAH-induced mutagenesis over extended exposure durations are not well characterized, limiting our understanding of the health risks associated with chronic PAH exposure.

Benzo[b]fluoranthene (BbF) is one of 16 priority PAHs identified by the U.S. Environmental Protection Agency because of its DNA-damaging, mutagenic, and carcinogenic potential (Keith & Telliard, 1979; Mass et al., 1996). Structurally similar to other prominent PAHs, such as benzo[*a*]pyrene, BbF is classified by the International Agency for Research on Cancer (IARC) as a Group 2B possible human carcinogen (IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, 2010; International Agency for Research on Cancer, 1985, 2023). BbF's lipophilicity and

physicochemical properties enable its systemic distribution and preferential partitioning into lipid rich tissues (Pastor-Belda et al., 2019). Toxicokinetic studies indicate that BbF undergoes biotransformation by Cytochrome P450 (CYP) enzymes, leading to the formation of reactive intermediates, including dihydrodiolepoxides and hydroquinones, that can induce DNA damage (Alexandrov et al., 2020; Harrigan, 2004; Liu et al., 2025; Yang et al., 2018; Yang et al., 2021). Our previous work demonstrated that oral exposure to BbF induces robust, dose-dependent increases in chromosome damage and mutation frequency (MF) in mouse somatic tissues after 28 days of exposure (Long et al., 2016; Schuster et al., 2024). However, it has not been determined if the mutation frequencies observed following standard 28 days of dosing are predictive of the mutagenesis induced by prolonged BbF exposure.

We hypothesize that prolonged exposure to BbF increases the mutational burden without changing the mutagenic mode of action. This assumption is based on the increased likelihood that induced DNA damage escapes repair mechanisms, resulting in progressive mutation accumulation and increased mutagenic potency over time (Cosentino & Heddle, 1999; Hashimoto et al., 2007; Krolewski et al., 1988). Moreover, we posit that tissue-specific biological characteristics such as proliferation rate, metabolic activity, and DNA repair capacity are key drivers that modulate MFs and expansion of mutant cells, contributing to tissue-specific responses (Hanahan & Weinberg, 2011; Schneider et al., 2017; Tomasetti & Vogelstein, 2015).

To test these hypotheses, we used Duplex Sequencing (DS), a sensitive error-corrected sequencing technology (LeBlanc et al., 2022; Schmitt et al., 2012), to evaluate the mutagenic impact and potency of BbF exposure in MutaMouse bone marrow (BM) and liver over increasing exposure durations (i.e., 28, 90, and 180 days). Our primary objectives were to: (1) quantify the dose- and time-dependent effects of BbF exposure on MF using DS; (2) assess the contribution of clonally derived mutations on the mutational burden over time; (3) evaluate changes in mutation types and localization within the 20 endogenous loci; (4) characterize the BbF trinucleotide mutation spectrum and identify similarities to known human cancer signatures from the Catalogue of Somatic Mutations in Cancer (COSMIC); and (5) apply benchmark dose (BMD) modeling to evaluate time-dependent changes in chemical potency. Together, these analyses provide critical insights into how exposure to a ubiquitous PAH reshapes the somatic mutational landscape in a tissue-specific manner over time.

3.3 Methods

3.3.1 Animal Exposure and Tissue Collection

MutaMouse animals were maintained and exposed at Health Canada. The Health Canada Ottawa Animal Care Committee approved the study (No: 2023-011), and all procedures adhered to the Canadian Council on Animal Care guidelines. Ten to fourteen-week-old males received daily oral gavage doses of BbF (CAS 205-99-2; Sigma-Aldrich, Oakville, ON, Canada) dissolved in olive oil. Concurrent vehicle controls received olive oil only. We exposed 48 animals with eight animals per dose group for each exposure duration resulting in a total of 144 mice for this study. The top dose for the 28-day exposure duration was chosen as 100 mg/kg body weight/day (hereafter mg/kg/day) based on a previous study (Long et al., 2016), while a dose-range finding study was conducted over 180 days to identify sustainable top doses for 90- and 180-day exposure durations. We used five dose groups alongside vehicle controls for each exposure duration. These dose groups were:

28 days of exposure: 0, 6.25, 12.5, 25, 50, or 100 mg/kg/day.

90 days of exposure: 0, 3.125, 6.25, 12.5, 25 or 50 mg/kg/day.

180 days of exposure: 0, 1.56, 3.125, 6.25, 12.5 or 25 mg/kg/day.

BM and livers were harvested from euthanized animals 3 days after the end of the 28 days of exposure, following OECD Test Guideline 488 (OECD, 2025d) and within 24 hours after the last exposure for the 90- (OECD, 2025b) and 180-day (OECD, 2018b) exposure durations. BM was collected by centrifuging the femurs after removing the distal ends, and the resulting cell pellet was resuspended in phosphate-buffered saline (Corning cellgro, Manassas, VA, USA), while livers were separated into right posterior, right anterior, median, caudate, and left lobes. Both BM and liver samples of each exposure duration were flash-frozen and stored at −80°C for further processing.

Animal body weights were measured daily throughout the exposure, and liver weights were collected during the necropsy. The percentage change in body weight (initial to final exposure day) was used as an indicator of systemic toxicity. Additionally, we analysed changes in absolute liver weight and the ratio of liver weight to body weight as indicators for BbF-induced hepatotoxicity.

Statistical analysis of changes in weight gain, absolute and relative liver weights assumed homogeneity of variance, and was confirmed using a Levene's test. The data were analysed using a one-way analysis of variance followed by a pairwise comparison using a Dunnett's test to determine statistically significant changes (defined as *$p < 0.05$ and **$p < 0.01$) of exposed cohorts compared to their respective control groups.

3.3.2 Histopathological Analysis

During the necropsy, left liver lobes were immediately fixed in 10% neutral buffered formalin for 24 to 48 hours. Fixed tissues were trimmed, dehydrated, cleared in xylene, and embedded in paraffin. Histological sections were cut at a nominal thickness of 3 µm and mounted on glass slides. The sections were deparaffinized and stained with hematoxylin and eosin (for routine microscopic examination). Non-blinded histopathological evaluation (with the knowledge of the treatment group) of all slides was evaluated by a board-certified veterinary pathologist via brightfield/light microscopy. Histopathological findings were graded on a semi-quantitative severity scale (e.g., 1 = minimal, 2 = mild, 3 = moderate, 4 = marked, 5 = severe). The assessment focused on the identification of test item-related hepatocellular changes, including hypertrophy, necrosis, cytoplasmic rarefaction (glycogen accumulation), and foci of cellular alteration. Animals with no observable microscopic lesions were recorded as within normal limits (WNL).

3.3.3 DNA Extraction from Bone Marrow and Liver Tissue

DNA was extracted from BM and the anterior liver lobe of four randomly selected animals per dose group using the QIAGEN DNeasy® blood and tissue kit protocol (Cat. #69504, QIAGEN, Hilden, Germany) with minor modifications: (a) tissue disruption by bead beating was performed twice at 4000 rpm for 30 s prior to DNA extraction; and (b) proteinase K digestion was prolonged with an incubation duration of 180 minutes at 37°C. DNA concentrations were measured using the Qubit 1x dsDNA Broad Range Assay Kit (Thermo Fisher Scientific Inc., Waltham, MA, USA), and DNA integrity was evaluated via Agilent TapeStation Genomic DNA ScreenTape assay (Agilent Technologies Inc., Santa Clara, CA, USA). All samples met the requirements for DS library preparation, exhibiting DNA integrity numbers above 7.0 and concentrations greater than 25 ng/µL.

3.3.4 Library Preparation and Sequencing

Library preparations were performed using the TwinStrand Duplex SequencingTM Mutagenesis Panel (Mouse-50) v1.0, following the DuplexSeqTM Kit Manual provided by TwinStrand Biosciences Inc. (Seattle, Washington, USA). In short, 500 ng of DNA per sample were enzymatically fragmented to a size of approximately 400 bp. The fragmented DNA was end-polished, adenylated, and ligated to DS adapters containing 12-nucleotide unique molecular identifiers (TwinStrand Biosciences Inc., Seattle, WA, USA). Library concentrations were measured using QubitTM 1x dsDNA High Sensitivity Assay Kits (Thermo Fisher, Waltham, MA, USA) and average fragment sizes determined using an Agilent TapeStation High Sensitivity D1000 Screen Tape assay (Agilent Technologies Inc., Santa Clara, CA, USA).

The Mouse Mutagenesis Panel comprises 20 endogenous loci (2.4 kb each) distributed across all mouse autosomes (with two loci on chromosome 1). These loci were selected to broadly represent the nucleotide composition, GC content, heterochromatin state, and transcriptional properties of the mouse genome, while excluding repetitive regions, pseudogenes, and selected genes (LeBlanc et al., 2022). This rigorous selection process ensures both efficient hybrid capture and an unbiased assessment of mutagenesis. The panel includes nine genic and eleven intergenic loci (as identified in Mouse GenCode v.M25) that were captured using 120-mer biotinylated oligo probes following the protocol described in the DuplexSeqTM Kit Manual by TwinStrand Biosciences.

Library sequencing was performed at Psomagen (Rockville, MD, USA) on NovaSeq 6000 S4 flow cells (Illumina, San Diego, CA, USA), targeting a minimum of 600 million raw reads passing filters per sample with a mean on-target sequencing depth greater than 10,000x. The raw sequencing data are accessible in the Sequence Read Archive under accession number PRJNA1131313 for the 28-, 90- and 180-exposure durations.

3.3.5 Data Processing and Error Correction

FASTQ files were processed using the TwinStrand DuplexSeq™ Mutagenesis App v3.4.2 (TwinStrand Biosciences). Data processing involved initial extraction of the 12-nucleotide unique molecular identifiers sequence and adapter trimming. The sequence was then parsed to create a duplex tag header. Following paired-end read

alignment and sorting, a single-strand consensus sequence was generated. The single-strand consensus sequences were subsequently grouped and analysed to yield the double-strand consensus sequences. This procedure, which leverages unique molecular identifiers, eliminates technical errors and false mutation discoveries, reducing the error rate to below one error per 108 nucleotides (Schmitt et al., 2012; Valentine et al., 2020), enabling highly sensitive rare event detection.

3.3.6 Tissue-Specific Mutation Frequency Calculation

MF was calculated from the DS data by aligning sequencing reads to the mouse reference genome (M. musculus mm10, GRCm38.p6) and identifying substitutions. Accounting for clonal expansion, where identical mutations may arise from a single event in a mutant cell, MFs were calculated under two assumptions (Dodge et al., 2023):

1. Mutation frequency minimum (MFmin): Identical mutations in a sample are counted as a single event.

2. Mutation frequency maximum (MFmax): All mutations in a sample are counted as independent events.

Mutations present in more than one consensus read per sample were categorized as clonal events. A variant allele fraction cutoff of 0.01 was applied to exclude potential germline mutations and low-level mosaicism.

Statistical analyses and mutation spectrum characterizations were performed using the MutSeqR software (github.com/EHSRB-BSRSE-Bioinformatics/MutSeqR) (Dodge et al., 2025). Dose-dependent MFmin and MFmax were estimated through generalized linear models with an over-dispersed binomial error distribution. MutSeqR utilizes the doBy R package for pairwise comparisons, calculating back-transformed standard errors via the delta method and adjusting p-values with a Sidak correction. For locus-specific MFs, the software uses generalized linear mixed models incorporated in the glmer function from the lme4 package, with pairwise comparisons similarly handled via doBy. Single-base substitution (SBS) spectra were derived by sequence comparison against the mm10 reference genome. The 28-day exposure

data from Schuster et al. (Schuster et al., 2024) were included as a baseline for comprehensive time-course comparisons.

3.3.7 Mutational Signature Analysis

Mutations were quantified in the 96-base context format, accounting for the immediate 5′ and 3′ flanking bases relative to the mutated nucleotide. A *de novo* mutational signature was extracted from the aggregated trinucleotide mutation data across all samples within each dose group. Signature assignment involved cross-referencing the extracted *de novo* signature against the COSMIC mouse SBS signatures catalog version 3.5 (https://cancer.sanger.ac.uk/signatures/sbs/). Signature extraction and reconstruction were performed on all dose groups and exposure durations of each tissue using the SigProfilerExtractor python package (Islam et al., 2022). Combining the four samples by dose group and exposure duration within each tissue makes the dataset more informative, while avoiding confounding effects from differences in BbF's mutagenic action between tissues. While we considered reconstructions with cosine values above 0.9 to be strong, we applied greater scrutiny to those below 0.8 due to lower confidence in their signature assignments (Islam et al., 2022).

3.3.8 Benchmark Dose Modeling

Benchmark dose (BMD) modeling was performed to determine the mutagenic potency of BbF. The BMD represents the dose at which a predefined change relative to background MF is expected. We used the PROAST v70.1 software package to analyse the MFmin and MFmax values for liver and BM, using the Akaike Information Criterion (AIC) to identify the best-fit models. As recommended by the European Food Safety Authority, an AIC difference of 2 units indicated a superior fit (EFSA Scientific Committee et al., 2017). The models applied during the BMD analyses included the 3- and 5-parameter Hill and Exponential, the Inverse Exponential, and the Log-Normal. The most probable BbF dose-response curve for the various endpoints was determined by averaging 200 bootstrap runs. AIC was calculated throughout all bootstrap runs, and models that failed to converge were excluded from the model averaging. A 50% increase compared to the vehicle control was used as the benchmark response (BMR) as recommended for mutagenicity (White et al., 2025).

The respective BMD lower and upper confidence limit values (BMDL and BMDU) were derived from the 90% confidence intervals.

3.4 Results

3.4.1 Assessment of Systemic and Hepatic Toxicity

Animal health and body weight gain were monitored throughout the study for signs of systemic toxicity (Figure 8 and Appendix D2 Table 1). With the exception of two dead animals at the 90-day exposure (3.125 and 6.25 mg/kg/day group) and a single mortality in the 180-day cohort (3.125 mg/kg/day group), unrelated to the exposure, all animals remained in good health. Short-term exposure to BbF did not cause notable effects on the well-being of the mice. In contrast, prolonged BbF exposure caused systemic toxicity, as evidenced by significant dose-dependent decreases in body weight gain compared to the vehicle controls after 90 and 180 days. Specifically, the two highest doses at 90 days and the highest dose at 180 days impaired the weight gain of the cohort (Figure 8 and Appendix D2 Table 1).

In addition, we analyzed both absolute liver weight and the liver weight to body weight ratio (relative liver weight) as indicators of hepatotoxicity (Figure 8). Absolute liver weights remained largely consistent across exposure doses within each time point. However, we observed significant increases in relative liver weights. After 28 days of exposure, relative liver weight increased at the two highest doses (50 and 100 mg/kg/day). In contrast, 90 and 180 days of exposure caused a statistically significant increase in relative liver weight only at the respective top doses of 50 and 25 mg/kg/day. We note that prolonged exposure to high BbF doses caused systemic toxicity that manifested in reduced body weight gain and adaptive increases in relative liver weight, while absolute liver mass remained constant, which is consistent with other studies (Liu et al., 2025).

3.4.2 Absence of Dose- or Time-Dependent Histopathological Changes

Histopathological examination of liver tissues revealed no evidence of overt test item-related toxicity following exposure to BbF for 28, 90, or 180 days. Except for a few cases, animals across all dose groups were classified as WNL without histopathological findings (Appendix D2 Table 2).

In the 28- and 90-day cohorts, lesions were sporadic and did not follow a dose-dependent pattern. Minimal (Grade 1) focal necrosis and mild (Grade 2) cytoplasmic rarefaction were noted in a few animals, but these findings appeared with similar frequency in control and low-dose groups (Appendix D2 Table 2). Importantly, the highest dose groups at both timepoints showed no increase in lesion severity or incidence compared to controls.

Following 180 days of chronic exposure, rare events of foci of cellular alteration (Grade 1; including one animal from the untreated control group) and cytoplasmic rarefaction (Grade 2–3) were observed in lower-dose animals (Appendix D2 Table 2); however, no histopathological abnormalities were observed in the highest dose group (25 mg/kg/day). The absence of such lesions at the maximum dose confirms that the findings at lower doses were incidental events unrelated to BbF treatment.

Collectively, these results confirm that the BbF treatment did not induce dose-dependent histopathological effects in the liver, even after extended exposure durations.

3.4.3 BbF Induces Non-Linear, Dose- and Time-Dependent Increases in MFmin

We used DS to quantify the effects of BbF dose and exposure duration on mutation induction in MutaMouse liver and BM, two tissues with distinct metabolic and proliferation profiles. Sequencing performance analysis confirmed that all samples met minimum quality thresholds, including at least 700 million informative duplex bases, mean on-target sequencing depth greater than 10,000x, and sufficient peak tag family size (Appendix D2 Table 3).

We first established MFmin in liver and BM across time points in vehicle controls. In BM, MFmin (×10-8) ± SEM were 4.1 ± 0.4, 3.9 ± 0.7, and 4.2 ± 0.3 after 28, 90, and 180 days of vehicle exposure, respectively (Supporting Figure 1). In liver, MFmin values were 5.3 ± 1.0, 5.0 ± 0.7, and 5.9 ± 0.7 after 28, 90, and 180 days, respectively (Appendix D2 Figure 1). MFmin in vehicle control animals did not differ significantly across exposure durations. Nonetheless, BbF-induced effects within an exposure duration were evaluated only against their concurrent control animals.

We observed significant BbF-induced dose-responses in MFmin in both tissues that increased with prolonged exposure durations (Figure 9). After 28 days of exposure, MFmin (×10-8) ± SEM reached 14.8 ± 0.1 in BM and 36.3 ± 2.5 in liver at the top dose (100 mg/kg/day) (Appendix D2 Table 4). After 90 days, MFmin was 18.4 ± 0.1 in BM and 87.4 ± 1.9 in liver at the top dose (50 mg/kg/day) (Appendix D2 Table 5). Finally, after 180 days (top dose of 25 mg/kg/day), MFmin reached 19.5 ± 1.1 in BM and 94.6 ± 7.1 in liver (Appendix D2 Table 6). The dose-response was non-linear, with significant increases in MFmin observed only above a dose- and time-dependent threshold. Specifically, the lowest statistically significant dose shifted from 12.5 to 6.25 and then to 3.125 mg/kg/day in BM, and from 25 to 12.5 and then 6.25 mg/kg/day in liver, for 28, 90, and 180 days of exposure respectively. The shift in lowest observable genotoxic effect levels over time demonstrates that the threshold for significant mutation induction is not static but rather decreases as BbF exposure is prolonged.

We next explored the mutagenic effects at 25 mg/kg/day, the highest dose group in common across all three exposure durations, to determine whether the observed increases in MFmin scaled with cumulative dose. Over 90 and 180 days of exposure, the cumulative dose increased by approximately 3.2- and 6.4-fold, respectively, relative to the 28-day study. However, we observed a statistically significant lower increase (Welch's t-test, $p < 0.05$) in MFmin in BM of only 1.7-fold from 28 days to 90 days and 2.1-fold from 28 days to 180 days at 25 mg/kg/day. In contrast, a more pronounced increase than expected based on cumulative dose was observed in liver, where MFmin increased 3.8-fold and 8.6-fold (Welch's t-test, $p < 0.05$) from 28 days to 90 and 180 days, respectively. Together, these findings demonstrate that the fold-change increase in MFmin does not align with the prediction according to the cumulative dose in either tissue.

In summary, BbF exposure caused tissue-specific, non-linear increases in MFmin, with prolonged exposure lowering the dose required for mutation induction and increasing maximal mutational burden, particularly in liver. These findings demonstrate a pronounced time dependence of BbF-induced mutagenesis beyond a simple cumulative dose effect.

3.4.4 Prolonged Exposure Promotes Clonal Expansion in BM

To understand the impact of the accumulation of mutant cell lineages on the total mutation burden over time following BbF exposure, we calculated MFmax. Unlike MFmin, MFmax does not collapse identical mutations; instead, it sums the total mutant read depth, counting every observation of a mutation as a contributor to the total burden. In BM, MFmax exceeded MFmin at all doses after 28 (Appendix D2 Table 4), 90 (Appendix D2 Table 5) and 180 days (Appendix D2 Table 6), consistent with the presence of clonally derived mutations, whereas the difference was mostly absent in liver (Appendix D2 Figure 2). After 90 days, MFmax (×10-8 ± SEM) reached 28.1 ± 2.8 in BM and 89.8 ± 12.3 in liver at 50 mg/kg/day. After 180 days, the highest MFmax was 40.1 ± 3.1 in BM and 98.8 ± 7.2 in liver at 25 mg/kg/day. In BM, the difference between MFmin and MFmax increased with prolonged exposure; for example, at 180 days, MFmax was approximately 2.1-fold higher than MFmin at the top dose. This divergence indicates that the physiological propagation of mutant cells is a major contributor to the total mutation burden in BM. Conversely, the majority of mutations across all exposure durations were unique in liver, indicating a minimal presence of clonally derived mutations in this tissue. This likely reflects differences in cellular turnover, where the rapid proliferation of bone marrow cells amplifies existing mutations, a phenomenon largely absent in liver.

Since MFmax reflects both the number of clonal mutations and the extent of expansion of individual mutations, we next quantified the alternate read depth (alt depth) supporting each mutant allele at individual genomic positions as a proxy for clonal size (Figure 10). In BM, the proportion of total mutations present as multiplets (exhibiting an alt depth greater than one) increased with longer exposure durations, whereas the maximum alt depth (i.e., number of multiples) was similar across durations, with 47 after 28 days (Appendix D2 Figure 3), 55 after 90 days (Appendix D2 Figure 4), and 46 after 180 days (Appendix D2 Figure 5). Although the number of recurring mutations increased with exposure duration in BM, no dominant mutation subtype was identified, as all subtypes were similarly affected by clonal expansion. Sporadic individual multiplets were also detected in liver, with maximum alt depths of 93 after 28 days (Appendix D2 Figure 3), 74 after 90 days (Appendix D2 Figure 4), and 33 after 180 days; however, these rare instances of individual larger clonal mutants were not dose-dependent (Appendix D2 Figure 5). These findings indicate that prolonged exposure

primarily increased the number of clonally derived mutations rather than the size of the largest individual clone.

In summary, the results demonstrate that clonally derived mutations accumulate over time and contribute to the total mutational burden only in BM. Consequently, the BbF's mutational landscape in BM reflects both the induction of new mutations and their persistence across cell lineages, whereas the mutational landscape in liver is driven predominantly by new mutation events.

3.4.5 Consistent Mutation Spectrum

To assess potential mechanistic differences, we characterized induced mutation subtypes. BbF-induced mutations remained highly consistent with the 28-day spectrum but were detectable at progressively lower doses with increasing exposure duration. In both BM and liver, BbF predominantly induced dose-dependent increases in C:G>A:T and C:G>T:A mutations across all exposure durations (Appendix D2 Table 7, Appendix D2 Table 8 and Appendix D2 Table 9). As for the overall MFs, the dose required to elicit statistically significant increases in these primary mutation subtypes decreased with exposure duration, indicating a time-dependent increase in mutagenic sensitivity. After 28 days, a significant increase in C:G>A:T transversions was first observed at 12.5 mg/kg/day in BM and 25 mg/kg/day in liver, whereas C:G>T:A transitions required a higher dose of 50 mg/kg/day to reach significance in both tissues (Appendix D2 Figure 6). After 90 days, C:G>A:T transversions were detected at 6.25 mg/kg/day in BM and 12.5 mg/kg/day in liver, whereas C:G>T:A transitions were first detected at 12.5 mg/kg/day in both tissues (Appendix D2 Figure 7). At 180 days, significant increases in C:G>A:T transversions were first observed at 3.125 mg/kg/day in BM and 6.25 mg/kg/day in liver, while C:G>T:A transitions occurred at 6.25 mg/kg/day in BM and 12.5 mg/kg/day in liver (Figure 11).

Dose-dependent increases in other mutation subtypes were also observed, although the range of the induced response differed between tissues and over time. Although the top dose at 28-days (100 mg/kg/day) induced significant increases in most single base substitution subtypes in BM, this profile narrowed over time. At 90 days, T:A>G:C transversions were no longer significantly increased at the adjusted top dose, and after 180 days, neither T:A>C:G nor T:A>G:C mutations were significantly elevated. No significant increases in insertions and deletions were detected in BM after any

exposure duration. In contrast, liver exhibited a broader mutational profile, with all six subtypes showing a significant increase at the top dose across all exposure durations. Additionally, after 180-days of BbF exposure, a significant increase in deletions was observed in liver. However, despite some tissue-specific differences in subtype diversity, the overall mutation spectrum remained dominated by C:G>A:T transversions across exposure durations, indicating a consistent BbF-associated mutational signature in both tissues.

3.4.6 BbF Induces Mutations Predominantly in Intergenic Loci

We then investigated whether BbF-induced mutations preferentially occur in specific genomic regions by analyzing MFs across the 20 endogenous loci. Although the magnitude of the response varied, BbF induced an increase in MFmin across all 20 loci in both tissues at the 180-day exposure duration. Furthermore, the rank order of the most responsive loci at the top dose remained largely consistent across all exposure durations in both tissues (Figure 12, Appendix D2 Table 10, Appendix D2 Table 11 and Appendix D2 Table 12).

Notably, the five most sensitive loci in both tissues at all exposure durations were intergenic. In BM, the mean MFmin ($\times 10^{-8}$ ± SEM) were 21.8 ± 3.8, 25.7 ± 4.4 and 27.9 ± 4.6 in intergenic loci and 6.7 ± 1.2, 10.2 ± 1.7 and 11.0 ± 2.0 in genic loci after 28, 90 and 180 days, respectively. In liver, we calculated the MFmin in intergenic loci as 40.5 ± 2.8, 102.3 ± 10.3, 122.5 ± 10.9 and in genic loci as 31.2 ± 2.9, 69.8 ± 9.3 and 64.9 ± 8.9 after 28, 90 and 180 days, respectively. Intergenic loci exhibited significantly higher MFs than genic loci across both tissues and all exposure durations (Welch's t-test, $p < 0.05$). Specifically, the fold changes of intergenic versus genic MFmin were: 3.3 and 1.3 after 28 days; 2.5 and 1.5 after 90 days; and 2.5 and 1.9 after 180 days, in BM and liver, respectively. This trend indicates heightened susceptibility of intergenic regions to BbF-induced mutagenesis.

3.4.7 Similarity between BbF Trinucleotide Mutation Pattern and Tobacco Smoke-Associated Mutation Signature

To identify the mutational processes driving BbF-induced mutagenesis, we quantified mutations in the trinucleotide context format and reconstructed the extracted

signature per dose for all 72 BM and 72 liver samples using the SigProfilerExtractor Python package and the COSMIC (v3.5) signatures (Appendix D2 Table 13).

The signature extraction and reconstruction of the trinucleotide mutation pattern found in BM revealed the contribution of SBS1, SBS4, SBS5, SBS17a/b and SBS18 (Figure 13). As expected, we detected the clock-like signatures SBS1, associated with deamination of 5-methylcysteine (Nik-Zainal et al., 2015), and SBS5 (Alexandrov et al., 2013) in treated and untreated samples independently. Additionally, we found a dose- and time-dependent increase in the contribution of SBS4, a tobacco-smoke associated signature (Alexandrov et al., 2013; Nik-Zainal et al., 2015). SBS4 contributed up to 62.1%, 65.0% and 69.6% of the signature reconstruction at the top dose after 28, 90 and 180 days, respectively (Appendix D2 Table 13). Notably, the lowest dose with contribution of SBS4 shifted from 25 to 12.5 to 6.25 mg/kg/day after 28, 90 and 180 days, respectively.

Minor, dose-independent contributions of SBS18, associated with reactive oxygen species (Alexandrov et al., 2013; Kucab et al., 2019, 2026), were detected in a majority of signature reconstructions. The reconstructions for BM at the top dose after 90 days and the two highest doses after 180 days reached a robust cosine similarity greater than 0.9. Eleven samples reached a cosine similarity greater than 0.8, and only 4 reconstructions were unreliable with a cosine similarity below 0.8.

Conversely, in liver, the higher number of mutations per sample enabled a more robust signature extraction and reconstruction with contributions of SBS1, SBS4, SBS5 and SBS40a (Figure 13). The clock-like signatures, SBS1 and SBS5, were detectable independent of dosing or exposure duration. With increasing doses and exposure duration, signature reconstruction revealed major contributions of the tobacco-smoke associated SBS4 (Alexandrov et al., 2013; Nik-Zainal et al., 2015). SBS4 contributed up to 70.3%, 86.5% and 84.4% at the top dose after 28, 90 and 180 days, respectively (Appendix D2 Table 13). Prolonged exposure enabled the detection of similarities to SBS4 even at lower doses. Additionally, the contribution of SBS40a, of unknown etiology (Senkin et al., 2024), was detected at all doses after 28 days and sporadically after prolonged exposure. The top two doses and the third highest dose after 90 and 180 days yielded reconstructions with cosine similarities exceeding 0.9. Only the

control and lowest exposure dose groups after 28 and 90 days did not reach the cosine similarity threshold of 0.8.

Together, these findings reveal the contribution of clock-like signatures in all reconstructions and a dose- and time-dependent contribution of the tobacco-smoke associated mutational signatures in both tissues, with confident reconstructions at higher doses and after prolonged exposure durations.

3.4.8 Benchmark Dose Modeling Reveals Increased BbF Mutagenic Potency

We applied BMD modeling to quantitatively assess how the mutagenic potency of BbF changes with exposure duration and tissue type. Using MFmin as the primary endpoint, BbF potency increased with exposure duration in both tissues, as reflected by declining BMD50 values from 28 to 180 days (Figure 14 and Appendix D2 Table 14).

In BM, the BMD50 confidence interval (CI; from lower [BMDL] to upper [BMDU]) shifted from 3.6 - 12.8 mg/kg/day at 28 days to 1.6 - 5.8 mg/kg/day at 90 days and to 2.3 - 3.8 mg/kg/day at 180 days. Although this pattern indicates increasing potency with exposure duration, the CIs between the three exposure durations overlapped.

In liver, the BMD CI decreased from 10.1 - 18.4 mg/kg/day at 28 days to 2.3 - 6.3 mg/kg/day at 90 days to 3.1 - 4.6 mg/kg/day at 180 days. Potency increased between 28 and 90 days, as indicated by non-overlapping CIs; however, no further shift was observed from 90 to 180 days of exposure, suggesting a plateau in potency in liver after 90 days of exposure. Despite the liver showing a higher maximum mutagenic response, the BMD50 confidence intervals for BM and liver overlapped at each exposure duration, indicating comparable mutagenic potency across these tissues at low doses. Lastly, we observed that the width of the CIs narrowed with increasing exposure duration, consistent with improved precision of potency estimates over longer exposure durations.

We next calculated BMD50 for MFmax to determine the influence of clonal expansion on potency estimates. While precision in potency estimation is influenced by various experimental factors, the generally wider confidence intervals observed for MFmin compared to MFmax could additionally be influenced by the greater inter-

animal variability associated with animal specific clonal expansion. In addition, the broader CI at 28 days overlaps with the CIs of prolonged exposures, indicating that inclusion of clonally derived multiplets reduces the power to detect potency increases.

Finally, we applied BMD modeling to MFmin for C:G>A:T transversions, the main mutation subtype induced by BbF, to assess the sensitivity of a mechanism-based endpoint (Appendix D2 Table 14). This approach yielded the lowest BMD50 values. At 90 days, the BMD50 CIs for C:G>A:T transversions were 0.2 - 2.1 mg/kg/day in BM and 1.0 - 2.4 mg/kg/day in liver. After 180 days of exposure, the CIs ranged from 0.9 - 2.2 mg/kg/day in BM and 2.1 - 3.7 mg/kg/day in liver.

Collectively, BMD modeling shows that BbF potency increases with exposure duration and that mechanism-specific mutational endpoints can provide more sensitive estimates of genotoxic potency.

3.5 Discussion

We demonstrate that prolonged oral exposure to BbF leads to pronounced dose-, time- and tissue-dependent effects on mutagenesis in mouse somatic tissues. Using DS, we show that mutation burden increases with exposure duration in both BM and liver, but through distinct biological mechanisms. In liver, a primary site of PAH metabolism, the occurrence of new mutations is much greater than in BM. In contrast, in the highly proliferative BM, mutation burden is shaped by the accumulation of clonally derived mutations, driving an approximate doubling in MFmax over MFmin at high doses and longer exposure durations. Despite these differences in mutational dynamics, mutation spectra and responding genomic loci are similar across tissues and exposure durations. Increasing mutation counts with prolonged exposures is associated with a higher contribution of COSMIC signature SBS4 to the mutation pattern in both tissues but being more prevalent in liver. This signature is associated with tobacco smoke-induced cancers and correlates with PAH exposure (Alexandrov et al., 2013; Nik-Zainal et al., 2015). BMD modeling revealed a time-dependent increase in BbF mutagenic potency from 28 to 90 days, but with no further declines in BMDs beyond 90 days suggesting maximum potency had been achieved. Despite differences in maximum responses, BMD50 in BM and liver are similar. Collectively, our work demonstrates that both biological context (e.g., tissue-specific proliferation, metabolism, and DNA repair characteristics) and exposure duration critically influence

the mutational consequences of BbF exposure. The study provides crucial data for refining human health risk assessment frameworks for chronic exposure to environmental mutagens.

The rate and magnitude of mutation accumulation differed between tissues, with liver consistently exhibiting enhanced mutagenicity relative to BM. These tissue-specific differences likely arise from differences in metabolic activity, cellular proliferation, and DNA repair capacities (Hoch, 2023). The liver expresses high levels of CYP450 enzymes and is a primary PAH-metabolizing organ, leading to more reactive BbF metabolites such as dihydrodiol epoxides and hydroquinones, that can induce DNA adducts and oxidative DNA damage (Amin et al., 1982, 1985; Yang, 1988). At the same time, the inherently high proliferation rate of BM makes it particularly susceptible to the fixation of mutations during DNA replication despite its lower metabolic activity for xenobiotics (Long et al., 2016; Pellegrini et al., 1995; Yeh et al., 2024). This balance between metabolic activation, DNA damage, and cell proliferation is likely driving the tissue-specific differences in susceptibility to BbF induced mutagenicity.

The use of long exposure durations allowed us to investigate the cumulative nature of mutagenicity and whether this differs between tissues. We observed duration- and tissue-dependent increases in MFmin, revealing the cumulative nature of mutagenicity; however, the magnitude of accumulation deviated from simple cumulative dose predictions. In the BM, MFmin increases were lower than the 3.6- and 6.4-fold elevations predicted by cumulative dose at 25 mg/kg/day for the prolonged exposure durations. This discrepancy likely stems from the continuous turnover and release of hematopoietic cells into the bloodstream (Wilson et al., 2008). In contrast, the liver exhibited an increased response, where the 8.6-fold increase at 25 mg/kg/day in MFmin at 180 days significantly exceeded cumulative dose expectations. This increased response in the liver is likely driven by physiological mechanisms such as saturation of DNA repair pathways by persistent damage (Klapacz et al., 2016), metabolic adaptation and enzyme upregulation caused by prolonged AhR activation (Liu et al., 2025; Stoddard et al., 2021), and/or potential underestimation of the MFmin after 28 days of exposure due to insufficient mutation fixation time in this slowly proliferating tissue (Marchetti et al., 2021; OECD, 2025d). These findings highlight that

short-term studies may significantly underestimate mutagenic risk in the metabolically active liver where accumulation dynamics are non-linear.

Overall, our study demonstrates that the combined influence of repeated DNA damage, metabolic adaptation, and mutation fixation kinetics must be considered when interpreting mutagenicity data. It also highlights the value of integrating mutagenicity testing into longer exposure duration studies and the importance of ensuring appropriate fixation periods when assessing environmental mutagens.

One of the most striking tissue-specific differences was the presence of pronounced numbers of clonally derived mutations in BM that was largely absent in liver. This pattern is consistent with the inherently high proliferation rate of BM. Because hematopoietic stem and progenitor cells undergo extensive cell division to maintain blood cell production (Fliedner, 2002; Jaiswal et al., 2014; Xie et al., 2014), mutations arising in progenitor cells have a high probability of appearing as recurring mutations, thereby elevating MFmax. Clonal expansion in BM was not limited to a few dominant clones but involved multiple independently expanded mutations. This pattern suggests ongoing mutagenesis and clonal drift over time rather than early-arising mutant cells that expanded by positive selection. Thus, the results are consistent with the interpretation that increasing mutation burden with exposure duration in BM is driven in part by numerous, temporally distinct clonal events.

In contrast, liver tissue showed minimal evidence of clonally derived mutations, even after 180 days of exposure. The relative absence of recurrent mutations in liver despite the higher overall mutation burden, indicates that increases in MFmax are more likely driven by clonal propagation of induced mutations rather than repeated independent mutations at specific genomic sites. Previous studies have demonstrated that clonal dynamics are limited in healthy liver tissues, whereas proliferation and subsequent clonal expansion increases in diseased liver (Brunner et al., 2019). Despite the observed increases in relative liver weight, we did not detect histopathological indications of morphological changes or increases in absolute liver weights between vehicle controls and exposed tissues. These findings suggest that, while the exposures induced mutations in the liver, they did not cause gross pathological changes at any doses or durations. The lower proliferation rate likely limits the opportunity for clonal expansion, leading to the absence of significant differences

between MFmin and MFmax. Our findings suggest that mitotic activity and the proliferative kinetics of the resident stem cells and respective descendant cells are central determinants of the occurrence and extent of clonal expansion, emphasizing their importance when selecting tissues for error-corrected sequencing-based mutagenicity studies.

The observed tissue-specific dynamics of clonal expansion may also point toward different tissue-specific mechanisms of carcinogenesis following chronic mutagen exposure. In a relatively quiescent tissue like the liver, cancer risk may be driven primarily by the accumulation of *de novo* mutations in target genes. In contrast, in a high-turnover tissue like BM, cancer risk may be driven by the clonal propagation of an early-arising mutation, resulting in a large pool of pre-mutated cells susceptible to additional mutational events. This aligns with the hypothesis that cancer risk is correlated with the total number of stem cell divisions (Tomasetti & Vogelstein, 2015), suggesting a mechanistic link between high turnover and clonal expansion. This accumulation of genetically distinct clones following persistent mutagen exposure is a defining feature of high-turnover tissues (Attolini et al., 2010; Parsons, 2018; Tan & Lynch, 2013). However, as our study focused on genomic loci with no known influences on cell growth, it remains unclear how these clonal dynamics would be influenced by mutations occurring in cancer driver genes, which could lead to positive selective pressure. Investigating this interaction will be an important focus of future work. Overall, the MFmax approach provides valuable insight into the biological processes and the dynamics of clonal expansion of mutations, although its potential use for risk assessment and chemical evaluation remains largely unexplored.

The mutational spectrum induced by BbF remained remarkably consistent across tissues and exposure durations, dominated by C:G>A:T transversions with a secondary contribution of C:G>T:A transitions. The prevalence of C:G>A:T mutations aligns with the expected formation of bulky DNA adducts at guanine bases, a well-established outcome of PAH metabolism through reactive dihydrodiol epoxide intermediates (Amin et al., 1982). We also observed modest, yet consistent, increases in all single base substitution mutation subtypes over time, suggesting the formation of various types of DNA adducts and damage. This pattern is consistent with the three main metabolic pathways through which BbF can induce DNA damage: formation of dihydrodiol epoxides catalyzed by CYP450 enzymes and epoxide hydrolase;

generation of PAH radical cations through CYP450 peroxidase-mediated oxidation; and formation of ortho-quinones through oxidation of catechols via dihydrodiol dehydrogenases (Amin et al., 1982; Guengerich, 2008; Shimada, 2006). The production of radical cations and redox cycling of quinones leads to an increase in reactive oxygen species, causing oxidative DNA damage that may manifest in C:G>T:A mutations (Moorthy et al., 2015). Although BbF metabolites cause C:G>T:A transitions, these substitutions also arise from spontaneous deamination of 5-methylcytosine, an endogenous process that may increase with age or inflammation, which may be promoted by BbF (Cagan et al., 2022; Liu et al., 2025; Wiebauer et al., 1993). Overall, the consistency in mutation spectrum observed across all exposure durations indicates that the underlying mutational mechanism of BbF remains stable over time, even as the total mutation burden increases, reinforcing previous findings on BbF's genotoxic profile (Schuster et al., 2024).

We extended our analyses on induced mutation subtypes by determining the trinucleotide mutation pattern and identifying similarities to COSMIC SBS signatures. The BbF-induced mutation pattern exhibited a strong similarity to SBS4, a signature linked to tobacco smoke exposure and characterized by a subset of C:G>A:T mutations (Alexandrov et al., 2013; LeBlanc et al., 2022; Meier et al., 2017; Nik-Zainal et al., 2015).

In BM, we detected a dose- and time-dependent increase in the contribution of SBS4 and a dose-independent similarity to SBS18. However, at 28 days of exposure, the reconstruction analysis was limited to a maximum of 456 mutations at the top dose and did not reach a cosine similarity exceeding 0.9. Consequently, while SBS4 may be present at lower doses, no reliable conclusion can be derived. Generally, the extraction of chemically induced signatures and their reconstruction via SigProfilerExtractor are strongly dependent on total mutation counts, which increase confidence in the estimated contribution of cancer signatures (Islam et al., 2022). Nevertheless, the detection of SBS4 and SBS18 aligns with the presence of BbF in tobacco smoke and its proposed metabolic activation (Amin et al., 1982).The formation of reactive epoxide metabolites that induce DNA damage in the form of bulky DNA adducts is a known origin of cancers with SBS4 signatures (Nik-Zainal et al., 2015), and hydroquinones that create reactive oxygen species align with the etiology of SBS18 (Kucab et al., 2019). Based on the biological plausibility and the robust reconstruction at prolonged

exposure durations, even cosine similarities greater than 0.8 but below 0.9 can provide insight into the possible carcinogenic properties of BbF.

In contrast, we sequenced more than 3000 mutations at the top dose in liver. This resulted in robust signature extraction and reconstruction, identifying great similarities between SBS4 and the BbF-induced trinucleotide pattern. With multiple cosine similarities greater than 0.9, we are confident in the association between BbF exposure and the detection similarities to SBS4 in liver. The notable difference between BM and liver suggest that studies relying solely on short-term exposures may fail to identify relevant mutational mechanisms, particularly when weaker mutagens or lower doses do not generate sufficient mutations for robust signature analysis. The similarity between the BbF-induced mutation pattern and the tobacco-smoke-associated signature SBS4 in both tissues corresponds with BbF's categorization as a possible human carcinogen (IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, 2010; International Agency for Research on Cancer, 1985, 2023) and aligns with tobacco smoke being a major source of BbF exposure (Hearn et al., 2018). Collectively, these results support the relevance of BbF-induced mutations for human carcinogenesis associated with PAH exposure.

Analysis of locus-specific MFs revealed that BbF induces mutations across all 20 endogenous loci, maintaining a relatively consistent rank order of susceptibility to mutagenesis in both BM and liver. Intergenic regions accumulated more mutations than genic regions in both tissues, despite their distinct metabolic and proliferative profiles. This divergence, which increased over time, is likely attributable to transcription-coupled nucleotide excision repair preferentially clearing DNA damage from actively transcribed genes (Hanawalt & Spivak, 2008).

These findings underscore that mutational responses to BbF vary across loci and genomic contexts, highlighting a critical limitation of current *in vivo* gold-standard assays, such as the Transgenic Rodent Somatic and Germ Cell Gene Mutation Assays (OECD, 2025d) and the Pig-a Gene Mutation Assay (OECD, 2025c). These assays rely on quantifying mutagenesis in single reporter genes that are either heavily methylated prokaryotic transgenes integrated into the genome e.g. *lacZ* or restricted to a single hemizygous gene e.g. Pig-a on the X chromosome (Guengerich, 2008; Olsen et al., 2017). Consequently, reliance on single-locus measurements may either

under- or over-estimate the genome-wide mutational burden, depending on the genomic context of the reporter gene. Assessing mutagenicity across a broader and more diverse cross-section of the genome provides a more representative and biologically informative evaluation of genotoxic potential.

Finally, BMD modeling revealed that the mutagenic potency of BbF is influenced by both exposure duration and the mutagenic endpoint analysed. As expected, lower doses were required to induce a 50% increase in MFmin and MFmax after extended exposure durations. These findings suggest that prolonged exposure increases the sensitivity of tissues to BbF-induced mutagenesis, potentially due to cumulative DNA damage, physiological effects, or declining DNA repair efficiency over time (Gorbunova et al., 2007).

Although modeling MFmax generally produced wider CIs than MFmin, likely due to greater variability introduced by clonal dynamics, both endpoints consistently supported a time-dependent increase in BbF potency. For regulatory applications, BMD modeling of MFmin may provide the most robust approach, as it yields narrower CIs and improved precision in potency estimation. Importantly, MFmin modeling is not influenced by stochastic clonal events that could inflate potency estimates.

Although BMD modeling of the predominant C:G>A:T subtype yielded the lowest BMD values, providing a conservative and mechanism-based estimate of potency, modeling specific subtypes or individual genomic loci remains exploratory. Focusing on hotspot regions, such as cancer driver genes, or transcriptionally active genes may reveal differential susceptibility within the genome and may be useful particularly when the response is weak. Nevertheless, modeling the overall MFmin provides an integrated, pragmatic approach to potency estimation and a biologically comprehensive measure of chemical potency that considers all mutation subtypes while maintaining higher confidence in the derived BMD. MFmin based BMD estimates offer a relevant and robust measure of mutagenic potency for subchronic and chronic exposure studies, consistent with emerging applications of error-corrected sequencing (Ashford et al., 2025; Bercu et al., 2023; Cho et al., 2023; Dodge et al., 2023; LeBlanc et al., 2022; Schuster et al., 2024). Overall, our findings align with the recent recommendation of the International Workshop on Genotoxicity Testing on the use of MFmin for potency estimation (Yauk et al., 2026).

3.6 Conclusion

In summary, we show that prolonged exposure to a mutagenic PAH leads to pronounced tissue-specific increases in somatic mutation burden, driven by both increased mutagenesis and time-dependent accumulation of mutations, and in BM but not liver, clonal expansion. These effects are shaped by tissue-specific biological characteristics, including metabolic activity and proliferation rate. Clonal mutation burden is highly dependent on tissue context and increases with exposure duration in rapidly proliferating tissues like BM, whereas mutation accumulation in liver is driven primarily by *de novo* mutagenesis.

Although the mutation spectrum of BbF remains remarkably consistent over time, progressive mutation accumulation with prolonged exposure enables more refined trinucleotide spectrum analyses, revealing an increasing similarity towards the tobacco-smoke-associated signature SBS4. Furthermore, BMD modeling shows that BbF mutagenic potency estimates are influenced by exposure duration, with lower points of departure derived from longer-term exposures compared with short-term exposure. This highlights the limitations of short-term study designs for predicting long-term mutagenic risk.

Together, our findings underscore the value of long-term, high-resolution mutagenicity studies using advanced sequencing-based tools such as DS, which can be readily integrated into existing 90- or 180-day studies to improve mechanistic understanding and regulatory assessment of environmental carcinogens under more realistic exposure scenarios.

3.7 Acknowledgements

C.L.Y., F.M., and P.A.W. hold an Innovation in Regulatory Science Award from the Burroughs Wellcome Fund (#1021737). C.L.Y. is grateful for funding from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2021-02806) with infrastructure support through the Canadian Foundation for Innovation John Evans Leaders Fund (#233109), the Ontario Research Fund (#40569), and the University of Ottawa. C.L.Y. acknowledges that this project was conducted, in part, thanks to funding provided through the Canada Research Chairs Program (CRC-2020-00060). Work at Health Canada was supported by the Chemicals Management Plan to F.M. and

Genomics Research and Development Initiative to F.M. and M.M. We thank James Alper Alcaraz for his support running the SigProfilerExtractor. We thank Dr. Vanessa Schumacher from Roche and Dr. Guosheng Chen and Dr. Eunnara Cho from Health Canada for their helpful review of the manuscript.

3.8 Declaration on Use of Artificial Intelligence

We used large language models (Gemini Advanced and ChatGPT4) to assist with language editing and sentence structuring. The authors reviewed all suggestions for accuracy and maintained full editorial control over the final output. The AI tools were not used to generate scientific hypotheses, interpret data, or formulate conclusions. The underlying scientific concepts and interpretations originated solely from the authors. The authors are taking full responsibility for the content of the published article.

3.9 Tables and Figures

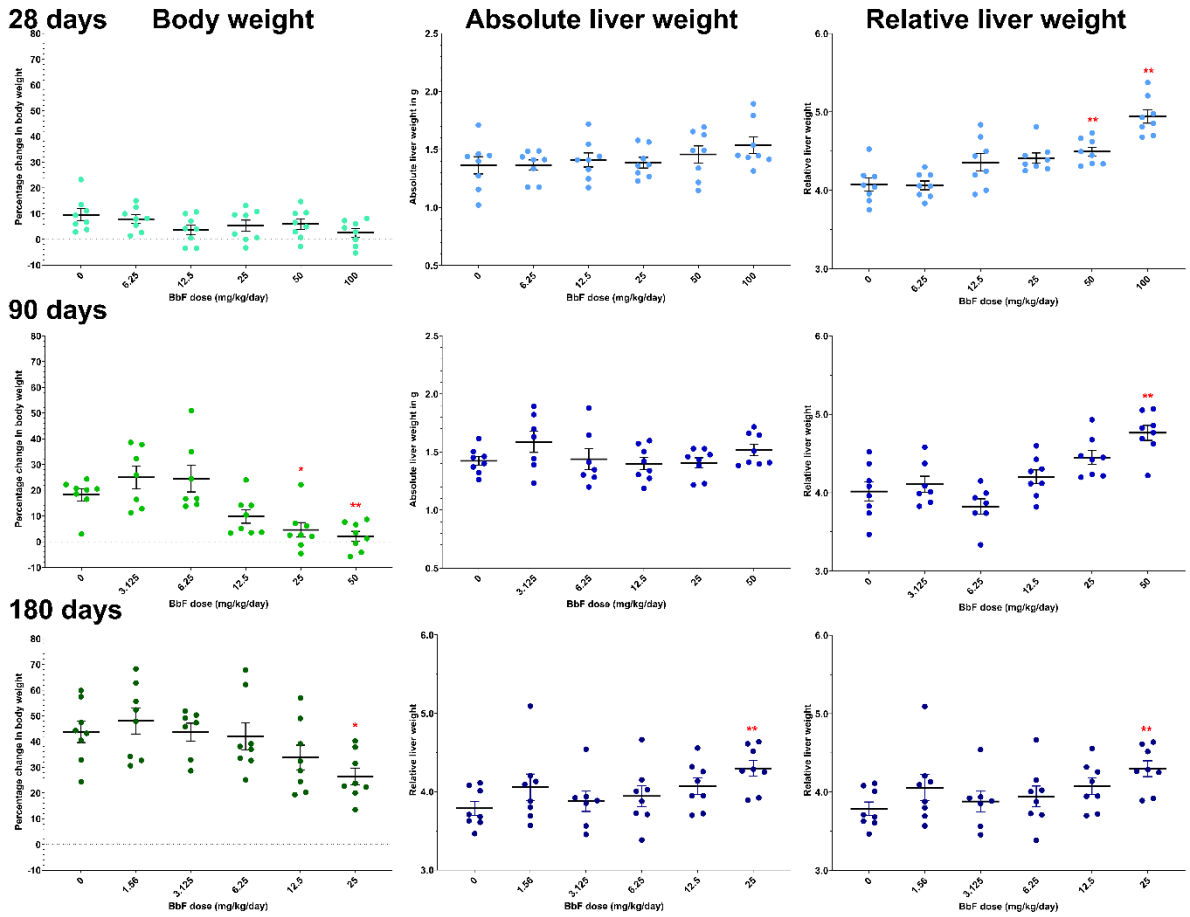


*Figure 8: Body weight per animal was measured at the beginning and end of the BbF exposure and used to calculate the average weight gain ± SEM (green) over time for each dose group after 28, 90 and 180 days. The liver weight (blue) per animal (absolute liver weight) was measured and used to calculate the relative liver weight compared to the final body weight per dose ± SEM over time for each dose group after 28, 90 and 180 days. Significant differences compared to respective vehicle control at each exposure duration were identified using a Dunnett's test and are indicated as * $p < 0.05$ and ** $p < 0.01$*

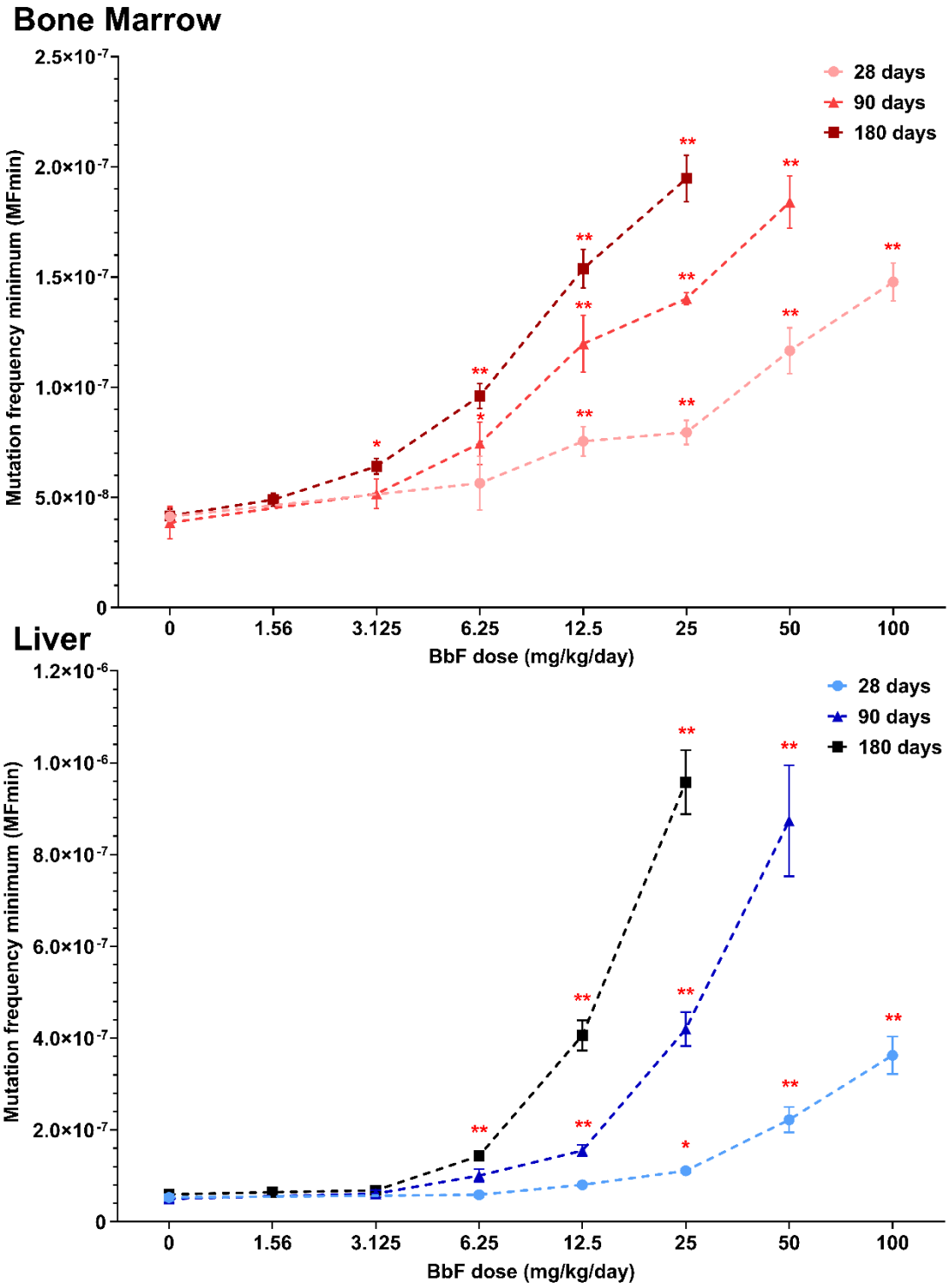


*Figure 9: Estimated vehicle control and BbF-induced mutation frequency minimum (MFmin ± SEM) in bone marrow (red) and liver (blue) of MutaMouse males after BbF exposure for 28, 90, and 180 days. Statistical testing was done using a generalized linear mixed model of each dose against their respective vehicle control; *p < 0.05 and **p < 0.01; n = 4 per dose group.*

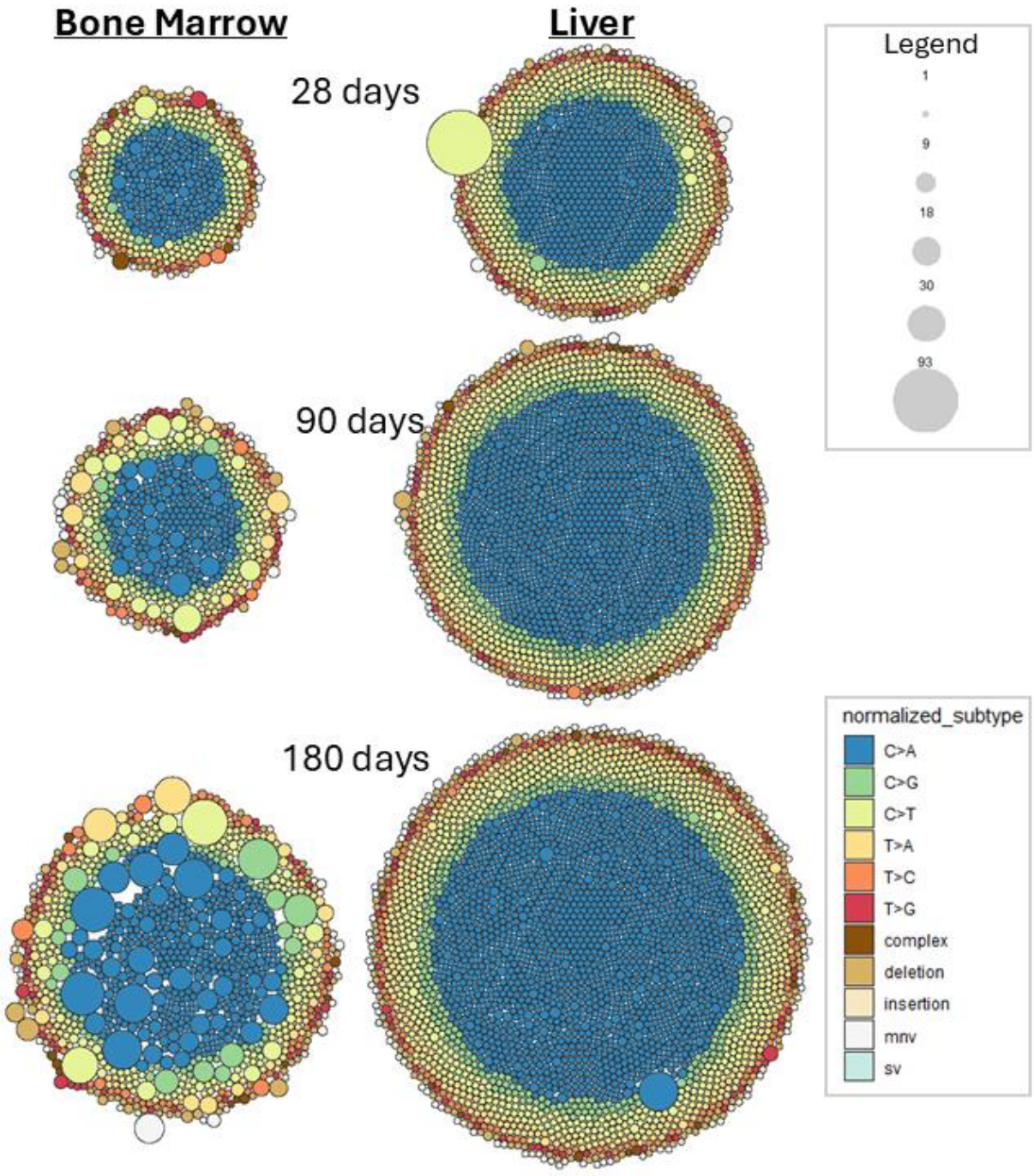


Figure 10: Bubble plots indicating the alt depth (size) of unique mutations (bubble) at the top dose after 28-, 90- and 180-days of exposure to BbF in bone marrow and liver. Colours are indicative of the normalized subtype of mutations according to the legend on the right.

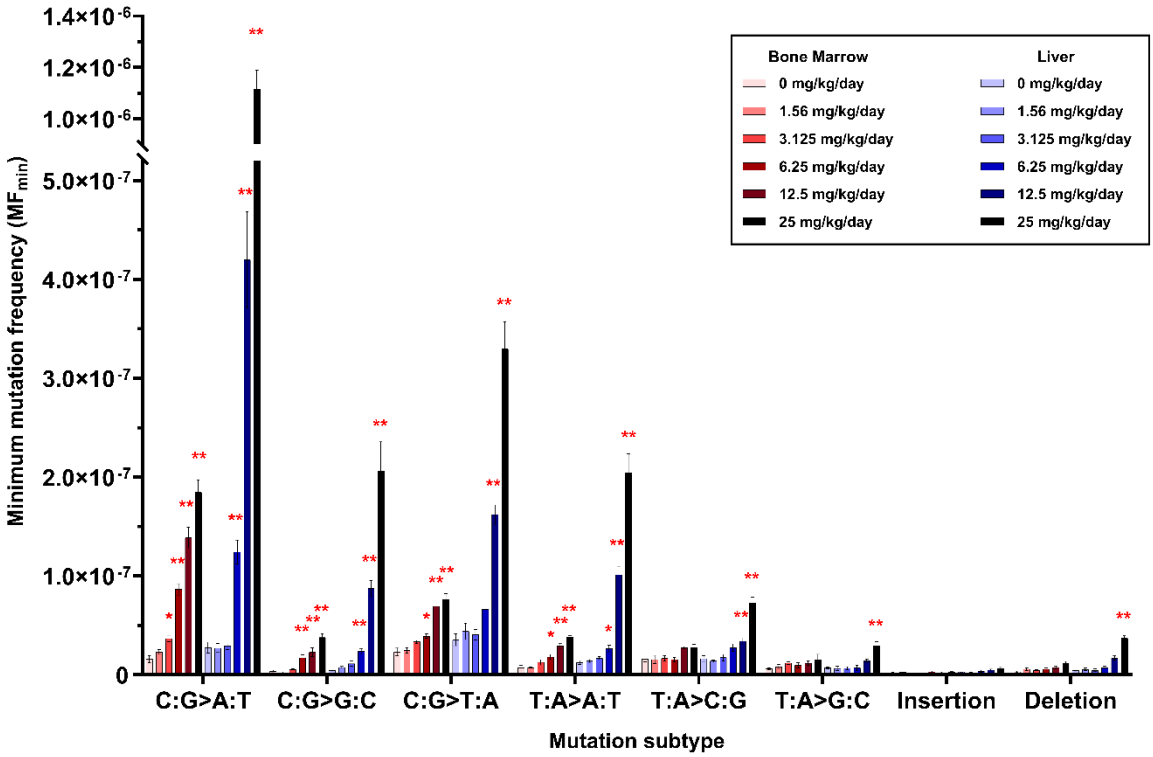


*Figure 11: Mutation frequency minimum (MFmin) with SEM per subtype in bone marrow (A; red) and liver (B; blue) after 180 days of exposure to olive oil or increasing doses of BbF. Statistical testing was done using generalized linear models for each dose and subtype. Significant differences between each dose and the vehicle controls are indicated through * $p < 0.05$ and ** $p < 0.01$.*

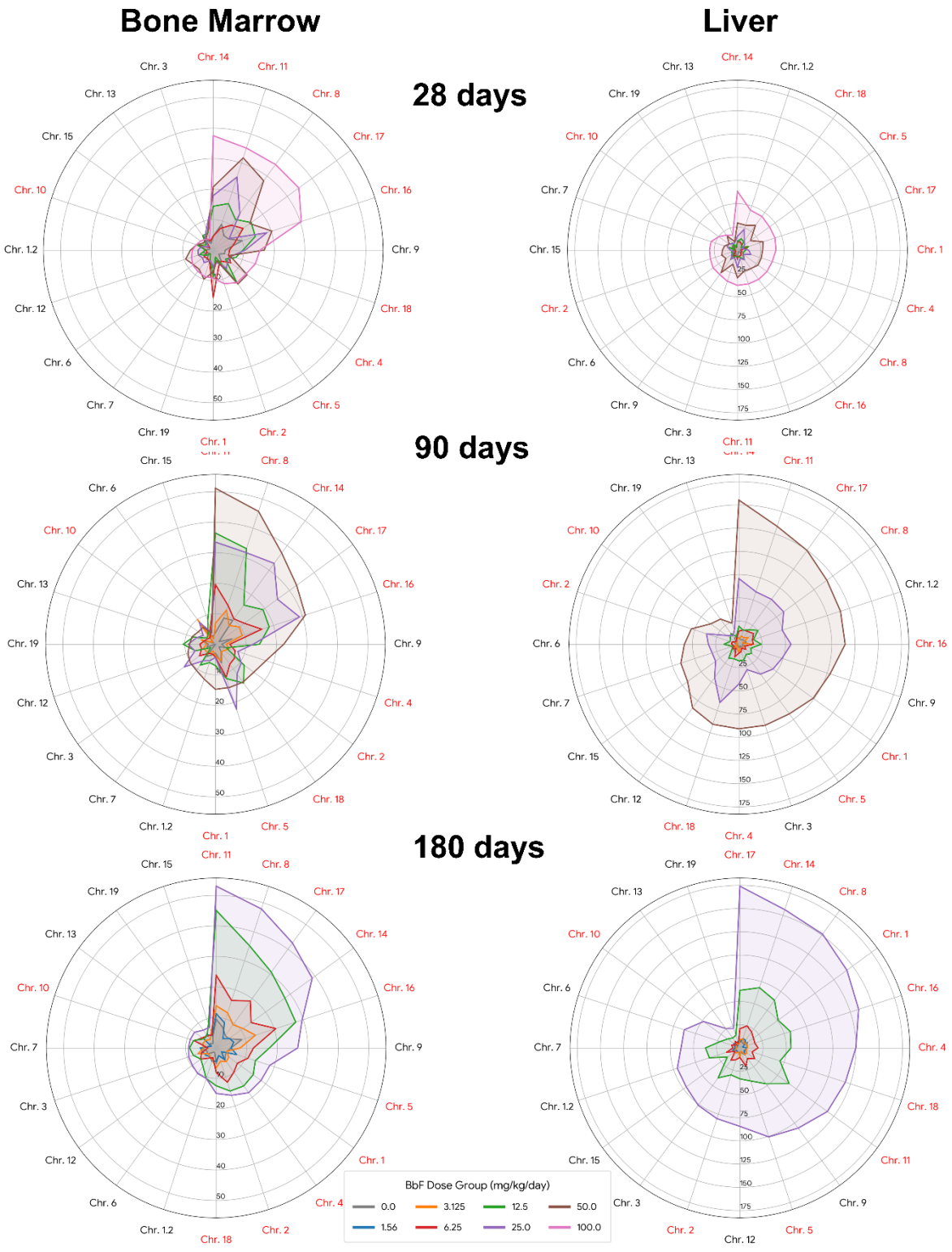


Figure 12: Mean mutation frequency minimum (MFmin) (x10^{-8}) per dose of each 20 loci captured within the Mouse Mutagenesis panel using Duplex Sequencing after 28, 90 and 180 days in bone marrow and liver. The loci are arranged from the highest to lowest MFmin at the top dose for each exposure duration and tissue, with target font colour indicating intergenic (red) and genic (black) loci according to Mouse GenCode v.M25. The colour indication is consistent throughout all radar plots per dose group, and the scale is identical within each tissue

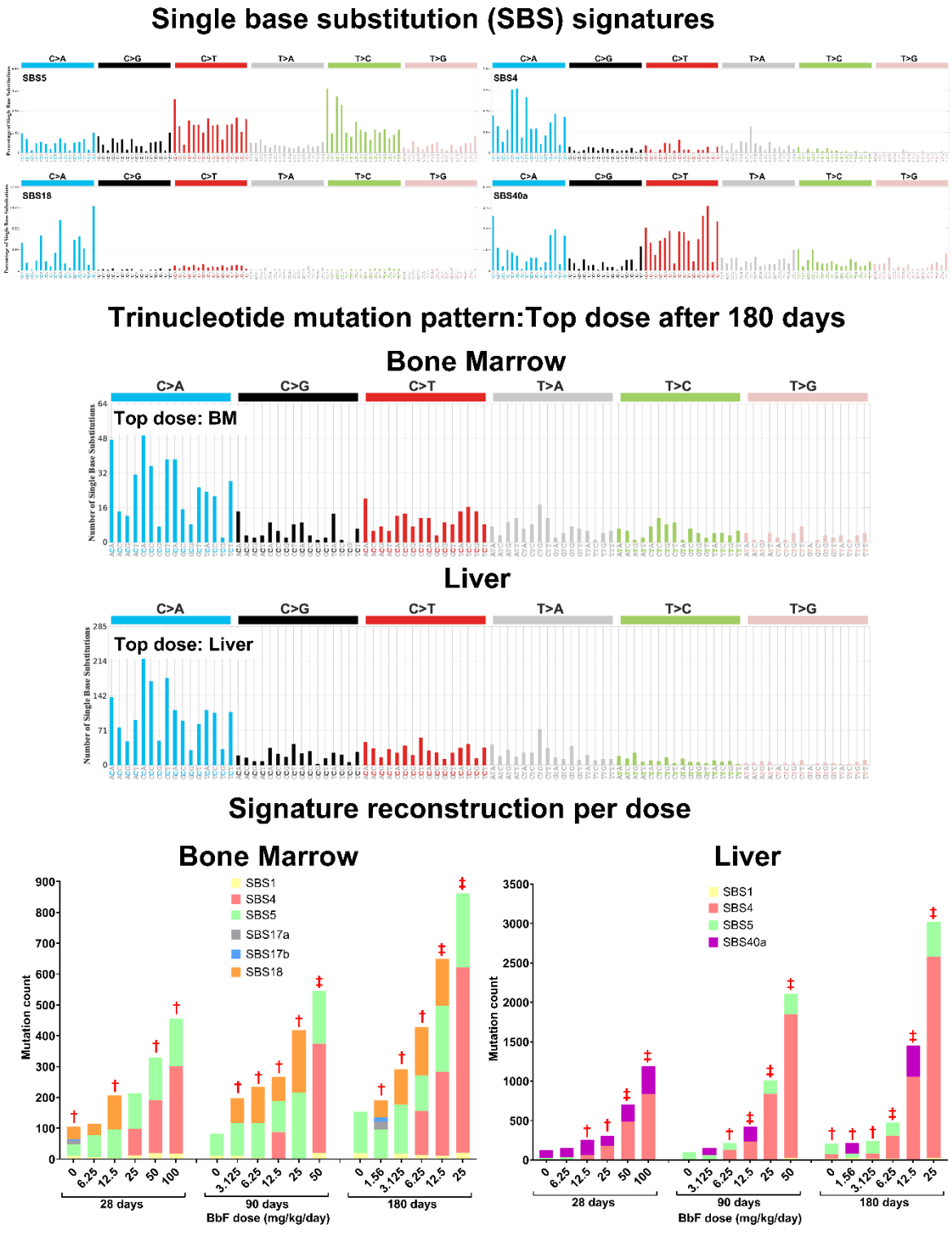


Figure 13: Overview of the C>A (blue), C>G (black), C>T (red), T>A (grey), T>C (green) and T>G (pink) mutation profiles of the mouse single base substitutions (SBS) SBS5, SBS4, SBS18 and SBS40a (top) and the trinucleotide mutation pattern (middle) of the bone marrow and liver top dose after 180 days of exposure to 25 mg/kg/day of BbF. The bottom panels display the results of the signature reconstruction of each bone marrow (left) and liver (right) dose, with indication of the total mutation count and the contribution of SBS1 (yellow), SBS4 (red), SBS5 (green), SBS17b (blue), SBS18 (orange), and SBS40a (purple). Cosine similarities are highlighted as † = greater than 0.8 and ‡ = greater than 0.9

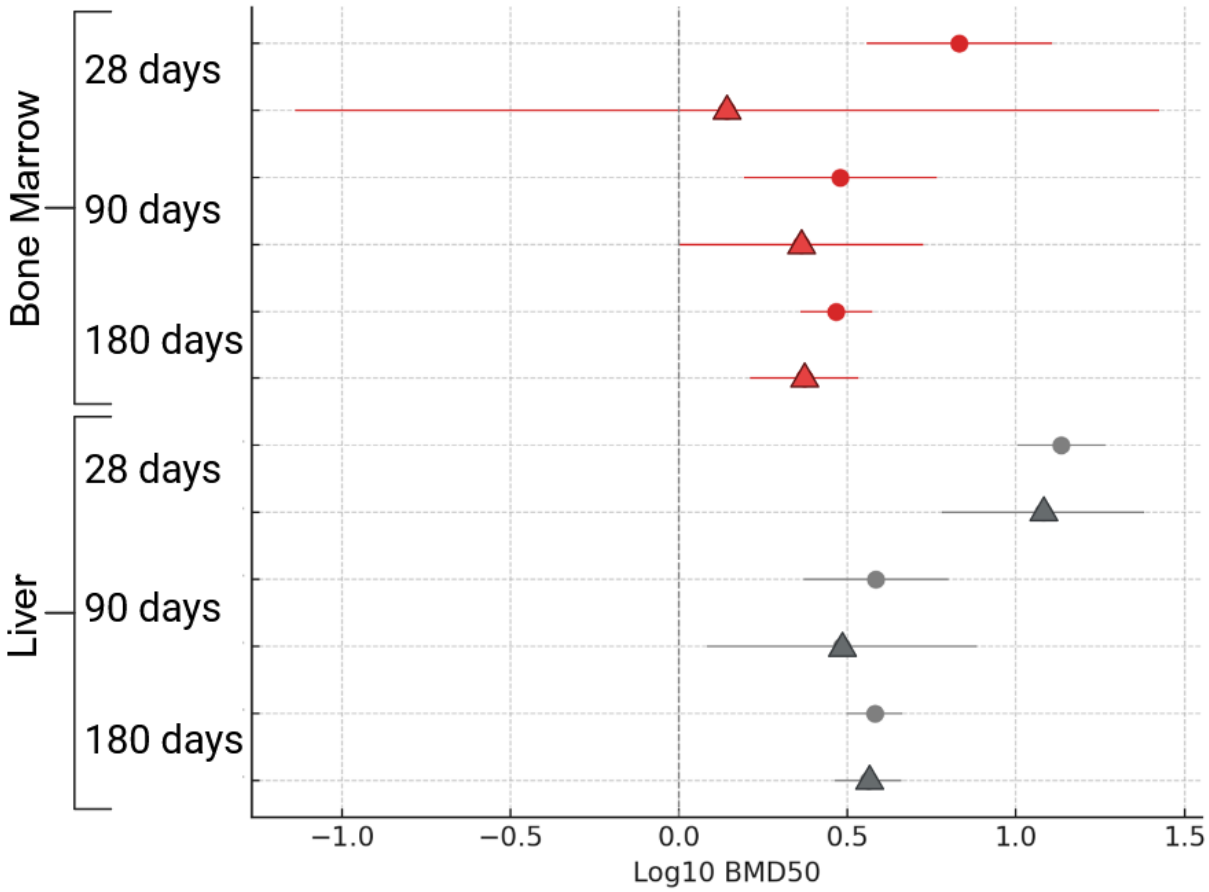


Figure 14: Log_{10} confidence intervals (BMDL to BMDU) of benchmark dose (BMD) modeling using a 50% benchmark response on the mutation frequency minimum (circle) and mutation frequency maximum (triangle). BMD modeling on bone marrow is indicated in red, and liver is indicated in grey.

Chapter 4: Comparative *In Vivo* Mutagenicity Assessment of Benzo[*b*]fluoranthene: Concordance of Error-Corrected Sequencing and OECD Mutation Assays

Modified from: David M. Schuster, Gu Zhou, Danielle LeBlanc, Matthew J. Meier, Paul A. White, Andrew Williams, Jesse J. Salk, Jake Higgins, Ann-Karin Olsen, Steven Dertinger, Jeffrey Bemis, Connie L. Chen, Carole L. Yauk and Francesco Marchetti. Comparative *In Vivo* Mutagenicity Assessment of Benzo[*b*]fluoranthene: Concordance of Error-Corrected Sequencing and OECD Mutation Assays. In preparation for submission to Archives of Toxicology.

4.1 Abstract

Conventional *in vivo* mutation assays are central to genetic toxicology, but their reliance on specialized transgenic models or phenotype-based endpoints can limit their integration into standard toxicology studies and their ability to provide mechanistic information. Error-corrected sequencing (ECS) has emerged as a powerful approach to detect *in vivo* mutagenesis while generating mechanistic information. We characterized the *in vivo* mutagenicity of benzo[*b*]fluoranthene (BbF) in MutaMouse males using ECS alongside established OECD test guideline methodologies. Specifically, we applied: the transgenic rodent (TGR) assay and Duplex Sequencing (DS) targeting the Mouse Mutagenesis Panel (MMP) or the *lacZ* transgene in bone marrow (BM) and liver to quantify mutagenic effects. We also performed DS on the *Pig-a* locus in BM and the phenotypic *Pig-a* assay in peripheral blood. BbF induced dose-related increases in the TGR in both BM and liver. Similarly, DS revealed dose-related increases in mutation frequency (MF) in the MMP and *lacZ* loci in both tissues. DS-derived MF at the MMP and *lacZ* loci strongly correlated with TGR mutant frequency, indicating close concordance between sequencing- and plaque-based endpoints. Benchmark dose confidence intervals overlapped across TGR and DS endpoints within each tissue, indicating similar points-of-departure and potency estimates. BbF-induced mutation spectra were heavily dominated by C:G>A:T substitutions, alongside smaller increases in C:G>T:A substitutions in both the MMP and *lacZ* locus, consistent with the expected profile of a metabolically activated polycyclic aromatic hydrocarbons. In contrast, neither the phenotypic *Pig-a* assay nor DS at the *Pig-a* locus exhibited a robust dose-related response indicating that *Pig-a* is

less informative within this study design. Ultimately, these findings show that DS reproduces the biological and quantitative hazard conclusions of conventional methodologies while providing sequence-resolved mechanistic information. These data robustly support the use of DS as an alternative approach for *in vivo* mutagenicity assessment in standard repeat-dose studies.

4.2 Introduction

4.2.1 Importance of *In Vivo* Mutagenicity Assessment

The accurate assessment of *in vivo* somatic mutagenicity is central to genetic toxicology and human health risk assessment (Albertini et al., 1993; Heflich et al., 2020). Mutations are permanent DNA sequence changes that arise when DNA damage escapes repair and becomes fixed (Chatterjee & Walker, 2017). Because somatic mutation accumulation is an initiating event in carcinogenesis (Parsons et al., 2025), robust *in vivo* methods are needed to measure mutation induction under physiologically realistic conditions and in relevant tissues (EFSA Scientific Committee, 2011). Mutagenic responses vary across tissues due to differences in exposure, metabolic activation, DNA repair, and cell proliferation. These biological variables can alter both the magnitude of the mutation burden and the tissue-specific potency of a chemical (Douglas et al., 2022; Long et al., 2016, 2018; Lynch et al., 2024). Accordingly, quantitative dose-response analysis of mutagenicity in target organs can support carcinogenic hazard identification, mode-of-action analysis, and point-of-departure estimation to inform human health risk assessment (Chepelev et al., 2023; Long et al., 2018; Parsons et al., 2025; Schuster et al., 2024; Vogelstein et al., 2013; White et al., 2025).

4.2.2 Conventional Approaches of *In Vivo* Mutagenicity Assessment

Conventional *in vivo* somatic mutagenicity assessment relies on well-established assays described in the Organisation for Economic Co-operation and Development (OECD) test guidelines (OECD, 2017, 2018a, 2025d, 2025c). These standardized assays are widely used for regulatory decision-making (EFSA Scientific Committee, 2011). The primary approaches for measuring *in vivo* gene mutations are the transgenic rodent (TGR) assay, OECD TG 488 (OECD, 2025d), and the erythrocyte phosphatidylinositol glycan class A (*Pig-a)* gene mutation assay, OECD TG 470 (OECD, 2025c).

The TGR assay quantifies mutations in recoverable reporter transgenes, such as *lacZ* or *cII*, that are integrated into the rodent genome (OECD, 2025d). Following chemical exposure, transgenes are recovered from genomic DNA isolated from target tissues and scored using phenotypic selection systems (Gingerich et al., 2014; OECD, 2025d). While effective for hazard identification, the TGR assay has inherent methodological constraints (Lambert et al., 2005). The assay measures an exogenous bacterial sequence with limited genomic context providing insufficient insight into genome-wide variability driven by transcriptional status, chromatin environment, and local sequence composition (Beal et al., 2015; Gingerich et al., 2014; Lambert et al., 2005). Furthermore, reliance on phenotypic selection obscures silent mutations (Beal et al., 2020; Lambert et al., 2005; OECD, 2025d). Finally, TGR mutant frequencies can be confounded by clonal expansion unless subsequent sequencing is performed to identify non-independent mutants (Beal et al., 2015; Lambert et al., 2005; Meier et al., 2019).

Conversely, the *Pig-a* assay measures mutagenesis in an endogenous X-linked locus (OECD, 2025c). This assay detects mutant erythrocytes based on the loss of glycosylphosphatidylinositol (GPI)-anchored surface markers using flow cytometry (Bryce et al., 2008; Dertinger et al., 2011; OECD, 2025c). However, the *Pig-a* assay is largely confined to the hematopoietic compartment and detects primarily loss-of-function mutations without providing inherent mutation-spectrum information (Dobrovolsky et al., 2010; Heflich et al., 2019). Additionally, mutant frequencies can be influenced by the expansion of viable *Pig-a*–mutant progenitors, particularly when mutations arise in self-renewing hematopoietic stem cells (Dertinger et al., 2021; Miura, 2014; Phonethepswath et al., 2010).

Together, the limitations of the TGR and *Pig-a* assays highlight the need for approaches that can directly measure mutations in endogenous DNA with higher sequence-level resolution (Menon & Brash, 2023; Salk & Kennedy, 2020; Yauk et al., 2026). Ideally, these new approaches should improve the capacity for quantitative comparisons across diverse tissues and endpoints while providing additional mechanistic information.

4.2.3 Mutagenicity Assessment Based on Error-Corrected Sequencing Technologies and Quantitative Modeling

Recent advances in error-corrected sequencing (ECS) provide a powerful alternative to conventional phenotype-based mutation assays. ECS enables the direct, high-resolution quantification of *in vivo* somatic mutation frequencies (MF) and spectra across diverse tissues (Maslov et al., 2022; Matsumura et al., 2019; Salk & Kennedy, 2020; Schmitt et al., 2012; Yauk et al., 2026; You et al., 2023). These approaches improve upon standard next-generation sequencing by incorporating molecular error correction strategies and to distinguish true biological mutations from technical artifacts (Valentine et al., 2020; Yauk et al., 2026). Consequently, background error rates are reduced to levels comparable to spontaneous MF (Kennedy et al., 2014; Schmitt et al., 2012). ECS therefore allows for the direct quantification of mutations in endogenous DNA without reliance on transgenic constructs or phenotypic selection (Marchetti et al., 2023b, 2023a; Yauk et al., 2026).

In addition to quantifying overall MF, ECS provides simultaneous data on mutation spectra and the frequency of clonally derived mutations (Dodge et al., 2023; Schuster et al., 2024; Smith-Roe et al., 2023). This sequence-resolved context enables the detection of a broad range of mutation subtypes across diverse genomic contexts and supports mechanistic interpretation of mutagenic processes (Kucab et al., 2026; Yauk et al., 2026).

ECS can also be applied to extracted DNA from essentially any tissue (Kennedy et al., 2014; Marchetti et al., 2023b). Thus, ECS-based mutation assessment can be readily integrated into standard repeat-dose toxicity studies, reducing the need for separate, standalone mutagenicity tests using transgenic models.

4.2.4 Validation Studies and Performance Evaluation

Substantial progress has been made in evaluating the performance of ECS for *in vivo* mutagenicity assessment. This progress is driven largely by empirical studies conducted alongside established regulatory assays (Ashford et al., 2025; Schuster et al., 2024; Yauk et al., 2026). To date, the majority of validation efforts have been generated using Duplex Sequencing (DS) (Esina et al., 2024; Valentine et al., 2020; Zhang et al., 2025). DS has been successfully applied across multiple rodent models,

tissues, exposure durations, and classes of mutagens (Ashford et al., 2025; Bercu et al., 2023; Dodge et al., 2023; LeBlanc et al., 2022; Schuster et al., 2024; Smith-Roe et al., 2023; Sun et al., 2026). These studies consistently demonstrate strong concordance between DS-derived MF and phenotypic TGR assay results (Ashford et al., 2025; Bercu et al., 2023). Furthermore, DS produces tissue-specific responses, dose–response relationships, and BMD estimates comparable to those derived from conventional endpoints (Ashford et al., 2025; Bercu et al., 2023; Long et al., 2016; Schuster et al., 2024). Collectively, these findings provide compelling evidence that DS reliably detects chemically induced somatic mutations *in vivo* with high sensitivity and reproducibility.

4.2.5 Existing Knowledge Gap for DS Validation

Most DS studies have used the Rat Mutagenesis or Mouse Mutagenesis Panel (MMP), a distributed set of endogenous loci designed to sample mutagenesis across the rodent genome (Bercu et al., 2023; Dodge et al., 2023; LeBlanc et al., 2022; Schuster et al., 2024; Smith-Roe et al., 2023).

Although recent studies have compared DS-derived MFs with conventional TGR endpoints, direct comparisons at identical reporter-gene targets remain limited (Ashford et al., 2025; Bercu et al., 2023; Sun et al., 2026). This distinction is important because regulatory decisions based on endogenous DS targets, such as the MMP, must be evaluated relative to the exogenous reporter transgenes used in established TGR assays. Moreover, available comparative datasets remain weighted toward alkylating agents and N-nitrosamines, highlighting the need to evaluate concordance across chemicals with diverse mutagenic mechanisms (Ashford et al., 2025; Bercu et al., 2023; Smith-Roe et al., 2023; Sun et al., 2026). Additional comparative studies are needed to bridge the biological and quantitative conclusions of established *in vivo* reporter-gene assays and DS, including their implications for hazard identification, potency ranking, point-of-departure determination, and mechanistic interpretation.

4.2.6 Benzo[*b*]fluoranthene as Model Compound for Comparative Evaluations

Polycyclic aromatic hydrocarbons (PAHs), including benzo[*a*]pyrene (BaP) and benzo[*b*]fluoranthene (BbF), have been used as model mutagens in DS validation studies (LeBlanc et al., 2022; Schuster et al., 2024). BbF is a ubiquitous PAH formed

by the incomplete combustion of organic material and environmental regulatory agencies designate BbF as a priority pollutant due to its widespread prevalence and significant toxicological concern (Keith, 2015; Zhang et al., 2020). Furthermore, the International Agency for Research on Cancer classifies BbF as a Group 2B possible human carcinogen (IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, 2010). BbF requires biotransformation by CYP450 enzymes to induce DNA damage. This metabolic activation generates highly reactive intermediates that form bulky covalent DNA adducts (Amin et al., 1982; Moorthy et al., 2015; Smith et al., 2022). Previous *in vivo* studies demonstrate that oral exposure to BbF induces robust, tissue-specific DNA damage and mutagenic responses in rodent models (Long et al., 2016; Schuster et al., 2024). Thus, BbF serves as model chemical for comparing the quantitative and biological performance of ECS against established phenotypic endpoints.

4.2.7 Study Objectives

In this study, we used a TG 488-compliant repeat-dose exposure design (OECD, 2025d) to evaluate BbF-induced somatic mutagenicity across conventional and sequence-based endpoints. We measured tissue-specific mutant frequency using the TGR assay in BM and liver and *Pig-a* mutant frequency in peripheral blood. We then compared these outcomes to DS-derived MF generated in endogenous MMP and in the *Pig-a* and *lacZ* loci. This design enabled direct comparison of DS with established reporter-gene and *Pig-a*-based endpoints, including comparisons at identical or conceptually matched loci.

The primary objectives of this study were to: (1) evaluate the concordance between conventional *in vivo* gene mutation assays and DS-derived mutation frequencies; (2) determine whether TGR and DS yield comparable quantitative potency estimates and points of departure using benchmark dose modeling; (3) assess how endogenous and exogenous target selection influences mutation detection and interpretation; and (4) characterize the predominant mutation types to determine if DS-derived mutation spectra support the expected BbF mutagenic mechanism and clarify locus-specific difference in mutagenic response. By integrating the two principal OECD guideline *in vivo* gene mutation assays with sequence-resolved mutation analyses

within a single experimental framework, our study provides empirical data to evaluate ECS as a valid alternative to conventional assays.

4.3 Methods

4.3.1 Animal Exposure and Tissue Collection

Male MutaMouse (*Mus musculus*, CD2F1 background) animals were handled and exposed at Health Canada. The study was approved by the Health Canada Ottawa Animal Care Committee (No: 2021-008). All procedures strictly adhered to Canadian Council on Animal Care guidelines. Mice were maintained under standard conditions (approximately 21 °C, ~50% humidity, 12h light/12h dark cycle) with food and water provided *ad libitum*. Eight- to ten-week-old males were exposed to BbF (CAS 205-99-2) dissolved in olive oil. The compound was administered once daily by oral gavage for 28 consecutive days. Animals (n = 8) were randomly assigned to receive 0, 6.25, 12.5, 25, 50, or 100 mg/kg body weight/day BbF based on previously established dose-response data (Long et al., 2016). In accordance with OECD TG 488 sampling recommendations, animals were euthanized three days after the final administration. Bone marrow (BM) and liver tissues were immediately collected, flash-frozen and stored for downstream analyses.

Peripheral blood for the *Pig-a* assay was collected from an independent, parallel cohort of MutaMouse males. These mice were exposed to BbF under an identical 28-day daily oral gavage study design. Due to adjusted dose group selection for prolonged exposure durations, this cohort included an additional 10 mg/kg/day dose group. Peripheral blood was collected from the facial vein on day 29 and shipped to Litron Laboratories for downstream analysis (Schuster et al., 2026, in preparation).

4.3.2 Transgenic Rodent Gene Mutation Assay

We used the TGR assay to quantify mutant frequencies, mutants per Plaque Forming Unit (pfu), in the *lacZ* transgene in accordance with OECD Test Guideline 488 (OECD, 2025d). Genomic DNA was isolated from flash-frozen BM and liver tissues by phenol-chloroform extraction (Gingerich et al., 2014). The bacteriophage shuttle vector carrying the *lacZ* reporter was rescued from the genomic DNA using packaging extracts from Agilent Technologies (Transpack Packaging Extract Catalog #200220). The resulting phages were used to infect selective *E. coli* host cells. Mutant plaques were identified using the phenyl-β-D-galactoside (P-Gal) positive selection system, as

described for the MutaMouse TGR assay (Douglas et al., 1997; Gingerich et al., 2014). For each sample, total plaque-forming units (pfu) were enumerated on non-selective titer plates. Concurrently, mutant plaques were enumerated on P-Gal selection plates. Mutant frequency was subsequently calculated as the ratio of mutant pfu to total pfu. Assay outputs are reported per tissue and dose group as mutant frequency with standard error of the mean (SEM), alongside plaque counts. To assess treatment-related effects, dose groups were compared with the concurrent vehicle control using one-way ANOVA followed by a Dunnett's post hoc test to control for multiple comparisons.

4.3.3 Phosphatidylinositol Glycan-Class A Gene Assay

Pig-a mutant frequencies were measured in peripheral blood using flow cytometry (Dertinger et al., 2011). This assay detects mutant erythrocytes based on the loss of glycosylphosphatidylinositol (GPI)-anchored surface markers, specifically CD24 in mice (OECD, 2025a). Flow cytometric analysis was conducted at Litron Laboratories (Rochester, NY) using MutaFlow® reagents. Briefly, peripheral blood samples were processed using immunomagnetic enrichment to concentrate reticulocytes and mutant cells prior to analysis (Dertinger et al., 2021). This enrichment facilitated the interrogation of at least 1×10^6 RET and RBC per animal, conforming to established statistical power recommendations for this assay (Bryce et al., 2008; Dertinger et al., 2021). Assay outputs included the percentage of reticulocytes (%RET), mutants per reticulocyte (mut/RET), and mutants per mature red blood cell (mut/RBC). Pig-a results were Box–Cox transformed to better satisfy homogeneity assumptions, and treatment groups were then compared with the concurrent vehicle control using one-way ANOVA followed by Dunnett's post hoc test to control for multiple comparisons (Dertinger et al., 2021; Igl et al., 2018).

4.3.4 Duplex Sequencing Library Preparation and Sequencing Metrics

DS was used to quantify MF in genomic DNA extracted from BM and liver tissues. This approach relies on error-corrected, double-strand consensus reads. Libraries were prepared from high-integrity genomic DNA with DNA integrity numbers greater than 8.0 and concentrations greater than 50 ng/µL.

Samples were ligated with DS adapters containing unique molecular identifiers. Adapter-ligated DNA subsequently underwent PCR indexing. Finally, target regions

were enriched via hybrid-capture using predefined DS mutagenesis panels (Dodge et al., 2023; LeBlanc et al., 2022; Schuster et al., 2024). DS was performed using three distinct target capture designs to facilitate direct method-to-method comparisons

1. Liver DS libraries were constructed using 500 ng of genomic DNA and the standard MMP. The MMP targets 20 endogenous loci (on chr1–chr19; 48kb of target DNA in total) to provide a distributed sampling of the mouse genome (LeBlanc et al., 2022).
2. 30 ng of either BM or liver genomic DNA was used to build libraries using a custom *lacZ* panel (pan00335). This panel targets the integrated λgt10 transgene locus, encompassing the *cII* region, flanking sequences, and the *lacZ* coding sequence. This specific design supports direct comparisons between phenotypic mutant frequencies and sequencing-based MF at the identical transgenic locus.
3. 500 ng of BM genomic DNA was used to build libraries on a custom MMP+*Pig-a* panel (pan00354). This panel captures the 20 endogenous MMP loci alongside the endogenous *Pig-a* locus on the X chromosome. This multiplexed approach enabled simultaneous DS-based MF estimation for both *Pig-a* and the MMP from the same DNA extract.

Libraries were sequenced on an Illumina NovaSeq S4 platform. Sequencing aimed to achieve sufficient depth across all targets to enable the robust detection of rare *de novo* mutations. Samples were required to meet established DS minimum performance criteria consistent with prior *in vivo* mutagenicity assessments (LeBlanc et al., 2022; Schuster et al., 2024; Valentine et al., 2020). These thresholds required at least 700 million informative duplex bases per sample and a mean on-target duplex depth greater than 10,000x (LeBlanc et al., 2022; Schuster et al., 2024).

4.3.5 Processing of Error-Corrected Sequencing Data using MutSeqR

Demultiplexed FASTQ files were processed using the DS bioinformatics workflow hosted on the DNAnexus platform. This pipeline computationally removed sequencing and amplification artifacts through double-strand consensus building. Following this error correction, variant calling was performed on the duplex consensus reads using standard pipeline settings. This workflow generated duplex consensus alignments and per-sample variant calls (Kennedy et al., 2014; Salk & Kennedy, 2020;

Schmitt et al., 2012). The resulting mutation output files were imported into the R statistical computing environment for downstream processing. This analysis used MutSeqR, an open-source package designed for the standardized analysis of ECS mutation data (Dodge et al., 2025).

Within MutSeqR, variants were first annotated based on the reference genome. These variants were categorized by mutation class (e.g., SNVs, insertions, deletions) to facilitate spectrum analyses. Variants were subsequently filtered to remove putative germline events and technical artifacts using criteria aligned with established workflows that employ variant allele frequency (VAF) thresholds and panel-specific filtering parameters (Dodge et al., 2025). Following filtration, MF were calculated as the number of mutations divided by the total number of informative duplex bases. To account for clonal expansion, two complementary metrics were reported. First, MFmin was calculated by collapsing identical mutations observed more than once within a single sample into one independent count. Second, MFmax was calculated by counting all repeated observations of a mutation within a sample as independent events (Dodge et al., 2023, 2025; LeBlanc et al., 2022; Schuster et al., 2024).

Dose effects on MF were evaluated using a generalized linear mixed model (Dodge et al., 2025). This model used a quasibinomial framework to account for overdispersion in the mutation count data. Pairwise comparisons against concurrent vehicle controls were conducted using appropriate multiple-testing adjustments. Mutation spectra were assessed using MutSeqR utilities. Specifically, subtype proportions were compared using modified contingency-table approaches designed for rare-event data and independent mutation counting (Dodge et al., 2025; Schuster et al., 2024).

4.3.6 Comparative Quantitative Analyses: Pearson Correlation and Benchmark Dose Modeling

Pearson correlation coefficients were calculated to assess the quantitative association between DS MF and TGR mutant frequencies (De Siqueira Santos et al., 2014; Mughal et al., 2010). Statistical significance was evaluated using a two-sided t-test. Analyses were performed on both raw and log10-transformed data. The logarithmic transformation was applied to specifically address heteroscedasticity and non-linearity within the datasets. Additionally, Spearman rank correlation was

calculated as a non-parametric measure of association independent of distributional assumptions. Confidence intervals for the correlations were derived using the Fisher z-transformation.

These statistical analyses used individual-animal data across the complete BbF dose range for both BM and liver. Correlations were calculated between the phenotypic TGR mutant frequencies and DS MF for both the MMP and *lacZ* loci. For DS, MFmin served as the primary metric for cross-method comparisons. We selected MFmin for our primary approach to mathematically reduce the confounding influence of clonally expanded events (Dodge et al., 2025; Schuster et al., 2024). To thoroughly evaluate the impact of clonal expansion on assay concordance, parallel correlation analyses were also performed using the MFmax metric. Concordance analyses involving the *Pig-a* endpoints were not performed because neither the phenotypic *Pig-a* assay nor DS at the endogenous *Pig-a* locus exhibited a robust dose-related response under the applied study conditions.

4.4 Results

4.4.1 Increase in Mutant Frequency Detectable using the TGR Assay

We quantified BbF-induced mutant frequencies in BM and liver using the phenotypic TGR assay following 28+3 days of oral exposure. This analysis established the primary regulatory reference endpoint against which DS-derived data were subsequently compared.

In BM, the vehicle control mutant frequency was $3.63 \pm 0.6 \times 10^{-5}$. A statistically significant increase to $10.0 \pm 0.4 \times 10^{-5}$ occurred by 12.5 mg/kg/day ($p < 0.05$). All higher dose groups exhibited significant increases ($p < 0.01$). At the maximum dose of 100 mg/kg/day, the mutant frequency reached $37.7 \pm 2.1 \times 10^{-5}$. This represents a 10.4-fold increase relative to the vehicle control (Figure 15 and Appendix D3 Table 1).

In liver, mutant frequency increased significantly from $6.20 \pm 0.6 \times 10^{-5}$ in vehicle controls to $22.9 \pm 2.0 \times 10^{-5}$ at 25 mg/kg/day ($p < 0.05$). All subsequent higher doses were statistically significant ($p < 0.01$). At 100 mg/kg/day, the mutant frequency increased to $86.8 \pm 14.2 \times 10^{-5}$, representing a 14.0-fold increase over the vehicle control (Figure 15 and Appendix D3 Table 1).These data confirm dose-dependent increases in *lacZ* mutant frequencies in BM and liver following oral exposure to BbF.

4.4.2 Absence of Mutant Frequency Induction in the *Pig-a* Assay

We next evaluated BbF-induced mutagenicity using the conventional *Pig-a* assay in peripheral blood, providing a second OECD gene mutation endpoint focused on the hematopoietic compartment. For this analysis, we assessed the percentage of reticulocytes (%RET), mutants per reticulocyte (mut/RET), and mutants per mature red blood cell (mut/RBC) with each dose group comprised of eight animals (n = 8).

The vehicle control %RET of 3.92 ± 0.275 increased to 5.25 ± 0.25 following treatment with 25 mg/kg/day of BbF ($p < 0.01$). Significant increases ($p < 0.01$) to 5.56 ± 0.35 and 5.73 ± 0.36 were also observed at 50 mg/kg/day and 100 mg/kg/day, respectively. At the maximum dose, %RET was 1.46-fold higher than the vehicle control (Appendix D3 Figure 1 and Appendix D3 Table 2).

For mut/RET, a significant increase only occurred at the maximum dose of 100 mg/kg/day. At this dose, mut/RET increased from $3.8 \pm 1.7 \times 10^{-7}$ in vehicle controls to $82.8 \pm 50.0 \times 10^{-7}$. However, we observed inter-animal variability, and statistical significance was primarily driven by a single elevated sample. Similarly, the vehicle control for mut/RBC increased significantly from $4.0 \pm 2.3 \times 10^{-7}$ in controls to $18.3 \pm 5.0 \times 10^{-7}$ at 100 mg/kg/day (Figure 16 and Appendix D3 Table 2). This response was also strongly influenced by the identical elevated sample.

Overall, BbF significantly increased %RET at doses ≥25 mg/kg/day. This indicates biologically meaningful exposure and toxicity to the hematopoietic compartment. Nevertheless, *Pig-a* mutant cell frequencies did not exhibit a robust, dose-related response under these study conditions.

4.4.3 Duplex Sequencing based Mutagenicity Assessment

We next used DS to examine whether BbF-induced mutagenicity was consistently detected across tissues and genomic targets, and whether the conventional TGR and Pig-a phenotypic responses were consistent with mutagenesis measured by DS at the endogenous MMP and Pig-a loci and the *lacZ* transgene. Libraries targeting the MMP and *lacZ* were analyzed in both BM and liver. The *Pig-a* locus was additionally analyzed in BM. Across all DS endpoints, four animals (n = 4) were analyzed per dose group.

In BM, DS-derived MF at endogenous MMP targets increased with BbF dose. The MMP MFmin in the vehicle control was 4.13 ± 0.13 × 10^{-8}. A statistically significant increase first occurred at 12.5 mg/kg/day ($p < 0.05$), with an MFmin of 7.54 ± 0.67 × 10^{-8}. All higher doses were statistically significant ($p < 0.01$). At the maximum dose of 100 mg/kg/day, MMP MFmin reached 14.8 ± 0.87 × 10^{-8}. This corresponds to a 3.6-fold increase relative to the vehicle control. MMP MFmax increased from 4.50 ± 0.26 × 10^{-8} in vehicle controls to 17.6 ± 1.3 × 10^{-8} at 100 mg/kg/day, with an increase of 3.9-fold at the top dose (Figure 17 and Appendix D3 Table 3).

MFmin at the *lacZ* locus also increased with dose in BM. The vehicle control *lacZ* MFmin was 18.7 ± 2.5 × 10^{-8}. A statistically significant increase first occurred at 50 mg/kg/day ($p < 0.01$), rising to 47.3 ± 5.8 × 10^{-8}. At 100 mg/kg/day, *lacZ* MFmin reached 82.0 ± 20.8 × 10^{-8}, corresponding to a 4.4-fold increase over the vehicle control. *LacZ* MFmax exhibited greater between-animal variability. MFmax reached statistical significance only at the 100 mg/kg/day dose, where it increased from 22.1 ± 3.2 × 10^{-8} in vehicle controls to 89.6 ± 23.9 × 10^{-8} (Figure 17 and Appendix D3 Table 4).

In contrast to the dose-dependent effects observed at MMP and *lacZ* targets, DS analysis of the endogenous *Pig-a* locus in BM showed no statistically significant increase in MF at any dose. *Pig-a* MFmin was 8.2 ± 1.0 × 10^{-8} in vehicle controls. At 100 mg/kg/day, it was 10.0 ± 1.7 × 10^{-8}, a 1.2-fold increase that was not statistically significant. Similarly, *Pig-a* MFmax was 130 ± 23.4 × 10^{-8} in vehicle controls and 123 ± 22.0 × 10^{-8} at the top dose (Figure 18 and Appendix D3 Table 5).

In liver, DS-derived MF in MMP targets showed a strong dose-related increase in BbF-treated mice. A statistically significant increase first occurred at 25 mg/kg/day ($p < 0.05$), where MFmin increased from 5.3 ± 0.74 × 10^{-8} in vehicles controls to 11.1 ± 0.43 × 10^{-8}. All higher doses showed significant increases ($p < 0.01$). At 100 mg/kg/day, liver MMP MFmin reached 36.3 ± 4.1 × 10^{-8}, corresponding to a 6.8-fold increase over the vehicle control. MMP MFmax increased from 6.97 ± 1.8 × 10^{-8} in vehicle controls to 39.8 ± 6.1 × 10^{-8} at 100 mg/kg/day (5.7-fold change) and followed the same significance pattern (Figure 17 and Appendix D3 Table 3).

Liver *lacZ* MFmin also increased with BbF dose, rising from 17.1 ± 1.6 × 10^{-8} in controls to 43.9 ± 3.3 × 10^{-8} at 50 mg/kg/day, the first dose with a significant increase ($p < 0.01$). At 100 mg/kg/day, *lacZ* MFmin reached 73.1 ± 10.5 × 10^{-8}, corresponding to a 4.3-fold increase. As observed in BM, *lacZ* MFmax exhibited greater variability. MFmax reached statistical significance only at the 100 mg/kg/day dose, where it increased from 19.0 ± 1.8 × 10^{-8} in vehicle controls to 76.1 ± 11.8 × 10^{-8} (Figure 17 and Appendix D3 Table 4).

Overall, DS MF measured at the MMP and *lacZ* loci exhibited clear, dose-related increases in both BM and liver. Conversely, the endogenous *Pig-a* locus in BM failed to show a dose-related mutational response.

4.4.4 BbF Mutation Spectrum Driven by C:G>A:T and C:G>T:A Mutations

We next examined DS-derived mutation spectra to determine whether BbF induced the expected PAH-associated mutation profile and whether this profile was consistent across tissues and genomic targets. Mutation subtype analyses were performed for the MMP and *lacZ* panels in BM and liver, and for the *Pig-a* locus in BM ($n = 4$ per group). Subtype analyses are reported using MFmin to inform unique chemically induced mutation for MMP and *lacZ*. Both MFmin and MFmax are reported for the Pig-a locus because the data showed a high contribution of indels, especially when repeated mutation calls were retained.

In BM, the BbF-induced mutation spectrum at the MMP targets was dominated by increases in C:G>A:T transversions, with a smaller increase in C:G>T:A transitions. In the vehicle control C:G>A:T MFmin was 2.1 ± 0.4 × 10^{-8}, rising significantly to 5.5 ± 0.8 × 10^{-8} at 12.5 mg/kg/day ($p < 0.05$). At the maximum dose of 100 mg/kg/day, C:G>A:T MFmin reached 12.8 ± 1.4 × 10^{-8}, corresponding to a 6.1-fold increase. C:G>T:A MFmin increased from 2.4 ± 0.2 × 10^{-8} in vehicle controls to 7.8 ± 0.8 × 10^{-8} at 100 mg/kg/day (3.3-fold increase). More broadly, all substitution subtypes were significantly increased across the MMP targets in BM, with the exception of insertions and deletions (Figure 19 and Appendix D3 Table 6).

C:G>A:T was also the dominant induced subtype in the *lacZ* locus of BM. This subtype increased significantly from the vehicle control C:G>A:T MFmin baseline of 5.9 ± 2.2 × 10^{-8} to 15.7 ± 1.9 × 10^{-8} at 12.5 mg/kg/day ($p < 0.05$). At 100 mg/kg/day, it

reached 43.3 ± 4.2 × 10^{-8}, corresponding to a 7.3-fold increase over the vehicle control. The secondary subtype was C:G>T:A. The C:G>T:A MFmin increased from 21.9 ± 2.7 × 10^{-8} in vehicle controls to 44.9 ± 11.2 × 10^{-8} at 100 mg/kg/day ($p < 0.05$). The BM *lacZ* panel showed significant increases in all substitution subtypes except C:G>G:C, insertions, and deletions (Figure 19 and Appendix D3 Table 7).

In contrast, the *Pig-a* mutation profiles in BM did not show dose-related increases for any subtype under either MFmin or MFmax assumptions. For MFmin, no single mutation subtype dominated the spectrum, while insertions and deletions had elevated MFmin even in vehicle controls. For example, insertions MFmin was 3.2 ± 0.4 × 10^{-8} in vehicle controls and 2.9 ± 0.7 × 10^{-8} at 100 mg/kg/day, while C:G>A:T MFmin was 1.8 ± 0.8 × 10^{-8} in vehicle controls and 3.0 ± 1.2 × 10^{-8} at the top dose (Figure 20 Appendix D3 Table 8). MFmax indels were more abundant than any other substitution classes across all doses. Insertion MFmax was 52.6 ± 9.2 × 10^{-8} and 54.5 ± 9.3 × 10^{-8}, and deletion MFmax was 75.3 ± 19.0 × 10^{-8} and 64.3 ± 13.8 × 10^{-8}, in vehicle controls and at 100 mg/kg/day, respectively (Appendix D3 Figure 4).

The MMP liver spectrum was dominated by C:G>A:T and C:G>T:A substitutions. C:G>A:T MFmin significantly increased from 1.2 ± 0.3 × 10^{-8} in vehicle controls to was 9.0 ± 0.8 × 10^{-8} at 25 mg/kg/day ($p < 0.05$), the lowest dose showing an effect. At 100 mg/kg/day, C:G>A:T MFmin reached 35.5 ± 5.6 × 10^{-8}, corresponding to a 29.3-fold increase over the vehicle control. The C:G>T:A MFmin increased from 4.8 ± 0.6 × 10^{-8} in vehicle controls to 16.5 ± 1.3 × 10^{-8} at 100 mg/kg/day (3.5-fold increase). As observed with the BM MMP, all substitution subtypes were significantly increased, whereas insertions and deletions were not increased (Figure 19 and Appendix D3 Table 6).

C:G>A:T was again the dominant induced subtype in the liver *lacZ* locus. The vehicle control C:G>A:T MFmin was 1.6 ± 0.2 × 10^{-8}. A statistically significant increase first occurred at 25 mg/kg/day ($p < 0.05$). At this dose, the C:G>A:T MFmin was 15.6 ± 1.7 × 10^{-8}. At 100 mg/kg/day, it reached 70.3 ± 9.9 × 10^{-8}. This corresponds to a 44.6-fold increase over the vehicle control. In contrast to the other responsive endpoints, liver *lacZ* did not show a statistically significant increase in the C:G>T:A subtype. Significant responses in these loci were limited to C:G>A:T, C:G>G:C and T:A>A:T substitutions (Figure 19 and Appendix D3 Table 7).

Overall, BbF-induced DS spectra in BM and liver were characterized by a dominant induction of C:G>A:T substitutions, with secondary induction of C:G>T:A substitutions. These patterns remained consistent across the MMP or *lacZ* panels. In contrast, the endogenous *Pig-a* locus in BM showed no significant subtype response. Furthermore, the Pig-a locus exhibited a comparatively stronger contribution of indels within the spectrum analysis.

4.4.5 Quantitative concordance between TGR and DS endpoints

We next evaluated quantitative concordance between DS-derived MF on the MMP and *lacZ* loci and the phenotypic TGR assay mutant frequencies. Correlation analyses were performed for DS at the MMP and lacZ targets, which showed dose-related responses suitable for quantitative comparison with TGR. Because neither the phenotypic Pig-a assay nor DS at the Pig-a locus exhibited a robust dose-related increase in MF, Pig-a-based endpoints were not included in this analysis.

In BM, the DS MFmin for the MMP was strongly correlated with TGR mutant frequencies (Pearson correlation r = 0.89; 95% CI: 0.77–0.95; p < 0.001) (Figure 21). This robust association was maintained after log10 transformation (r = 0.84, 95% CI: 0.67–0.93, p < 0.001). Furthermore, Spearman rank analysis confirmed a strong monotonic relationship (ρ = 0.88). A positive association was also observed between the DS *lacZ* panel and the TGR endpoint. The linear Pearson correlation was r = 0.70 (95% CI: 0.43–0.85, p < 0.001) (Figure 21). Log10 transformation increased this correlation to r = 0.82 (95% CI: 0.63–0.92, p < 0.001). Spearman correlation similarly confirmed a strong monotonic relationship for this target (ρ = 0.89) (Appendix D3 Table 9).

Parallel analyses using the MFmax approach revealed reduced concordance in BM. For the MMP, the MFmax Pearson correlation fell to r = 0.46 (95% CI: 0.06–0.73, p = 0.025) on the linear scale and r = 0.59 (95% CI: 0.24–0.80, p = 0.003) after log10 transformation. The Spearman coefficient indicated a moderate monotonic relationship (ρ = 0.70). For the DS *lacZ* target, the MFmax approach demonstrated weak or absent linear associations. The Pearson correlation was r = 0.09 (95% CI: −0.32–0.48, p = 0.67) on the linear scale (Supporting Figure 3) and r = 0.44 (95% CI: 0.05–0.72, p = 0.031) after log10 transformation. Nevertheless, a monotonic relationship was preserved (ρ = 0.58) (Appendix D3 Table 9).

In the liver, correlations between the two methodologies were similarly strong. The Pearson correlation for the DS MMP MFmin versus the TGR mutant frequency was r = 0.91 (95% CI: 0.80–0.96, p < 0.001) on the linear scale (Figure 21). Log10 transformation slightly increased this correlation (r = 0.93, 95% CI: 0.84–0.97, p < 0.001). Spearman analysis confirmed a strong monotonic association (ρ = 0.90). The DS *lacZ* loci also exhibited excellent agreement with the TGR endpoint. The Pearson correlation was r = 0.87 (95% CI: 0.70–0.94, p < 0.001) on the linear scale (Figure 21) and r = 0.93 (95% CI: 0.84–0.97, p < 0.001) after log10 transformation. Spearman analysis confirmed this strong relationship (ρ = 0.93) (Appendix D3 Table 9).

In the liver, concordance was maintained when evaluated under the MFmax approach for the MMP. The MFmax Pearson correlations were r = 0.88 (95% CI: 0.73–0.95, p < 0.001) on the linear scale and r = 0.87 (95% CI: 0.71–0.94, p < 0.001) after log10 transformation. The Spearman coefficient confirmed a strong relationship (ρ = 0.82). However, the DS *lacZ* loci evaluated under MFmax assumption showed a loss of linearity. The Pearson correlation dropped to r = 0.07 (95% CI: −0.34–0.46, p = 0.74) on the linear scale and r = 0.49 (95% CI: 0.10–0.74, p = 0.016) after log10 transformation. Despite this loss of linearity, a significant Spearman correlation remained (ρ = 0.74) (Appendix D3 Table 9).

.

4.4.6 Benchmark Dose Modeling to Compare Mutagenic Potency Across Endpoints

We next used benchmark dose (BMD) modeling to determine whether responsive DS and TGR endpoints produced comparable point-of-departure estimates within each tissue. Models were fit using a benchmark response (BMR) of 50%. Results are reported as 90% confidence intervals representing the lower and upper bounds (BMDL–BMDU) in mg/kg/day. As above, BMD modeling was not applied to the phenotypic *Pig-a* assay or the DS *Pig-a* target because neither exhibited a robust, dose-related increase in MF.

Within each tissue, the BMD confidence intervals overlapped across all modeled endpoints. In BM, the BMD confidence interval for the DS MMP endpoint was

3.60 to 12.8 mg/kg/day. For the DS *lacZ* loci, the interval was 7.62 to 32.5 mg/kg/day. For the phenotypic TGR assay, the interval was 1.11 to 7.79 mg/kg/day (Figure 22 and Appendix D3 Table 10). In liver, the BMD confidence interval for the DS MMP endpoint was 10.1 to 18.4 mg/kg/day. For the DS *lacZ* loci, the interval was 10.6 to 29.7 mg/kg/day. For the phenotypic TGR assay, the interval was 5.08 to 15.4 mg/kg/day (Figure 22 and Appendix D3 Table 10). This overlap indicates similar quantitative potency estimates for both the DS and TGR methodologies under this modeling framework.

4.5 Discussion

Our study shows that the mutagenic activity of BbF is robustly captured by sequence-resolved mutation analysis and that the responses observed carry a clear biological signature. Across BM and liver, DS detected dose-related increases in MF at both endogenous MMP targets and the *lacZ* transgene, closely matching changes measured using the conventional TGR assay. Importantly, the mutation spectra were highly consistent across responsive tissues and target classes, with BbF inducing a profile dominated by C>A transversions and smaller increases in C>T transitions. This spectrum is consistent with the expected mutagenic activity of a metabolically activated PAH and demonstrates the added value of DS for linking MF measurements to mutational mechanism. Quantitative modeling showed overlapping BMD confidence intervals within tissues for DS and TGR, indicating that they produce concordant estimates of mutagenic potency and points of departure. However, the Pig-a endpoints told a different biological story: the phenotypic Pig-a assay produced only a weak and variable response, while DS at the endogenous Pig-a locus failed to recover the characteristic BbF mutation spectrum and instead showed a high contribution of insertion and deletion calls. Together, these findings show that DS can reproduce the biological and quantitative conclusions of established *in vivo* mutation assays while also revealing how mutagenic responses vary by tissue, target, and locus-specific sequence context.

Oral exposure to BbF induced robust, dose-related increases in somatic mutation burdens in both BM and liver tissues. This mutagenic response was consistently captured across multiple analytical platforms, including the conventional TGR assay and error-corrected DS targeting both the exogenous *lacZ* locus and the

endogenous MMP loci. The magnitude of mutagenesis directly reflected the distinct biological characteristics of the target tissues. In the highly proliferative BM, significant increases in mutation frequency occurred at lower administered doses, consistent with the rapid fixation of DNA damage during frequent cell division (Dertinger et al., 2012; Douglas et al., 2022; Yeh et al., 2024). Conversely, the liver exhibited a greater maximal fold-increase at the highest administered doses. This pronounced hepatic response aligns with the liver's capacity for xenobiotic biotransformation, which generates reactive intermediates required for PAH mutagenesis (Amin et al., 1982; Long et al., 2016; Mass et al., 1996; Moorthy et al., 2015; Schuster et al., 2024). These findings confirm that BbF is a potent *in vivo* mutagen and demonstrate that the observed inter-tissue variation is driven by physiological differences such as metabolic activation, DNA repair, and cellular proliferation kinetics rather than a lack of mutagenic potential (Chatterjee & Walker, 2017; Douglas et al., 2022; Marchetti et al., 2021).

Sequence-based analysis directly reproduced the phenotypic hazard identification, with DS yielding biological conclusions identical to the conventional TGR assay. To enable a direct cross-platform comparison and mitigate the confounding influence of localized clonal expansion, DS data were evaluated MFmin, an approach formally recommended for the interpretation of error-corrected sequencing data, in comparable endpoints (Yauk et al., 2026). Both the endogenous MMP and the *lacZ* locus exhibited stable dose-response relationships that aligned directly with the plaque-based TGR mutant frequencies. This strong correlation between the phenotypic reporter readout and direct sequence-resolved analysis confirms that DS reliably captures *in vivo* mutagenicity.

Despite the strong concordance between the assays, target sequence selection influenced the sensitivity of the mutational response. The endogenous MMP exhibited less inter-animal variability and detected statistically significant increases at lower doses compared to the DS *lacZ* panel. This variance is attributable to the fundamental biological composition of the *lacZ* transgene (OECD, 2022; Shwed et al., 2010). As a heavily methylated, transcriptionally silent multi-copy array located within heterochromatin (Shwed et al., 2010), the *lacZ* locus is not subjected to endogenous transcription-coupled DNA repair (Lefkofsky et al., 2015). This reduced activity of targeted repair increases the accumulation of both spontaneous and chemically induced mutations in the transgene compared to actively transcribed endogenous loci

(Meier et al., 2019; Skopek et al., 1996). However, the *lacZ* transgene contains a sufficient representation of the specific sequence contexts targeted by BbF (Beal et al., 2015, 2023; Schuster et al., 2024). Consequently, the robust agreement between DS and TGR assay demonstrates that locus-specific biological biases or differing sequence compositions do not distort the overall qualitative hazard identification of BbF.

Sequence-resolved mutation analysis of the MMP and *lacZ* locus provided direct mechanistic insights into BbF-induced mutagenicity. Across both targeted panels and tissues, the mutation spectrum was dominated by C:G>A:T transversions, accompanied by smaller but consistent increases in C:G>T:A transitions. Crucially, this specific mutational profile was identical between the distributed endogenous MMP and the localized exogenous *lacZ* transgene. This reproducibility confirms that the observed mutational signal is determined by the specific mode of action of the exposed chemical rather than an endpoint-specific artifact or locus-specific bias. However, the magnitude of C:G>A:T substitution induction was greater in liver compared to BM (Amin et al., 1982; Moorthy et al., 2015; Wu et al., 2025). Nevertheless, the detection of this identical mutation spectrum in BM hints towards a systemic exposure of BbF or the reactive BbF metabolites (Ginsberg & Atherholt, 1989; LaVole et al., 1982).

The observed predominance of C:G>A:T substitutions aligns with established PAH mutagenesis (LeBlanc et al., 2022; Schuster et al., 2024). Similar to other PAHs, BbF undergoes metabolic activation to form reactive intermediates that subsequently form bulky covalent DNA adducts, preferentially at guanine bases. During DNA replication, these specific bulky lesions drive the formation of C:G>A:T transversions. The secondary induction of C:G>T:A transitions is also a characteristic signature of PAH exposure, which may reflect the formation of alternative DNA lesion classes or the secondary cellular processing of existing DNA damage (Amin et al., 1982; Long et al., 2016; Zhang et al., 2020). Therefore, the distinct mutational profile detected by DS provides direct molecular evidence that the measured response reflects canonical BbF mutagenesis.

This high mechanistic resolution highlights a fundamental advantage of error-corrected sequencing over conventional phenotypic mutation assays. While traditional assays are restricted to measuring overall phenotypic mutant frequencies, DS directly

characterizes mutation spectra and specific sequence contexts. By confirming that BbF induces the expected PAH-like mutational signature across both endogenous and transgenic loci, DS provides essential spectral evidence that strengthens the biological interpretation of the quantitative hazard identification

Quantitative modeling showed that DS and TGR produced comparable point-of-departure estimates within tissues, supporting DS as a quantitative as well as mechanistic alternative. BMD modeling demonstrated that the enhanced mechanistic resolution of sequence-based endpoints does not alter the overall interpretation of mutagenic potency. Within both bone marrow and liver tissues, the BMD confidence intervals derived from the sequence-resolved endogenous MMP and *lacZ* loci consistently overlapped with the intervals derived from the conventional phenotypic TGR assay. This overlapping response confirms that the two distinct analytical platforms yield highly comparable quantitative point-of-departure (PoD) estimates for BbF-induced mutagenicity.

This quantitative concordance aligns directly with previous comparative evaluations that have consistently demonstrated close agreement between DS- and TGR-derived potency estimates across different chemical classes(Ashford et al., 2025; Bercu et al., 2023; Dodge et al., 2023; Valentine et al., 2020). Crucially, the concordant BMD intervals were consistently maintained despite fundamental differences in target sequence composition and the observation that DS often exhibited lower within-group variability and greater sensitivity at lower administered doses.

These findings establish that DS serves as a robust quantitative alternative to conventional "gold-standard" reporter-gene assays. By yielding identical quantitative PoD estimates while simultaneously providing direct molecular evidence of the mutation spectrum, this approach increases the utility of *in vivo* mutagenicity data. Ultimately, these data support the integration of error-corrected sequencing endpoints into quantitative dose-response frameworks to refine human health risk assessment and regulatory decision-making

In contrast to the robust mutagenicity detected at the MMP and *lacZ* loci, neither the phenotypic Pig-a assay nor sequence-resolved analysis at the endogenous Pig-a locus exhibited a clear dose-related mutational response. The phenotypic Pig-a assay

revealed significant increases in the percentage of reticulocytes (%RET) at doses ≥25 mg/kg/day, confirming that BbF reached the erythropoietic compartment and induced systemic bone marrow toxicity (Bemis et al., 2019; Chikura et al., 2019; Dertinger et al., 2021; Dobrovolsky et al., 2010; Phonethepswath et al., 2010). However, this toxicity did not translate into a concomitant increase in mutant cell frequencies. The absence of a phenotypic response is constrained by the biological requirements of the assay, which exclusively captures mutations resulting in a loss-of-function phenotype (Heflich et al., 2019; OECD, 2025a) and strictly depends on the survival, maturation, and persistence of mutant erythroid cells in peripheral blood (Dertinger et al., 2021; Heflich et al., 2019; Hu et al., 2005; Ohtani et al., 2012).

Direct sequence analysis of the endogenous Pig-a locus from bone marrow DNA similarly failed to detect a BbF-induced dose response. Because identical tissue extracts yielded robust dose-dependent mutation increases at the distributed MMP and exogenous *lacZ loci*, the negative Pig-a findings do not reflect an absence of BbF mutagenicity in the bone marrow overall. Instead, sequence-resolved data revealed locus-specific constraints. The mutation spectrum at the Pig-a target lacked the characteristic C:G>A:T transversions indicative of PAH exposure observed in the other loci (LeBlanc et al., 2022; Schuster et al., 2024). Instead, the Pig-a locus exhibited a high background contribution of insertions and deletions. This elevated indel background is likely driven by the presence of homopolymeric sequence tracts within the Pig-a region (Liu et al., 2015; Yang et al., 2018), which are highly prone to alignment and variant-calling artifacts during sequencing (Dunn et al., 2023; Feng et al., 2016).

Consequently, the lack of a detectable mutational signal at the Pig-a locus is attributable to a combination of phenotypic assay limitations and locus-specific sequence constraints rather than an absence of bone marrow mutagenicity. Under these study conditions, Pig-a-based methodologies served primarily as an informative contrast, proving significantly less reliable for hazard identification and mechanistic characterization compared to the robust responses captured by the distributed endogenous MMP and exogenous *lacZ* locus.

4.6 Conclusion

BbF induced clear *in vivo* mutagenicity in MutaMouse BM and liver. These mutagenic responses were consistently and concordantly detected by both the TGR

assay and DS. DS at the endogenous MMP and exogenous *lacZ* loci reproduced the overall biological conclusions of the conventional TGR mutation assay. Furthermore, DS identified a distinct mutation spectrum heavily dominated by C:G>A:T substitutions, alongside secondary increases in C:G>T:A substitutions. This spectrum is highly consistent with the expected mutational profile of a metabolically activated PAH. Conversely, *Pig-a*-based endpoints were less informative under the specific conditions of this study. This lack of response reflects endpoint-specific biological constraints rather than a contradiction of the positive TGR and DS findings. Taken together, these data robustly support the implementation of DS as an alternative approach for *in vivo* mutagenicity assessment in repeat-dose studies. This is particularly true when mechanistic interpretation and flexible tissue selection are critical for human health risk assessment.

4.7 Acknowledgements

C.L.Y., F.M., and P.A.W. hold an Innovation in Regulatory Science Award from the Burroughs Wellcome Fund (#1021737). C.L.Y. is grateful for funding from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2021-02806) with infrastructure support through the Canadian Foundation for Innovation John Evans Leaders Fund (#233109), the Ontario Research Fund (#40569), and the University of Ottawa. C.L.Y. acknowledges that this project was conducted, in part, thanks to funding provided through the Canada Research Chairs Program (CRC-2020-00060). Work at Health Canada was supported by the Chemicals Management Plan to F.M. and Genomics Research and Development Initiative to F.M. and M.M.

4.8 Declaration on Use of Artificial Intelligence

We used large language models (Gemini 3.1 Pro and Microsoft Copilot Pro) to assist with language editing and sentence structuring. The authors reviewed all suggestions for accuracy and maintained full editorial control over the final output. The AI tools were not used to generate scientific hypotheses, interpret data, or formulate conclusions. The underlying scientific concepts and interpretations originated solely from the authors and reviewers. The authors are taking full responsibility for the data and conclusion presented in this study.

4.9 Tables and Figures

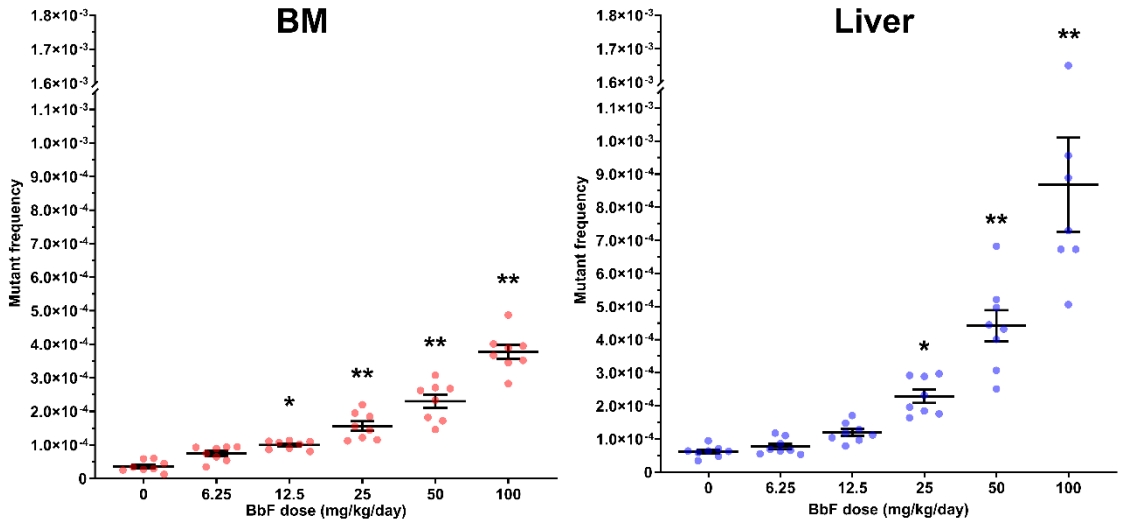


*Figure 15: Mutant frequency (mut/pfu) ± SEM measured using the TGR assay in BM (left) and liver (right) from MutaMouse males exposed to increasing doses of BbF or olive oil for 28+3 days via oral gavage (n = 8 per dose group). Statistically significant increases relative to the vehicle control were determined by one-way ANOVA with post-hoc Dunnett test and are indicated as * p < 0.05 and ** p < 0.01.*

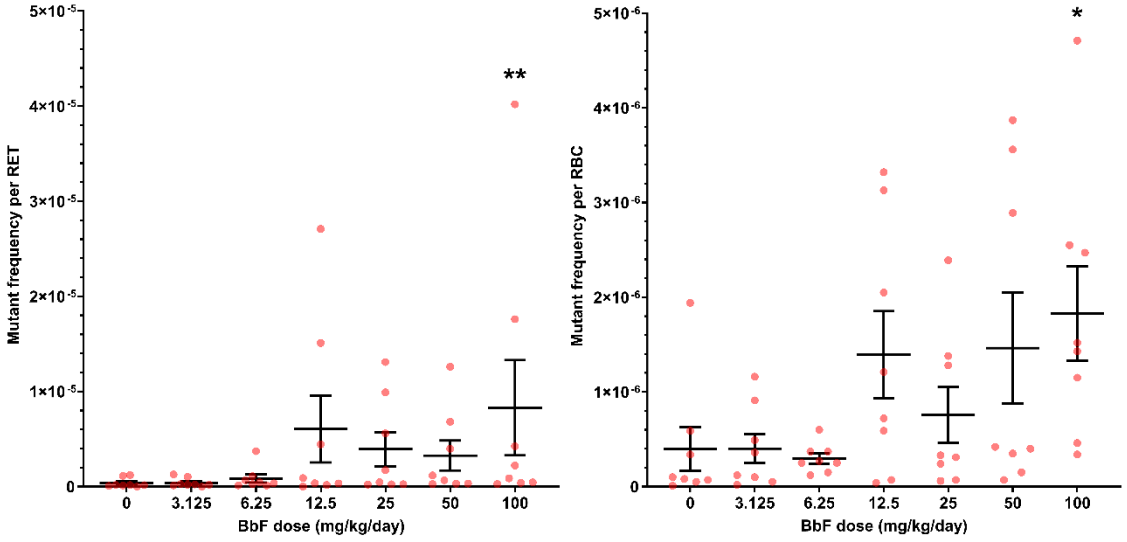


*Figure 16: Mutant frequency per reticulocyte (RET; left) and per red blood cell (RBC; right) ± SEM measured using the Pig-a assay on male MutaMouse blood extracted after 28-days of exposure to increasing doses of BbF or olive oil via oral gavage. Statistically significant increases compared to the respective vehicle controls were determined after Box-Cox transformation of the mutant frequencies followed by ANOVA with post-hoc Dunnett test and are indicated as * $p < 0.05$ and ** $p < 0.01$.*

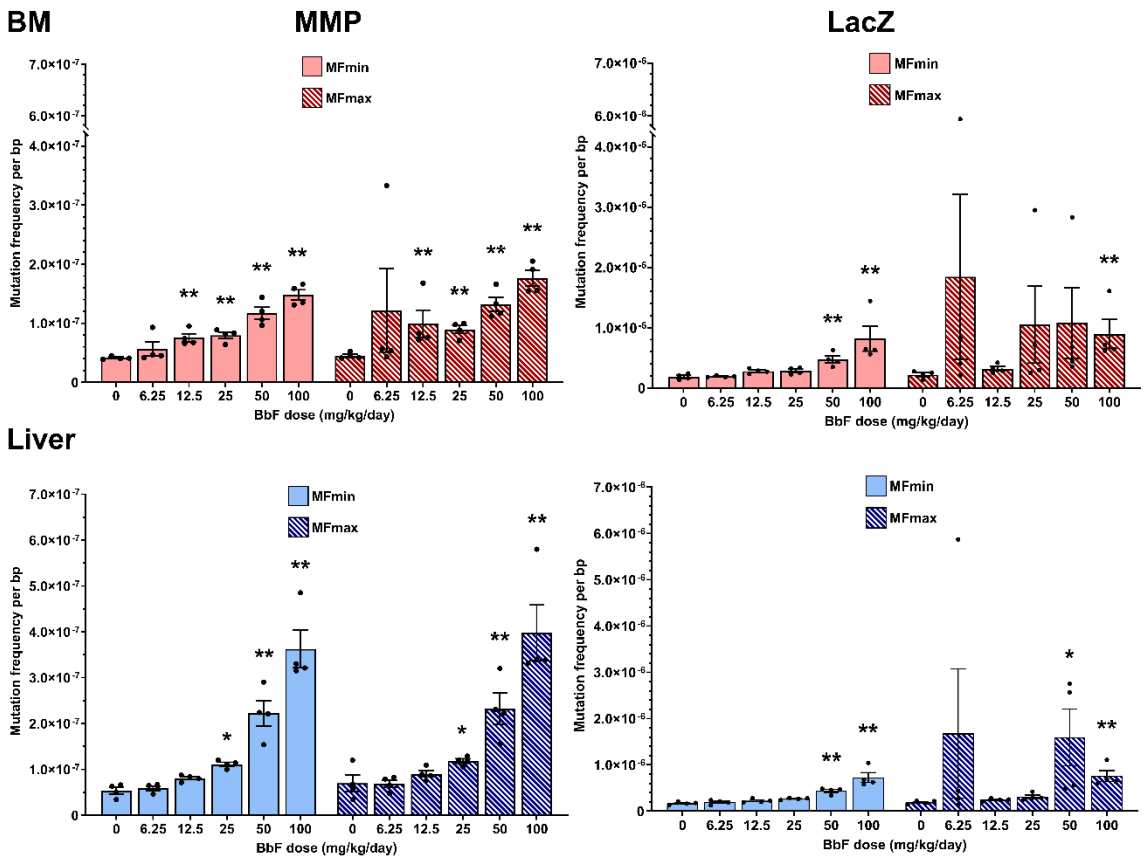


*Figure 17: Mutation frequency minimum (MFmin; filled bar) and maximum (MFmax; striped bar) ± SEM estimates measured by duplex sequencing in MutaMouse males exposed to increasing doses of BbF or olive oil for 28+3 days via oral gavage. Results are shown for the mouse mutagenesis panel (MMP) (left) and lacZ locus (right), in bone marrow (BM: top) and Liver (bottom) with n = 4 per group. Statistically significant increases compared to respective vehicle controls were determined using generalized linear mixed models and are indicated as * p < 0.05 and ** p < 0.01.*

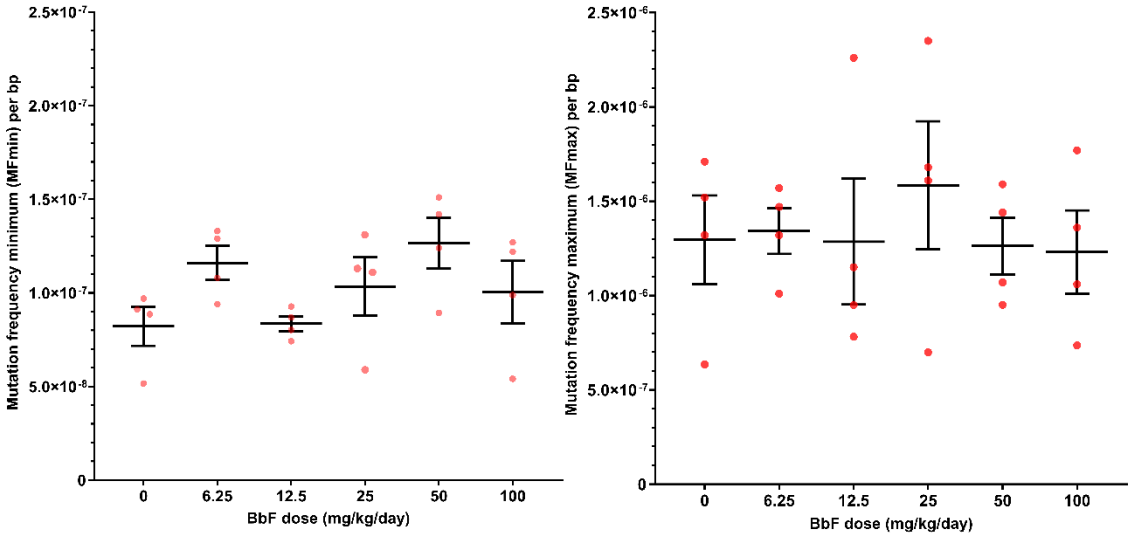


Figure 18: Mutation frequency minimum (MFmin; left) and maximum (MFmax; right) ± SEM measured by Duplex Sequencing in MutaMouse males exposed to increasing doses of BbF or olive oil for 28+3 days via oral gavage. Results are shown for the Pig-a locus with n = 4 per group. No dose showed a statistically significant increase relative to respective vehicle controls, as determined using generalized linear mixed models.

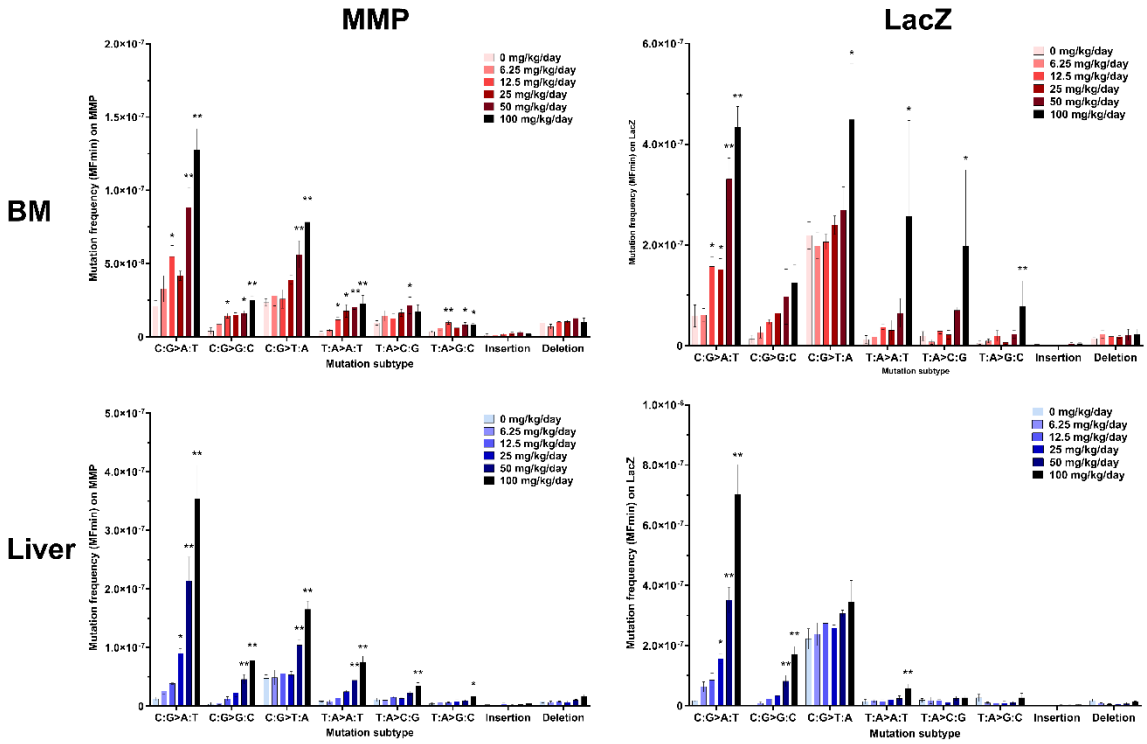


*Figure 19: Mutation frequency minimum (MFmin) ± SEM estimates per mutation subtype measured by duplex sequencing in MutaMouse males exposed to increasing doses of BbF or olive oil for 28+3 days via oral gavage. Results are shown for the mouse mutagenesis panel (MMP) (left) and lacZ locus (right), in bone marrow (BM: top) and Liver (bottom) with n = 4 per group. Statistically significant increases compared to respective vehicle controls were determined using generalized linear mixed models and are indicated as * p < 0.05 and ** p < 0.01.*

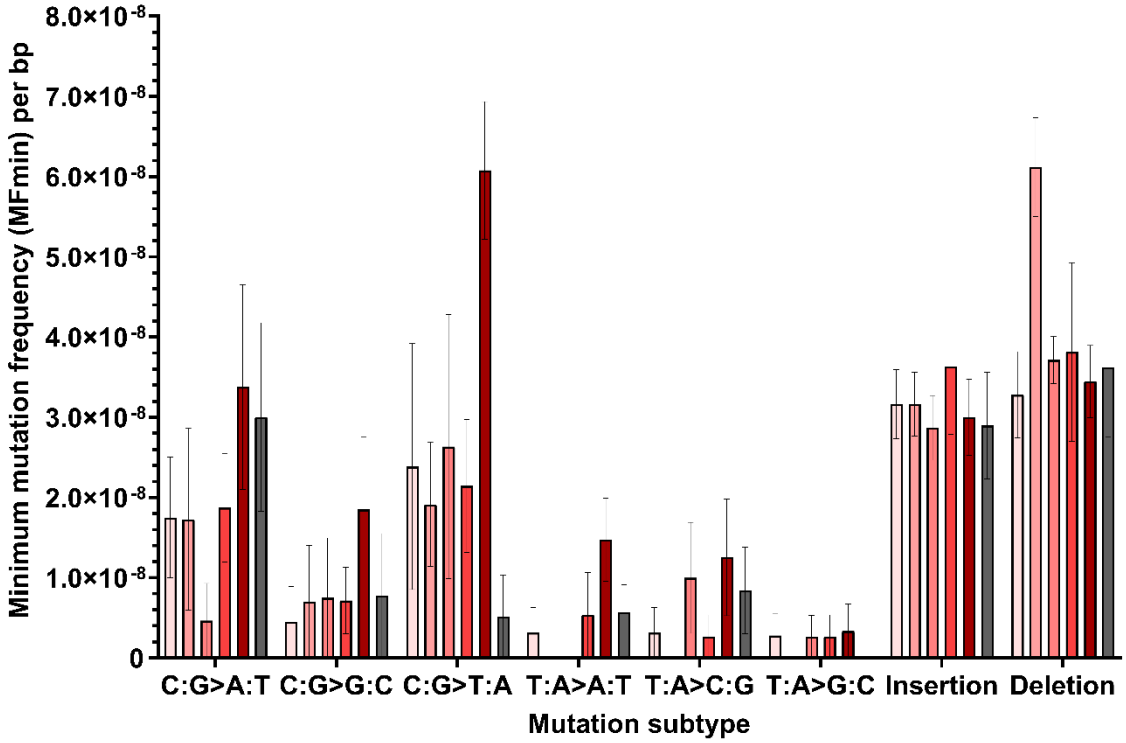


*Figure 20: Mutation frequency minimum (MFmin) ± SEM estimates per mutation subtype measured by duplex sequencing in MutaMouse males exposed to increasing doses of BbF or olive oil for 28+3 days via oral gavage. Results are shown for the Pig-a locus in bone marrow (BM: top) with n = 4 per group. Statistically significant increases compared to respective vehicle controls were determined using generalized linear mixed models and are indicated as * $p < 0.05$ and ** $p < 0.01$.*

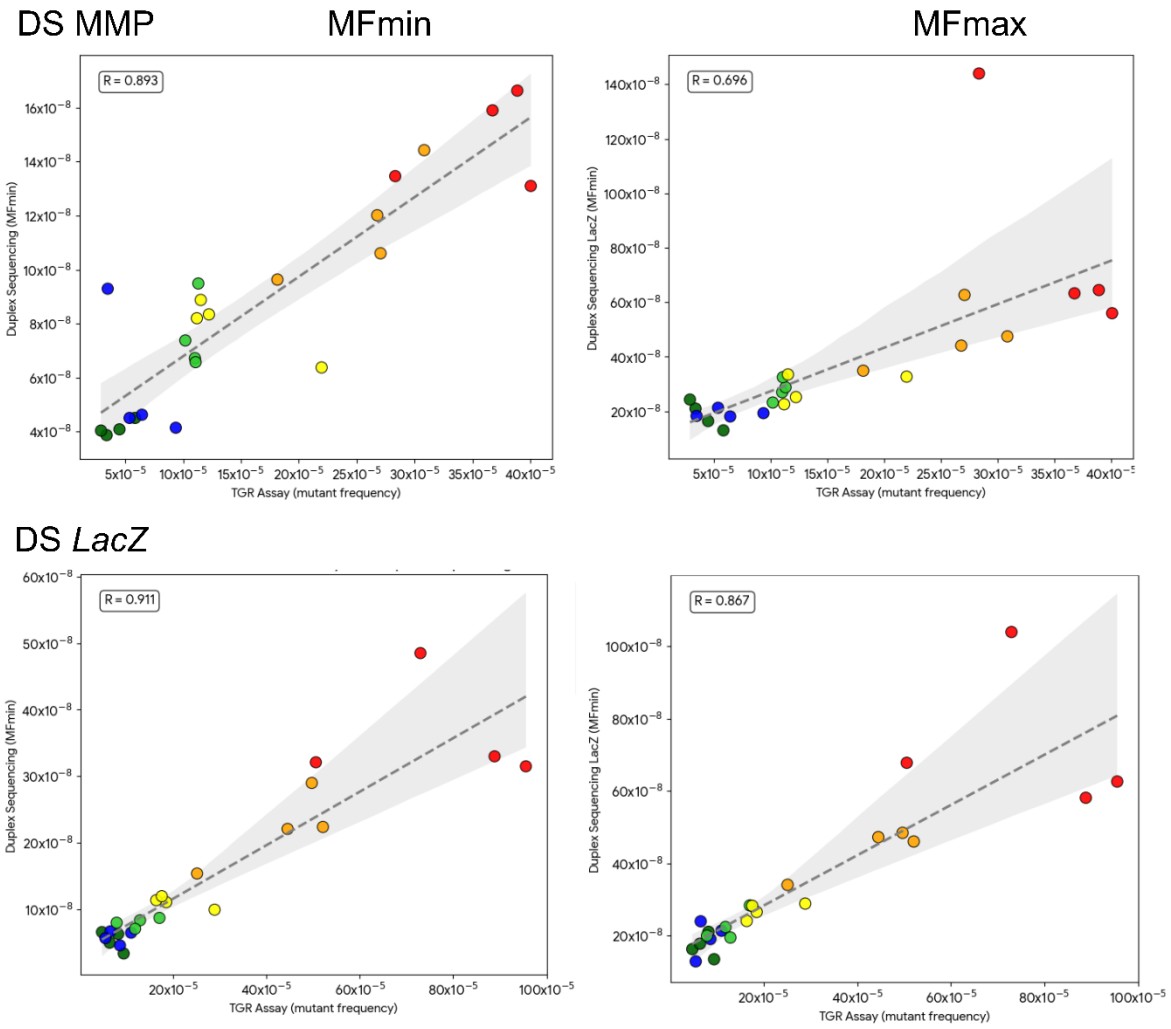


Figure 21: Pearson correlation analysis between mutation frequency minimum of duplex sequencing using the MMP (left) or the lacZ-transgene (right) versus the mutant frequencies obtained through the phenotypic TGR assay. The dose groups are indicated by colour for 0 (olive), 6.25 (blue), 12.5 (green), 25 (yellow), 50 (orange) and 100 (red) mg/kg/day. The 95% confidence level is visualized in grey.

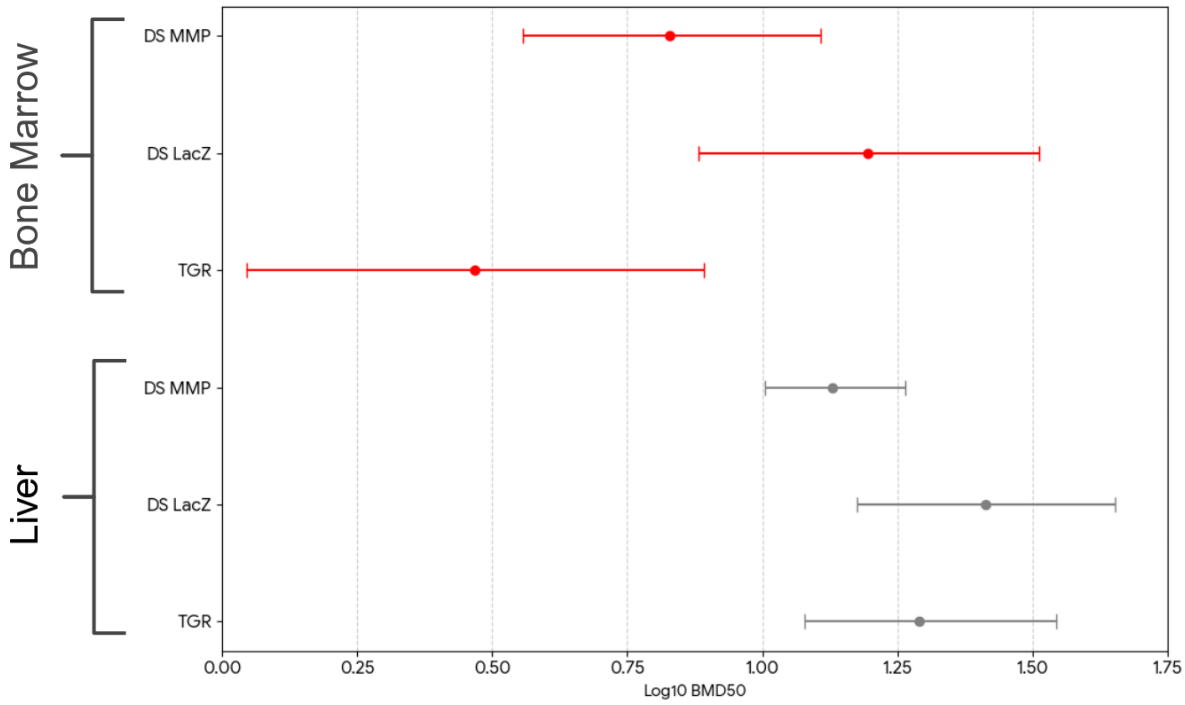


Figure 22: Log_{10} confidence intervals (BMDL to BMDU) of benchmark dose (BMD) modeling using a 50% benchmark response on the mutation frequency minimum for the MMP (top) and lacZ loci (middle) and mutant frequency of the TGR assay (bottom) measured in MutaMouse males exposed to increasing doses of BbF or olive oil for 28+3 days via oral gavage. BMD modeling on bone marrow is indicated in red, and liver is indicated in grey.

Chapter 5: Discussion

5.1 Summary of Thesis Outcomes

My thesis was designed around four overarching goals: (1) characterize the genotoxic capacities of the priority pollutant BbF, (2) advancing fundamental knowledge of chemically induced mutagenesis, (3) validate DS as a modern tool for mutagenicity assessment, and (4) advancing *in vivo* mutagenicity assessment based on quantitative and mechanistic data. The three main objectives supported these goals:

1. Apply DS to deeply characterize the dose-dependent mutagenic effects of short-term BbF oral exposure in MutaMouse males.
2. Determine how prolonged exposure durations and tissue-specific characteristics influence the accumulation of mutations, clonal expansion, and shifts in mutagenic potency.
3. Evaluate the qualitative and quantitative concordance of DS against established conventional *in vivo* mutation assays and inform on endpoint-specific drivers of mutagenesis.

5.1.1 Objectives and Outcomes of Chapter 2

In Chapter 2, my specific objectives were to:

1. Quantify the dose-dependent effects of BbF on chromosome damage and MF.
2. Determine the mutation spectra in control and BbF-treated animals.
3. Characterize variations in the mutagenic response between somatic tissues and across 20 endogenous genomic regions.
4. Extract the BbF trinucleotide signature and identify similarities to human cancer signatures.
5. Compare tissue- and endpoint-specific BbF mutagenic potency.

My hypothesis was that BbF induces dose-dependent mutagenicity in the liver and BM, with mutation spectral shifts consistent with polycyclic aromatic hydrocarbon exposure and enhanced mutagenicity in intergenic regions. My objectives were met, establishing a comprehensive *in vivo* mutagenic profile of BbF under a standard 28-day repeat-dose design. I demonstrated that BbF caused clear, dose-dependent increases in micronucleus frequency in peripheral blood and MF in both BM and liver. By applying

DS, I revealed that the mutation spectrum was predominantly driven by C:G>A:T transversions. Furthermore, I found that mutagenicity was significantly higher in intergenic regions relative to genic regions, supporting a protective role for transcription-coupled repair. Mutational signature analysis directly linked the BbF-induced trinucleotide patterns to tobacco-associated human cancer signatures, specifically SBS4. Finally, while the maximum mutagenic response was higher in the highly metabolic liver compared to the BM, BMD modeling revealed that the overall mutagenic potency was comparable between the two tissues and endpoints. Ultimately, Chapter 2 not only defined the baseline hazard and potency of BbF but also proved that DS is highly informative under standard 28-day regulatory study designs.

5.1.2 Objectives and Outcomes of Chapter 3

In Chapter 3, my specific objectives were to:

1. Quantify the dose- and time-dependent effects of BbF exposure on MF using DS.
2. Assess the contribution of clonally derived mutations to the mutational burden over time.
3. Evaluate time-dependent changes in mutation types and localization within 20 endogenous loci.
4. Characterize shifts in the BbF trinucleotide mutation spectrum and identify similarities to known human cancer signatures from COSMIC over time.

I hypothesized that prolonged exposure to BbF increases the overall mutational burden without changing the underlying mutagenic mechanism. Furthermore, I posited that tissue-specific biological characteristics, such as proliferation rate, metabolic activity, and DNA repair capacity, act as key drivers that modulate MF and the expansion of mutant cells. I successfully met the chapter's objectives, demonstrating that mutational burden increased with exposure duration in a tissue- and time dependent manner. The underlying accumulation dynamics fundamentally differed: in the BM, prolonged exposure heavily drove the clonal expansion of mutant cell lineages, whereas in the liver, prolonged exposure produced a much larger increase in unique *de novo* mutations. Despite these quantitative differences, I found that the dominant C:G>A:T mutation spectrum remained consistent over time and that mutations were preferentially induced in intergenic regions across all exposure durations. Furthermore,

trinucleotide spectrum analysis revealed a dose- and time-dependent increase in similarities to the tobacco smoke-associated human cancer signature SBS4 in both tissues. Finally, BMD modeling demonstrated that the mutagenic potency of BbF increased with prolonged exposure, yielding lower point-of-departure estimates over time. Ultimately, I conclude that *in vivo* MF is a complex endpoint, and that prolonged exposure fundamentally alters both the composition of mutational response and its mutagenic potency estimates.

5.1.3 Objectives and Outcomes of Chapter 4

In Chapter 4, my specific objectives were to:

1. Directly evaluate concordance between conventional *in vivo* gene mutation assays and DS approaches for BbF-induced somatic mutagenicity.
2. Determine whether these methods yield comparable mutagenic potency estimates and points of departure using BMD modeling.
3. Inform how sequence context affects mutagenesis.
4. Identify predominant mutation types that drive endpoint-specific mutagenicity.

I hypothesized that ECS produces qualitative and quantitative mutagenicity conclusions that are concordant with established phenotypic *in vivo* assays. Furthermore, I posited that it simultaneously provides superior mechanistic resolution regarding mutation spectra and genomic context. I met my objectives by demonstrating that DS yields qualitative hazard conclusions. Additionally, mutagenic potency estimates directly overlapped between DS and the TGR assay. Specifically, BMD modeling revealed overlapping confidence intervals across the TGR and DS endpoints in both the BM and liver. By applying sequence-resolved analysis, I identified that the BbF-induced mutagenicity was driven by C:G>A:T and C:G>T:A substitutions. Crucially, this predominant mutation spectrum was identical across both the endogenous MMP and the exogenous *lacZ* loci, confirming that target-specific sequence contexts did not distort hazard identification. Conversely, I found that neither the phenotypic Pig-a assay nor DS at the endogenous Pig-a locus exhibited a robust dose-response. DS revealed that this lack of concordance was driven by endpoint-specific sequence constraints, as the Pig-a locus accumulated insertion and deletion artifacts that obscured the characteristic BbF mutational signature. Ultimately, I conclude that DS matches the qualitative and quantitative hazard conclusions of the

"gold-standard" TGR assay in responsive tissues, bypassing the need for phenotypic selection while informing on mutation spectra, location, and clonal expansion

5.2 Contributions to Scientific Knowledge

5.2.1 Characterization of the Genotoxic Capacities of BbF

Prior to this work, the mutagenic mechanisms of BbF on endogenous mammalian DNA were largely uncharacterized. My thesis provides the first comprehensive, sequence-resolved *in vivo* genotoxic profile for this priority PAH. I demonstrated that oral BbF exposure induces clear, dose-dependent increases in chromosome damage, measured via the micronucleus assay. I detected dose- and exposure duration-dependent increases in somatic MF following 28-, 90-, and 180-day exposures. Despite an absence of overt histopathological changes, the metabolically active liver accumulated a greater mutational burden than the BM. Crucially, sequence-level resolution revealed distinct tissue-specific mutation accumulation dynamics. Prolonged exposure drove the clonal expansion of mutant cell lineages in the BM. Conversely, the mutational burden in the liver was driven almost entirely by unique *de novo* mutations. These findings establish that mutagenicity cannot be characterized solely from external dose, total MF, or a single tissue; rather, mutational burden reflects complex tissue-specific biology, including metabolic activation, cellular proliferation, and clonal expansion, alongside the critical consideration of exposure duration.

Furthermore, I found that mutations were preferentially localized within intergenic loci across all exposure durations. This observation supports a protective role for TC-NER in actively clearing bulky DNA lesions from transcribed genic regions (Hess et al., 1997; Shuck et al., 2008). This implies that mutagenicity results must be interpreted within their genomic context, as overall mutation frequencies can mask localized regional susceptibilities and the targeted protective effects of DNA repair mechanisms.

The BbF mutation spectrum was dominated by C:G>A:T transversions, which reflects the expected formation of bulky DNA adducts. I also observed secondary increases in C:G>T:A transitions, that are predominantly caused by oxidative DNA damage and partially by bulky DNA adducts. This mutational pattern exhibited a dose-dependent increase in similarity to the human cancer signature SBS4. This demonstrates that this mouse exposure model, combined with error-corrected

sequencing, can be utilized to experimentally identify characteristic mutational fingerprints of PAHs that correspond directly to those found in human cancers resulting from tobacco smoke exposure.

Finally, I integrated these sequence datasets with BMD modeling to estimate BbF's mutagenic potency in different tissues and loci, across different endpoints and over varying exposure durations. This modeling revealed that BbF potency increases is affected by exposure duration. Consequently, point-of-departure estimates decrease with prolonged exposure. Both the BMD modeling and the induced mutations types showed strong similarities to well-studied PAH BaP mutation spectra (LeBlanc et al., 2022; Long et al., 2016). These similarities imply that BbF acts through a comparable mutagenic mechanism to BaP, demonstrating how high-resolution profiling and potency data can be used to validate the grouping of PAHs and complex PAH mixtures for human health risk assessments.

Overall, this work provides an in-depth understanding of PAH mutagenesis that can be used to help evaluate their mutagenic potential across species, tissues, loci, and exposure duration. By transitioning BbF into a well-characterized priority pollutant, this thesis provides a robust framework to inform the potential mutagenic hazards and, ultimately, the carcinogenic mechanisms of environmental contaminants.

5.2.2 Advancing Knowledge on Chemically Induced Mutagenesis

My research demonstrates that *in vivo* MF is not a direct reflection of external dose; rather, it is a complex composite endpoint. Mutagenesis is shaped by exposure duration and tissue-specific biology including biological factors such as metabolic capacity and cellular lifespan. By tracking mutagenesis over 180 days, representing the longest sequence-resolved *in vivo* mutagenicity study conducted to date, I identified a distinct contrast in how different tissues accumulate mutations. In the highly proliferative BM, prolonged exposure drove the mutational burden primarily through the clonal expansion of mutant cell lineages. In the metabolically active liver, this burden was driven almost entirely by the continuous accumulation of unique *de novo* mutations. As shown, the physiological propagation of mutant cells contributes fundamentally differently to the overall mutation burden depending on the target tissue's turnover rate, which must be critically evaluated when selecting tissues for mutagenicity risk assessments (Nangalia et al., 2019; Parsons et al., 2025).

Furthermore, the untreated control cohorts provided critical baseline data for a collaborative study characterizing age-related increases in spontaneous MF. We found that significant increases in background MF only occur in the liver and BM of geriatric MutaMouse males (e.g., >75 weeks of age), with no spontaneous mutation accumulation occurring during the shorter study durations (Esina, manuscript in preparation). For standard subchronic or chronic exposure protocols, the natural aging process does not significantly confound chemical mutagenicity assessments, implying that age does not need to be mathematically modeled as a confounding variable when assessing chemical risk.

Addressing a critical regulatory challenge, I generated concurrent DS and TGR data across 28-, 90-, and 180-day timepoints to enable the mathematical modeling of exposure duration-dependent shifts in chemical potency. Currently, human risk assessment heavily relies on fixed, default uncertainty factors, such as a standard 10-fold adjustment, to extrapolate risk from short-term assays to chronic, life-long exposures because empirical data on long-term mutation accumulation are lacking (Escher et al., 2020; Pieters et al., 1998). By explicitly modeling shifts in BMD over time, my work provides the rare empirical data required to verify existing assumptions with precise, data-derived extrapolation factors, significantly strengthening the robustness and accuracy of human health risk assessments.

Ultimately, elucidating the underlying biological processes required the high sensitivity of ECS. ECS provided the necessary single-nucleotide resolution to decipher the mutation profiles and investigate cofounding factors of mutagenesis. Consequently, this work moves genetic toxicology beyond basic hazard identification and directly supports the transition toward mechanistic, quantitative decision-making frameworks.

5.2.3 Validation of DS as Reproducible and Robust Tool for *In Vivo* Mutagenicity Assessment

I began the validation of DS by establishing the required statistical power and optimized study design (Appendix A: Power Analyses to Inform Duplex Sequencing Study Designs for MutaMouse Liver and Bone Marrow). I provided foundational *in vivo* sequencing data and performed statistical calculations for a power analysis to mathematically formalize the statistical efficiency of DS. This study demonstrated that

DS detects biologically meaningful changes in MF using smaller group sizes. Specifically, a sample size of four animals per dose group provides greater than 80% power to detect a two-fold change in MF when paired with sufficient sequencing depth. I applied this optimized study design to execute my primary thesis experiments and successfully detected dose-related increases in MF using four animals per group. In contrast, the concurrent conventional MN and TGR assays required eight animals. This confirms that DS can significantly reduce animal use in genetic toxicology testing, directly supporting 3R-aligned study designs (Kroeger, 2006; Russell & Burch, 1959) while providing the foundational statistical framework necessary to execute all subsequent inter-laboratory validation and empirical experiments.

I continued supporting the validation of DS by participating in a multi-laboratory reconstruction study to establish the technical reproducibility of the assay (Appendix B: Transferability, Reproducibility and Sensitivity of Mutation Quantification by Duplex Sequencing). This collaborative study distributed rat liver DNA samples with predefined mutation frequencies across eight independent laboratories. The samples included DNA mixtures from BaP and N-ethyl-N-nitrosourea exposed animals, which were formulated to represent 1.2-, 1.5-, and 2.0-fold increases in mutation frequency relative to untreated controls. I contributed directly as the operator for one participating laboratory and supported the collective data analysis. The study demonstrated that DS generates nearly identical mutation frequencies and mutational spectra across different facilities, regardless of prior operator experience. This robust inter-laboratory concordance proves that DS is a highly transferable and standardized tool, providing the technical reliability required for regulatory agencies to confidently accept its data for human health risk assessments.

Directly comparing DS performance against established methodologies using matched genomic endpoints, including the endogenous MMP, the exogenous *lacZ* transgene, and the endogenous *Pig-a* locus further confirmed DS as a promising tool for *in vivo* mutagenicity assessment. This sequence-resolved comparison confirms that DS produces identical qualitative hazard calls to the OECD-approved TGR and Pig-a assays in both responsive and non-responsive targets. I observed strong Pearson correlations between the DS MFs and the conventional plaque-based mutant frequencies. BMD modeling confirmed this quantitative alignment. The modeling yielded overlapping confidence intervals and comparable potency estimates between

DS and the TGR assay across tissues. This direct quantitative and qualitative concordance validates DS as a robust alternative to current gold-standard methods, proving that regulatory agencies can confidently rely on sequence-derived potency estimates for human health risk assessments without requiring specialized transgenic animal models.

I established the mechanistic validity of DS by identifying the mutational subtypes driving the BbF response. DS successfully identified the expected BbF-associated mutation spectrum, which was dominated primarily by C:G>A:T transversions and secondary C:G>T:A transitions. In addition, DS enabled the precise differentiation between unique *de novo* mutations and recurring events, ensuring that clonal expansion does not bias the spectrum analysis or hazard conclusions. The recent International Workshops on Genotoxicity Testing consensus statement highlights the regulatory value of this sequence-level resolution, noting that dose-dependent changes in a single mutation subtype can be sufficient for a positive mutagenicity call even if the overall MF does not meet positivity criteria (Yauk et al., 2026). By providing biologically interpretable mutation spectra rather than a simple aggregated MF, DS delivers the mechanistic evidence required to confidently link chemical exposures directly to specific mutagenic mechanisms for regulatory risk assessments.

Therefore, by demonstrating statistical power, reproducibility, concordance, and mechanistic validity, my independent work establishes a modern approach for ECS based mutagenicity assessment.

5.2.4 Contributing to Modern, Mechanistic *In Vivo* Assessment.

My work establishes an empirical foundation for modernizing *in vivo* mutagenicity assessment. Currently, regulatory *in vivo* mutagenicity testing often relies on specialized transgenic models in standalone studies (OECD, 2025d, 2025c). In my data chapters, I demonstrated that ECS seamlessly integrates into existing standard toxicity studies, including prolonged 90- and 180-day exposure protocols. ECS uses endogenous DNA and eliminates the need for independent transgenic animal cohorts. This integration directly supports the 3R principles of animal testing by significantly reducing overall animal use while maximizing the mutagenicity data obtained from standard safety evaluations.

This thesis demonstrates the value of ECS-based approaches that quantify mutagenicity while simultaneously informing on underlying mechanisms of mutagenesis. My chapters heavily relied on quantitative data rather than binary qualitative hazard calls. I used benchmark dose modeling to estimate the dose at which a 50% increase in MF occurs (White et al., 2025), rather than providing a simple yes or no answer regarding hazard. This quantitative approach identifies the specific dose at which mutagenic effects occur, providing the exact metric required to accurately understand and calculate human health risk.

Specifically, the benchmark dose lower confidence limits derived from the sequence-resolved datasets serve as robust PoDs for quantitative human health risk assessment. For example, the standard 28-day BM BMDL50 of 3.6 mg/kg/day can be integrated with traditional uncertainty factors (UFs) to calculate a permissible daily intake, while the chronic 90- and 180-day exposure datasets provide the first empirical evidence to refine the default 10-fold subchronic-to-chronic duration UF by modeling actual time-dependent mutation plateauing. The PoDs can also be used to calculate the margin of exposure against human exposure levels, such as the typical dietary BbF intake of approximately 0.09 µg/day (Martorell et al., 2010), to evaluate whether current environmental exposures present a significant mutagenic concern. Furthermore, because humans are exposed to PAHs as complex environmental mixtures, these quantitative data offer a powerful framework for refining toxic equivalency factors. Direct comparison of our 28-day bone marrow BbF BMD50 (3.6 to 12.8 mg/kg/day) with the bone marrow BaP BMD50 (4.5 to 7.5 mg/kg/day) from (LeBlanc et al. 2022) reveals overlapping confidence intervals and comparable *in vivo* mutagenic potencies. This equivalent potency (~1:1 ratio) directly challenges conventional regulatory toxic equivalency factors that assume BbF is tenfold less toxic than BaP (i.e., a toxic equivalency factors of 0.1), demonstrating that current frameworks may underestimate the cumulative mutagenic risk of BbF within complex environmental PAH mixtures.

Transparent and standardized data processing is required to ensure regulatory agencies can utilize complex sequence datasets. The lack of bioinformatic workflows previously posed a major barrier to regulatory adoption and data comparison. I promoted essential standardization by contributing to the development of MutSeqR (Appendix C: MutSeqR: An Open-Source R Package for Standardized Analysis of Error-Corrected Next-Generation Sequencing Data in Genetic Toxicology), an open-

source data processing R package (Dodge et al., 2025). MutSeqR addresses this bioinformatic barrier and provides a reproducible pipeline for variant filtering, mutation spectrum analysis, and BMD modeling across different sequencing technologies. My specific contributions included providing *in vivo* test datasets, writing and testing the code, designing data visualization strategies, and co-authoring the software manual. By establishing a transparent and harmonized analytical pipeline, this collaborative software development directly removes a critical technical barrier preventing the global regulatory adoption of sequence-resolved methodologies.

The empirical data generated in my primary data chapters and collaborative support ongoing efforts to amend OECD Test Guideline 488 to formally include error-corrected sequencing as an alternative to the TGR assay for *in vivo* mutagenicity assessment (OECD, 2025d). By demonstrating reproducibility, providing open-source analytical tools, and generating high-resolution empirical data, this thesis delivers the evidence required to transition genetic toxicology towards a modernized, animal-sparing regulatory framework.

5.3 Scientific Questions Resulting from my Thesis

5.3.1 Linking Endpoints to Upstream Biological Determinants

How do specific BbF DNA lesions and targeted DNA repair mechanisms quantitatively translate into the mutation burdens detected by DS?

Across my data chapters, I demonstrated that BbF induces dose-dependent increases in mutagenicity. I consistently observed significantly higher MF in intergenic loci compared to genic loci, suggesting that transcription-coupled repair actively protects transcribed genes. However, sequence data alone cannot establish the precise relationship between the initial DNA damage and the resulting mutations.

Future work could integrate sequence-resolved mutation data with direct measurements of upstream key events. Mass spectrometry can be used to quantify bulky DNA adducts and identify the exact BbF metabolites that bind to endogenous DNA. An Enzyme-Linked Immunosorbent Assay (e.g. 8-OHdG ELISA) or a modified Comet assay could quantify oxidative DNA damage (AbuArrah, 2021; Cordelli et al., 2021). These targeted analyses cover both reactive metabolites created during the metabolism of BbF and provide a direct *in vivo* follow-up to previous metabolite

characterization efforts (Amin et al., 1982; Yang et al., 2023). The direct comparison between DNA damage and resulting mutagenicity would provide crucial information on DNA damage repair, mutation fixation and dose relationships between the two endpoints. Additionally, transcriptomic analyses on biological samples of this project could measure the activation of targeted DNA repair pathways, such as genes involved in TC-NER (Adimoolam & Ford, 2002; Lefkofsky et al., 2015).

This proposed work expands on two main contributions of this thesis. First, it strengthens the toxicological profile of BbF by defining its specific *in vivo* metabolites and induced DNA damage. Second, it advances fundamental knowledge on mutagenesis by elucidating how specific DNA lesions translate into mutations through targeted DNA repair pathways.

5.3.2 Selection of Suitable Sequencing Targets

How does the endogenous mutational profile established for BbF compare to mutation accumulation under positive selective pressure and what is the most suitable sequencing target?

This thesis established a high-resolution profile of BbF-induced mutagenicity by targeting endogenous loci captured in the MMP. Additionally, I used custom panels to evaluate induced mutagenicity in the exogenous *lacZ* transgene and the endogenous *Pig-a* gene. The *lacZ* panel yielded identical qualitative hazard calls and mutation spectra to the MMP. Thus, targeting the *lacZ* transgene in ECS studies provides no substantial benefit over assessing the twenty endogenous loci captured by the MMP. DS on the *Pig-a* locus failed to produce a positive mutagenicity call in this study design. This failure highlights the limitations of relying on a single gene endpoint for hazard identification and mutagenicity assessment. The reported data in Chapter 4 directly triggered a follow-up study applying the *Pig-a* assay and DS *Pig-a* sequencing in the rat model using allometrically scaled BbF doses. The Pig-a assay is more extensively optimized for rats; thus, this follow-up study was able to detect a dose-dependent increase in MF and mutant frequencies in the *Pig-a* gene. Nevertheless, this discrepancy highlights the importance of a target selection that is applicable to various model systems. Assessing a diverse cross-section of the endogenous genome provides a more robust and biologically representative evaluation of genotoxic

potential, protecting against locus-specific artifacts or species-specific assay limitations.

The loci captured within the MMP are not under selective pressure. Contrarily, carcinogenesis relies on selective pressure and promotion of clonal expansion of cells harboring mutations in specific cancer driver genes. This work encouraged the Health and Environmental Sciences Institute to conduct a study comparing the endogenous data presented in this thesis against mutation accumulation within cancer driver genes. Applying DS to a customizable cancer driver gene panel will evaluate how selective pressure and sequence context influence clonal expansion and mutagenicity.

Exploring customizable panels raises the question of optimal target size. A multi-locus endogenous panel provides a comprehensive representation of physiological conditions. It captures diverse sequence characteristics and avoids single-locus biases while eliminating the need for specialized transgenic animal models. Expanding the panel beyond 48 kilobases, or applying whole genome sequencing, could determine if a wider sequence analysis provides more mechanistic information without diminishing returns. Conversely, reducing the panel size could optimize efficiency by decreasing sequencing costs. Exploring this balance will identify the optimal target size for routine, cost-effective mutagenicity assessment. Defining this optimal balance is critical to standardize sequencing workflows, ensuring they are both economically viable for routine industrial screening and scientifically rigorous for regulatory safety evaluations.

5.3.3 Informing Uncertainty Factors for Risk Assessment

Are existing default uncertainty factors for short-term to life-long and in vitro to in vivo extrapolations applicable to quantitative, sequence-resolved mutation data?

Current regulatory frameworks rely heavily on default uncertainty factors. A 10-fold adjustment is typically used to extrapolate risk from short-term tests to chronic, life-long human exposures (Dankovic et al., 2015; Dourson et al., 1996; Nessel et al., 1995). I observed fold-change increases in MF at identical exposure doses across the 28-, 90-, and 180-day time points. However, these increases did not align proportionally with the cumulative administered dose. This suggests that mutagenic potency increases non-linearly with prolonged exposure and is influenced by tissue characteristics. Consequently, assuming static mutation rates over a life-long exposure

period introduces uncertainty into human health risk assessments. Additionally, to the DS MF data after prolonged BbF exposure, I generated TGR data for the 90- and 180-day time points. These extended-duration datasets can be used to explicitly model shifts in BMD over time. Defining the quantitative relationship between short-term mutation induction and prolonged accumulation provides the empirical data necessary to derive precise uncertainty factors for exposure duration. Future work could expand this approach. Time-dependent BMD shifts could be mathematically modeled to derive broadly applicable uncertainty factors.

With prolonged exposure, mutation accumulation does not scale predictably with cumulative dose. Therefore, selecting the appropriate tissue is essential for accurate risk assessment. International regulatory agencies aim to eliminate mammalian animal testing by 2035 in favor of non-animal alternatives. Thus, default uncertainty factors routinely applied for animal-to-human and *in vitro* to *in vivo* extrapolation need to be validated for quantitative mutation data (Bell et al., 2018). Future studies could apply DS to established human cell models, such as HepaRG, HepG2, and TK6 (with S9) cells (Chepelev et al., 2023; Cho et al., 2023; Conway et al., 2025). DS could also be applied to emerging *in vitro* platforms like organ-on-a-chip systems (D. Singh et al., 2022). This approach aligns directly with the 3Rs principles of replacement, reduction, and refinement of animal use (Kroeger, 2006; Russell & Burch, 1959). Using human-derived *in vitro* models allows for the mechanistic investigation of mutagenesis under highly controlled metabolic and proliferative conditions (Conway et al., 2025). Ultimately, integrating extended-duration data with human-relevant *in vitro* models will allow regulators to replace default assumptions with data-driven models. This integration ensures the accurate extrapolation of testing data to protect human health against chronic, life-long exposures.

5.3.4 Constraints of ECS Acceptance for Regulatory Decision-Making

What additional empirical and methodological milestones are required to facilitate the regulatory adoption of ECS for routine mutagenicity assessment?

This thesis provides data supporting the regulatory application of DS. The global uptake of the technology remains gradual. Recent simulation analyses using vehicle control data demonstrate that DS maintains appropriate type 1 error rates and successfully identifies true negative responses (Williams et al., 2026). These findings

establish the specificity of the assay and alleviate concerns regarding false-positive rates. Confirming a low false-positive rate ensures that regulatory agencies can rely on ECS to correctly identify non-mutagenic compounds, preventing the unwarranted elimination of safe chemicals from development pipelines (Williams et al., 2026).

Regulatory acceptance of OECD assays previously relied heavily on establishing stringent historical control distributions. Building robust historical databases for commonly targeted tissues and models would be beneficial for a broader adoption of ECS platforms but would require global efforts. The International Workshops on Genotoxicity Testing consensus dictates that data interpretation should be based on comparing overall MF directly to concurrent vehicle controls (Yauk et al., 2026). Historical control data can provide valuable context for interpreting borderline results when available. Thus, standardized reference DNA samples with predefined MF should be established. These reference materials allow new laboratories to establish technical proficiency and ensure inter-laboratory reproducibility before conducting official regulatory studies. By prioritizing concurrent controls and standardized reference materials over exhaustive historical databases, the field can accelerate the global implementation of ECS across diverse toxicological models.

The field should systematically compare DS against other emerging ECS technologies. Alternative platforms include single-molecule mutation sequencing (SMM-seq), PacBio HiFi sequencing, and Concatenating Original Duplex for Error Correction (CODEC). Collaborative studies could compare these methods using the shared reference DNA samples mentioned above. These comparisons would evaluate differences in error-correction strategies, genomic coverage, and sequencing costs. This direct comparison will define the specific technical limitations and advantages of each platform. Clarifying these distinctions will allow researchers and regulators to apply different ECS technologies in a fit-for-purpose manner for specific safety evaluation contexts. Fulfilling these methodological milestones will support the transition of ECS from a promising research methodology into a fully adopted, standardized regulatory tool.

5.4 Concluding Remarks

This thesis addressed the three main objectives established at the beginning of this chapter. First, I characterized the genotoxic effects of short-term BbF exposure.

Second, I determined how prolonged exposure durations and tissue-specific characteristics influence mutagenesis. Third, I evaluated the concordance of DS against established conventional *in vivo* mutation assays.

These empirical data extend beyond the hazard identification of a priority pollutant. I demonstrated that somatic mutagenesis is shaped by tissue-specific biology, clonal expansion, and prolonged exposure durations. Consequently, standard short-term testing protocols may underestimate the mutagenic risk of chronic chemical exposures. This empirical evidence addresses a gap in current regulatory frameworks. Specifically, it provides the biological foundation required to integrate prolonged exposure metrics and empirically derived uncertainty factors into human health risk extrapolation.

Furthermore, this thesis advances the methodological approach to *in vivo* genetic toxicology. Historically, the field relied on specialized transgenic animal models to generate binary hazard calls. I demonstrated that DS addresses these limitations by directly quantifying mutations within the endogenous mammalian genome. I demonstrated that DS matches the qualitative hazard calls and quantitative potency conclusions of established OECD assays. This ECS-based approach utilizes endogenous DNA and can be integrated directly into standard repeat-dose toxicity studies. This capability supports the 3R principles and maximizes mechanistic data acquisition while reducing overall animal use.

Ultimately, this work supports the transition in genetic toxicology toward quantitative, mechanism-based risk assessment. By combining ECS with BMD modeling, I demonstrated a modern approach for chemical evaluation. This framework provides quantitative potency estimates and identifies specific mechanisms through mutational spectra analysis. It also characterizes clonal expansion dynamics from a single sequencing endpoint. Equipping regulatory agencies with these sequence-resolved testing strategies will accelerate the evaluation of priority chemicals. Consequently, this thesis provides the empirical foundation required to modernize chemical safety guidelines and protect human health from mutagenic hazards.

References

Abbasi, A., & Alexandrov, L. B. (2021). Significance and limitations of the use of next-generation sequencing technologies for detecting mutational signatures. *DNA Repair*, *107*, 103200. https://doi.org/10.1016/j.dnarep.2021.103200

AbuArrah, M. (2021). 8-Hydroxy-2-Deoxyguanosine as Oxidative DNA Damage Biomarker of Medical Ionizing Radiation: A Scoping Review. *Journal of Biomedical Physics and Engineering*, *11*(3). https://doi.org/10.31661/jbpe.v0i0.2101-1258

Adesina, O. A., Olowolafe, T. I., & Igbafe, A. (2022). Levels of polycyclic aromatic hydrocarbon from mainstream smoke of tobacco products and its risks assessment. *Journal of Hazardous Materials Advances*, *5*, 100053. https://doi.org/10.1016/j.hazadv.2022.100053

Adimoolam, S., & Ford, J. M. (2002). P53 and DNA damage-inducible expression of the xeroderma pigmentosum group C gene. *Proceedings of the National Academy of Sciences*, *99*(20), 12985–12990. https://doi.org/10.1073/pnas.202485699

Albertini, R. J., Nicklas, J. A., & O'Neill, J. P. (1993). Somatic cell gene mutations in humans: Biomarkers for genotoxicity. *Environmental Health Perspectives*, *101*(suppl 3), 193–201. https://doi.org/10.1289/ehp.93101s3193

Alexandrov, L. B., Kim, J., Haradhvala, N. J., Huang, M. N., Tian Ng, A. W., Wu, Y., Boot, A., Covington, K. R., Gordenin, D. A., Bergstrom, E. N., Islam, S. M. A., Lopez-Bigas, N., Klimczak, L. J., McPherson, J. R., Morganella, S., Sabarinathan, R., Wheeler, D. A., Mustonen, V., PCAWG Mutational Signatures Working Group, … Von Mering, C. (2020). The repertoire of mutational

signatures in human cancer. *Nature*, *578*(7793), 94–101. https://doi.org/10.1038/s41586-020-1943-3

Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Campbell, P. J., & Stratton, M. R. (2013). Deciphering Signatures of Mutational Processes Operative in Human Cancer. *Cell Reports*, *3*(1), 246–259. https://doi.org/10.1016/j.celrep.2012.12.008

Amend, S. R., Valkenburg, K. C., & Pienta, K. J. (2016). Murine Hind Limb Long Bone Dissection and Bone Marrow Isolation. *Journal of Visualized Experiments*, (110), 53936. https://doi.org/10.3791/53936

Amin, S., Huie, K., & Hecht, S. S. (1985). Mutagenicity and tumor initiating activity of methylated benzo[b]fluoranthenes. *Carcinogenesis*, *6*(7), 1023–1025. https://doi.org/10.1093/carcin/6.7.1023

Amin, S., LaVoie, E. J., & Hecht, S. S. (1982). Identification of metabolites of benzo[b]fluoranthene. *Carcinogenesis*, *3*(2), 171–174. https://doi.org/10.1093/carcin/3.2.171

Armijo, A. L., Thongararm, P., Fedeles, B. I., Yau, J., Kay, J. E., Corrigan, J. J., Chancharoen, M., Chawanthayatham, S., Samson, L. D., Carrasco, S. E., Engelward, B. P., Fox, J. G., Croy, R. G., & Essigmann, J. M. (2023). Molecular origins of mutational spectra produced by the environmental carcinogen *N* -nitrosodimethylamine and SN1 chemotherapeutic agents. *NAR Cancer*, *5*(2), zcad015. https://doi.org/10.1093/narcan/zcad015

Ashford, A. L., Nachmanson, D., Wills, J. W., Higgins, J. E., Smith, T. H., Vavra, K. C., Dahalan, F. A., Howe, J., Elloway, J. M., Salk, J. J., Doherty, A., & Lynch, A. M. (2025). Alignment between Duplex Sequencing and transgenic rodent mutation

assay data in the assessment of in vivo NDMA-induced mutagenesis. *Archives of Toxicology*, *99*(10), 4227–4242. https://doi.org/10.1007/s00204-025-04121-0

Attolini, C. S.-O., Cheng, Y.-K., Beroukhim, R., Getz, G., Abdel-Wahab, O., Levine, R. L., Mellinghoff, I. K., & Michor, F. (2010). A mathematical framework to determine the temporal sequence of somatic genetic events in cancer. *Proceedings of the National Academy of Sciences*, *107*(41), 17604–17609. https://doi.org/10.1073/pnas.1009117107

Bae, J. H., Liu, R., Roberts, E., Nguyen, E., Tabrizi, S., Rhoades, J., Blewett, T., Xiong, K., Gydush, G., Shea, D., An, Z., Patel, S., Cheng, J., Sridhar, S., Liu, M. H., Lassen, E., Skytte, A.-B., Grońska-Pęski, M., Shoag, J. E., … Adalsteinsson, V. A. (2023). Single duplex DNA sequencing with CODEC detects mutations with high sensitivity. *Nature Genetics*, *55*(5), 871–879. https://doi.org/10.1038/s41588-023-01376-0

Beal, M. A., Gagné, R., Williams, A., Marchetti, F., & Yauk, C. L. (2015). Characterizing Benzo[a]pyrene-induced lacZ mutation spectrum in transgenic mice using next-generation sequencing. *BMC Genomics*, *16*(1), 812. https://doi.org/10.1186/s12864-015-2004-4

Beal, M. A., Meier, M. J., Dykes, A., Yauk, C. L., Lambert, I. B., & Marchetti, F. (2023). The functional mutational landscape of the lacZ gene. *iScience*, *26*(12), 108407. https://doi.org/10.1016/j.isci.2023.108407

Beal, M. A., Meier, M. J., LeBlanc, D. P., Maurice, C., O'Brien, J. M., Yauk, C. L., & Marchetti, F. (2020). Chemically induced mutations in a MutaMouse reporter gene inform mechanisms underlying human cancer mutational signatures. *Communications Biology*, *3*(1), 438. https://doi.org/10.1038/s42003-020-01174-y

Beal, M. A., Yauk, C. L., & Marchetti, F. (2017). From sperm to offspring: Assessing the heritable genetic consequences of paternal smoking and potential public health impacts. *Mutation Research. Reviews in Mutation Research*, *773*, 26–50. https://doi.org/10.1016/j.mrrev.2017.04.001

Bell, S. M., Chang, X., Wambaugh, J. F., Allen, D. G., Bartels, M., Brouwer, K. L. R., Casey, W. M., Choksi, N., Ferguson, S. S., Fraczkiewicz, G., Jarabek, A. M., Ke, A., Lumen, A., Lynn, S. G., Paini, A., Price, P. S., Ring, C., Simon, T. W., Sipes, N. S., … Kleinstreuer, N. C. (2018). In vitro to in vivo extrapolation for high throughput prioritization and decision making. *Toxicology in Vitro*, *47*, 213–227. https://doi.org/10.1016/j.tiv.2017.11.016

Bemis, J. C., Hall, N. E., & Dertinger, S. D. (2019). Erythrocyte-Based Pig-a Gene Mutation Assay. In A. Dhawan & M. Bajpayee (Eds.), *Genotoxicity Assessment* (Vol. 2031, pp. 29–57). Springer New York. https://doi.org/10.1007/978-1-4939-9646-9_2

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bercu, J. P., Zhang, S., Sobol, Z., Escobar, P. A., Van, P., & Schuler, M. (2023). Comparison of the transgenic rodent mutation assay, error corrected next generation duplex sequencing, and the alkaline comet assay to detect dose-related mutations following exposure to N-nitrosodiethylamine. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, *891*, 503685. https://doi.org/10.1016/j.mrgentox.2023.503685

Bergstrom, E. N., Huang, M. N., Mahto, U., Barnes, M., Stratton, M. R., Rozen, S. G., & Alexandrov, L. B. (2019). SigProfilerMatrixGenerator: A tool for visualizing and exploring patterns of small mutational events. *BMC Genomics*, *20*(1), 685. https://doi.org/10.1186/s12864-019-6041-2

Bernard, A., & Dudler, V. (2022). Health Risk Assessment of Dermal Exposure to Polycyclic Aromatic Hydrocarbons from the Use of Infant Diapers. *International Journal of Environmental Research and Public Health*, *19*(22), 14760. https://doi.org/10.3390/ijerph192214760

Besaratinia, A., Li, H., Yoon, J.-I., Zheng, A., Gao, H., & Tommasi, S. (2012). A high-throughput next-generation sequencing-based method for detecting the mutational fingerprint of carcinogens. *Nucleic Acids Research*, *40*(15), e116–e116. https://doi.org/10.1093/nar/gks610

Besaratinia, A., Zheng, A., Bates, S. E., & Tommasi, S. (2018). Mutation Analysis in Cultured Cells of Transgenic Rodents. *International Journal of Molecular Sciences*, *19*(1), 262. https://doi.org/10.3390/ijms19010262

Bessems, J. G. M., & Geraets, L. (2013). Proper knowledge on toxicokinetics improves human hazard testing and subsequent health risk characterisation. A case study approach. *Regulatory Toxicology and Pharmacology*, *67*(3), 325–334. https://doi.org/10.1016/j.yrtph.2013.08.010

Błaszczyk, E., & Mielżyńska-Švach, D. (2017). Polycyclic aromatic hydrocarbons and PAH-related DNA adducts. *Journal of Applied Genetics*, *58*(3), 321–330. https://doi.org/10.1007/s13353-016-0380-3

Blokzijl, F., de Ligt, J., Jager, M., Sasselli, V., Roerink, S., Sasaki, N., Huch, M., Boymans, S., Kuijk, E., Prins, P., Nijman, I. J., Martincorena, I., Mokry, M.,

Wiegerinck, C. L., Middendorp, S., Sato, T., Schwank, G., Nieuwenhuis, E. E. S., Verstegen, M. M. A., … van Boxtel, R. (2016). Tissue-specific mutation accumulation in human adult stem cells during life. *Nature*, *538*(7624), 260–264. https://doi.org/10.1038/nature19768

Boström, C.-E., Gerde, P., Hanberg, A., Jernström, B., Johansson, C., Kyrklund, T., Rannug, A., Törnqvist, M., Victorin, K., & Westerholm, R. (2002). Cancer risk assessment, indicators, and guidelines for polycyclic aromatic hydrocarbons in the ambient air. *Environmental Health Perspectives*, *110*(suppl 3), 451–488. https://doi.org/10.1289/ehp.110-1241197

Boysen, G., Alexandrov, L. B., Rahbari, R., Nookaew, I., Ussery, D., Chao, M.-R., Hu, C.-W., & Cooke, M. S. (2025). Investigating the origins of the mutational signatures in cancer. *Nucleic Acids Research*, *53*(1), gkae1303. https://doi.org/10.1093/nar/gkae1303

Brambilla, G., Cavanna, M., & Parodi, S. (1978). Evaluation of DNA damage and repair in mammalian cells exposed to chemical carcinogens. *Pharmacological Research Communications*, *10*(8), 693–717. https://doi.org/10.1016/S0031-6989(78)80040-X

Brunner, S. F., Roberts, N. D., Wylie, L. A., Moore, L., Aitken, S. J., Davies, S. E., Sanders, M. A., Ellis, P., Alder, C., Hooks, Y., Abascal, F., Stratton, M. R., Martincorena, I., Hoare, M., & Campbell, P. J. (2019). Somatic mutations and clonal dynamics in healthy and cirrhotic human liver. *Nature*, *574*(7779), 538–542. https://doi.org/10.1038/s41586-019-1670-9

Bryce, S. M., Bemis, J. C., & Dertinger, S. D. (2008). In vivo mutation assay based on the endogenous *Pig-a* locus. *Environmental and Molecular Mutagenesis*, *49*(4), 256–264. https://doi.org/10.1002/em.20379

Cagan, A., Baez-Ortega, A., Brzozowska, N., Abascal, F., Coorens, T. H. H., Sanders, M. A., Lawson, A. R. J., Harvey, L. M. R., Bhosle, S., Jones, D., Alcantara, R. E., Butler, T. M., Hooks, Y., Roberts, K., Anderson, E., Lunn, S., Flach, E., Spiro, S., Januszczak, I., … Martincorena, I. (2022). Somatic mutation rates scale with lifespan across mammals. *Nature*, *604*(7906), 517–524. https://doi.org/10.1038/s41586-022-04618-z

Cartiser, N., Bévalot, F., Fanton, L., Gaillard, Y., & Guitton, J. (2011). State-of-the-art of bone marrow analysis in forensic toxicology: A review. *International Journal of Legal Medicine*, *125*(2), 181–198. https://doi.org/10.1007/s00414-010-0525-6

Chakraborty, A., Tapryal, N., Islam, A., Mitra, S., & Hazra, T. (2021). Transcription coupled base excision repair in mammalian cells: So little is known and so much to uncover. *DNA Repair*, *107*, 103204. https://doi.org/10.1016/j.dnarep.2021.103204

Chatterjee, N., & Walker, G. C. (2017). Mechanisms of DNA damage, repair, and mutagenesis. *Environmental and Molecular Mutagenesis*, *58*(5), 235–263. https://doi.org/10.1002/em.22087

Chepelev, N., Long, A. S., Beal, M., Barton-Maclaren, T., Johnson, G., Dearfield, K. L., Roberts, D. J., Van Benthem, J., & White, P. (2023). Establishing a quantitative framework for regulatory interpretation of genetic toxicity dose–response data: Margin of exposure case study of 48 compounds with both in vivo mutagenicity and carcinogenicity dose–response data. *Environmental and Molecular Mutagenesis*, *64*(1), 4–15. https://doi.org/10.1002/em.22517

Cheung, J., Dobo, K., Zhang, S., Nudelman, R., Schmidt, F., Wenzel, J., Czich, A., & Schuler, M. (2024). Evaluation of the nitrosamine impurities of ACE inhibitors using computational, in vitro, and in vivo methods demonstrate no genotoxic

potential. *Environmental and Molecular Mutagenesis*, *65*(6–7), 203–221. https://doi.org/10.1002/em.22618

Chikura, S., Kimoto, T., Itoh, S., Sanada, H., Muto, S., & Horibata, K. (2019). Standard protocol for the total red blood cell Pig-a assay used in the interlaboratory trial organized by the Mammalian Mutagenicity Study Group of the Japanese Environmental Mutagen Society. *Genes and Environment*, *41*(1), 5. https://doi.org/10.1186/s41021-019-0121-z

Chiruvella, K. K., Liang, Z., & Wilson, T. E. (2013). Repair of Double-Strand Breaks by End Joining. *Cold Spring Harbor Perspectives in Biology*, *5*(5), a012757–a012757. https://doi.org/10.1101/cshperspect.a012757

Cho, E., Swartz, C. D., Williams, A., Rivas, M., Recio, L., Witt, K. L., Schmidt, E. K., Yaplee, J., Smith, T. H., Van, P., Lo, F. Y., Valentine, C. C., Salk, J. J., Marchetti, F., Smith-Roe, S. L., & Yauk, C. L. (2023). *Error-corrected Duplex Sequencing enables direct detection and quantification of mutations in human TK6 cells with remarkable inter-laboratory consistency*. Genomics. https://doi.org/10.1101/2023.02.22.529418

Conway, G. E., Chavanel, B., Virard, F., Shah, U.-K., Burgum, M. J., Evans, S. J., Korenjak, M., Thomas, L. E., Jenkins, G. J., Zavadil, J., & Doak, S. H. (2025). Harnessing the power of an advanced *in vitro* 3D liver model and error-corrected duplex sequencing for the detection of mutational signatures. *Mutagenesis*, geaf015. https://doi.org/10.1093/mutage/geaf015

Cordelli, E., Bignami, M., & Pacchierotti, F. (2021). Comet assay: A versatile but complex tool in genotoxicity testing. *Toxicology Research*, *10*(1), 68–78. https://doi.org/10.1093/toxres/tfaa093

Cosentino, L., & Heddle, J. A. (1999). Effects of extended chronic exposures on endogenous and transgenic loci: Implications for low-dose extrapolations. *Environmental and Molecular Mutagenesis*, *34*(2–3), 208–215.

Costello, M., Pugh, T. J., Fennell, T. J., Stewart, C., Lichtenstein, L., Meldrim, J. C., Fostel, J. L., Friedrich, D. C., Perrin, D., Dionne, D., Kim, S., Gabriel, S. B., Lander, E. S., Fisher, S., & Getz, G. (2013). Discovery and characterization of artifactual mutations in deep coverage targeted capture sequencing data due to oxidative DNA damage during sample preparation. *Nucleic Acids Research*, *41*(6), e67–e67. https://doi.org/10.1093/nar/gks1443

Dankovic, D. A., Naumann, B. D., Maier, A., Dourson, M. L., & Levy, L. S. (2015). The Scientific Basis of Uncertainty Factors Used in Setting Occupational Exposure Limits. *Journal of Occupational and Environmental Hygiene*, *12*(sup1), S55–S68. https://doi.org/10.1080/15459624.2015.1060325

David, S. S., O'Shea, V. L., & Kundu, S. (2007). Base-excision repair of oxidative DNA damage. *Nature*, *447*(7147), 941–950. https://doi.org/10.1038/nature05978

De Siqueira Santos, S., Takahashi, D. Y., Nakata, A., & Fujita, A. (2014). A comparative study of statistical methods used to identify dependencies between gene expression signals. *Briefings in Bioinformatics*, *15*(6), 906–918. https://doi.org/10.1093/bib/bbt051

Dertinger, S. D., Bhalli, J. A., Roberts, D. J., Stankowski, L. F., Gollapudi, B. B., Lovell, D. P., Recio, L., Kimoto, T., Miura, D., & Heflich, R. H. (2021). Recommendations for conducting the rodent erythrocyte *Pig-a* assay: A report from the HESI GTTC *Pig-a* Workgroup. *Environmental and Molecular Mutagenesis*, *62*(3), 227–237. https://doi.org/10.1002/em.22427

Dertinger, S. D., Phonethepswath, S., Avlasevich, S. L., Torous, D. K., Mereness, J., Bryce, S. M., Bemis, J. C., Bell, S., Weller, P., & MacGregor, J. T. (2012). Efficient Monitoring of In Vivo Pig-a Gene Mutation and Chromosomal Damage: Summary of 7 Published Studies and Results From 11 New Reference Compounds. *Toxicological Sciences*, *130*(2), 328–348. https://doi.org/10.1093/toxsci/kfs258

Dertinger, S. D., Torous, D. K., Hayashi, M., & MacGregor, J. T. (2011). Flow cytometric scoring of micronucleated erythrocytes: An efficient platform for assessing in vivo cytogenetic damage. *Mutagenesis*, *26*(1), 139–145. https://doi.org/10.1093/mutage/geq055

Díaz-Gay, M., Vangara, R., Barnes, M., Wang, X., Islam, S. M. A., Vermes, I., Duke, S., Narasimman, N. B., Yang, T., Jiang, Z., Moody, S., Senkin, S., Brennan, P., Stratton, M. R., & Alexandrov, L. B. (2023). Assigning mutational signatures to individual samples and individual somatic mutations with SigProfilerAssignment. *Bioinformatics*, *39*(12), btad756. https://doi.org/10.1093/bioinformatics/btad756

Dixit, R., Riviere, J., Krishnan, K., & Andersen, M. (2003). Toxicokinetics and Physiologically Based Toxicokinetics in Toxicology and Risk Assessment. *Journal of Toxicology and Environmental Health, Part B*, *6*(1), 1–40. https://doi.org/10.1080/10937400306479

Dobrovolsky, V. N., Miura, D., Heflich, R. H., & Dertinger, S. D. (2010). The in vivo *pig-a* gene mutation assay, a potential tool for regulatory safety assessment. *Environmental and Molecular Mutagenesis*, *51*(8–9), 825–835. https://doi.org/10.1002/em.20627

Dodge, A. E., LeBlanc, D. P. M., Zhou, G., Williams, A., Meier, M. J., Van, P., Lo, F. Y., Valentine Iii, C. C., Salk, J. J., Yauk, C. L., & Marchetti, F. (2023). Duplex sequencing provides detailed characterization of mutation frequencies and spectra in the bone marrow of MutaMouse males exposed to procarbazine hydrochloride. *Archives of Toxicology*, *97*(8), 2245–2259. https://doi.org/10.1007/s00204-023-03527-y

Dodge, A. E., Williams, A., LeBlanc, D. P. M., Schuster, D. M., Esina, E., Valentine, C. C., Salk, J. J., Maslov, A. Y., Bradley, C., Yauk, C. L., Marchetti, F., & Meier, M. J. (2025). MutSeqR: An open source R package for standardized analysis of error-corrected next-generation sequencing data in genetic toxicology. *Bioinformatics Advances*, *5*(1), vbaf265. https://doi.org/10.1093/bioadv/vbaf265

Douglas, G. R., Beevers, C., Gollapudi, B., Keig-Shevlin, Z., Kirkland, D., O'Brien, J. M., Van Benthem, J., Yauk, C. L., Young, R. R., & Marchetti, F. (2022). Impact of sampling time on the detection of mutations in rapidly proliferating tissues using transgenic rodent gene mutation models: A review. *Environmental and Molecular Mutagenesis*, *63*(8–9), 376–388. https://doi.org/10.1002/em.22514

Dourson, M. L., Felter, S. P., & Robinson, D. (1996). Evolution of Science-Based Uncertainty Factors in Noncancer Risk Assessment. *Regulatory Toxicology and Pharmacology*, *24*(2), 108–120. https://doi.org/10.1006/rtph.1996.0116

EFSA Scientific Committee. (2011). Scientific opinion on genotoxicity testing strategies applicable to food and feed safety assessment. *EFSA Journal*, *9*(9). https://doi.org/10.2903/j.efsa.2011.2379

EFSA Scientific Committee, Hardy, A., Benford, D., Halldorsson, T., Jeger, M. J., Knutsen, K. H., More, S., Mortensen, A., Naegeli, H., Noteborn, H., Ockleford, C., Ricci, A., Rychen, G., Silano, V., Solecki, R., Turck, D., Aerts, M., Bodin, L.,

Davis, A., … Schlatter, J. R. (2017). Update: Use of the benchmark dose approach in risk assessment. *EFSA Journal*, *15*(1). https://doi.org/10.2903/j.efsa.2017.4658

EFSA Scientific Committee, More, S. J., Bampidis, V., Benford, D., Bragard, C., Halldorsson, T. I., Hernández-Jerez, A. F., Bennekou, S. H., Koutsoumanis, K., Lambré, C., Machera, K., Mennes, W., Mullins, E., Nielsen, S. S., Schrenk, D., Turck, D., Younes, M., Aerts, M., Edler, L., … Schlatter, J. (2022). Guidance on the use of the benchmark dose approach in risk assessment. *EFSA Journal*, *20*(10). https://doi.org/10.2903/j.efsa.2022.7584

Erickson, R. P. (2003). Somatic gene mutation and human disease other than cancer. *Mutation Research/Reviews in Mutation Research*, *543*(2), 125–136. https://doi.org/10.1016/S1383-5742(03)00010-3

Escher, S. E., Mangelsdorf, I., Hoffmann-Doerr, S., Partosch, F., Karwath, A., Schroeder, K., Zapf, A., & Batke, M. (2020a). Time extrapolation in regulatory risk assessment: The impact of study differences on the extrapolation factors. *Regulatory Toxicology and Pharmacology*, *112*, 104584. https://doi.org/10.1016/j.yrtph.2020.104584

Esina, E., Dodge, A. E., Williams, A., Schuster, D. M., LeBlanc, D. P. M., Marchetti, F., & Yauk, C. L. (2024). Power analyses to inform Duplex Sequencing study designs for MUTAMOUSE liver and bone marrow. *Environmental and Molecular Mutagenesis*, *65*(8), 234–242. https://doi.org/10.1002/em.22619

Evstyukhina, T. A., Alekseeva, E. A., Peshekhonov, V. T., Skobeleva, I. I., Fedorov, D. V., & Korolev, V. G. (2023). The Role of Chromatin Assembly Factors in Induced Mutagenesis at Low Levels of DNA Damage. *Genes*, *14*(6), 1242. https://doi.org/10.3390/genes14061242

Fliedner, M. C. (2002). Research within the field of blood and marrow transplantation nursing: How can it contribute to higher quality of care? *International Journal of Hematology*, *76*(S2), 289–291. https://doi.org/10.1007/BF03165135

Franchin, A., Gregoris, E., Sorarù, L., Stocco, E., Vindigni, V., Porzionato, A., Brunelli, D., Macchi, V., Gambaro, A., & Roman, M. (2026). First Application of QuEChERS-GC-MS Analysis for Polycyclic Aromatic Hydrocarbons Detection in Human Adipose Tissue. *ACS Omega*, *11*(9), 14287–14295. https://doi.org/10.1021/acsomega.5c06719

Frankish, A., Diekhans, M., Ferreira, A.-M., Johnson, R., Jungreis, I., Loveland, J., Mudge, J. M., Sisu, C., Wright, J., Armstrong, J., Barnes, I., Berry, A., Bignell, A., Carbonell Sala, S., Chrast, J., Cunningham, F., Di Domenico, T., Donaldson, S., Fiddes, I. T., … Flicek, P. (2019). GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Research*, *47*(D1), D766–D773. https://doi.org/10.1093/nar/gky955

Fung, K. Y., Douglas, G. R., & Krewski, D. (1998). Statistical analysis of lacZ mutant frequency data from MutaMouse mutagenicity assays. *Mutagenesis*, *13*(3), 249–255. https://doi.org/10.1093/mutage/13.3.249

Galvan, N., Teske, D., Zhou, G., Moorthy, B., Macwilliams, P., Czuprynski, C., & Jefcoate, C. (2005). Induction of CYP1A1 and CYP1B1 in liver and lung by benzo(a)pyrene and 7,12-d imethylbenz(a)anthracene do not affect distribution of polycyclic hydrocarbons to target tissue: Role of AhR and CYP1B1 in bone marrow cytotoxicity. *Toxicology and Applied Pharmacology*, *202*(3), 244–257. https://doi.org/10.1016/j.taap.2004.06.026

García-Nieto, P. E., Morrison, A. J., & Fraser, H. B. (2019). The somatic mutation landscape of the human body. *Genome Biology*, *20*(1), 298. https://doi.org/10.1186/s13059-019-1919-5

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (1995a). *Bayesian Data Analysis*. Chapman and Hall/CRC. https://doi.org/10.1201/9780429258411

Gingerich, J. D., Soper, L., Lemieux, C. L., Marchetti, F., & Douglas, G. R. (2014). Transgenic Rodent Gene Mutation Assay in Somatic Tissues. In L. M. Sierra & I. Gaivão (Eds.), *Genotoxicity and DNA Repair* (pp. 305–321). Springer New York. https://doi.org/10.1007/978-1-4939-1068-7_18

Ginsberg, G. L., & Atherholt, T. B. (1989). Transport of DNA-adducting metabolites in mouse serum following benzo [*a*]pyrene administration. *Carcinogenesis*, *10*(4), 673–679. https://doi.org/10.1093/carcin/10.4.673

Godschalk, R. W. L., Yauk, C. L., Van Benthem, J., Douglas, G. R., & Marchetti, F. (2020). *In utero* Exposure to Genotoxicants Leading to Genetic Mosaicism: An Overlooked Window of Susceptibility in Genetic Toxicology Testing? *Environmental and Molecular Mutagenesis*, *61*(1), 55–65. https://doi.org/10.1002/em.22347

Gollapudi, B. B., Johnson, G. E., Hernandez, L. G., Pottenger, L. H., Dearfield, K. L., Jeffrey, A. M., Julien, E., Kim, J. H., Lovell, D. P., MacGregor, J. T., Moore, M. M., Van Benthem, J., White, P. A., Zeiger, E., & Thybaud, V. (2013). Quantitative approaches for assessing dose–response relationships in genetic toxicology studies. *Environmental and Molecular Mutagenesis*, *54*(1), 8–18. https://doi.org/10.1002/em.21727

Gollapudi, B. B., Lynch, A. M., Heflich, R. H., Dertinger, S. D., Dobrovolsky, V. N., Froetschl, R., Horibata, K., Kenyon, M. O., Kimoto, T., Lovell, D. P., Stankowski, L. F., White, P. A., Witt, K. L., & Tanir, J. Y. (2015). The in vivo Pig-a assay: A report of the International Workshop On Genotoxicity Testing (IWGT) Workgroup. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, *783*, 23–35. https://doi.org/10.1016/j.mrgentox.2014.09.007

Gorbunova, V., Seluanov, A., Mao, Z., & Hine, C. (2007). Changes in DNA repair during aging. *Nucleic Acids Research*, *35*(22), 7466–7474. https://doi.org/10.1093/nar/gkm756

Gowda, A. S. P., Krzeminski, J., Amin, S., Suo, Z., & Spratt, T. E. (2017). Mutagenic Replication of N^2 -Deoxyguanosine Benzo[*a*]pyrene Adducts by *Escherichia coli* DNA Polymerase I and *Sulfolobus solfataricus* DNA Polymerase IV. *Chemical Research in Toxicology*, *30*(5), 1168–1176. https://doi.org/10.1021/acs.chemrestox.6b00466

Greaves, M., & Maley, C. C. (2012). Clonal evolution in cancer. *Nature*, *481*(7381), 306–313. https://doi.org/10.1038/nature10762

Guengerich, F. P. (2008). Cytochrome P450 and Chemical Toxicology. *Chemical Research in Toxicology*, *21*(1), 70–83. https://doi.org/10.1021/tx700079z

Gustafson, P., Östman, C., & Sällsten, G. (2008). Indoor Levels of Polycyclic Aromatic Hydrocarbons in Homes with or without Wood Burning for Heating. *Environmental Science & Technology*, *42*(14), 5074–5080. https://doi.org/10.1021/es800304y

Halekoh, U., & Højsgaard, S. (2006). *doBy: Groupwise Statistics, LSmeans, Linear Estimates, Utilities* (p. 4.7.1) [Dataset]. https://doi.org/10.32614/CRAN.package.doBy

Halekoh, U., & Højsgaard, S. (2024). *doBy: Groupwise Statistics, LSmeans, Linear Estimates, Utilities* (Version 4.6.21) [Computer software]. https://cran.r-project.org/web/packages/doBy/index.html

Hanahan, D., & Weinberg, R. A. (2011a). Hallmarks of Cancer: The Next Generation. *Cell*, *144*(5), 646–674. https://doi.org/10.1016/j.cell.2011.02.013

Hanawalt, P. C., & Spivak, G. (2008). Transcription-coupled DNA repair: Two decades of progress and surprises. *Nature Reviews Molecular Cell Biology*, *9*(12), 958–970. https://doi.org/10.1038/nrm2549

Harrigan, J. A. (2004). DNA Adduct Formation in Precision-Cut Rat Liver and Lung Slices Exposed to Benzo[a]pyrene. *Toxicological Sciences*, *77*(2), 307–314. https://doi.org/10.1093/toxsci/kfh030

Harvey, R. G. (Ed.). (1985). *Polycyclic Hydrocarbons and Carcinogenesis* (Vol. 283). American Chemical Society. https://doi.org/10.1021/bk-1985-0283

Hashimoto, A. H., Amanuma, K., Hiyoshi, K., Sugawara, Y., Goto, S., Yanagisawa, R., Takano, H., Masumura, K., Nohmi, T., & Aoki, Y. (2007). Mutations in the lungs of *gpt* delta transgenic mice following inhalation of diesel exhaust. *Environmental and Molecular Mutagenesis*, *48*(8), 682–693. https://doi.org/10.1002/em.20335

Hayashi, M. (1984). Kinetics of micronucleus formation in relation to chromosomal aberrations in mouse bone marrow. *Mutation Research/Fundamental and*

Molecular Mechanisms of Mutagenesis, *127*(2), 129–137. https://doi.org/10.1016/0027-5107(84)90014-9

Hearn, B. A., Ding, Y. S., Watson, C. H., Johnson, T. L., Zewdie, G., Jeong-Im, J. H., Walters, M. J., Holman, M. R., & Rochester, C. G. (2018). Multi-year Study of PAHs in Mainstream Cigarette Smoke. *Tobacco Regulatory Science*, *4*(3), 96–106. https://doi.org/10.18001/TRS.4.3.9

Heddle, J. A. (1999). On clonal expansion and its effects on mutant frequencies, mutation spectra and statistics for somatic mutations in vivo. *Mutagenesis*, *14*(3), 257–260. https://doi.org/10.1093/mutage/14.3.257

Heflich, R. H., Dertinger, S. D., Dobrovolsky, V. N., Bhalli, J. A., Kenyon, M. O., Kimoto, T., Miura, D., Horibata, K., MacGregor, J. T., Roberts, D. J., Igl, B.-W., Lovell, D. P., Mittelstaedt, R. A., Shemansky, J. M., Wynne, R. A., & McDaniel, L. P. (2019). *The in vivo erythrocyte Pig-a gene mutation assay Part 1: Detailed review paper and performance assessment*.

Heflich, R. H., Johnson, G. E., Zeller, A., Marchetti, F., Douglas, G. R., Witt, K. L., Gollapudi, B. B., & White, P. A. (2020). Mutation as a Toxicological Endpoint for Regulatory Decision-Making. *Environmental and Molecular Mutagenesis*, *61*(1), 34–41. https://doi.org/10.1002/em.22338

Hess, M. T., Gunz, D., Luneva, N., Geacintov, N. E., & Naegeli, H. (1997). Base Pair Conformation-Dependent Excision of Benzo[*a*]Pyrene Diol Epoxide-Guanine Adducts by Human Nucleotide Excision Repair Enzymes. *Molecular and Cellular Biology*, *17*(12), 7069–7076. https://doi.org/10.1128/MCB.17.12.7069

Hoch, N. C. (2023). Tissue Specificity of DNA Damage and Repair. *Physiology*, *38*(5), 231–241. https://doi.org/10.1152/physiol.00006.2023

Højsgaard, S., & Halekoh, U. (2018). *doBy: Groupwise Statistics, LSmeans, Linear Estimates, Utilities* (Version 4.6.10) [Computer software].

Hou, L., Zhang, X., Wang, D., & Baccarelli, A. (2012). Environmental chemical exposures and human epigenetics. *International Journal of Epidemiology*, *41*(1), 79–105. https://doi.org/10.1093/ije/dyr154

Hu, R., Mukhina, G. L., Piantadosi, S., Barber, J. P., Jones, R. J., & Brodsky, R. A. (2005). PIG-A mutations in normal hematopoiesis. *Blood*, *105*(10), 3848–3854. https://doi.org/10.1182/blood-2004-04-1472

IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. (2010). Some non-heterocyclic polycyclic aromatic hydrocarbons and some related exposures. *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*, *92*, 1–853.

Igl, B.-W., Dertinger, S. D., Dobrovolsky, V. N., Raschke, M., Sutter, A., & Vonk, R. (2018). A statistical approach for analyzing data from the in vivo Pig-a gene mutation assay. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, *831*, 33–44. https://doi.org/10.1016/j.mrgentox.2018.05.007

International Agency for Research on Cancer. (1983). *Polynuclear Aromatic Compounds, Part 1: Chemical, Environmental and Experimental Data* (Vol. 32). IARC Press.

International Agency for Research on Cancer. (1984). *Polynuclear Aromatic Compounds*. IARC Press.

International Agency for Research on Cancer. (1985). *Tobacco Habits Other than Smoking; Betel-Quid and Areca-Nut Chewing; and Some Related Nitrosamines* (Vol. 37). IARC Press.

International Agency for Research on Cancer (Ed.). (2008). *1,3-Butadiene, ethylene oxide and vinyl halides (vinyl fluroide, vinyl chloride and vinyl bromide): This publication represents the views and expert opinions of an IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, which met in Lyon, 5 - 12 June 2007*. IARC.

International Agency for Research on Cancer. (2023). *Occupational Exposure as a Firefighter* (Vol. 132). IARC Press.

Islam, S. M. A., Díaz-Gay, M., Wu, Y., Barnes, M., Vangara, R., Bergstrom, E. N., He, Y., Vella, M., Wang, J., Teague, J. W., Clapham, P., Moody, S., Senkin, S., Li, Y. R., Riva, L., Zhang, T., Gruber, A. J., Steele, C. D., Otlu, B., … Alexandrov, L. B. (2022). Uncovering novel mutational signatures by de novo extraction with SigProfilerExtractor. *Cell Genomics*, *2*(11), 100179. https://doi.org/10.1016/j.xgen.2022.100179

Jackson, M., Marks, L., May, G. H. W., & Wilson, J. B. (2018). The genetic basis of disease. *Essays in Biochemistry*, *62*(5), 643–723. https://doi.org/10.1042/EBC20170053

Jaiswal, S., Fontanillas, P., Flannick, J., Manning, A., Grauman, P. V., Mar, B. G., Lindsley, R. C., Mermel, C. H., Burtt, N., Chavez, A., Higgins, J. M., Moltchanov, V., Kuo, F. C., Kluk, M. J., Henderson, B., Kinnunen, L., Koistinen, H. A., Ladenvall, C., Getz, G., … Ebert, B. L. (2014). Age-Related Clonal Hematopoiesis Associated with Adverse Outcomes. *New England Journal of Medicine*, *371*(26), 2488–2498. https://doi.org/10.1056/NEJMoa1408617

Johnson, R. D., & Jasin, M. (2000). Sister chromatid gene conversion is a prominent double-strand break repair pathway in mammalian cells. *The EMBO Journal*, *19*(13), 3398–3407. https://doi.org/10.1093/emboj/19.13.3398

Jung, K. H., Patel, M. M., Moors, K., Kinney, P. L., Chillrud, S. N., Whyatt, R., Hoepner, L., Garfinkel, R., Yan, B., Ross, J., Camann, D., Perera, F. P., & Miller, R. L. (2010). Effects of Heating Season on Residential Indoor and Outdoor Polycyclic Aromatic Hydrocarbons, Black Carbon, and Particulate Matter in an Urban Birth Cohort. *Atmospheric Environment*, *44*(36), 4545–4552. https://doi.org/10.1016/j.atmosenv.2010.08.024

Kane, D. P., & Shcherbakova, P. V. (2014). A common cancer-associated DNA polymerase ε mutation causes an exceptionally strong mutator phenotype, indicating fidelity defects distinct from loss of proofreading. *Cancer Research*, *74*(7), 1895–1901. https://doi.org/10.1158/0008-5472.CAN-13-2892

Keith, L. H. (2015). The Source of U.S. EPA's Sixteen PAH Priority Pollutants. *Polycyclic Aromatic Compounds*, *35*(2–4), 147–160. https://doi.org/10.1080/10406638.2014.892886

Keith, L., & Telliard, W. (1979). ES&T Special Report: Priority pollutants: I-a perspective view. *Environmental Science & Technology*, *13*(4), 416–423. https://doi.org/10.1021/es60152a601

Kelly, J. M., Ivatt, P. D., Evans, M. J., Kroll, J. H., Hrdina, A. I. H., Kohale, I. N., White, F. M., Engelward, B. P., & Selin, N. E. (2021). Global Cancer Risk From Unregulated Polycyclic Aromatic Hydrocarbons. *GeoHealth*, *5*(9), e2021GH000401. https://doi.org/10.1029/2021GH000401

Kennedy, S. R., Schmitt, M. W., Fox, E. J., Kohrn, B. F., Salk, J. J., Ahn, E. H., Prindle, M. J., Kuong, K. J., Shen, J.-C., Risques, R.-A., & Loeb, L. A. (2014). Detecting ultralow-frequency mutations by Duplex Sequencing. *Nature Protocols*, *9*(11), 2586–2606. https://doi.org/10.1038/nprot.2014.170

Kennedy, S. R., Zhang, Y., & Risques, R. A. (2019). Cancer-Associated Mutations but No Cancer: Insights into the Early Steps of Carcinogenesis and Implications for Early Cancer Detection. *Trends in Cancer*, *5*(9), 531–540. https://doi.org/10.1016/j.trecan.2019.07.007

Khandekar, A., Vangara, R., Barnes, M., Díaz-Gay, M., Abbasi, A., Bergstrom, E. N., Steele, C. D., Pillay, N., & Alexandrov, L. B. (2023). Visualizing and exploring patterns of large mutational events with SigProfilerMatrixGenerator. *BMC Genomics*, *24*(1), 469. https://doi.org/10.1186/s12864-023-09584-y

Kirkland, D., Uno, Y., Luijten, M., Beevers, C., Van Benthem, J., Burlinson, B., Dertinger, S., Douglas, G. R., Hamada, S., Horibata, K., Lovell, D. P., Manjanatha, M., Martus, H.-J., Mei, N., Morita, T., Ohyama, W., & Williams, A. (2019). In vivo genotoxicity testing strategies: Report from the 7th International workshop on genotoxicity testing (IWGT). *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, *847*, 403035. https://doi.org/10.1016/j.mrgentox.2019.03.008

Klapacz, J., Pottenger, L. H., Engelward, B. P., Heinen, C. D., Johnson, G. E., Clewell, R. A., Carmichael, P. L., Adeleye, Y., & Andersen, M. E. (2016). Contributions of DNA repair and damage response pathways to the non-linear genotoxic responses of alkylating agents. *Mutation Research/Reviews in Mutation Research*, *767*, 77–91. https://doi.org/10.1016/j.mrrev.2015.11.001

Kokoska, R. J., Bebenek, K., Boudsocq, F., Woodgate, R., & Kunkel, T. A. (2002). Low Fidelity DNA Synthesis by a Y Family DNA Polymerase Due to Misalignment in the Active Site. *Journal of Biological Chemistry*, *277*(22), 19633–19638. https://doi.org/10.1074/jbc.M202021200

Kostecka, A., Nowikiewicz, T., Olszewski, P., Koczkowska, M., Horbacz, M., Heinzl, M., Andreou, M., Salazar, R., Mair, T., Madanecki, P., Gucwa, M., Davies, H., Skokowski, J., Buckley, P. G., Pęksa, R., Śrutek, E., Szylberg, Ł., Hartman, J., Jankowski, M., … Piotrowski, A. (2022). High prevalence of somatic PIK3CA and TP53 pathogenic variants in the normal mammary gland tissue of sporadic breast cancer patients revealed by duplex sequencing. *Npj Breast Cancer*, *8*(1), 1–10. https://doi.org/10.1038/s41523-022-00443-9

Krewski, D., Acosta, D., Andersen, M., Anderson, H., Bailar, J. C., Boekelheide, K., Brent, R., Charnley, G., Cheung, V. G., Green, S., Kelsey, K. T., Kerkvliet, N. I., Li, A. A., McCray, L., Meyer, O., Patterson, R. D., Pennie, W., Scala, R. A., Solomon, G. M., … Staff Of Committee On Toxicity Testing And Assessment Of Env. (2010). Toxicity Testing in the 21st Century: A Vision and a Strategy. *Journal of Toxicology and Environmental Health, Part B*, *13*(2–4), 51–138. https://doi.org/10.1080/10937404.2010.483176

Kroeger, M. (2006). How omics technologies can contribute to the '3R' principles by introducing new strategies in animal testing. *Trends in Biotechnology*, *24*(8), 343–346. https://doi.org/10.1016/j.tibtech.2006.06.003

Krolewski, B., Little, J. B., & Reynolds, R. J. (1988). Effect of duration of exposure to benzo(a)pyrene diol-epoxide on neoplastic transformation, mutagenesis, cytotoxicity, and total covalent binding to DNA of rodent cells. *Teratogenesis, Carcinogenesis, and Mutagenesis*, *8*(3), 127–136. https://doi.org/10.1002/tcm.1770080302

Kucab, J. E., Nandi, S. P., Al-Serori, H., Dunstone, E., Beck, R. S. S., Caipa Garcia, A. L., Wilde, E. C., Saeed, S., Francies, H., Garnett, M. J., Fu, B., Yang, F., Saeb-Parsy, K., Huch, M., Drost, J., Zilbauer, M., Humphreys, L., Kisby, G., Arlt, V. M.,

… Phillips, D. H. (2026). Mutational signatures of environmental carcinogens in human tissue organoids revealed by duplex sequencing. *Cell Reports*, *45*(6), 117406. https://doi.org/10.1016/j.celrep.2026.117406

Kucab, J. E., Zou, X., Morganella, S., Joel, M., Nanda, A. S., Nagy, E., Gomez, C., Degasperi, A., Harris, R., Jackson, S. P., Arlt, V. M., Phillips, D. H., & Nik-Zainal, S. (2019). A Compendium of Mutational Signatures of Environmental Agents. *Cell*, *177*(4), 821-836.e16. https://doi.org/10.1016/j.cell.2019.03.001

Lai, Z., Markovets, A., Ahdesmaki, M., Chapman, B., Hofmann, O., McEwen, R., Johnson, J., Dougherty, B., Barrett, J. C., & Dry, J. R. (2016). VarDict: A novel and versatile variant caller for next-generation sequencing in cancer research. *Nucleic Acids Research*, *44*(11), e108–e108. https://doi.org/10.1093/nar/gkw227

Lambert, I. B., Singer, T. M., Boucher, S. E., & Douglas, G. R. (2005). Detailed review of transgenic rodent mutation assays. *Mutation Research/Reviews in Mutation Research*, *590*(1–3), 1–280. https://doi.org/10.1016/j.mrrev.2005.04.002

LaVole, E. J., Amin, S., Hecht, S. S., Furuya, K., & Hoffmann, D. (1982). Tumour initiating activity of dihydrodiols of benzo[b]fluoranthene, benzo[j]fluoranthene, and benzo[k]fluoranthene. *Carcinogenesis*, *3*(1), 49–52. https://doi.org/10.1093/carcin/3.1.49

Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M. T., & Carey, V. J. (2013). Software for Computing and Annotating Genomic Ranges. *PLoS Computational Biology*, *9*(8), e1003118. https://doi.org/10.1371/journal.pcbi.1003118

LeBlanc, D. P. M., Meier, M., Lo, F. Y., Schmidt, E., Valentine, C., Williams, A., Salk, J. J., Yauk, C. L., & Marchetti, F. (2022). Duplex sequencing identifies genomic features that determine susceptibility to benzo(a)pyrene-induced in vivo mutations. *BMC Genomics*, *23*(1), 542. https://doi.org/10.1186/s12864-022-08752-w

Lefkofsky, H. B., Veloso, A., & Ljungman, M. (2015). Transcriptional and post-transcriptional regulation of nucleotide excision repair genes in human cells. *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis*, *776*, 9–15. https://doi.org/10.1016/j.mrfmmm.2014.11.008

Li, R., Di, L., Li, J., Fan, W., Liu, Y., Guo, W., Liu, W., Liu, L., Li, Q., Chen, L., Chen, Y., Miao, C., Liu, H., Wang, Y., Ma, Y., Xu, D., Lin, D., Huang, Y., Wang, J., … Wu, C. (2021). A body map of somatic mutagenesis in morphologically normal human tissues. *Nature*, *597*(7876), 398–403. https://doi.org/10.1038/s41586-021-03836-1

Liamin, M., Boutet-Robinet, E., Jamin, E. L., Fernier, M., Khoury, L., Kopp, B., Le Ferrec, E., Vignard, J., Audebert, M., & Sparfel, L. (2017). Benzo[a]pyrene-induced DNA damage associated with mutagenesis in primary human activated T lymphocytes. *Biochemical Pharmacology*, *137*, 113–124. https://doi.org/10.1016/j.bcp.2017.04.025

Liu, B., Xue, Q., Tang, Y., Cao, J., Guengerich, F. P., & Zhang, H. (2016). Mechanisms of mutagenesis: DNA replication in the presence of DNA damage. *Mutation Research/Reviews in Mutation Research*, *768*, 53–67. https://doi.org/10.1016/j.mrrev.2016.03.006

Liu, X., Zhang, X., Zhu, J., Zou, W., Liang, L., Zhang, J., Wen, C., Li, Y., Liu, G., & Xu, X. (2025). BbF-induced liver injury in Balb/c mice: AhR activation as the

conductor of metabolism, oxidative stress, lipid metabolism disorder, and inflammatory response. *Free Radical Biology and Medicine*, *241*, 617–630. https://doi.org/10.1016/j.freeradbiomed.2025.10.008

Long, A. S., Lemieux, C. L., Arlt, V. M., & White, P. A. (2016). Tissue-specific in vivo genetic toxicity of nine polycyclic aromatic hydrocarbons assessed using the Muta™Mouse transgenic rodent assay. *Toxicology and Applied Pharmacology*, *290*, 31–42. https://doi.org/10.1016/j.taap.2015.11.010

Long, A. S., Wills, J. W., Krolak, D., Guo, M., Dertinger, S. D., Arlt, V. M., & White, P. A. (2018). Benchmark dose analyses of multiple genetic toxicity endpoints permit robust, cross-tissue comparisons of MutaMouse responses to orally delivered benzo[a]pyrene. *Archives of Toxicology*, *92*(2), 967–982. https://doi.org/10.1007/s00204-017-2099-2

Luo, K., Zeng, Y., Li, M., Man, Y., Zeng, L., Zhang, Q., Luo, J., & Kang, Y. (2021). Inhalation bioacessibility and absorption of polycyclic aromatic hydrocarbons (PAHs) in indoor PM2.5 and its implication in risk assessment. *Science of The Total Environment*, *774*, 145770. https://doi.org/10.1016/j.scitotenv.2021.145770

Lynch, A. M., Howe, J., Hildebrand, D., Harvey, J. S., Burman, M., Harte, D. S. G., Chen, L., Kmett, C., Shi, W., McHugh, C. F., Patel, K. K., Junnotula, V., Kenny, J., Haworth, R., & Wills, J. W. (2024). N-Nitrosodimethylamine investigations in Muta™Mouse define point-of-departure values and demonstrate less-than-additive somatic mutant frequency accumulations. *Mutagenesis*, *39*(2), 96–118. https://doi.org/10.1093/mutage/geae001

Ma, X., Shao, Y., Tian, L., Flasch, D. A., Mulder, H. L., Edmonson, M. N., Liu, Y., Chen, X., Newman, S., Nakitandwe, J., Li, Y., Li, B., Shen, S., Wang, Z., Shurtleff, S.,

Robison, L. L., Levy, S., Easton, J., & Zhang, J. (2019). Analysis of error profiles in deep next-generation sequencing data. *Genome Biology*, *20*(1), 50. https://doi.org/10.1186/s13059-019-1659-6

MacGregor, J. T., Frötschl, R., White, P. A., Crump, K. S., Eastmond, D. A., Fukushima, S., Guérard, M., Hayashi, M., Soeteman-Hernández, L. G., Johnson, G. E., Kasamatsu, T., Levy, D. D., Morita, T., Müller, L., Schoeny, R., Schuler, M. J., & Thybaud, V. (2015). IWGT report on quantitative approaches to genotoxicity risk assessment II. Use of point-of-departure (PoD) metrics in defining acceptable exposure limits and assessing human risk. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, *783*, 66–78. https://doi.org/10.1016/j.mrgentox.2014.10.008

Mackenzie, K. M., & Murray Angevine, D. (1981). Infertility in Mice Exposed in utero to Benzo(a)pyrene1. *Biology of Reproduction*, *24*(1), 183–191. https://doi.org/10.1095/biolreprod24.1.183

Maertens, R. M., Yang, X., Zhu, J., Gagne, R. W., Douglas, G. R., & White, P. A. (2008). Mutagenic and Carcinogenic Hazards of Settled House Dust I: Polycyclic Aromatic Hydrocarbon Content and Excess Lifetime Cancer Risk from Preschool Exposure. *Environmental Science & Technology*, *42*(5), 1747–1753. https://doi.org/10.1021/es702449c

Manhart, C. M., & Alani, E. (2017). DNA replication and mismatch repair safeguard against metabolic imbalances. *Proceedings of the National Academy of Sciences*, *114*(22), 5561–5563. https://doi.org/10.1073/pnas.1705971114

Mao, P., & Wyrick, J. J. (2019). Organization of DNA damage, excision repair, and mutagenesis in chromatin: A genomic perspective. *DNA Repair*, *81*, 102645. https://doi.org/10.1016/j.dnarep.2019.102645

Marchetti, F., Cardoso, R., Chen, C. L., Douglas, G. R., Elloway, J., Escobar, P. A., Harper Jr, T., Heflich, R. H., Kidd, D., Lynch, A. M., Myers, M. B., Parsons, B. L., Salk, J. J., Settivari, R. S., Smith-Roe, S. L., Witt, K. L., Yauk, C., Young, R. R., Zhang, S., & Minocherhomji, S. (2023a). Error-corrected next-generation sequencing to advance nonclinical genotoxicity and carcinogenicity testing. *Nature Reviews Drug Discovery*, *22*(3), 165–166. https://doi.org/10.1038/d41573-023-00014-y

Marchetti, F., Cardoso, R., Chen, C. L., Douglas, G. R., Joanne Elloway, Escobar, P. A., Harper, T., Heflich, R. H., Kidd, D., Myers, M. B., Parsons, B. L., Salk, J. J., Settivari, R. S., Smith-Roe, S. L., Witt, K. L., Yauk, C. L., Young, R., Zhang, S., & Minocherhomji, S. (2023b). Error-corrected next generation sequencing – Promises and challenges for genotoxicity and cancer risk assessment. *Mutation Research - Reviews in Mutation Research*, *792*, 108466. https://doi.org/10.1016/j.mrrev.2023.108466

Marchetti, F., Douglas, G. R., & Yauk, C. L. (2020). A Return to the Origin of the EMGS: Rejuvenating the Quest for Human Germ Cell Mutagens and Determining the Risk to Future Generations. *Environmental and Molecular Mutagenesis*, *61*(1), 42–54. https://doi.org/10.1002/em.22327

Marchetti, F., Zhou, G., LeBlanc, D., White, P. A., Williams, A., Yauk, C. L., & Douglas, G. R. (2021). The 28 + 28 day design is an effective sampling time for analyzing mutant frequencies in rapidly proliferating tissues of MutaMouse animals. *Archives of Toxicology*, *95*(3), 1103–1116. https://doi.org/10.1007/s00204-021-02977-6

Marra, G., & Schär, P. (1999). Recognition of DNA alterations by the mismatch repair system. *The Biochemical Journal*, *338 (Pt 1)*(Pt 1), 1–13.

Martincorena, I. (2019). Somatic mutation and clonal expansions in human tissues. *Genome Medicine*, *11*(1), 35. https://doi.org/10.1186/s13073-019-0648-4

Martínez-Jiménez, F., Muiños, F., Sentís, I., Deu-Pons, J., Reyes-Salazar, I., Arnedo-Pac, C., Mularoni, L., Pich, O., Bonet, J., Kranas, H., Gonzalez-Perez, A., & Lopez-Bigas, N. (2020). A compendium of mutational cancer driver genes. *Nature Reviews Cancer*, *20*(10), 555–572. https://doi.org/10.1038/s41568-020-0290-x

Martorell, I., Perelló, G., Martí-Cid, R., Castell, V., Llobet, J. M., & Domingo, J. L. (2010). Polycyclic aromatic hydrocarbons (PAH) in foods and estimated PAH intake by the population of Catalonia, Spain: Temporal trend. *Environment International*, *36*(5), 424–432. https://doi.org/10.1016/j.envint.2010.03.003

Maslov, A. Y., Makhortov, S., Sun, S., Heid, J., Dong, X., Lee, M., & Vijg, J. (2022). Single-molecule, quantitative detection of low-abundance somatic mutations by high-throughput sequencing. *Science Advances*, *8*(14), eabm3259. https://doi.org/10.1126/sciadv.abm3259

Mass, M. J., Abu-Shakra, A., Roop, B. C., Nelson, G., Galati, A. J., Stoner, G. D., Nesnow, S., & Ross, J. A. (1996). Benzo[*b*]fluoranthene: Tumorigenicity in strain A/J mouse lungs, DNA adducts and mutations in the Ki- *ras* oncogene. *Carcinogenesis*, *17*(8), 1701–1704. https://doi.org/10.1093/carcin/17.8.1701

Matsumura, S., Sato, H., Otsubo, Y., Tasaki, J., Ikeda, N., & Morita, O. (2019). Genome-wide somatic mutation analysis via Hawk-Seq^{TM} reveals mutation profiles associated with chemical mutagens. *Archives of Toxicology*, *93*(9), 2689–2701. https://doi.org/10.1007/s00204-019-02541-3

McMullin, D. R., Kirkland, A. K., Rehman, I., Kovesi, T., Mallach, G., & Miller, J. D. (2025). Polycyclic aromatic hydrocarbons from environmental tobacco smoke and wood stoves dominate in settled house dust from Northwestern Ontario First Nations communities. *International Journal of Circumpolar Health*, *84*(1), 2457786. https://doi.org/10.1080/22423982.2025.2457786

Meier, M. J., Beal, M. A., Schoenrock, A., Yauk, C. L., & Marchetti, F. (2019). Whole Genome Sequencing of the Mutamouse Model Reveals Strain- and Colony-Level Variation, and Genomic Features of the Transgene Integration Site. *Scientific Reports*, *9*(1), 13775. https://doi.org/10.1038/s41598-019-50302-0

Meier, M. J., O'Brien, J. M., Beal, M. A., Allan, B., Yauk, C. L., & Marchetti, F. (2017). *In Utero* Exposure to Benzo[*a*]Pyrene Increases Mutation Burden in the Soma and Sperm of Adult Mice. *Environmental Health Perspectives*, *125*(1), 82–88. https://doi.org/10.1289/EHP211

Menon, V., & Brash, D. E. (2023). Next-generation sequencing methodologies to detect low-frequency mutations: "Catch me if you can." *Mutation Research - Reviews in Mutation Research*, *792*, 108471. https://doi.org/10.1016/j.mrrev.2023.108471

Menz, J., Götz, M. E., Gündel, U., Gürtler, R., Herrmann, K., Hessel-Pras, S., Kneuer, C., Kolrep, F., Nitzsche, D., Pabel, U., Sachse, B., Schmeisser, S., Schumacher, D. M., Schwerdtle, T., Tralau, T., Zellmer, S., & Schäfer, B. (2023). Genotoxicity assessment: Opportunities, challenges and perspectives for quantitative evaluations of dose–response data. *Archives of Toxicology*, *97*(9), 2303–2328. https://doi.org/10.1007/s00204-023-03553-w

Milholland, B., Dong, X., Zhang, L., Hao, X., Suh, Y., & Vijg, J. (2017). Differences between germline and somatic mutation rates in humans and mice. *Nature Communications*, *8*(1), 15183. https://doi.org/10.1038/ncomms15183

Miranda, J. A., Fenner, K., McKinzie, P. B., Dobrovolsky, V. N., & Revollo, J. R. (2023). Unbiased whole genome detection of ultrarare off-target mutations in genome-edited cell populations by HiFi sequencing. *Environmental and Molecular Mutagenesis*, *64*(7), 374–381. https://doi.org/10.1002/em.22566

Miura, D. (2014). The In Vivo Pig-a Gene Mutation Assay. *Genes and Environment*, *36*(4), 169–173. https://doi.org/10.3123/jemsge.2014.027

Mocquet, V., Kropachev, K., Kolbanovskiy, M., Kolbanovskiy, A., Tapias, A., Cai, Y., Broyde, S., Geacintov, N. E., & Egly, J. (2007). The human DNA repair factor XPC-HR23B distinguishes stereoisomeric benzo[a]pyrenyl-DNA lesions. *The EMBO Journal*, *26*(12), 2923–2932. https://doi.org/10.1038/sj.emboj.7601730

Moorthy, B., Chu, C., & Carlin, D. J. (2015). Polycyclic Aromatic Hydrocarbons: From Metabolism to Lung Cancer. *Toxicological Sciences*, *145*(1), 5–15. https://doi.org/10.1093/toxsci/kfv040

Morita, T., MacGregor, J. T., & Hayashi, M. (2011). Micronucleus assays in rodent tissues other than bone marrow. *Mutagenesis*, *26*(1), 223–230. https://doi.org/10.1093/mutage/geq066

Mughal, A., Vikram, A., Ramarao, P., & Jena, G. B. (2010). Micronucleus and comet assay in the peripheral blood of juvenile rat: Establishment of assay feasibility, time of sampling and the induction of DNA damage. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, *700*(1–2), 86–94. https://doi.org/10.1016/j.mrgentox.2010.05.014

Müller, L., Kikuchi, Y., Probst, G., Schechtman, L., Shimada, H., Sofuni, T., & Tweats, D. (1999). ICH-Harmonised guidances on genotoxicity testing of pharmaceuticals: Evolution, reasoning and impact. *Mutation Research/Reviews in Mutation Research*, *436*(3), 195–225. https://doi.org/10.1016/S1383-5742(99)00004-6

Mulugeta, S., Marchetti, F., Chen, G., Chen, C., Puglisi, R., & White, P. A. (2026). Benchmark Response Values for Error-Corrected Sequencing Mutagenicity Assessment Technologies. *Environmental and Molecular Mutagenesis*, *67*(4), e70051. https://doi.org/10.1002/em.70051

Nangalia, J., Mitchell, E., & Green, A. R. (2019). Clonal approaches to understanding the impact of mutations on hematologic disease development. *Blood*, *133*(13), 1436–1445. https://doi.org/10.1182/blood-2018-11-835405

Nebert, D. W., Dalton, T. P., Okey, A. B., & Gonzalez, F. J. (2004). Role of aryl hydrocarbon receptor-mediated induction of the CYP1 enzymes in environmental toxicity and cancer. *The Journal of Biological Chemistry*, *279*(23), 23847–23850. https://doi.org/10.1074/jbc.R400004200

Nessel, C. S., Lewis, S. C., Stauber, K. L., & Adgate, J. L. (1995). Subchronic to chronic exposure extrapolation: Toxicologic evidence for a reduced uncertainty factor. *Human and Ecological Risk Assessment: An International Journal*, *1*(5), 516–526. https://doi.org/10.1080/10807039509380043

Nguengang Wakap, S., Lambert, D. M., Olry, A., Rodwell, C., Gueydan, C., Lanneau, V., Murphy, D., Le Cam, Y., & Rath, A. (2020). Estimating cumulative point prevalence of rare diseases: Analysis of the Orphanet database. *European Journal of Human Genetics*, *28*(2), 165–173. https://doi.org/10.1038/s41431-019-0508-0

Nik-Zainal, S., Kucab, J. E., Morganella, S., Glodzik, D., Alexandrov, L. B., Arlt, V. M., Weninger, A., Hollstein, M., Stratton, M. R., & Phillips, D. H. (2015). The genome as a record of environmental exposure. *Mutagenesis*, gev073. https://doi.org/10.1093/mutage/gev073

Nishino, H., Schaid, D. J., Buettner, V. L., Haavik, J., & Sommer, S. S. (1996). Mutation frequencies but not mutant frequencies in Big Blue mice fit a Poisson distribution. *Environmental and Molecular Mutagenesis*, *28*(4), 414–417. https://doi.org/10.1002/(SICI)1098-2280(1996)28:4%3C414::AID-EM16%3E3.0.CO;2-I

OECD. (1995). *Revised Guides for Compliance Monitoring Procedures for Good Laboratory Practice*. Organisation for Economic Co-operation and Development (OECD). https://www.oecd.org/en/publications/1995/06/revised-guides-for-compliance-monitoring-procedures-for-good-laboratory-practice_g1gh32ea.html

OECD. (2002). *Guidance Document for the Development of OECD Guidelines for Testing of Chemicals*. OECD. https://doi.org/10.1787/9789264077928-en

OECD. (2016a). *Test No. 474: Mammalian Erythrocyte Micronucleus Test*. OECD. https://doi.org/10.1787/9789264264762-en

OECD. (2016b). *Test No. 476: In Vitro Mammalian Cell Gene Mutation Tests using the Hprt and xprt genes*. OECD. https://doi.org/10.1787/9789264264809-en

OECD. (2017). *Overview of the set of OECD Genetic Toxicology Test Guidelines and updates performed in 2014-2015—Second edition*. OECD. https://doi.org/10.1787/61eca5cd-en

OECD. (2018a). *OECD GUIDELINE FOR THE TESTING OF CHEMICALS: COMBINED CHRONIC TOXICITY\CARCINOGENICITY STUDIES*.

OECD. (2018b). *Test No. 452: Chronic Toxicity Studies*. OECD. https://doi.org/10.1787/9789264071209-en

OECD. (2020a). *Test No. 471: Bacterial Reverse Mutation Test*. OECD Publishing. https://www.oecd.org/en/publications/2020/06/test-no-471-bacterial-reverse-mutation-test_g1gh295b.html

OECD. (2020b). *Test No. 471: Bacterial Reverse Mutation Test*. OECD. https://doi.org/10.1787/9789264071247-en

OECD. (2022). *Test No. 488: Transgenic Rodent Somatic and Germ Cell Gene Mutation Assays*. Organisation for Economic Co-operation and Development. https://www.oecd-ilibrary.org/environment/test-no-488-transgenic-rodent-somatic-and-germ-cell-gene-mutation-assays_9789264203907-en

OECD. (2023). *Test No. 487: In Vitro Mammalian Cell Micronucleus Test*. OECD Publishing. https://www.oecd.org/en/publications/2023/07/test-no-487-in-vitro-mammalian-cell-micronucleus-test_g1g6fb2a.html

OECD. (2025a). *Test No. 407: Repeated Dose 28-day Oral Toxicity Study in Rodents*. OECD Publishing. https://doi.org/10.1787/9789264070684-en

OECD. (2025b). *Test No. 408: Repeated Dose 90-Day Oral Toxicity Study in Rodents*. OECD Publishing. https://doi.org/10.1787/9789264070707-en

OECD. (2025c). *Test No. 470: Mammalian Erythrocyte Pig-a Gene Mutation Assay*. OECD Publishing. https://doi.org/10.1787/4faea90e-en

OECD. (2025d). *Test No. 488: Transgenic Rodent Somatic and Germ Cell Gene Mutation Assays*. OECD Publishing. https://doi.org/10.1787/9789264203907-en

Ohtani, S., Unno, A., Ushiyama, A., Kimoto, T., Miura, D., & Kunugita, N. (2012). *The in vivo Pig-a* gene mutation assay is useful for evaluating the genotoxicity of ionizing radiation in mice. *Environmental and Molecular Mutagenesis*, *53*(8), 579–588. https://doi.org/10.1002/em.21724

Olsen, A., Dertinger, S. D., Krüger, C. T., Eide, D. M., Instanes, C., Brunborg, G., Hartwig, A., & Graupner, A. (2017). The *Pig-a* Gene Mutation Assay in Mice and Human Cells: A Review. *Basic & Clinical Pharmacology & Toxicology*, *121*(S3), 78–92. https://doi.org/10.1111/bcpt.12806

Otsubo, Y., Matsumura, S., Ikeda, N., & Morita, O. (2021). Hawk-Seq™ differentiates between various mutations in *Salmonella typhimurium* TA100 strain caused by exposure to Ames test-positive mutagens. *Mutagenesis*, *36*(3), 245–254. https://doi.org/10.1093/mutage/geab006

Paigen, B., Havens, M. B., & Morrow, A. (1985). *Effect of 3-Methylcholanthrene on the Development of Aortic Lesions in Mice*. *45*.

Parsons, B. L. (2018). Multiclonal tumor origin: Evidence and implications. *Mutation Research/Reviews in Mutation Research*, *777*, 1–18. https://doi.org/10.1016/j.mrrev.2018.05.001

Parsons, B. L., Beal, M. A., Dearfield, K. L., Douglas, G. R., Gi, M., Gollapudi, B. B., Heflich, R. H., Horibata, K., Kenyon, M., Long, A. S., Lovell, D. P., Lynch, A. M., Myers, M. B., Pfuhler, S., Vespa, A., Zeller, A., Johnson, G. E., & White, P. A. (2025). SEVERITY OF EFFECT CONSIDERATIONS REGARDING THE USE OF MUTATION AS

A TOXICOLOGICAL ENDPOINT FOR RISK ASSESSMENT: A REPORT FROM THE 8TH INTERNATIONAL WORKSHOP ON GENOTOXICITY TESTING (IWGT). *Environmental and Molecular Mutagenesis*, *66*(S2), 121–143. https://doi.org/10.1002/em.22599

Pastor-Belda, M., Campillo, N., Arroyo-Manzanares, N., Torres, C., Pérez-Cárceles, M. D., Hernández-Córdoba, M., & Viñas, P. (2019). Bioaccumulation of Polycyclic Aromatic Hydrocarbons for Forensic Assessment Using Gas Chromatography–Mass Spectrometry. *Chemical Research in Toxicology*, *32*(8), 1680–1688. https://doi.org/10.1021/acs.chemrestox.9b00213

Patel, A. B., Shaikh, S., Jain, K. R., Desai, C., & Madamwar, D. (2020). Polycyclic Aromatic Hydrocarbons: Sources, Toxicity, and Remediation Approaches. *Frontiers in Microbiology*, *11*, 562813. https://doi.org/10.3389/fmicb.2020.562813

Paules, R. S., Cordeiro-Stone, M., Mass, M. J., Poirier, M. C., Yuspa, S. H., & Kaufman, D. G. (1988). Benzo[alpha]pyrene diol epoxide I binds to DNA at replication forks. *Proceedings of the National Academy of Sciences*, *85*(7), 2176–2180. https://doi.org/10.1073/pnas.85.7.2176

Pedone, E., Olteanu, V.-A., Marucci, L., Muñoz-Martin, M. I., Youssef, S. A., De Bruin, A., & Cosma, M. P. (2017). Modeling Dynamics and Function of Bone Marrow Cells in Mouse Liver Regeneration. *Cell Reports*, *18*(1), 107–121. https://doi.org/10.1016/j.celrep.2016.12.008

Pelkonen, O., & Raunio, H. (1997). Metabolic activation of toxins: Tissue-specific expression and metabolism in target organs. *Environmental Health Perspectives*, *105*(suppl 4), 767–774. https://doi.org/10.1289/ehp.97105s4767

Pellegrini, W., Facchetti, F., Marocolo, D., Salvi, L., Capucci, A., Tironi, A., & Rossi, G. (1995). Assessment of cell proliferation in normal and pathological bone marrow biopsies: A study using double sequential immunophenotyping on paraffin sections. *Histopathology*, *27*(5), 397–405. https://doi.org/10.1111/j.1365-2559.1995.tb00302.x

Peruzzi, B., Araten, D. J., Notaro, R., & Luzzatto, L. (2010). The use of PIG-A as a sentinel gene for the study of the somatic mutation rate and of mutagenic agents in vivo. *Mutation Research*, *705*(1), 3–10. https://doi.org/10.1016/j.mrrev.2009.12.004

Phillips, D. H. (2018). Mutational spectra and mutational signatures: Insights into cancer aetiology and mechanisms of DNA damage and repair. *DNA Repair*, *71*, 6–11. https://doi.org/10.1016/j.dnarep.2018.08.003

Phonethepswath, S., Franklin, D., Torous, D. K., Bryce, S. M., Bemis, J. C., Raja, S., Avlasevich, S., Weller, P., Hyrien, O., Palis, J., MacGregor, J. T., & Dertinger, S. D. (2010). Pig-a Mutation: Kinetics in Rat Erythrocytes Following Exposure to Five Prototypical Mutagens. *Toxicological Sciences*, *114*(1), 59–70. https://doi.org/10.1093/toxsci/kfp289

Piegorsch, W. W., & Bailer, A. J. (1994). Statistical approaches for analyzing mutational spectra: Some recommendations for categorical data. *Genetics*, *136*(1), 403–416. https://doi.org/10.1093/genetics/136.1.403

Pieters, M. N., Kramer, H. J., & Slob, W. (1998). Evaluation of the Uncertainty Factor for Subchronic-to-Chronic Extrapolation: Statistical Analysis of Toxicity Data. *Regulatory Toxicology and Pharmacology*, *27*(2), 108–111. https://doi.org/10.1006/rtph.1997.1196

Poon, S. L., McPherson, J. R., Tan, P., Teh, B. T., & Rozen, S. G. (2014). Mutation signatures of carcinogen exposure: Genome-wide detection and new opportunities for cancer prevention. *Genome Medicine*, *6*(3), 24. https://doi.org/10.1186/gm541

R: The R Project for Statistical Computing. (n.d.). Retrieved May 7, 2024, from https://www.r-project.org/

Ramesh, A., Walker, S. A., Hood, D. B., Guillén, M. D., Schneider, K., & Weyand, E. H. (2004). Bioavailability and Risk Assessment of Orally Ingested Polycyclic Aromatic Hydrocarbons. *International Journal of Toxicology*, *23*(5), 301–333. https://doi.org/10.1080/10915810490517063

Ren, N., Atyah, M., Chen, W.-Y., & Zhou, C.-H. (2017). The various aspects of genetic and epigenetic toxicology: Testing methods and clinical applications. *Journal of Translational Medicine*, *15*(1), 110. https://doi.org/10.1186/s12967-017-1218-4

Russell, W. M. S., & Burch, R. L. (1959). *The Principles of Humane Experimental Technique*. Methuen & Co Ltd.

Sadler, N. C., Nandhikonda, P., Webb-Robertson, B.-J., Ansong, C., Anderson, L. N., Smith, J. N., Corley, R. A., & Wright, A. T. (2016). Hepatic Cytochrome P450 Activity, Abundance, and Expression Throughout Human Development. *Drug Metabolism and Disposition*, *44*(7), 984–991. https://doi.org/10.1124/dmd.115.068593

Sale, J. E. (2013). Translesion DNA synthesis and mutagenesis in eukaryotes. *Cold Spring Harbor Perspectives in Biology*, *5*(3), a012708. https://doi.org/10.1101/cshperspect.a012708

Salk, J. J., & Kennedy, S. R. (2020). Next-Generation Genotoxicology: Using Modern Sequencing Technologies to Assess Somatic Mutagenesis and Cancer Risk. *Environmental and Molecular Mutagenesis*, *61*(1), 135–151. https://doi.org/10.1002/em.22342

Salk, J. J., Schmitt, M. W., & Loeb, L. A. (2018). Enhancing the accuracy of next-generation sequencing for detecting rare and subclonal mutations. *Nature Reviews Genetics*, *19*(5), 269–285. https://doi.org/10.1038/nrg.2017.117

Schmitt, M. W., Kennedy, S. R., Salk, J. J., Fox, E. J., Hiatt, J. B., & Loeb, L. A. (2012). Detection of ultra-rare mutations by next-generation sequencing. *Proceedings of the National Academy of Sciences*, *109*(36), 14508–14513. https://doi.org/10.1073/pnas.1208715109

Schneider, G., Schmidt-Supprian, M., Rad, R., & Saur, D. (2017). Tissue-specific tumorigenesis: Context matters. *Nature Reviews Cancer*, *17*(4), 239–253. https://doi.org/10.1038/nrc.2017.5

Schreck, I., Chudziak, D., Schneider, S., Seidel, A., Platt, K. L., Oesch, F., & Weiss, C. (2009). Influence of aryl hydrocarbon- (Ah) receptor and genotoxins on DNA repair gene expression and cell survival of mouse hepatoma cells. *Toxicology*, *259*(3), 91–96. https://doi.org/10.1016/j.tox.2009.02.006

Schuster, D., LeBlanc, D. P. M., Zhou, G., Meier, M. J., Dodge, A. E., White, P. A., Long, A. S., Williams, A., Hobbs, C., Diesing, A., Smith-Roe, S. L., Salk, J. J., Marchetti, F., & Yauk, C. L. (2024). Dose-Related Mutagenic and Clastogenic Effects of Benzo[*b*]fluoranthene in Mouse Somatic Tissues Detected by Duplex Sequencing and the Micronucleus Assay. *Environmental Science & Technology*, *58*(49), 21450–21463. https://doi.org/10.1021/acs.est.4c07236

Senkin, S., Moody, S., Díaz-Gay, M., Abedi-Ardekani, B., Cattiaux, T., Ferreiro-Iglesias, A., Wang, J., Fitzgerald, S., Kazachkova, M., Vangara, R., Le, A. P., Bergstrom, E. N., Khandekar, A., Otlu, B., Cheema, S., Latimer, C., Thomas, E., Atkins, J. R., Smith-Byrne, K., … Brennan, P. (2024). Geographic variation of mutagenic exposures in kidney cancer genomes. *Nature*, *629*(8013), 910–918. https://doi.org/10.1038/s41586-024-07368-2

Seo, J.-E., Le, Y., Revollo, J., Miranda-Colon, J., Xu, H., McKinzie, P., Mei, N., Chen, T., Heflich, R. H., Zhou, T., Robison, T., Bonzo, J. A., & Guo, X. (2024). Evaluating the mutagenicity of N-nitrosodimethylamine in 2D and 3D HepaRG cell cultures using error-corrected next generation sequencing. *Archives of Toxicology*, *98*(6), 1919–1935. https://doi.org/10.1007/s00204-024-03731-4

Seplyarskiy, V. B., Soldatov, R. A., Koch, E., McGinty, R. J., Goldmann, J. M., Hernandez, R. D., Barnes, K., Correa, A., Burchard, E. G., Ellinor, P. T., McGarvey, S. T., Mitchell, B. D., Vasan, R. S., Redline, S., Silverman, E., Weiss, S. T., Arnett, D. K., Blangero, J., Boerwinkle, E., … Sunyaev, S. (2021). Population sequencing data reveal a compendium of mutational processes in the human germ line. *Science*, *373*(6558), 1030–1035. https://doi.org/10.1126/science.aba7408

Shiizaki, K., Kawanishi, M., & Yagi, T. (2017). Modulation of benzo[a]pyrene–DNA adduct formation by CYP1 inducer and inhibitor. *Genes and Environment*, *39*(1), 14. https://doi.org/10.1186/s41021-017-0076-x

Shimada, T. (2006). Xenobiotic-Metabolizing Enzymes Involved in Activation and Detoxification of Carcinogenic Polycyclic Aromatic Hydrocarbons. *Drug Metabolism and Pharmacokinetics*, *21*(4), 257–276. https://doi.org/10.2133/dmpk.21.257

Shuck, S. C., Short, E. A., & Turchi, J. J. (2008). Eukaryotic nucleotide excision repair: From understanding mechanisms to influencing biology. *Cell Research*, *18*(1), 64–72. https://doi.org/10.1038/cr.2008.2

Shwed, P. S., Crosthwait, J., Douglas, G. R., & Seligy, V. L. (2010). Characterisation of MutaTMMouse gt10-lacZ transgene: Evidence for in vivo rearrangements. *Mutagenesis*, *25*(6), 609–616. https://doi.org/10.1093/mutage/geq048

Sidak, Z. (1967). Rectangular Confidence Regions for the Means of Multivariate Normal Distributions. *Journal of the American Statistical Association*, *62*(318), 626. https://doi.org/10.2307/2283989

Simon, S., Schlingemann, J., Johnson, G., Brenneis, C., Guessregen, B., Kostal, J., & Dieckhoff, J. (2025). Deriving safe limits for N-nitroso-bisoprolol by error-corrected next-generation sequencing (ecNGS) and benchmark dose (BMD) analysis, integrated with QM modeling and CYP-docking analysis. *Archives of Toxicology*, *99*(10), 3935–3962. https://doi.org/10.1007/s00204-025-04103-2

Singh, D., Mathur, A., Arora, S., Roy, S., & Mahindroo, N. (2022). Journey of organ on a chip technology and its role in future healthcare scenario. *Applied Surface Science Advances*, *9*, 100246. https://doi.org/10.1016/j.apsadv.2022.100246

Singh, R., Kaur, B., Kalina, I., Popov, T. A., Georgieva, T., Garte, S., Binkova, B., Sram, R. J., Taioli, E., & Farmer, P. B. (2007). Effects of environmental air pollution on endogenous oxidative DNA damage in humans. *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis*, *620*(1–2), 71–82. https://doi.org/10.1016/j.mrfmmm.2007.02.024

Skopek, T. R., Kort, K. L., Marino, D. R., Mittal, L. V., Umbenhauer, D. R., Laws, G. M., & Adams, S. P. (1996). System Issues: Mutagenic response of the

endogenoushprt gene andlacI transgene in benzo[a]pyrene-treated Big Blue™ B6C3F1 mice. *Environmental and Molecular Mutagenesis*, *28*(4), 376–384. https://doi.org/10.1002/(SICI)1098-2280(1996)28:4%3C376::AID-EM11%3E3.0.CO;2-C

Slob, W. (2017). A general theory of effect size, and its consequences for defining the benchmark response (BMR) for continuous endpoints. *Critical Reviews in Toxicology*, *47*(4), 342–351. https://doi.org/10.1080/10408444.2016.1241756

Smith, J. N., Gaither, K. A., & Pande, P. (2022). Competitive Metabolism of Polycyclic Aromatic Hydrocarbons (PAHs): An Assessment Using In Vitro Metabolism and Physiologically Based Pharmacokinetic (PBPK) Modeling. *International Journal of Environmental Research and Public Health*, *19*(14), 8266. https://doi.org/10.3390/ijerph19148266

Smith-Roe, S. L., Hobbs, C. A., Hull, V., Todd Auman, J., Recio, L., Streicker, M. A., Rivas, M. V., Pratt, G. A., Lo, F. Y., Higgins, J. E., Schmidt, E. K., Williams, L. N., Nachmanson, D., Valentine Iii, C. C., Salk, J. J., & Witt, K. L. (2023). Adopting duplex sequencing technology for genetic toxicity testing: A proof-of-concept mutagenesis experiment with N-ethyl-N-nitrosourea (ENU)-exposed rats. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, *891*, 503669. https://doi.org/10.1016/j.mrgentox.2023.503669

Sørensen, M., Autrup, H., Møller, P., Hertel, O., Jensen, S. S., Vinzents, P., Knudsen, L. E., & Loft, S. (2003). Linking exposure to environmental pollutants with biological effects. *Mutation Research/Reviews in Mutation Research*, *544*(2–3), 255–271. https://doi.org/10.1016/j.mrrev.2003.06.010

Spivak, G. (2016). Transcription-coupled repair: An update. *Archives of Toxicology*, *90*(11), 2583–2594. https://doi.org/10.1007/s00204-016-1820-x

Srogi, K. (2007). Monitoring of environmental exposure to polycyclic aromatic hydrocarbons: A review. *Environmental Chemistry Letters*, *5*(4), 169–195. https://doi.org/10.1007/s10311-007-0095-0

Staresincic, L., Fagbemi, A. F., Enzlin, J. H., Gourdin, A. M., Wijgers, N., Dunand-Sauthier, I., Giglia-Mari, G., Clarkson, S. G., Vermeulen, W., & Schärer, O. D. (2009). Coordination of dual incision and repair synthesis in human nucleotide excision repair. *The EMBO Journal*, *28*(8), 1111–1120. https://doi.org/10.1038/emboj.2009.49

Stepanov, I., Villalta, P. W., Knezevich, A., Jensen, J., Hatsukami, D., & Hecht, S. S. (2010). Analysis of 23 Polycyclic Aromatic Hydrocarbons in Smokeless Tobacco by Gas Chromatography−Mass Spectrometry. *Chemical Research in Toxicology*, *23*(1), 66–73. https://doi.org/10.1021/tx900281u

Stinson, B. M., Moreno, A. T., Walter, J. C., & Loparo, J. J. (2020). A Mechanism to Minimize Errors during Non-homologous End Joining. *Molecular Cell*, *77*(5), 1080-1091.e8. https://doi.org/10.1016/j.molcel.2019.11.018

Stoddard, E. G., Nag, S., Martin, J., Tyrrell, K. J., Gibbins, T., Anderson, K. A., Shukla, A. K., Corley, R., Wright, A. T., & Smith, J. N. (2021). Exposure to an Environmental Mixture of Polycyclic Aromatic Hydrocarbons Induces Hepatic Cytochrome P450 Enzymes in Mice. *Chemical Research in Toxicology*, *34*(9), 2145–2156. https://doi.org/10.1021/acs.chemrestox.1c00235

Suciu, I., Pamies, D., Peruzzo, R., Wirtz, P. H., Smirnova, L., Pallocca, G., Hauck, C., Cronin, M. T. D., Hengstler, J. G., Brunner, T., Hartung, T., Amelio, I., & Leist, M. (2023). G × E interactions as a basis for toxicological uncertainty. *Archives of Toxicology*, *97*(7), 2035–2049. https://doi.org/10.1007/s00204-023-03500-9

Sun, S., Osterman, M. D., & Li, M. (2019). Tissue specificity of DNA damage response and tumorigenesis. *Cancer Biology & Medicine*, *16*(3), 396–414. https://doi.org/10.20892/j.issn.2095-3941.2019.0097

Sun, X., Dertinger, S. D., Elhajouji, A., Roberts, D. J., Van Benthem, J., Schuler, M., Roy, S., Taraboletti, A., & Chen, C. L. (2025). Critical Evaluation of Methods for the Identification of Aneugens. *Environmental and Molecular Mutagenesis*, *66*(8), 412–424. https://doi.org/10.1002/em.70023

Sun, X., Schuler, M., Zhang, S., Kostal, J., Cheung, J. R., Kenyon, M., Watt, E., & Dobo, K. (2026). Complex versus simple *N* -nitrosamines: Comprehensive genotoxicity and in silico carcinogenicity assessment toward future testing paradigms. *Science Advances*, *12*(22), eaee8680. https://doi.org/10.1126/sciadv.aee8680

Sundberg, C. D., & Hankinson, O. (2019). A CRISPR/Cas9 Whole-Genome Screen Identifies Genes Required for Aryl Hydrocarbon Receptor-Dependent Induction of Functional CYP1A1. *Toxicological Sciences: An Official Journal of the Society of Toxicology*, *170*(2), 310–319. https://doi.org/10.1093/toxsci/kfz111

Tan, D., & Lynch, H. T. (Eds.). (2013). *Principles of Molecular Diagnostics and Personalized Cancer Medicine*. Lippincott Williams & Wilkins.

Thurston, S. W., Ryan, L., Christiani, D. C., Snow, R., Carlson, J., You, L., Cui, S., Ma, G., Wang, L., Huang, Y., & Xu, X. (2000). Petrochemical exposure and menstrual disturbances. *American Journal of Industrial Medicine*, *38*(5), 555–564. https://doi.org/10.1002/1097-0274(200011)38:5%3C555::AID-AJIM8%3E3.0.CO;2-E

Thybaud, V., Lorge, E., Levy, D. D., van Benthem, J., Douglas, G. R., Marchetti, F., Moore, M. M., & Schoeny, R. (2017). Main issues addressed in the 2014-2015 revisions to the OECD Genetic Toxicology Test Guidelines. *Environmental and Molecular Mutagenesis*, *58*(5), 284–295. https://doi.org/10.1002/em.22079

Tomasetti, C., & Vogelstein, B. (2015). Variation in cancer risk among tissues can be explained by the number of stem cell divisions. *Science*, *347*(6217), 78–81. https://doi.org/10.1126/science.1260825

Tomei, G., Ciarrocca, M., Fortunato, B. R., Capozzella, A., Rosati, M. V., Cerratti, D., Tomao, E., Anzelmo, V., Monti, C., & Tomei, F. (2006). Exposure to traffic pollutants and effects on 17-β-estradiol (E2) in female workers. *International Archives of Occupational and Environmental Health*, *80*(1), 70–77. https://doi.org/10.1007/s00420-006-0105-8

Ueda, S., Yamashita, S., Nakajima, M., Kumamoto, T., Ogawa, C., Liu, Y., Yamada, H., Kubo, E., Hattori, N., Takeshima, H., Wakabayashi, M., Iida, N., Shiraishi, Y., Noguchi, M., Sato, Y., & Ushijima, T. (2022). A quantification method of somatic mutations in normal tissues and their accumulation in pediatric patients with chemotherapy. *Proceedings of the National Academy of Sciences*, *119*(31), e2123241119. https://doi.org/10.1073/pnas.2123241119

Valentine, C. C., Young, R. R., Fielden, M. R., Kulkarni, R., Williams, L. N., Li, T., Minocherhomji, S., & Salk, J. J. (2020). Direct quantification of in vivo mutagenesis and carcinogenesis using duplex sequencing. *Proceedings of the National Academy of Sciences*, *117*(52), 33414–33425. https://doi.org/10.1073/pnas.2013724117

Verma, N., Pink, M., Rettenmeier, A. W., & Schmitz-Spanke, S. (2012). Review on proteomic analyses of benzo[a]pyrene toxicity. *PROTEOMICS*, *12*(11), 1731–1755. https://doi.org/10.1002/pmic.201100466

Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz, L. A., & Kinzler, K. W. (2013). Cancer Genome Landscapes. *Science*, *339*(6127), 1546–1558. https://doi.org/10.1126/science.1235122

Vorrink, S., Hudachek, D., & Domann, F. (2014). Epigenetic Determinants of CYP1A1 Induction by the Aryl Hydrocarbon Receptor Agonist 3,3',4,4',5-Pentachlorobiphenyl (PCB 126). *International Journal of Molecular Sciences*, *15*(8), 13916–13931. https://doi.org/10.3390/ijms150813916

Vu, A. T., Taylor, K. M., Holman, M. R., Ding, Y. S., Hearn, B., & Watson, C. H. (2015). Polycyclic Aromatic Hydrocarbons in the Mainstream Smoke of Popular U.S. Cigarettes. *Chemical Research in Toxicology*, *28*(8), 1616–1626. https://doi.org/10.1021/acs.chemrestox.5b00190

Wang, J., Li, C., Han, J., Xue, Y., Zheng, X., Wang, R., Radak, Z., Nakabeppu, Y., Boldogh, I., & Ba, X. (2025). Reassessing the roles of oxidative DNA base lesion 8-oxoGua and repair enzyme OGG1 in tumorigenesis. *Journal of Biomedical Science*, *32*(1), 1. https://doi.org/10.1186/s12929-024-01093-8

Wang, Y., Mittelstaedt, R. A., Wynne, R., Chen, Y., Cao, X., Muskhelishvili, L., Davis, K., Robison, T. W., Sun, W., Schmidt, E. K., Smith, T. H., Norgaard, Z. K., Valentine, C. C., Yaplee, J., Williams, L. N., Salk, J. J., & Heflich, R. H. (2021). Genetic toxicity testing using human in vitro organotypic airway cultures: Assessing DNA damage with the CometChip and mutagenesis by Duplex Sequencing. *Environmental and Molecular Mutagenesis*, *62*(5), 306–318. https://doi.org/10.1002/em.22444

Wenger, A. M., Peluso, P., Rowell, W. J., Chang, P.-C., Hall, R. J., Concepcion, G. T., Ebler, J., Fungtammasan, A., Kolesnikov, A., Olson, N. D., Töpfer, A., Alonge, M., Mahmoud, M., Qian, Y., Chin, C.-S., Phillippy, A. M., Schatz, M. C., Myers, G., DePristo, M. A., … Hunkapiller, M. W. (2019). Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. *Nature Biotechnology*, *37*(10), 1155–1162. https://doi.org/10.1038/s41587-019-0217-9

Wheeler, M. W., Blessinger, T., Shao, K., Allen, B. C., Olszyk, L., Davis, J. A., & Gift, J. S. (2020). Quantitative Risk Assessment: Developing a Bayesian Approach to Dichotomous Dose–Response Uncertainty. *Risk Analysis*, *40*(9), 1706–1722. https://doi.org/10.1111/risa.13537

Wheeler, M. W., Cortiñas Abrahantes, J., Aerts, M., Gift, J. S., & Allen Davis, J. (2022). Continuous model averaging for benchmark dose analysis: Averaging over distributional forms. *Environmetrics*, *33*(5), e2728. https://doi.org/10.1002/env.2728

Wheeler, M. W., Lim, S., House, J. S., Shockley, K. R., John Bailer, A., Fostel, J., Yang, L., Talley, D., Raghuraman, A., Gift, J. S., Allen Davis, J., Auerbach, S. S., & Motsinger-Reif, A. A. (2023). ToxicR: A computational platform in R for computational toxicology and dose–response analyses. *Computational Toxicology*, *25*, 100259. https://doi.org/10.1016/j.comtox.2022.100259

White, P. A., Chen, G., Chepelev, N., Bell, M. A., Gallant, L. R., Johnson, G. E., Zeller, A., Beal, M. A., & Long, A. S. (2025). Benchmark Response (BMR) Values for In Vivo Mutagenicity Endpoints. *Environmental and Molecular Mutagenesis*, *66*(4), 172–184. https://doi.org/10.1002/em.70006

White, P. A., & Johnson, G. E. (2016). Genetic toxicology at the crossroads—From qualitative hazard evaluation to quantitative risk assessment. *Mutagenesis*, *31*(3), 233–237. https://doi.org/10.1093/mutage/gew011

White, P. A., Long, A. S., & Johnson, G. E. (2020). Quantitative Interpretation of Genetic Toxicity Dose-Response Data for Risk Assessment and Regulatory Decision-Making: Current Status and Emerging Priorities. *Environmental and Molecular Mutagenesis*, *61*(1), 66–83. https://doi.org/10.1002/em.22351

White, P. A., Luijten, M., Mishima, M., Cox, J. A., Hanna, J. N., Maertens, R. M., & Zwart, E. P. (2019). In vitro mammalian cell mutation assays based on transgenic reporters: A report of the International Workshop on Genotoxicity Testing (IWGT). *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, *847*, 403039. https://doi.org/10.1016/j.mrgentox.2019.04.002

Whitehead, T. P., Metayer, C., Petreas, M., Does, M., Buffler, P. A., & Rappaport, S. M. (2013). Polycyclic aromatic hydrocarbons in residential dust: Sources of variability. *Environmental Health Perspectives*, *121*(5), 543–550. https://doi.org/10.1289/ehp.1205821

Wiebauer, K., Neddermann, P., Hughes, M., & Jiricny, J. (1993). The repair of 5-methylcytosine deamination damage. In J.-P. Jost & H.-P. Saluz (Eds.), *DNA Methylation* (pp. 510–522). Birkhäuser Basel. https://doi.org/10.1007/978-3-0348-9118-9_23

Wilk, A., Waligórski, P., Lassak, A., Vashistha, H., Lirette, D., Tate, D., Zea, A. H., Koochekpour, S., Rodriguez, P., Meggs, L. G., Estrada, J. J., Ochoa, A., & Reiss, K. (2013). Polycyclic aromatic hydrocarbons—Induced ROS accumulation enhances mutagenic potential of T-antigen from human

polyomavirus JC. *Journal of Cellular Physiology*, *228*(11), 2127–2138. https://doi.org/10.1002/jcp.24375

Williams, A., Zhang, S., Li, D., Sun, W., Schuler, M. J., Marchetti, F., & Yauk, C. L. (2026). Evaluation of Type 1 Error Rates in Duplex Sequencing for Mutagenicity Testing Using Vehicle Control Data and Simulation Analyses. *Environmental and Molecular Mutagenesis*, *67*(5–7), e70067. https://doi.org/10.1002/em.70067

Wills, J. W., Johnson, G. E., Doak, S. H., Soeteman-Hernández, L. G., Slob, W., & White, P. A. (2016). Empirical analysis of BMD metrics in genetic toxicology part I: *In vitro* analyses to provide robust potency rankings and support MOA determinations. *Mutagenesis*, *31*(3), 255–263. https://doi.org/10.1093/mutage/gev085

Wilson, A., Laurenti, E., Oser, G., Van Der Wath, R. C., Blanco-Bose, W., Jaworski, M., Offner, S., Dunant, C. F., Eshkind, L., Bockamp, E., Lió, P., MacDonald, H. R., & Trumpp, A. (2008). Hematopoietic Stem Cells Reversibly Switch from Dormancy to Self-Renewal during Homeostasis and Repair. *Cell*, *135*(6), 1118–1129. https://doi.org/10.1016/j.cell.2008.10.048

Wright, S. P. (1992). Adjusted P-Values for Simultaneous Inference. *Biometrics*, *48*(4), 1005. https://doi.org/10.2307/2532694

Wu, J., Cui, S., Li, X., Zhang, X., Yang, S., Sun, J., & Jiang, X. (2025). Association between polycyclic aromatic hydrocarbons exposure and metabolic dysfunction-associated steatotic liver disease in US adults. *Frontiers in Public Health*, *13*, 1540357. https://doi.org/10.3389/fpubh.2025.1540357

Wyrick, J. J., & Roberts, S. A. (2015). Genomic approaches to DNA repair and mutagenesis. *DNA Repair*, *36*, 146–155. https://doi.org/10.1016/j.dnarep.2015.09.018

Xie, M., Lu, C., Wang, J., McLellan, M. D., Johnson, K. J., Wendl, M. C., McMichael, J. F., Schmidt, H. K., Yellapantula, V., Miller, C. A., Ozenberger, B. A., Welch, J. S., Link, D. C., Walter, M. J., Mardis, E. R., Dipersio, J. F., Chen, F., Wilson, R. K., Ley, T. J., & Ding, L. (2014). Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nature Medicine*, *20*(12), 1472–1478. https://doi.org/10.1038/nm.3733

Yang, J., Wang, Q., Zhuo, Q., Tian, H., Li, W., Luo, F., Zhang, J., Bi, D., Peng, J., Zhou, D., & Xin, H. (2018). A likely pathogenic variant putatively affecting splicing of *PIGA* identified in a multiple congenital anomalies hypotonia-seizures syndrome 2 (MCAHS 2) family pedigree via whole-exome sequencing. *Molecular Genetics & Genomic Medicine*, *6*(5), 739–748. https://doi.org/10.1002/mgg3.428

Yang, L., Zhang, H., Zhang, X., Xing, W., Wang, Y., Bai, P., Zhang, L., Hayakawa, K., Toriba, A., & Tang, N. (2021). Exposure to Atmospheric Particulate Matter-Bound Polycyclic Aromatic Hydrocarbons and Their Health Effects: A Review. *International Journal of Environmental Research and Public Health*, *18*(4), 2177. https://doi.org/10.3390/ijerph18042177

Yang, M., Mao, K., Cao, X., Liu, H., Wang, X., Mao, W., & Hao, L. (2023). Metabolism and elimination kinetics of mono-hydroxylated PAHs metabolites following single exposure to different combinations of PAH4 in rats. *Journal of Applied Toxicology*, *43*(11), 1594–1603. https://doi.org/10.1002/jat.4497

Yang, S. K. (1988). Stereoselectivity of cytochrome P-450 isozymes and epoxide hydrolase in the metabolism of polycyclic aromatic hydrocarbons. *Biochemical Pharmacology*, *37*(1), 61–70. https://doi.org/10.1016/0006-2952(88)90755-1

Yang, W., & Woodgate, R. (2007). What a difference a decade makes: Insights into translesion DNA synthesis. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(40), 15591–15598. https://doi.org/10.1073/pnas.0704219104

Yauk, C. L., Lynch, A. M., Dobrovolsky, V. N., Schuler, M., Smith-Roe, S. L., Fitzgerald, D., Higgins, J., Honarvar, N., Le Curieux, F., Matsumura, S., Minocherhomji, S., Recio, L., Salk, J. J., Sugiyama, K.-I., Suzuki, T., Wills, J. W., & Marchetti, F. (2026). Application of error-corrected sequencing technologies for in vivo regulatory mutagenicity assessment. *Regulatory Toxicology and Pharmacology: RTP*, *164*, 105985. https://doi.org/10.1016/j.yrtph.2025.105985

Yeh, S.-H., Yu, J.-H., Chou, P.-H., Wu, S.-H., Liao, Y.-T., Huang, Y.-C., Chen, T.-M., & Wang, J.-P. (2024). Proliferation and Differentiation Potential of Bone Marrow–Derived Mesenchymal Stem Cells From Children With Polydactyly and Adults With Basal Joint Arthritis. *Cell Transplantation*, *33*, 09636897231221878. https://doi.org/10.1177/09636897231221878

You, X., Cao, Y., Suzuki, T., Shao, J., Zhu, B., Masumura, K., Xi, J., Liu, W., Zhang, X., & Luan, Y. (2023). Genome-wide direct quantification of *in vivo* mutagenesis using high-accuracy paired-end and complementary consensus sequencing. *Nucleic Acids Research*, *51*(21), e109–e109. https://doi.org/10.1093/nar/gkad909

Zeller, A., Duran-Pacheco, G., & Guérard, M. (2017). An appraisal of critical effect sizes for the benchmark dose approach to assess dose–response relationships in

genetic toxicology. *Archives of Toxicology*, *91*(12), 3799–3807. https://doi.org/10.1007/s00204-017-2037-3

Zhang, L., Jin, Y., Huang, M., & Penning, T. M. (2012). The Role of Human Aldo-Keto Reductases in the Metabolic Activation and Detoxication of Polycyclic Aromatic Hydrocarbons: Interconversion of PAH Catechols and PAH o-Quinones. *Frontiers in Pharmacology*, *3*, 193. https://doi.org/10.3389/fphar.2012.00193

Zhang, S., Coffing, S. L., Gunther, W. C., Homiski, M. L., Spellman, R. A., Van, P., & Schuler, M. (2024). Assessing the genotoxicity of *N* -nitrosodiethylamine with three in vivo endpoints in male Big Blue® transgenic and wild-type C 57 BL /6 N mice. *Environmental and Molecular Mutagenesis*, *65*(6–7), 190–202. https://doi.org/10.1002/em.22615

Zhang, S., Parsons, B. L., Fitzgerald, D., Ashford, A., Auman, J. T., Chen, T., Dodge, A., Elhajouji, A., Pfaller, L., Harris, S., Higgins, J., Hobbs, C. A., Marchetti, F., Meier, M. J., Myers, M. B., Salk, J., Sahroui, R., Schuster, D., Settivari, R., … Chen, C. L. (2025). Transferability, Reproducibility and Sensitivity of Mutation Quantification by Duplex Sequencing. *Environmental and Molecular Mutagenesis*, *66*(6–7), 311–326. https://doi.org/10.1002/em.70020

Zhang, X., Yang, L., Zhang, H., Xing, W., Wang, Y., Bai, P., Zhang, L., Hayakawa, K., Toriba, A., Wei, Y., & Tang, N. (2021). Assessing Approaches of Human Inhalation Exposure to Polycyclic Aromatic Hydrocarbons: A Review. *International Journal of Environmental Research and Public Health*, *18*(6), 3124. https://doi.org/10.3390/ijerph18063124

Zhang, Y., Ding, J., Shen, G., Zhong, J., Wang, C., Wei, S., Chen, C., Chen, Y., Lu, Y., Shen, H., Li, W., Huang, Y., Chen, H., Su, S., Lin, N., Wang, X., Liu, W., & Tao, S. (2014). Dietary and inhalation exposure to polycyclic aromatic hydrocarbons

and urinary excretion of monohydroxy metabolites—A controlled case study in Beijing, China. *Environmental Pollution*, *184*, 515–522. https://doi.org/10.1016/j.envpol.2013.10.005

Zhang, Y., Liu, D., & Liu, Z. (2020). The benzo[b]fluoranthene in the atmospheric fine particulate matter induces mouse glomerular podocytes injury via inhibition of autophagy. *Ecotoxicology and Environmental Safety*, *195*, 110403. https://doi.org/10.1016/j.ecoenv.2020.110403

Zhivagui, M., Hoda, A., Valenzuela, N., Yeh, Y.-Y., Dai, J., He, Y., Nandi, S. P., Otlu, B., Van Houten, B., & Alexandrov, L. B. (2023). DNA damage and somatic mutations in mammalian cells after irradiation with a nail polish dryer. *Nature Communications*, *14*(1), 276. https://doi.org/10.1038/s41467-023-35876-8

Zhong, Q., Amin, S., Lazarus, P., & Spratt, T. E. (2010). Differential repair of polycyclic aromatic hydrocarbon DNA adducts from an actively transcribed gene. *DNA Repair*, *9*(9), 1011–1016. https://doi.org/10.1016/j.dnarep.2010.06.004

Zou, X., Koh, G. C. C., Nanda, A. S., Degasperi, A., Urgo, K., Roumeliotis, T. I., Agu, C. A., Badja, C., Momen, S., Young, J., Amarante, T. D., Side, L., Brice, G., Perez-Alonso, V., Rueda, D., Gomez, C., Bushell, W., Harris, R., Choudhary, J. S., … Nik-Zainal, S. (2021). A systematic CRISPR screen defines mutational mechanisms underpinning signatures caused by replication errors and endogenous DNA damage. *Nature Cancer*, *2*(6), 643–657. https://doi.org/10.1038/s43018-021-00200-0

Appendix A: Power Analyses to Inform Duplex Sequencing Study Designs for MutaMouse Liver and Bone Marrow

Modified from: Elena Esina, Annette E Dodge , Andrew Williams , **David M Schuster**, Danielle P M LeBlanc, Francesco Marchetti, Carole L Yauk*. Power analyses to inform Duplex Sequencing study designs for MutaMouse liver and bone marrow. Environ Mol Mutagen. 2024 Oct;65(8):234-242. doi: 10.1002/em.22619.

A.1 Abstract

Regulatory genetic toxicology testing is essential for identifying potentially mutagenic hazards. Duplex sequencing (DS) is an error-corrected next-generation sequencing technology that provides advantages for mutation analysis over conventional mutagenicity assays including improved accuracy of mutation detection, ability to measure changes in mutation spectrum, and applicability across diverse biological models. To apply DS for regulatory toxicology testing, power analyses are required to determine suitable sample sizes and study designs. In this study, we explored study designs to achieve sufficient power for various effect sizes in chemical mutagenicity assessment. We collected data from MutaMouse bone marrow and liver samples that were analyzed by DS using TwinStrand's Mouse Mutagenesis Panel. Average duplex reads achieved in two separate studies on liver and bone marrow were 8.4×10^8 ($\pm 7.4 \times 10^7$) and 9.5×10^8 ($\pm 1.0 \times 10^8$), respectively. Baseline mean mutation frequencies (MF) were 4.6×10^{-8} ($\pm 6.7 \times 10^{-9}$) and 4.6×10^{-8} ($\pm 1.1 \times 10^{-8}$), with estimated standard deviations for the animal-to-animal random effect of 0.15 and 0.20, for liver and bone marrow, respectively. We conducted simulation analyses based on these empirically derived parameters. We found that a sample size of four animals per group is sufficient to obtain over 80% power to detect a two-fold change in MF relative to baseline. In addition, we estimated the minimal total number of informative duplex bases sequenced with different sample sizes required to retain power for various effect sizes. Our work provides foundational data for establishing suitable study designs for mutagenicity testing using DS.

A.2 Introduction

Conventional *in vivo* mutagenicity tests rely on individual reporter genes for mutation quantification. The most widely used mutagenicity test is the transgenic

rodent (TGR) assay (OECD, 2022). This assay relies on bacterial reporter genes for mutation quantification by using viral phages to encapsulate these genes and infect bacteria; mutations are identified by plaque formation in bacterial lawns grown under selective conditions (Lambert et al., 2005). These reporter-based assays have several important limitations: (a) they quantify mutations in single bacterial genes that do not accurately represent the entirety of the mammalian genome; (b) they are impractical for the analysis of mutation spectra and thus cannot effectively inform on mutagenic mechanism (Beal et al., 2020); and (c) they rely on specific rodent strains and cannot be integrated with test methods using other model organisms. Development of methods to circumvent these limitations would allow for better quantification and understanding of mutagenesis to protect human health.

Duplex Sequencing (DS), an error-corrected next-generation DNA sequencing technology, provides an alternative method to quantify mutations that addresses the limitations of reporter-based assays (Valentine et al., 2020). DS quantifies mutation frequency (MF; mutations per bp) with an extremely low error rate of less than 1 in 10^7 sequenced nucleotides (Kennedy et al., 2014) in any tissue type and model organism, while simultaneously informing mutagenic mechanisms (Marchetti et al., 2023a) . DS evaluates mutagenicity across custom panels that capture different endogenous genomic loci to reflect genome-wide changes. These panels are composed of twenty 2.4 kb genic and intergenic loci spread across the autosomes and are commercially available for mutagenesis studies in mouse (LeBlanc et al., 2022), rat (Smith-Roe et al., 2023), and human (Cho et al., 2023) tissues. Overall, DS has emerged as a transformative tool for mutation analysis with potential to revolutionize the way mutagenicity assessments are conducted (Marchetti et al., 2023b).

DS accurately detects and characterizes mutations by creating consensus sequences from both strands of DNA (Kennedy et al., 2014). Briefly, DS labels fragmented DNA molecules with unique molecular identifiers so that, following PCR amplification and sequencing, reads from both strands may be assembled and grouped into tag families. Each read in a tag family is compared to create single strand consensus sequences (SSCS), which are then grouped for a second round of error correction to produce duplex consensus sequences (DCS). True mutations are scored if they are present at the same position in all reads of a tag family and on both strands of the DNA, as it is unlikely that a sequencing or PCR error would occur at the exact

same location on the sister strand. Thus, DS eliminates almost all sequencing errors, increasing sensitivity for *de novo* mutation detection.

DS has been effective at measuring dose-dependent increases in MF following exposure to well-established mutagens. Prior research has analyzed benzo[*a*]pyrene (LeBlanc et al., 2022), procarbazine hydrochloride (Dodge et al., 2023), *N*-nitrosodiethylamine (Bercu et al., 2023), and *N*-ethyl-N-nitrosourea (Smith-Roe et al., 2023) exposures in transgenic rodent models. These studies demonstrate the potential for DS in regulatory testing. However, every new assay must be evaluated to determine the minimum detectable 'signal' and the optimal approach to achieving it. Power analyses are necessary to establish acceptable study design parameters to advance the use of DS in mutagenicity assessment.

We conducted simulation studies to evaluate sample size and required number of informative duplex bases sequenced per sample for different effect sizes in bone marrow (BM) and liver. These tissues are widely used in regulatory mutagenicity testing and have very different characteristics. BM is highly proliferative, allowing DNA damage to be quickly converted into mutations and fixed in the genome. In contrast, liver is only moderately proliferative but is the most metabolically active tissue and the first pass organ for detoxication. We simulated different study designs to evaluate how power changes with different parameters and draw conclusions on feasible sample size and depth of sequencing.

A.3 Methods

A.3.1 Baseline Data

The DS data utilised in this study are from two studies on the livers and BM of MutaMouse males exposed daily for 28 days by oral gavage to olive oil (vehicle control: VC) or toxicants of interest (LeBlanc et al., 2022; Schuster et al., 2024)). These studies will henceforth be referred to as **“Study A”** and **“Study B.”** The methodology for DNA extraction, DS library preparation, and DS data analysis was as described previously (Dodge et al., 2023), except that DNA was enzymatically fragmented, not ultrasonically sheared, during library preparation. The data from these studies were used to identify a reasonable estimate of baseline MF and the number of informative duplex bases retrieved for mouse BM and liver DNA analyzed with a 500ng DNA input by DS. The standard deviations used for the power analysis cover a range of values observed in

the literature and in our own studies (Schuster et al., 2024). These estimates served as parameters that could be varied in simulation experiments for power analyses.

A.3.2 DS Data Interpretation

Mutation scoring in the studies used was done under the assumption that all identical mutations within a sample are the result of clonal expansion from a single mutational event and so are counted once (Dodge et al., 2023). Germline mutations (with variant allele frequency (VAF) > 0.01) were filtered out and only true *de novo* mutations were retained.

Sequencing results also give a measurement for the total number of informative duplex bases sequenced for each sample. These values can be converted into a target number of required raw reads. The number of unprocessed reads roughly making up one duplex read was empirically derived by Kennedy et al. (Kennedy et al., 2014). In optimal conditions, six raw reads are needed to make one SSCS for DS, and an average of six SSCS make one DCS. Therefore, ideally, we need thirty-six raw reads to make one DCS. Four more SSCSs were added to account for error and low-quality reads that fail to form duplex sequences, making a total of forty raw reads to make a DCS. Thus, the following formula can be used to convert duplex bases into the number of raw reads needed to achieve a certain power: (1)

$$n = \frac{40DG}{R}$$

Where n is the number of raw paired-end reads per sample, D is the average depth of coverage, G is the genome target size in base pairs, and R is the post-analysis read length in bases (300 bp for all of the datasets). The multiplication of DG is the number of informative duplex bases sequenced. Thus, this formula can be simplified to: (2)

$$n = \frac{40S}{R}$$

Where S is the total number of informative duplex bases sequenced per sample.

A.3.3 Power Analysis

The first step of simulation-based power analysis is to simulate a large number of data sets. When simulating responses from a generalized linear mixed model (GLMM), we must assume values for the intercept and the regression coefficients (the fixed effects) and the variances and covariances of the random effects. Estimates for these parameters, or a range of plausible values, are often available from previous studies or from a pilot study. If no data are available, a plausible range of parameter values could be justified based on prior knowledge of the study system.

To generate random samples from an over-dispersed binomial distribution in R (*R: The R Project for Statistical Computing*, n.d.) using the rbinom() function, three arguments are needed: the individual p_i, the number of informative duplex bases (or trials), and n, the number of random samples (which would be set to 1 for each i). The individual p_i's are obtained using the following model: (3)

$$logit(p_i) = \log\left(\frac{p_i}{1-p_i}\right) = \mu + Z_i$$

where $Z_i \sim N(0, \sigma^2)$. This formula can be re-written such that:

$$p_i = \frac{exp(\mu + Z_i)}{1 + exp(\mu + Z_i)} \approx exp(\mu + Z_i) = MF \times \exp(Z_i) \quad (4)$$

As $1 + exp(\mu + Z_i) \approx 1$ for small MF. To simulate mutation data for the treatment group, the p_i's are multiplied by a specified fold change.

We performed a power analysis at varying sample sizes and total number of informative duplex bases and degree of animal-to-animal variation. Simulated data for hypothetical treatment groups were randomly generated to produce defined increases in MF (e.g., 1.2-fold, 1.5-fold, and 2-fold). A general linear model (GLM) was used to calculate significance using the Wald statistic estimated using the doBy R library (Halekoh & Højsgaard, 2024). Experiments were simulated at least 1000 times for each scenario and the power was estimated as the frequency of experiments resulting in a significant difference between simulated treated and control groups at a significance level of 0.05.

A.4 Results and Discussion

The experimental DS data used for the power analysis is summarised in Table 1. The four datasets are described by the following parameters:

- **The number of test animals used (sample size) per dose group.** Study A and Study B used 4 animals per dose group. The data from these experiments were combined and simulated to produce different sample sizes for the power analysis (described below).

- **Background MF.** The VC MF was 5.33×10^{-8} ($\pm 7.35 \times 10^{-9}$) and 3.79×10^{-8} ($\pm 3.35 \times 10^{-9}$) mutations per bp for liver in Study A and Study B, respectively, and 4.13×10^{-8} ($\pm 1.35 \times 10^{-9}$) and 5.02×10^{-8} ($\pm 1.69 \times 10^{-8}$) mutations per bp for BM in Study A and Study B, respectively. The average pooled VC MF (i.e., combining both studies; n =8 per tissue) for liver and BM was 4.56×10^{-8} ($\pm 6.75 \times 10^{-9}$) and 4.58×10^{-8} ($\pm 1.14 \times 10^{-8}$) mutations per bp, respectively. As the averages for both tissues are very similar, an overall average of 4.57×10^{-8} mutations per bp was used in the subsequent power analyses.

- **Informative duplex bases sequenced per sample.** The number of raw reads for each dataset was 3.57×10^{8}, 2.47×10^{8}, 3.48×10^{8}, and 3.36×10^{8} for liver and BM in Study A and B, respectively. The average total number of informative duplex bases achieved for each sample was 8.34×10^{8} ($\pm 9.21 \times 10^{7}$) and 8.51×10^{8} ($\pm 6.05 \times 10^{7}$) duplex bases in the liver studies, and 8.32×10^{8} ($\pm 7.51 \times 10^{7}$) and 1.12×10^{9} ($\pm 7.39 \times 10^{7}$) duplex bases in the BM studies. Note that these numbers are lower than what is expected based on the conversion formula (equation 2) because the studies achieved peak tag family sizes that were larger than the optimal study described therein; peak tag family sizes in these experiments was a minimum of 9.

These empirical parameters were used to generate a range of hypothetical characteristics for potentially feasible DS experiments: sample sizes of 3-8 per dose group; a background MF of 4.57×10^{-8} mutations per bp consistent across liver and BM; an average total informative duplex bases sequenced per animal of up to 2×10^{9} duplex bases; a fold change of 1.5 from the VC group; and sample standard deviations of 0.1, 0.15, 0.2, and 0.3. The standard deviation in this context represents the sample

variability of MF on the ln (natural logarithm) scale as the logit link function was used in the GLMM to model the probability of observing a mutation. For instance, a MF of 4.57×10^{-8} is -16.90 on the ln scale; a standard deviation of 0.3 would then create a 95% confidence interval ranging from -16.39 to -17.41. This interval translates back to a MF range of 2.75×10^{-8} to 7.59×10^{-8}. Appendix Figure A1 represents the distribution of the animal-to-animal random effect for the other standard deviation values. As the chosen test parameters are based on empirical data, they enhance the biological relevance of the analysis.

We first simulated data to determine power to detect a 1.5-fold change in MF between control and exposed groups as the test parameters vary (Appendix Figure A2). We found that when the sample standard deviation is 0.1, 3 animals per dose group is insufficient to achieve 80% power with 2×10^9 informative duplex bases sequenced per animal – the maximum investigated in this study. However, increasing sample size to 4 achieves over 95% power with this number of informative duplex bases and 80% power with only 1×10^9 informative duplex bases. Increasing the sample size further decreases the required number of duplex bases for any given power value. When the sample standard deviation is 0.15, 80% power is achieved with a minimum sample size of 5 animals per dose group and 1×10^9 informative duplex bases. With a standard deviation of 0.2, 80% power is achieved with a minimum sample size of 6 animals per dose group and 1.5×10^9 informative duplex bases; lower samples sizes yield < 71% power. When the sample standard deviation is 0.3, 80% power could not be achieved with any combination of sample size and required number of duplex bases investigated here. Overall, increasing either the sample size or the number of informative duplex bases sequenced per animal increases the power of the study design. However, while larger sample sizes notably enhance power, increasing the number of informative duplex bases beyond about 1.2×10^9 duplex bases yields minimal additional power. This convergence occurs when you have sequenced deeply enough to have an accurate measure of changes in MF. Thus, increasing the number of samples has more impact on power than duplex depth by this point. Our analysis sheds light on the dynamics of statistical power as sample size and sequencing depth were varied.

We ran additional modeling experiments to determine the minimum sequencing depth and sample size needed to observe a significant increase in MF with a power of

80% and a p-value of <0.05 for different effect sizes and variances. For this analysis, the investigated parameter range was expanded: sample sizes were varied from 2 to 20 animals per dose group; background MF was kept constant at 4.57 x 10^{-8} mutations per bp; the average total number of informative duplex bases sequenced per animal was modeled up to 2 x 10^9 duplex bases; sample standard deviations ranged from 0.05 to 0.3, with increments of 0.05; and expected fold changes of 1.2, 1.5, and 2 were considered from the VC group to cover a wide range of responses. For each combination of sample size, variance, and expected fold change, we varied the number of informative duplex bases sequenced per animal until we found the minimum value needed to observe a significant increase in MF with a power of 80% (± 0.01%). This analysis provides a foundation to establish study designs required to achieve statistically significant results in diverse mutagenicity studies.

Our results demonstrate that several feasible study designs are suitable across the above range of parameters (Appendix Figure A3). A 1.2-fold change is only detectible with a minimum of 8 animals per dose group and an average of 1.6 x 10^9 informative duplex bases sequenced per sample if the sample standard deviation is extremely low – 0.05. Further sample size increases lower this required number of informative duplex bases; however, even with as many as 20 animals per dose group, study designs with sample standard deviation >0.15 cannot detect 1.2-fold changes, given that the required number of duplex bases modeled did not exceed 2 x 10^9 informative duplex bases. Detecting a 1.5-fold change is more feasible. Studies with a sample standard deviation of 0.05 are sufficiently powered with at least 3 animals per dose group, requiring a corresponding 1.2 x 10^9 informative duplex bases per sample. Studies where the sample standard deviation is 0.3 – the highest investigated variance – are only sufficiently powered with 12 animals per dose group and 8.9 x 10^8 informative duplex bases. Lastly, detecting a 2-fold change requires a sample size of 2 and 1.2 x 10^9 informative duplex bases when the sample standard deviation is 0.05, and a sample size of 5 and 7.0 x 10^8 informative duplex bases when the sample standard deviation is 0.3. Overall, our analysis reveals study designs that are amenable to detecting different fold changes using different sample sizes and sequencing depths.

Achieving a balance between the various study parameters is essential to effectively design sufficiently-powered studies. Some designs require more samples

per dose group than ethically permissible to achieve sufficient power. One way to reduce sample size requirements is to increase the number of informative duplex bases sequenced per sample. However, with a fixed sample size, power tends to plateau with an increasing number of informative duplex bases (Appendix Figure A2). Where this happens depends on the characteristics of the dataset. Therefore, increasing the number of informative duplex bases to use smaller sample sizes might not always be possible. Thus, trade-off between increasing sample size or the number of informative duplex bases must be carefully considered when designing studies to accommodate high sample deviations and small effect sizes.

Another consideration when increasing the number of informative duplex bases is the biological limit of how deeply we can sequence our samples. There is a limited availability of unique DNA molecules in a given amount of input DNA – known as DNA library complexity. Complexity must be balanced against the number of informative duplex bases achieved. Starting with a more complex sample (i.e., with a larger amount of input DNA) would allow us to increase the number of informative duplex bases. This area presents an opportunity for future research exploration to determine optimal DNA inputs and the number of informative duplex bases for specific DS applications.

Our analysis enabled us to consider acceptable study designs based on two main factors: the minimum sample size per dose group and the most feasible number of informative duplex bases sequenced per sample. In our laboratory, we typically sequence to a depth of 1.2×10^9 informative duplex bases per sample, which is economically feasible at present. Since the required number of informative duplex bases per sample and sample size are inversely correlated, several study designs that minimize sample size are outside the scope of what we can realistically sequence. Given this restriction, we considered study designs for DS that require $\leq 2.0 \times 10^9$ informative duplex bases (Appendix Figure A3). Using this restriction, detecting a 1.2-fold change in MF is challenging; large sample sizes are required to compensate for power loss due the lower number of informative duplex bases per sample and the low expected fold change in MF. In contrast, the detection of a 2-fold change is well within our scope. With a sample size of 4, we achieve sufficient power to analyze samples across most variances and at an easily attainable number of informative duplex bases. Increasing the sample size to 5 extends this capability to include samples with a standard deviation of 0.3. Overall, our results suggest that samples sizes of 4-5 are

sufficient to detect at least a 1.5-fold change with a feasible number of informative duplex bases even with high sample deviation using the present enzymatic fragmentation protocol in MutaMouse liver and BM samples. The TGR test guideline indicates that a minimum of 5 mice per dose group should be used to ensure sufficient statistical power to detect at least a 2-fold change (OECD, 2022). Thus, DS has a higher power to detect the required effect sizes recommended in mutation analyses than the TGR assay.

Given that the parameter range investigated in this study did not encompass all possible values, study parameters in other experiments may differ, potentially altering the recommended study designs. Sample deviation, background MF, and effect size are not under control of the researcher; they are determined by the study compound characteristics and the biology of the organism. However, researchers can adjust sample size and sequencing depth to optimize the power of the study design. Our results provide a foundation for strategic optimization of study designs for mutation analysis.

A.5 Conclusions

Study designs should be based on the biological characteristics of the test system and the objective of the mutagenicity assessment. With DS, ethical considerations about animal use must be balanced against sequencing (cost) feasibility, anticipated effect size, and group variances; there is no singular option that would work for everything. Our power analysis provides a basis for study design selection for liver and BM tissues in MutaMouse animals. We show that a variety of study designs can work to detect fold changes of 1.5 and 2. However, the routine detection of a 1.2-fold change is likely beyond our present capabilities. Curation of historical controls will be critical for DS analyses across different tissues. This will enable us to achieve greater accuracy in our estimates of background MF and sample deviation. Once enough data are collected, a similar power analysis could be done for mutation spectra. Overall, our study informs sample size and the required number of informative duplex bases needed per sample for DS study designs in MutaMouse BM and liver tissue to promote regulatory testing using this powerful new technology.

A.6 Acknowledgements

CLY is grateful for funding from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2021-02806), with infrastructure support through the Canadian Foundation of Innovation John Evan's Leadership Fund (#233109), the Ontario Research Fund (#40569), and uOttawa. CLY acknowledges that this project was conducted thanks, in part, to funding provided through the Canada Research Chairs Program (CRC-2020-00060). CLY and FM are thankful for funding through the Burroughs Wellcome Innovations in Regulatory Sciences Award (#1021737). FM acknowledges support from Health Canada's Genomics Research and Development Initiative and the Chemicals Management Plan.

A.7 Figures and Tables

Supplementary data A of Appendix A are available at *Environmental and Molecular Mutagenesis* online.

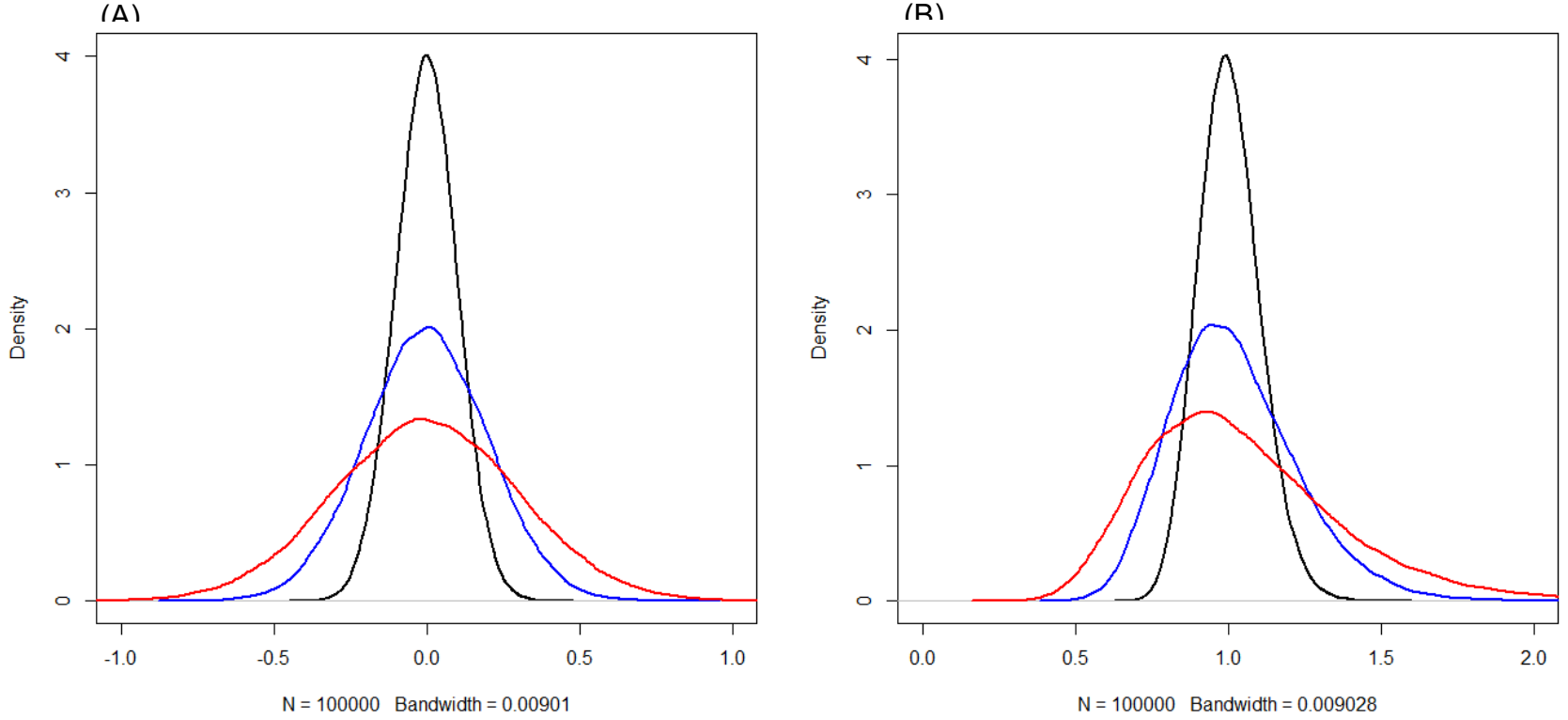
Table A1 Overview of the four DS datasets that served as the foundation for the simulation studies.

Tissue	Study	Sample size per dose group [a]	Average background MF [b] mutations per bp (± *SE*)	Average informative duplex bases per sample [c] (± *SE*)
Liver	A	4	5.33×10^{-8} *(± 7.35 x 10^{-9})*	8.51×10^{8} *(± 6.05 x 10^{7})*
	B	4	3.79×10^{-8} *(± 3.53 x 10^{-9})*	8.34×10^{8} *(± 9.21 x 10^{7})*
Bone marrow	A	4	4.13×10^{-8} *(± 1.35 x 10^{-9})*	8.32×10^{8} *(± 7.51 x 10^{7})*
	B	4	5.02×10^{-8} *(± 1.69 x 10^{-8})*	1.12×10^{9} *(± 7.39 x 10^{7})*

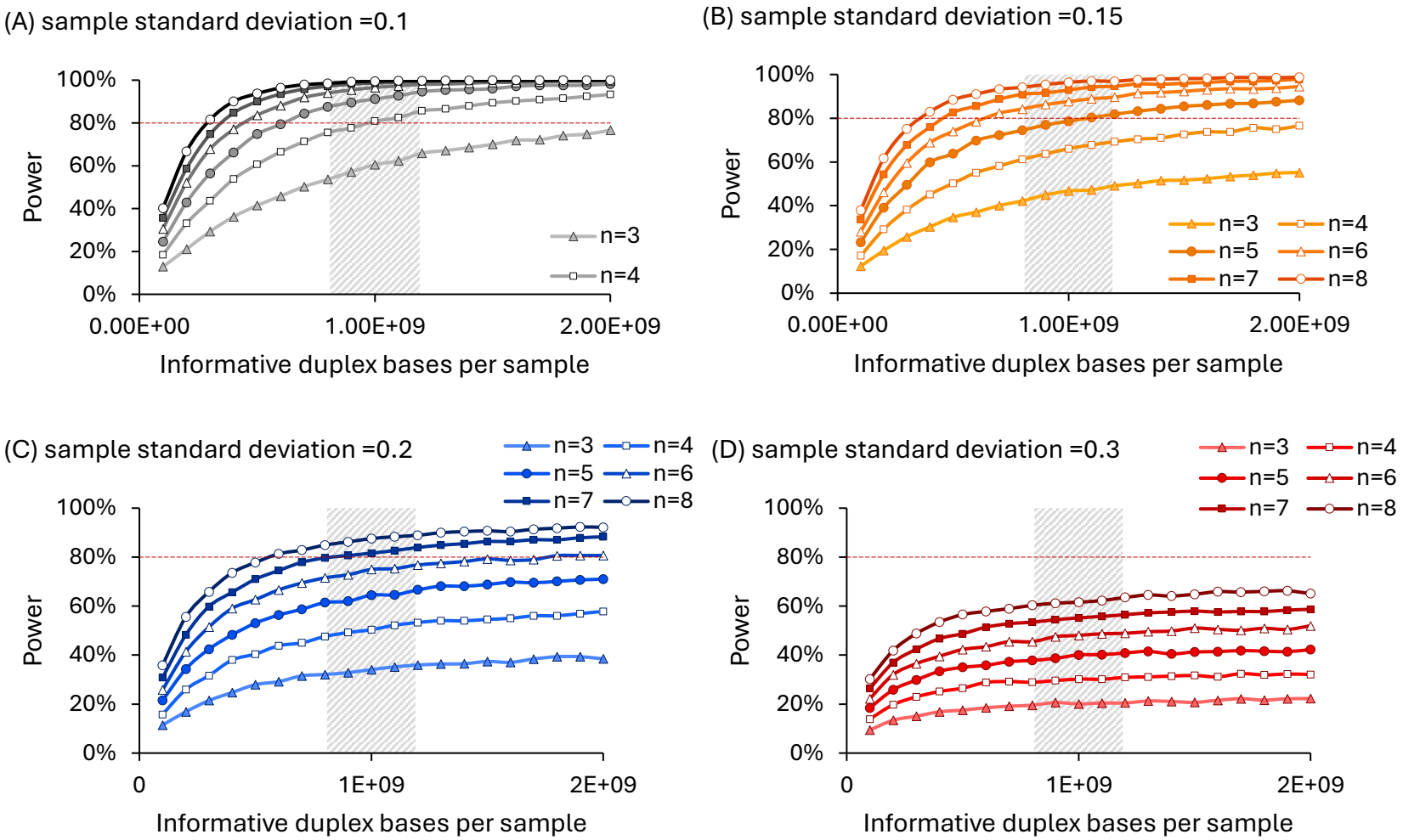
[a] *Male mice were exposed for 28 days by oral gavage to olive oil (vehicle control) or toxicants of interest (Study A and Study B).*

[b] *MF of the vehicle controls in each study.*

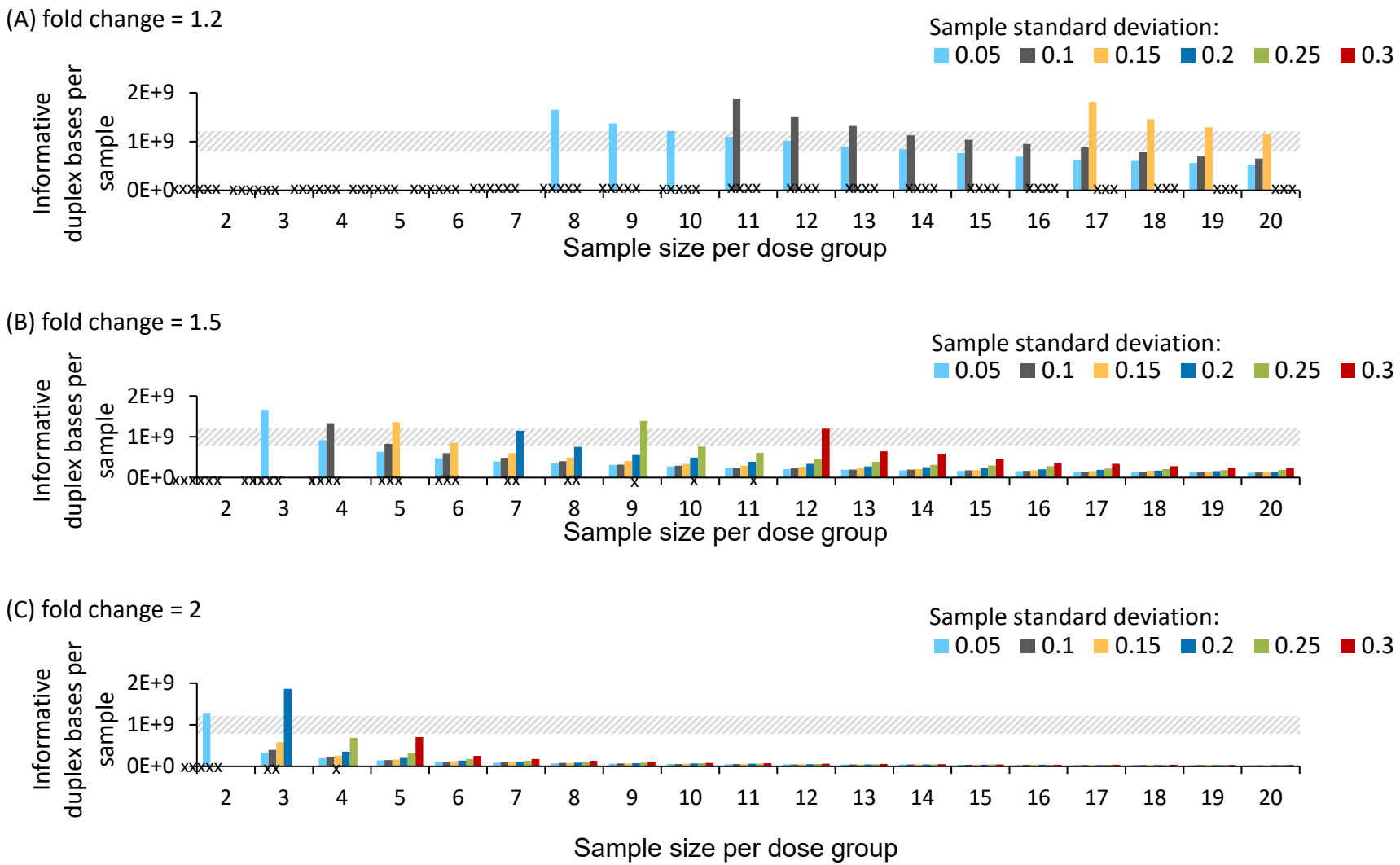
[c] *Number of informative duplex bases is averaged across all samples used in the experiments (i.e., both exposed and vehicle controls).*



Appendix Figure A1: The distribution of the animal-to-animal variability for standard deviations 0.1 (black), 0.2 (blue) and 0.3 (red), or $N(0, 0.1^2)$, $N(0, 0.2^2)$, *and* $N(0, 0.3^2)$ *(A), adjusted to the individual MFs (i.e.* $MF \times exp(Z_i)$ *(B).*



Appendix Figure A2: The estimated power to detect a 1.5-fold change in MF between control and exposed groups, with sample sizes of 3-8 animals per group, and sample standard deviations of 0.1 (A), 0.15 (B), 0.2 (C), and 0.3 (D). The grey highlighted zone shows the typical number of informative duplex bases sequenced per sample for our laboratory (between 8.0 x 10^8 and 1.2 x 10^9 duplex bases).



Appendix Figure A3 Minimum sequencing depth needed to detect a fold change of 1.2 (A), 1.5 (B), or 2 (C) given a background mutation frequency of 4.57 x 10-8 mutations per bp and a range of sample standard deviation. The grey highlighted zone shows the typical number of informative duplex bases sequenced per sample for our laboratory (between 8.0 x 108 and 1.2 x 109 duplex bases). x Value of required sequencing depth exceeds 2 x 109 duplex bases per sample.

Appendix B: Transferability, Reproducibility and Sensitivity of Mutation Quantification by Duplex Sequencing

Modified from Shaofei Zhang, Barbara L. Parsons, Devon Fitzgerald; Anne Ashford , James Todd Auman, Tao Chen, Annette Dodge, Azeddine Elhajouji, Lena Pfaller, Shawn Harris, Jake Higgins, Cheryl A. Hobbs, Francesco Marchetti, Matthew J. Meier, Meagan B. Myers, Jesse Salk, Rebecca Sahroui, **David M. Schuster**, Raja Settivari, Stephanie L. Smith-Roe, Carole L. Yauk, Jian Yan, Andrew Williams, Connie L. Chen. Transferability, Reproducibility and Sensitivity of Mutation Quantification by Duplex Sequencing. Environ Mol Mutagen. 2025 Aug;66(6-7):311-326. doi: 10.1002/em.70020.

B.1 Abstract

Duplex Sequencing (DS) is an ultra-accurate, error-corrected next generation sequencing technology that provides a powerful tool for mutation analysis. A working group (WG) within the Health and Environmental Sciences Institute's Genetic Toxicology Technical Committee is investigating the suitability of DS for regulatory mutagenicity testing. Initial steps to promote acceptance require demonstrating technical reproducibility across DS-experienced and inexperienced laboratories and establishing the method's sensitivity relative to conventional tests. Thus, the WG conducted a 'reconstruction experiment' to evaluate the transferability, reproducibility, and sensitivity of DS. TwinStrand Biosciences first applied DS to establish mutation frequency (MF) in DNA samples extracted from the livers of an untreated Sprague Dawley rat, or rats treated with either 100 mg/kg/day benzo[*a*]pyrene (BaP) for ten days or 40 mg/kg/day N-ethyl-N-nitrosourea (ENU) for three days. Using the measured MF in these original samples, mixtures were then constructed using the BaP- and ENU-treated samples to create "MF standards" with target MFs 1.2-, 1.5-, and 2-fold greater than the untreated control. Aliquots of these standards were distributed to seven laboratories in North America and Europe. DS libraries were prepared by each laboratory and TwinStrand. All eight laboratories met library preparation and assay performance metrics to yield high quality sequencing data with MF in the expected 'MF standard' range. The measured MF and mutation spectra were nearly identical across the laboratories and a 2-fold increase in MF could readily be identified in all labs relative

to the untreated controls. The results confirm the high reproducibility and sensitivity of DS for mutagenicity assessment.

B.2 Introduction

Next-generation DNA sequencing (NGS) has made it possible to analyze genetic alterations in millions of nucleotides of DNA at a time, but limitations in accuracy (error-rates on the order of one-in-one-hundred to one-in-one-thousand) preclude measurement of the much lower mutation frequencies (MFs) required for genetic toxicology applications. Duplex Sequencing (DS) is a double-stranded tag-based error correction technology that improves the accuracy of NGS by more than 10,000-fold and allows for sensitive detection of extremely rare mutations (Salk et al., 2018), such as those induced by chemical mutagens (Salk & Kennedy, 2020). Multiple studies provide evidence of DS utility in the detection of induced mutation (Cho et al., 2023; Seo et al., 2024; Smith-Roe et al., 2023; Wang et al., 2021), while others have evaluated DS equivalence with the transgenic rodent (TGR) mutation assays currently used in regulatory genetic safety assessment (Bercu et al., 2023; Cheung et al., 2024; Dodge et al., 2025; Valentine et al., 2020). In aggregate, these studies identified mutation induction by at least eight known mutagens in at least six tissues across four strains of rodents, as well as in human cells in culture and organoids. Additionally, this body of work collectively demonstrates that mutational sensitivities vary by genomic locus and DNA strand, based on contributing factors that may include transcription, methylation and chromatin state(Dodge et al., 2023; LeBlanc et al., 2022; Valentine et al., 2020). Given the apparent success of these early studies, an error-corrected NGS workgroup (ecNGS WG) under the auspices of the Health and Environmental Sciences Institute's (HESI) Genetic Toxicology Technical Committee (GTTC) enumerated the potential advantages of DS over current genetic toxicology testing methods and described opportunities for DS to advance other areas of mutation research (Marchetti et al., 2023b, 2023a).

The DS approach has a number of advantages over other conventional technologies with OECD test guidelines in terms of the richness of data generated and the broad applicability across species and tissues. Due to the very low rate at which even mutagen-exposed cells accumulate mutations, nearly all conventional OECD assays to date have relied on biological selection systems. Specifically, mutations in defined reporter genes that alter a phenotype to promote preferential growth over non-

mutated cells can be quantified using various assay methods. These include the bacterial reverse mutation (Ames) test (bacteria *in vitro*) (OECD, 2020b), HPRT/APRT/HGPRT (mammalian *in vitro*) (OECD, 2016b), and transgenic reporter systems like Big Blue and MutaMouse (rodent *in vivo*) (OECD, 2025d). However, several limitations to these conventional tests exist. Bacteria are convenient and inexpensive to work with but are far from perfect surrogates for mammals. Transgenic rodent models make it possible to test mutagenicity *in vivo* (Lambert et al., 2005); however, the availability of transgenic rodents can be an issue, a complex *ex vivo* laboratory component is required for its readout, and overall, the approach is expensive. Moreover, determining the spectra of mutations induced is very labor-intensive and impractical outside a research setting (Beal et al., 2020; Besaratinia et al., 2012). *HPRT* and related assays rely on genes natively present in immortalized mammalian cells and make it possible to directly detect mutation induction in human cells, yet the selective growth condition required to obtain single cell-derived colonies is not amenable to the majority of human cell types, and the very small reporter locus limits the generalizability of spectral determination and ability to extrapolate to the rest of the genome. Because DS measures chemically induced mutations directly and identifies the resultant mutational signatures using native DNA without the need for phenotypic selection, it enables analysis in any cell type from any organism with a reference genome, be it cells in culture or tissues from rodents or humans.

Recognizing the promise of DS, the ecNGS WG considered the data needed for adoption by regulatory agencies. Before a technology can be formally investigated using appropriate procedures (i.e., ring trial as per OECD), its methodology must first be demonstrated to be robust, sufficiently detailed, and validated (OECD, 2002). Although groups of investigators have used DS successfully for mutational analyses, a large inter-laboratory study demonstrating successful technology transfer and reproducibility had not been performed. Goals of this study, therefore, were to demonstrate the transferability and reproducibility of DS among laboratories with and without previous DS experience. Additional goals were to establish consensus on critical methodological parameters and gather information on technical sensitivity; this information is necessary to evaluate its equivalence with TGR methods. To accomplish this, a set of rat DNA samples with defined MFs was prepared and shared among investigators, who independently prepared libraries, submitted their libraries for sequencing, and jointly evaluated the study results.

While the DS validation study underscores the promising potential of ecNGS approaches in genetic toxicology, the discontinuation of DS kits presents a significant hurdle for this particular technology. Promptly addressing this issue and exploring alternative ecNGS methodologies will be crucial for the continued advancement of these technologies. The current work offers a potential model for assessing transferability, reproducibility, and sensitivity of other ecNGS technologies, aligning with the WG's broader efforts to promote the adoption of ecNGS methods in regulatory testing.

B.3 Material and Methods

B.3.1 Chemicals

N-Ethyl-*N*-nitrosourea (ENU; CAS #759-73-9) and BaP (CAS #50–32-8) were obtained from Sigma–Aldrich (St. Louis, MO). Formulations of ENU were prepared daily in phosphate buffered saline (Nova-Tech, Kingwood, TX) at a concentration of 8 mg/mL. BaP formulations were prepared weekly in sesame oil (Spectrum Chemical, New Brunswick, NJ) at a concentration of 20 mg/mL.

B.3.2 Ethics Statement

The in-life portion of this study was conducted by Integrated Laboratory Systems (ILS; Research Triangle Park, NC) with prior approval by the ILS Institutional Animal Care and Use Committee. All procedures complied with the Animal Welfare Act Regulations, 9 CFR 1–4 and animals were handled and treated according to the *Guide for the Care and Use of Laboratory Animals* (National Research Council 2011).

B.3.3 Animal Treatment and Tissue Collection

Male Sprague Dawley rats (Hsd:Sprague Dawley; Envigo Laboratories, Frederick, MD) were maintained in a Specific Pathogen-Free facility in constant temperature rooms (20°C–25°C) with a relative humidity of 30%–70% on a 12-h light/12-h dark cycle. Animals were housed in polycarbonate cages with micro-isolator tops containing absorbent heat-treated hardwood bedding (Northeastern Products Corp., Warrensburg, NY). Certified Purina Pico Chow No. 5001 (Ralston Purina Co., St. Louis, MO) and reverse osmosis treated tap water were provided *ad libitum*. The animals were assigned to a dose group such that the mean body weight of each group was not statistically different from any other group. At 8–10 weeks of age, three animals per group were left untreated or administered 40 mg/kg/day ENU or 100 mg/kg/day BaP by oral gavage daily for 3 or 10 consecutive weekdays, respectively. The dose levels were selected to maximize the mutational load under this dosing regimen. To be

consistent with the TG 488 tissue harvest protocol, animals were humanely euthanized at Day 31 following initiation of dosing using carbon dioxide asphyxiation, and death was confirmed by exsanguination.

All lobes of the liver were harvested and rinsed in cold mincing solution (Mg^{+2} and Ca^{+2} free Hanks Balanced Salt Solution, 10% v/v DMSO, and 20 mM EDTA pH 7.4–7.7) to remove residual blood. Tissue was cut into ~5 × 5 × 5 mm sections and placed into tubes embedded in dry ice. Tubes were transferred to a − 80°C freezer until shipped on dry ice to TwinStrand Biosciences (Seattle, WA). Care was taken to avoid cross-contamination of cells between different organ types within each animal and between animals. Clean instruments and gloves were used for each animal.

B.3.4 DNA Extraction and Mixture Preparation

DNA was extracted using a modified “salting out” method (DNA Extraction Kit, Agilent Technologies, Santa Clara, CA), according to the manufacturer's protocol for extraction from whole tissue. Extractions were performed with approximately 250 mg of liver tissue and overnight proteinase digestion at 37°C. Final DNA pellets were resuspended in IDTE (10 mM Tris, 0.1 mM EDTA) (Integrated DNA Technologies, Coralville, IA). To yield sufficient DNA, between 4 and 12 tissue sections were extracted per rat, and the resulting DNA was pooled together after initial quality assessment. Pooled DNA was concentrated using AMPure XP beads (Beckman Coulter, Brea, CA) at a 1:1 ratio. DNA samples were quantified using a Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, Waltham, MA) and quality was assessed using a Genomic DNA ScreenTape Assay (Agilent Technologies, Santa Clara, CA). All pooled DNA samples had concentrations > 50 ng/μL and DNA integrity numbers (DIN) > 7.

For each of the three pure DNA samples, DS libraries were prepared according to the standard protocol (see below). For each sample, three replicate libraries were prepared with 1000 ng of DNA input mass per library. MFs from replicate libraries were averaged to estimate the true MF for each pure sample. Average MFs were as follows: 8.12×10^{-8} for untreated, 4.36×10^{-7} for BaP-treated, and 1.82×10^{-6} for ENU-treated rats.

DNA extractions and mixture preparations were performed at TwinStrand Biosciences, and aliquots of pure and mixed DNA samples were distributed to seven other laboratories for DS library preparation (Appendix Figure B1A).

Mixtures of DNA from untreated and mutagenized rats were made to simulate samples with MFs 1.2× (120%), 1.5× (150%), and 2× (200%) higher than the pure untreated sample. These will be referred to as BaP (or ENU) 1.2× mixture, 1.5× mixture, and 2× mixture. One series of mixtures was made by combining DNA from untreated and BaP-treated rats and another by combining DNA from untreated and ENU-treated rats (Appendix Figure B1B). The quantities of control DNA, BaP pure DNA, and ENU pure DNA used in mixtures to achieve the desired MFmin are detailed in Supplementary Data B: Table S1.

B.3.5 Library Preparation and Sequencing

DS libraries were prepared by seven participating labs independently using TwinStrand DuplexSeq Mutagenesis kits according to the manufacturer's protocol. Briefly, 1000 ng or 500 ng of genomic DNA were digested to DNA fragments with a median size of about 250 to 300 base pairs using TwinStrand DuplexSeq Kit-Enzymatic Fragmentation Module, followed by end repair, A-tailing and DuplexSeq Adapters ligation reactions. A unique set of indexing primers used in the indexing reaction were pre-assigned to each participating lab. The rat mutagenesis hybrid capture panel (TwinStrand product #06–1007-02) includes 20 genomic targets (Table S2), each 2.4 kb (48 kb total), spread across the rat genome. The panel was designed to provide a representative sampling of the rat genome regarding GC% and trinucleotide content, genic/intergenic regions, and excluding repetitive elements and loci known to drive positive or negative selection. The size of final libraries was checked using the Agilent TapeStation system and libraries were quantified with the Qubit fluorometer. The set of libraries generated by the first three participating labs (A, B, and C) were pooled and sequenced on site at TwinStrand on an Illumina (San Diego, CA) NovaSeq 6000 system using an S4 flow cell with 2 × 150 bp read length. Libraries from the remaining four labs (D, E, F, and G) were pooled and sequenced at Psomagen on another S4 flow cell with 2 × 150 bp read length. The DS libraries prepared from each lab included one set of duplicate libraries made from the untreated rat DNA sample (DNA 01). The eighth set of libraries previously prepared and sequenced by TwinStrand (TS) were used as a reference for comparison with results from the seven other labs.

B.3.6 Duplex Consensus Making and Variant Calling

Raw sequencing data for all libraries were processed simultaneously using the TwinStrand DuplexSeq FASTQ to VCF Parallel App (version 3.21.0) in the DNAnexus cloud platform. The app performed consensus making, consensus read

postprocessing, and variant calling according to previously described methods (Kennedy et al., 2014; LeBlanc et al., 2022). When a clear duplex consensus base call could not be made, an N was called at that base in that molecule's duplex consensus sequence.

B.3.7 Calculation of MF and Spectrum

Total and subtype MF calculations, hierarchical clustering of base substitution spectra, and trinucleotide spectra generation were performed for all libraries simultaneously using the TwinStrand DuplexSeq Mutagenesis Report App (version 4.3.0) on DNAnexus. Only somatic variants with variant allele fraction (VAF) < 0.01 were included in the analysis. Analysis was further restricted to a custom reportable region, which excluded the following loci. First, any position with an apparent germline variant (VAF > 0.3) in one or more of the pure samples was excluded from analysis in all samples. This was done so that germline variants diluted to low frequency during DNA mixture preparation would not impact MFs and spectra. Additionally, an approximately 100 bp region within the target locus on chromosome 20 was excluded due to a previous observation of a very high number of artifactual variant calls flanking a cluster of closely spaced germline variants.

Unless otherwise noted, MFs were calculated using the minimum method (MFmin) with the assumption that multiple observations of the same mutation result from cell division rather than multiple independent mutation events, as described previously (Dodge et al., 2023; Valentine et al., 2020). Briefly, each unique rare variant was counted once, regardless of how many consensus reads supported the variant, and the total number of unique variants was divided by total duplex bases (excluding Ns). Subtype MFmin values were calculated similarly, but within context-specific total duplex bases. Alternatively, MFmax was calculated by counting all occurrences of each unique rare variant. Wilson binomial confidence intervals (95%) were calculated for all MFs.

For hierarchical clustering by base substitution spectra, simple base substitution variants were converted into pyrimidine context, and the proportion of each substitution type was calculated for each library. A cosine similarity matrix was calculated on substitution type proportions and unsupervised hierarchical clustering was performed using the Ward clustering algorithm.

To generate trinucleotide spectra, the frequency of each base substitution type in each trinucleotide context was derived by dividing the count of each substitution type by the relative abundance (total duplex depth, excluding Ns) of that context in the reportable region. Frequencies were normalized by the sum of the total single base substitution frequency, so that the proportions for each sample sum to one. Wilson binomial confidence intervals (95%) were calculated for context-specific substitution frequencies and were normalized by the sum of the total single base substitution frequencies.

B.3.8 Generalized Linear Modeling for Analyses of MFs

MFs were analyzed using a generalized linear model (GLM) with extra-binomial dispersion, assuming sample independence. All analyses were performed in the R statistical environment. The GLM included main effects of laboratory and treatment, along with a laboratory-by-treatment interaction, using the quasibinomial family of models. Mean estimates and pairwise comparisons were conducted using the utility R library for groupwise statistics, LSmeans, and linear estimates were called with doBy (Halekoh & Højsgaard, 2006). As this family of models employs a log link function, the estimates from the model parameters obtained through the esticon() function were exponentiated (i.e., back-transformed). Estimates of the back-transformed standard errors were approximated using the delta method. The *p*-values for the fold change estimates were adjusted for multiple comparisons using the Sidak method (Sidak, 1967; Wright, 1992). Fold changes between treatment and controls were calculated independently within each laboratory.

For group-based analysis of MFmin, libraries prepared from the same DNA samples were treated as replicates. A quasi-Poisson generalized linear model was used to perform all possible pairwise comparisons between sample groups, and Benjamini-Hochberg false-discovery rate adjustment (Benjamini & Hochberg, 1995) was used to correct for multiple comparisons.

B.3.9 Sampling-Based Replicate Analysis

To assess whether a ~ 1.2-fold increase in MFmin could be detected using fewer than eight technical replicates, we randomly selected replicates to make smaller groups and repeated the group-based analysis. We tested group sizes between 2 and 7 and took 100 random samples of replicates at each group size. Libraries generated with 500 ng of input DNA (Lab B) were excluded from this analysis because of the lower

data yield of these libraries. Sampling and pairwise statistical tests were performed separately for the BaP 1.2×, ENU 1.2×, BaP 1.5×, and ENU 1.5× mixtures compared to the untreated control sample, without correction for multiple comparisons. A similar analysis was repeated for each of the six single base substitution (SBS) subtype MFs.

B.3.10 Mutation Spectra

The trinucleotide spectra were analyzed using the likelihood ratio test as described in Piegorsch and Bailer (Piegorsch & Bailer, 1994). Pairwise comparisons were conducted on 2×96 contingency tables of counts for each trinucleotide as well as on 2×10 contingency tables of counts for each variant type and SNV (single-nucleotide variant) subtypes normalized to its pyrimidine context. These comparisons included treatments to controls and comparisons across laboratories for each treatment. Adjusted *p*-values were obtained using the Sidak method (Sidak, 1967) which was applied independently to each family of tests. This family of tests were defined as within laboratories and across laboratories for each treatment independently.

B.3.11 Power Analysis

A power analysis was conducted to determine the minimum sample size and informative duplex bases required to observe a significant increase in MF. Random samples were generated repeatedly to simulate experiments using estimates from this study. Over-dispersed binomial samples were randomly generated using the rbinom() function with an MF of 2×10^{-9} and overdispersion assuming a normal distribution with a mean of zero and standard deviations ranging from 0.05 to 0.15. A GLM was employed to calculate the significance between the control and treated samples. The bisection method was used to estimate the minimum detectable fold change of the simulated experiment, yielding a power of 80%. For each iteration of the bisection method, the power was estimated by simulating 2000 data sets and calculating the proportion of times a significant result was achieved at the 0.05 significance level. The sample size, sample variance, and number of informative duplex bases were varied, and the results from these analyses were used to produce power curves.

B.4 Results

B.4.1 Mortalities, Clinical Signs, and Body Weights

All animals survived to the scheduled termination; clinical observations indicated no signs of toxicity during the study. Lower terminal body weight as compared to

untreated control animals was exhibited in rats administered the test articles (−10.3% and − 7.8% for ENU and BaP, respectively).

B.4.2 QC Metrics Overview of Sequencing Data from Eight Participating Labs

Eight labs each constructed ten libraries from the nine DNA samples provided, with the untreated control prepared in duplicate by each lab. Libraries were sequenced across three different flow cells with each library allocated ~700 million raw paired-end reads (696,136,363 ± 26%) or ~350 million sequencing clusters (Appendix D2 Table 3: DNA Sequencing Metrics). Most libraries were prepared using 1000 ng of input DNA per library, but Lab B used 500 ng of input DNA per library. Because DS data yield depends greatly on the input DNA mass, with a roughly linear relationship, performance of the 500 ng and 1000 ng libraries were considered separately. Libraries prepared with 500 ng of input DNA yielded ~1.2 billion informative duplex bases each (1,222,407,630 ± 11%, Appendix Figure B2A, left) and libraries prepared with 1000 ng of input DNA yielded ~2.2 billion informative duplex bases each (2,233,908,655 ± 14%, Appendix Figure B2A, right). These total data yields equate to ~20,000× or ~37,000× average duplex molecular depth across the target regions for 500 ng and 1000 ng inputs, respectively (19,646% ± 10%, 36,810% ± 13%,Appendix Figure B2B). Peak tag family size (PTFS) is the modal number of reads representing each strand in duplex consensus families, with PTFS of 10 representing a balance between optimizing data yield and sequencing cost. All libraries had PTFS > 8 indicating sufficient sequencing and libraries prepared with 500 ng of input DNA had higher PTFS values (48.30% ± 9%) than the 1000 ng libraries (16.36% ± 26%), as expected (Appendix Figure B2C). Additional assay performance metrics are summarized in Table S3. Overall, all eight labs generated high quality DS libraries from the provided DNA samples. Data yields per library were relatively consistent within and between labs and scaled with input DNA mass, as expected.

B.4.3 Reproducibility Measured by MFs (MFmin)

Rigorous evaluation of DS sensitivity and reproducibility is needed to support its use as a regulatory test for mutagenicity. To evaluate reproducibility, eight labs analyzed the same set of DNA samples. Technical sensitivity of DS was analyzed concurrently by assessing the lowest quantity of treated rat DNA, when spiked into untreated rat DNA, could be detected as significantly increased MFmin relative to the untreated control. The DS MFmin measurements from the analysis of the set of 10 samples by each of the eight labs are shown in Appendix Figure B3. It is apparent that

the results from the different labs are highly reproducible. Importantly, the results from each lab showed comparable and expected increases in MF based on the amounts of pure DNA from mutagenized rats spiked into the different samples. MFmin is the preferred endpoint for the DS analysis of mutagenesis because it eliminates possible confounding due to clonal expansion of individual mutants. However, high interlaboratory reproducibility was also observed based on quantitation using MFmax (Supplementary Data B: Figure S1).

Some participating laboratories had considerable previous experience using DS. Other laboratories were trained immediately preceding library preparation of study samples. Given that reproducible results were obtained irrespective of extent of laboratory experience, it is concluded that the DS technology (including kit design, instruction, and support) has great across-laboratory transferability.

B.4.4 Reproducibility Measured by MF

A generalized linear model was used to estimate the effects of laboratory and treatment on MFmin. Mean estimates (Appendix Figure B4A) did not differ significantly between any of the eight laboratories for any of the treatment groups ($p > 0.05$). Fold changes from control were consistent with the expected fold increases in MFmin based on the amount of treated DNA in each sample across all labs (Appendix Figure B4B). Fold changes across laboratories ranged from 1.1–1.3×, 1.4–2.3×, 1.8–2.8×, and 5.3–7.2× for BaP 1.2× mixture, 1.5× mixture, 2× mixture, and pure BaP treated samples, respectively. For the ENU treated samples or mixtures, fold changes ranged from 1.1–1.4×, 1.4–2.1×, 1.9–3.0×, and 22-33× for ENU 1.2× mixture, 1.5× mixture, 2× mixture, and pure ENU treated samples, respectively. Significant increases from control ($p < 0.05$) were detected for most individual laboratories (two control replicates vs. one mixture replicate) starting at the 1.5× mixtures for both BaP and ENU. Laboratory D was able to detect significant increases from control at the 1.2× mixtures for both chemicals. All laboratories detected significant increases in MF at 2× mixtures and pure BaP and ENU treated samples.

B.4.5 Technical Sensitivity of DS

To further examine the technical sensitivity of DS, we treated all libraries prepared from the same DNA sample as technical replicates and performed group-based analysis (Appendix Figure B5). Considering libraries from all labs, both 1.2× mixtures had significantly increased MFmin relative to the untreated control (FDR-

adjusted p values, BaP 1.2× mixture $p = 1.82 \times 10^{-3}$, ENU 1.2 × mixture $p = 6.63 \times 10^{-4}$). The 1.5× and 2× mixtures were significantly increased as well ($p < 1 \times 10^{-8}$).

To assess whether a ~ 1.2-fold increase in MFmin could be detected using fewer than 8 replicates, we randomly selected replicates to make smaller groups and repeated the group-based analysis (Appendix Figure B5B and C). For the BaP 1.2× mixture, 97% of randomly sampled groups of $n = 6$ yielded a significant test result ($p < 0.05$) and for the ENU 1.2× mixture, 98% of randomly sampled groups of $n = 5$ yielded a significant test result ($p < 0.05$). This analysis suggests that five to six 1000 ng libraries provide enough statistical power to detect an ~1.2-fold change with high confidence, assuming a similar baseline MF and that the variability between replicates is similar to what is observed in inter-lab technical replicates in this study. The analysis was repeated for the 1.5× mixtures, with the results suggesting that 2–3 replicates per group provides sufficient power to detect an ~1.5-fold increase in MFmin (Supplementary Data B: Figure S2).

B.4.6 Single Base Substitution Spectra Analyses

Unsupervised hierarchical clustering was performed to classify libraries based on SBS spectra (Appendix Figure B6). Libraries prepared with pure DNA samples from untreated, BaP-treated, and ENU-treated rats showed distinct SBS spectra and formed discrete clusters. Libraries prepared with DNA mixture samples also resolved into near-perfect clusters by DNA sample, with the lower fold-change mixtures being more similar to the pure untreated rat libraries and the higher fold-change mixtures being more similar to the pure BaP- or ENU-treated rat samples. While the clustering analysis does not directly measure reproducibility across different labs, it shows that the variation within a specific mix rate (such as 1.2× or 1.5×) is less than the variation observed between different mix rates or treatments.

B.4.7 Subtype MF Analysis

Next, we assessed whether focusing on mutagen-specific mutation subtypes could boost sensitivity of DS. We repeated the replicate-sampling analysis using sub-type MFmin instead of total MFmin for the BaP and ENU 1.2× mixtures (Appendix Figure B7). For the BaP 1.2× mixture, 92% of randomly sampled groups of $n = 3$ showed a significant increase in C>A MF (Appendix Figure B7A) and 99% of randomly sampled groups of $n = 5$ showed a significant increase in C>G MF (Appendix Figure

B7B), compared to $n = 6$ needed to see a significant increase in total MFmin (Appendix Figure B5B). For the ENU 1.2× mixture, 100% of randomly sampled groups of $n = 4$ showed a significant increase in T>A MF (Appendix Figure B7C) and 100% of randomly sampled groups of $n = 3$ showed a significant increase in T>C MF (Appendix Figure B7D), compared to $n = 5$ needed to see a significant increase in total MFmin (Appendix Figure B5C). For all other mutation subtypes, either no significant change was detectable with any number of replicates or more replicates were needed than for total MFmin (Supplementary Data B: Figure S3).

B.4.8 Simple and Trinucleotide Spectra Analysis

Normalized proportions of mutation subtypes were compared across treatment groups and across laboratories using the likelihood ratio test within contingency tables of mutation counts (min). Comparisons were first made based on the simple mutation spectra (Appendix Figure B8A), including all variant types and SNV subtypes. Categories included MNV (multiple-nucleotide variants), insertions, deletions, complex variants, symbolic/structural variants, and the six SNV subtypes. Proportions for each mutation subtype were normalized to the sequencing depth. There was no significant difference ($p > 0.05$) between laboratories for any of the mixtures or pure samples except for laboratory C, pure ENU treated samples which was significantly different from all other laboratories. The small relative increase in C>A mutation found in sample from laboratory C is not fully understood but may arise from the process of sample preparation or library constructions. Expected differences in mutation proportions were observed between mixtures (or pure samples) and controls within each laboratory. Significant shifts in the mutation spectra were detected starting at BaP 1.5× mixture for six out of the eight labs. Labs B and E detected significant shifts in the spectrum starting at BaP 2× mixture. Significant differences in the mutation spectra were detected starting at ENU 1.5× mixture for all laboratories. Labs A and F were further able to detect significant differences in the mutation spectra for the ENU 1.2× mixture (Table S4).

Similar results were found when the 96-base trinucleotide spectrum was compared across treatment groups and laboratories (Appendix Figure B8B). The trinucleotide spectrum consists of the normalized proportions of the six SNV subtypes (pyrimidine notation) within the context of their two flanking nucleotides. Again, the proportions of mutation subtypes did not differ across laboratories for any treatment groups except pure ENU, which differed between lab C and all other laboratories. The

pure ENU treatment groups had many more mutations than any other group; therefore, it is possible that only small differences in the spectra are being detected in this case. Shifts in the mutation spectra between the treatment groups and the control were consistent across labs. Four out of the eight laboratories detected a significant shift in the trinucleotide spectrum starting at BaP 2× mixture. Lab A was able to detect a significant difference in BaP 1.5× mixture. Labs B, G, and TS were only able to detect a significant difference in trinucleotide spectra at pure BaP treated samples. Five out of eight laboratories detected a significant shift in the trinucleotide spectrum starting at ENU 2× mixture. Labs A and D detected a significant difference at ENU 1.5× mixture. Finally, lab B only detected a significant difference at pure ENU treated samples.

B.4.9 Power Analysis

The issue with empirical evaluations is that the underlying distributions of the data may not reflect true experimental conditions such as only evaluating technical variability. Simulation is a tool often used to test new methodologies based on estimates from the empirical evaluation. Simulating data from known distributions allows us to measure the performance of the test system under different scenarios.

Using the binomial variation observed in this study, a MF of 6.7×10^{-8} and an average number of informative bases (1.71×10^{9}), our simulation shows that DS can detect a 1.5-fold increase with an n of four with 80% power (the red line in Appendix Figure B9A). As n increases, DS can detect a 1.2-fold change with an n of 8 with a Bonferroni adjusted significance level of 0.05 with eight comparisons (i.e., eight labs). As the number of informative bases decreases from 100% to 80% (orange), 60% (green), 40% (blue) and 20% (purple), the minimal detectable fold change increases. With only 3.42×10^{8} informative duplex bases, to detect a 1.5-fold change would require an n of 8 per group.

Panels B, C and D in Appendix Figure B9 illustrate the minimal detectable fold change of DS as over dispersion or additional binomial variation is present. These analyzes attempt to simulate biological variability using estimates of standard deviations that have been previously published or cover a range of values observed in unpublished in-house data. As the additional binomial variation increases the minimal detectable fold change also increases. For example, with a standard deviation of 0.15, the sample size to detect a 1.5-fold change would be seven per group, the point at which the red line crosses the gray line in panel D of Appendix Figure B9.

B.5 Discussion

Recognition that ecNGS approaches bring improved precision and a broader range of applicability to the characterization of genomic change created enthusiasm for their further development and validation. Indeed, practitioners have described the potential for ecNGS methods to revolutionize the fields of genetic toxicology and cancer risk assessment (Marchetti et al., 2023b, 2023a).Establishing the robustness of ecNGS methods is a critical step toward regulatory adoption. As noted in the roadmap by Marchetti et al. (Marchetti et al., 2023a), essential steps toward the regulatory adoption of ecNGS test methods must include favorable demonstration of assay sensitivity, transferability, and inter-laboratory reproducibility through a ring trial (OECD series on testing and assessment –number 34, https://ntp.niehs.nih.gov/sites/default/files/iccvam/suppdocs/feddocs/oecd/oecd-gd34.pdf).

DS is one type of ecNGS that has been used in various studies to assess low-frequency mutations, both *in vitro* and *in vivo* (Armijo et al., 2023; Cho et al., 2023; Smith-Roe et al., 2023; Wang et al., 2021; S. Zhang et al., 2024). Given that the number of laboratories investigating ecNGS for genetic toxicology applications was greater for DS than for other ecNGS approaches (e.g., PacBio Hifi sequencing (Miranda et al., 2023), SMM-seq (Maslov et al., 2022), and Hawk-Seq (Matsumura et al., 2019; Otsubo et al., 2021)), investigating the technical sensitivity, transferability, and reproducibility of DS was the goal of this international collaboration.

The results from the eight laboratories participating in this study demonstrate the outstanding reproducibility of DS in the measurement of MF and mutation spectra. Indeed, the results of the current study robustly meet the criteria for OECD pre-validation of a test method. Further, the interlaboratory reproducibility (Appendix Figure B3 and Appendix Figure B6) supports the conclusion that the methods employed in the current study are sufficiently robust to be considered as a ring trial.

Importantly, several of the laboratories entered this study with minimal or no prior experience building libraries for DS. With appropriate training by TwinStrand personnel, these laboratories were able to produce results, both MF and mutation spectrum, nearly identical to those generated by laboratories with more experience. Thus, the DS method was established as having high transferability. Our observations regarding DS reproducibility and transferability support the conclusion that

interlaboratory variability is unlikely to limit the regulatory use of this method for MF determination.

One of the interesting observations of this study is that the use of 500 ng input DNA by lab B generated results very similar to the seven other labs that used 1000 ng of input DNA. Thus, the method appears to have good tolerance regarding variation in the quantity of input DNA. Both quantities of input DNA yielded large numbers of total informative duplex bases. Comparison of the performance metrics shows that libraries prepared using 500 ng of input DNA had a mean PTFS of 48, as compared to a mean PTFS of 16 for libraries prepared using 1000 ng input DNA. However, 500 ng input DNA generated only about half of the total informative duplex bases generated using 1000 ng input DNA with the same sequencing settings. Generally, a PTFS of 10 to 15 is considered ideal for high quality DS data. Although there are no negative impacts on the mutational analysis, greater PTFS values may signal the creation of unnecessary raw reads (and sequencing cost), without a proportional increase in informative duplex bases. Therefore, this study documented the relationship between the input DNA and raw reads as critical parameters of DS study design that must be considered to achieve optimal PTFS.

MF data can be expressed as MFmin (mutations recovered more than once within the same sample are counted only once) or as MFmax (all mutations are counted). Sequencing of mutants is sometimes performed as part of a TGR assay, particularly when high inter-individual variability is observed because “sequencing can be used to rule out the possibility of jackpots or clonal events by identifying the proportion of unique mutants from a particular tissue” (OECD, 2025d). In assessment of mutagenesis, this conveys a recognized preference for the analysis of unique mutations as compared to mutation data confounded by clonal events. The current study investigated MFs calculated as MFmin and MFmax. Initially some variation in MFmax was observed across replicated samples. Some of this variation was due to a few variants with germline frequencies in the pure samples erroneously left out of the initial list intended to mask these positions from analysis. This resulted in apparent large clones in some mixture samples. Furthermore, in mixtures where those variants were present at about ~1%, they fell below the threshold in some samples, so contributed 100 s of variant counts to the MFmax calculation and fell above the threshold in others so were excluded from MFmax calculations. When the reportable

region was updated to exclude these germline variants, the variability in MFmax greatly decreased. This demonstrates that in samples with true clones, that both the presence of clones and the exact size of the clones (barely above or below the somewhat arbitrary threshold of 1%) can have a dramatic impact on MFmax. Specifically, the dilution necessary for the construction of the DNA mixtures caused germline mutations to appear as large clones in some but not all samples. This example highlights the need to carefully scrutinize DS data. After addressing the artifactual variant calls, the MFmin and MFmax metrics both yielded reproducible results across the eight participating laboratories. Nevertheless, based on this experience and existing TGR guidance, MFmin is recommended as the appropriate metric to use for ecNGS assessment of mutagenicity, whereas MFmax may be a more relevant metric for carcinogenicity testing.

It is well known that different mutagens and DNA repair defects may leave unique signatures in the mutational spectra (Boysen et al., 2025; Kucab et al., 2019; Poon et al., 2014). In addition to the measurement of MF, DS yields a comprehensive view of mutation types, thereby providing insights into mutagenesis mechanisms. Identification of changes in mutational spectra could also provide additional confidence in the mutagenic effects observed from a specific treatment. In this study, significant changes in the simple mutation spectra were observed between control and the BaP 1.5× mixture (six out of eight laboratories) and between control and ENU 1.5× mixture (eight out of eight laboratories). The sensitivity to detect a shift based on the trinucleotide spectra was not as great as that observed using the simple mutation spectra. In comparisons with control, it was possible to detect a significant shift in the trinucleotide spectra for the BaP 2× mixture (four out of eight laboratories) and the ENU 2× mixture (five out of eight laboratories). Although regulatory applications do not mandate assessment of a shift in the mutation spectrum caused by a test article, detecting significant changes in simple and/or trinucleotide spectrum may help identify a positive response in conjunction with the MF analysis.

As a first step toward systematically evaluating DS sensitivity, this study developed data on the technical sensitivity of DS. Technical sensitivity was evaluated using replicate analyses of constructed and shared DNA mixtures ($n = 1$, without biological variability) designed to represent a range of MFs. Using the library preparations and sequencing approaches described, the results indicate a MF 1.5×

greater than control was statistically distinguishable from control MF. When data from different laboratories were pooled, a MF 1.2× greater than control was statistically distinguishable from control MF. This sensitivity was augmented by focusing on mutagen-specific mutation subtypes. For example, three samples were sufficient to demonstrate a statistically significant increase in C>A MF in the BaP 1.2× mixture versus control in 92% of randomly selected sets of samples whereas six samples were needed to detect a significant increase based on total MF. Similarly, in the comparison of control and the ENU 1.2× mixture three samples were sufficient to detect a significant increase in T>C MF in 100% of randomly selected set of samples, whereas five samples were required to observe a significant increase based on total MF. This demonstration of the ability of DS to detect subtle elevations in the induced total or subtype MF over the background level with relatively small number of replicates should not be construed as a recommendation that a 1.2-fold increase should be used as a cutoff for positive/negative calling in mutation testing because such a small increase may not be biologically relevant. Instead, this result means that the technical sensitivity of DS is unlikely to hinder its ability to detect biologically relevant changes in MF.

The simulation conducted in this paper provides further evidence to support the findings of the empirical evaluation and to explore the sensitivity of the method when incorporating different levels of biological variability, overdispersion, or additional binomial variation. The power analysis shows that it is possible to detect a 2-fold change using only three samples per group, 1.71×10^9 informative bases, and assuming a standard deviation of 0.1. Currently, a minimum of five animals per group is required for the TGR assay according to OECD Test Guideline 488 (OECD, 2025d). The requirement for five animals is largely due to the relatively high inter-animal variability observed with TGR assays. A recent study, which included a power analysis, supported the use of as few as three animals to detect a 1.5-fold increase in MF due to the very low inter-animal variability observed using DS (Dodge et al., 2023). If the possibility that fewer animals can be used in DS studies than TGR studies is confirmed, mutagenicity testing by DS would be consistent with the 3R principles of reducing, refining, and replacing animal use in experimentation (Kroeger, 2006).

One potential limitation of our study design is that it would have been possible to sequence each lab's libraries as discrete pools in separate lanes or on separate days. However, practical/real world sequencing analyses are expected to involve

pooling of libraries for sequencing on a given flow cell or “lane”. While theoretically, it cannot be ruled out that concurrent sequencing of the libraries prepared as in this study minimized potential variation, it is generally understood that library quality rather than competently performed input of samples for sequencing is the primary driver of sequence quality.

Recent studies have compared the performance of DS and TGR assays in transgenic animals and proven the equivalency of the two methods in mutagenicity testing (Bercu et al., 2023; Dodge et al., 2023; LeBlanc et al., 2022). Notably, the current study demonstrates the applicability of DS to wild-type animals, laying the groundwork for its application as an alternative to the conventional TGR assay. Overall, our data from eight participating laboratories demonstrated that DS has high technical sensitivity, as well as great reproducibility and transferability. The data shown in this study will be useful in the design and interpretation of future interlaboratory studies that examine DS reproducibility and sensitivity in the context of test-article exposed and unexposed rodents. As an intermediate step, however, additional studies are needed to establish the acceptability criteria for DS studies, i.e., the minimal number of animals required per group for *in vivo* studies, sufficiency in terms of target sequences analyzed, PTFS range, and/or number of informative duplex bases.

In summary, this DS validation study was undertaken with the view that ecNGS approaches may yield improved precision and greater flexibility in the characterization of genomic changes. A significant advantage of DS over TGR assays is its flexibility to be performed using DNA isolated from any tissue of any model organism. This study's results expand existing evidence indicating DS may improve the precision of the mutational analyses used in regulatory testing and the perception that ecNGS methods, in general, may revolutionize the field of genetic toxicology. However, to the authors' knowledge at the time of manuscript submission, DS kits and onsite trainings are not currently commercially available, and it is unclear when they will be once again. Long-term unavailability of DS kits would impose obvious limitations on future validation studies and widespread adoption of this particular ecNGS technology. This will become a significant roadblock for analysis of *in vivo* mutation by DS if the current situation is not resolved. Further, it highlights the need to characterize/validate multiple ecNGS methodologies. The present paper presents a possible model of how transferability, reproducibility, and sensitivity could be assessed for other ecNGS

methods. Independent replication of this study using alternative sequencing providers, reagents, and bioinformatic tools would be beneficial for further validation.

B.6 Acknowledgments

The experimental work described in this paper was performed under the auspices of the HESI GTTC. HESI is a non-profit scientific organization that facilitates public and private partnerships in human and environmental health in a pre-competitive space. The GTTC aims to improve the scientific basis of the interpretation of results from genetic toxicology tests for purposes of more accurate hazard identification and assessment of human risk; to develop follow-up strategies for determining the relevance of test results to human health; to provide a framework for integration of testing results into a risk-based assessment of the effects of chemical exposures on human health; to promote the integration and use of new techniques and scientific knowledge in the evaluation of genetic toxicology; and to monitor and promote the development of innovative tests and testing strategies. HESI GTTC provided funding for the library sequencing for laboratories D–G. All other experimental work was supported by in-kind contributions (from public and private sector participants) of time, expertise, and experimental and/or development effort from the participating laboratories. These contributions are supplemented by direct funding (that primarily supports program infrastructure and management and some project-related direct expenses) provided by HESI's corporate sponsors. A list of supporting organizations (public and private) is available at http://hesiglobal.org. Work performed for the National Toxicology Program, National Institutes of Environmental Health Sciences, National Institutes of Health, US Department of Health and Human Services, USA, was conducted under contract 75N96020C00001 (genetic toxicity testing) and ES103378-01.

B.7 Disclosure

TwinStrand Biosciences Inc. *has developed DS technology, provided key reagents and analysis tools for this study and may financially benefit from this assay and the findings of this study. No authors are current employees, but former employees J. H. and J. S. are private equity holders*. Shaofei Zhang is an employee of, and owns shares in, Pfizer Inc. Anne Ashford is an employee of, and owns shares in, AstraZeneca plc.

B.8 Data Availability Statement

The data that support the findings of this study are openly available in Sequence Read Archive at https://www.ncbi.nlm.nih.gov/sra, reference number BioProject PRJNA1206152.

B.9 Figure and Tables

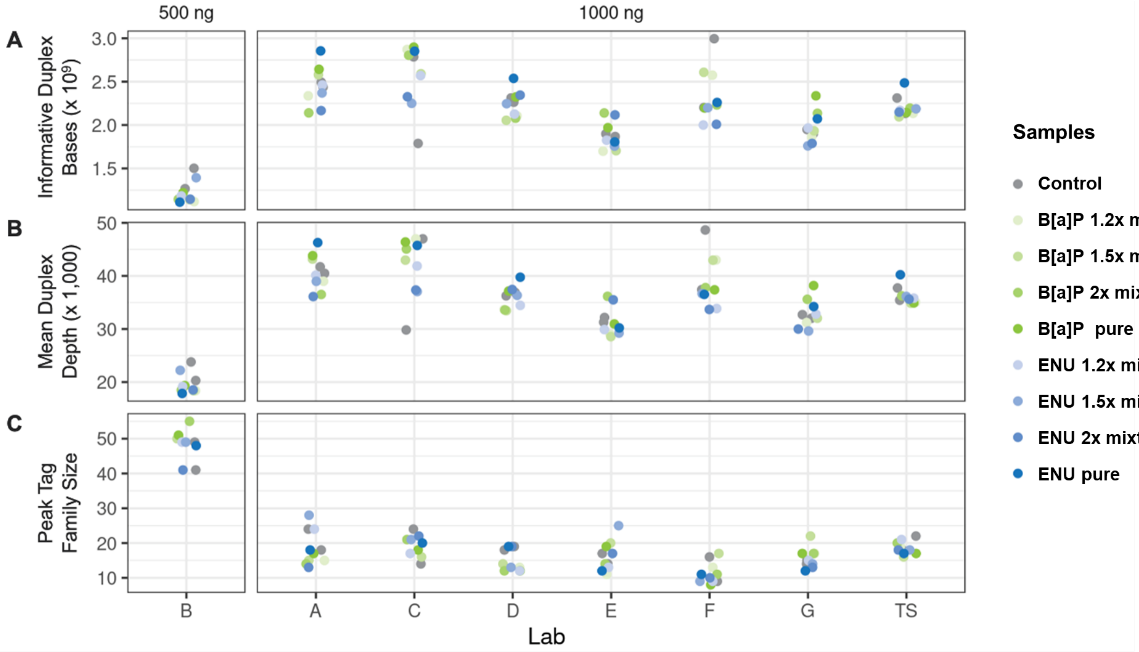
Supplementary data B of Appendix B are available at *Environmental and Molecular Mutagenesis* online.

A

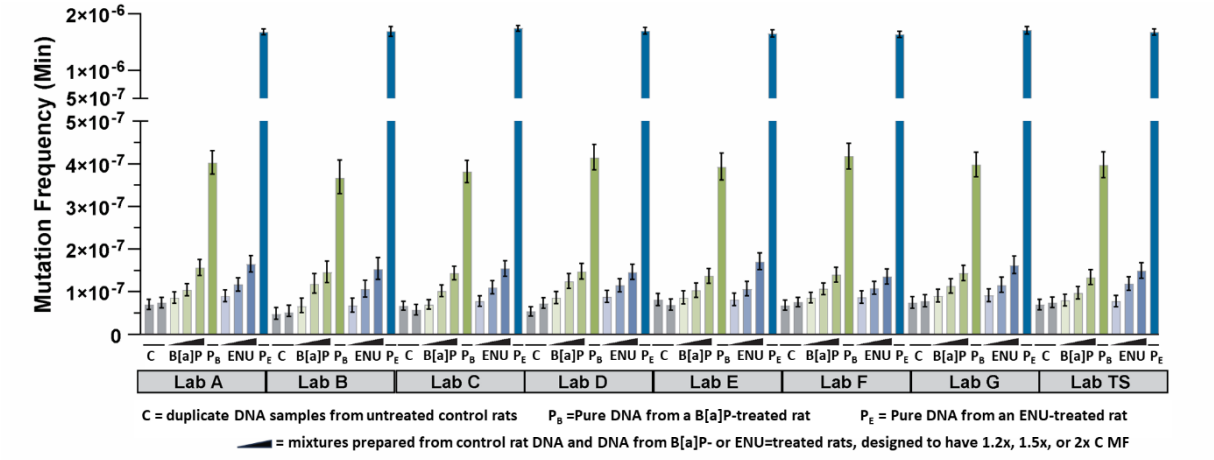
B

	Sample	MF_{min}
	Control	8.12×10^{-8}
	B[a]P 1.2x mixture	*9.85×10^{-8}*
	B[a]P 1.5x mixture	*1.23×10^{-7}*
	B[a]P 2x mixture	*1.64×10^{-7}*
	B[a]P pure	4.36×10^{-7}
	ENU 1.2x mixture	*9.85×10^{-8}*
	ENU 1.5x mixture	*1.23×10^{-7}*
	ENU 2x mixture	*1.64×10^{-7}*
	ENU pure	1.82×10^{-6}

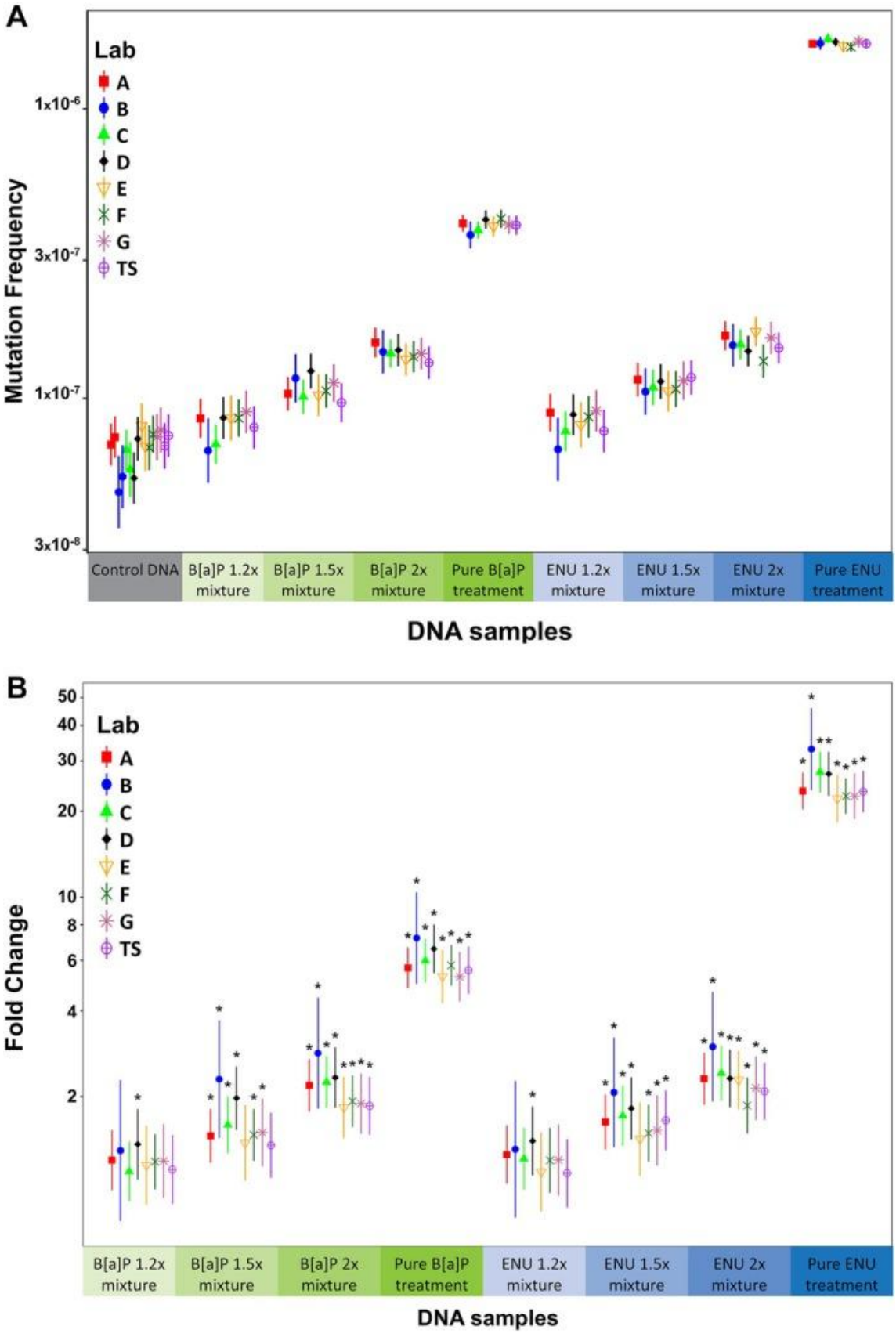
Appendix Figure B1: Study site locations and sample descriptions. A. Locations of eight study sites spanning four countries and two continents. B. Pure and mixture samples used in study and previously measured or expected mutation frequencies (MFmin). Expected MFmin values targeted by the DNA mixtures are italicized. Colored boxes at left of table indicate sample color coding used in subsequent figures. See Methods for details regarding preparation of mixtures.



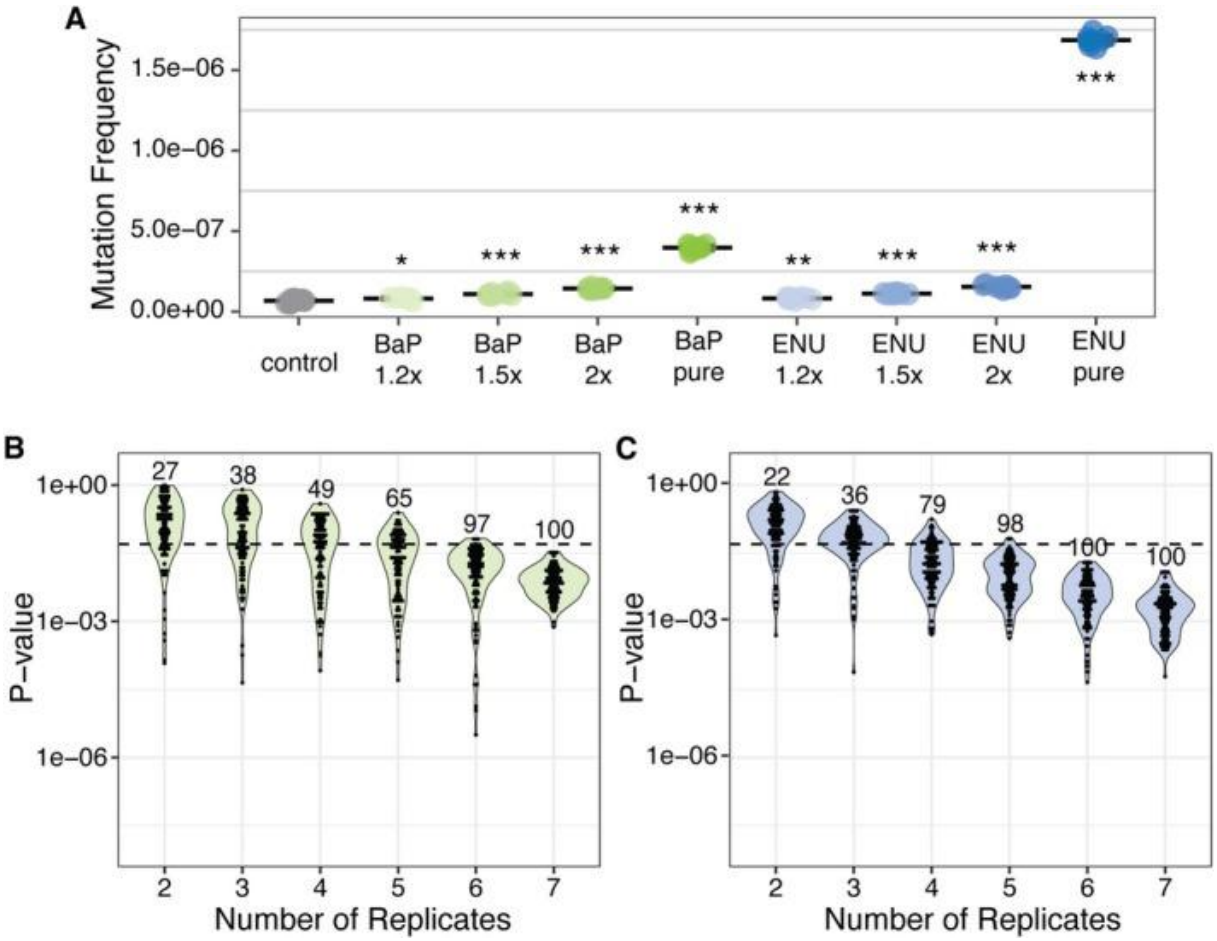
Appendix Figure B2: Assay performance matrix across eight laboratories. Per-library informative duplex bases (A), mean duplex depth (B), and peak tag family size (C) are plotted as points, grouped by lab along the x-axis and color-coded by sample type. Each plot is divided into sub-panels for libraries prepared with 500 ng (left) or 1000 ng (right) of input mass.



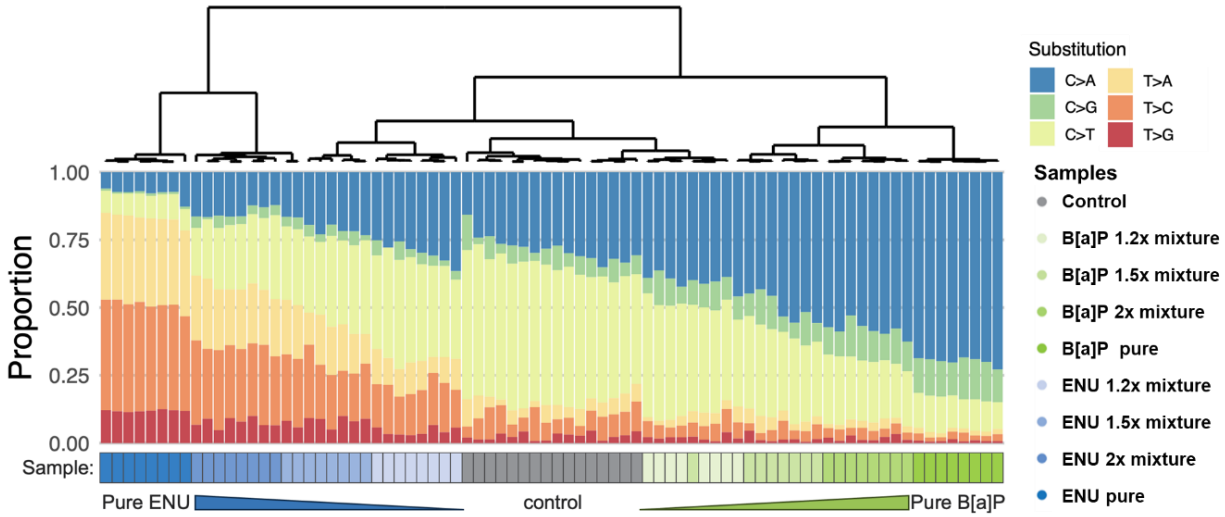
Appendix Figure B3: High interlaboratory reproducibility demonstrated in DS analyses of shared DNA samples. MFmin was calculated as the unique mutation count divided by the total duplex bases (excluding no-calls). Error bars represent 95% binomial proportion Wilson score intervals calculated from the mutation frequency data.



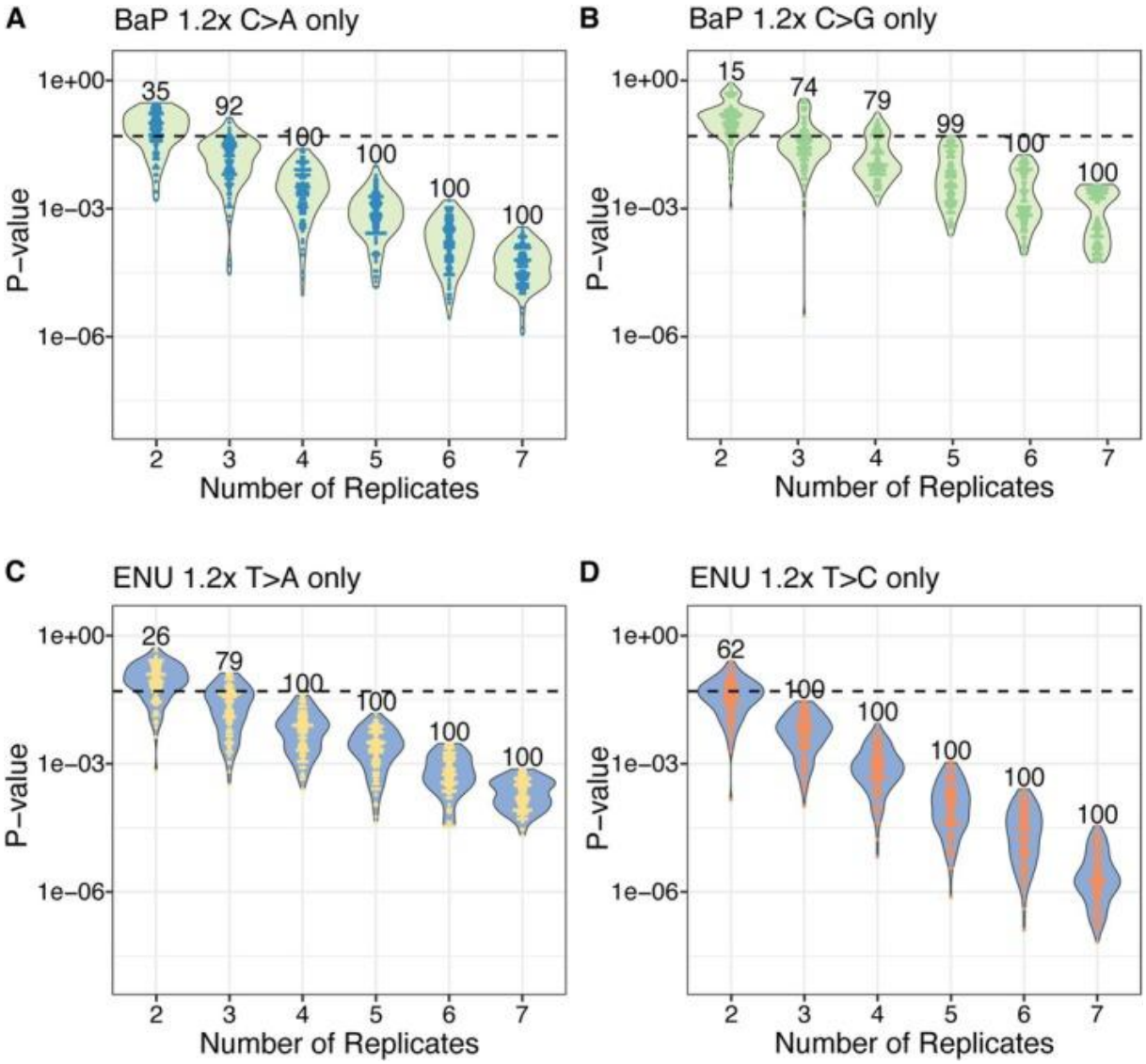
Appendix Figure B4: Highly reproducible results across labs for both B[a]P and ENU mixtures as measured by mutation frequencies. A. MFs for each sample estimated using a generalized linear model. Error bars represent the standard error. B. Fold change in MFs from control for each mixture or pure sample across labs. Error bars represent the standard error. Asterisks indicate significance at 0.05 level.



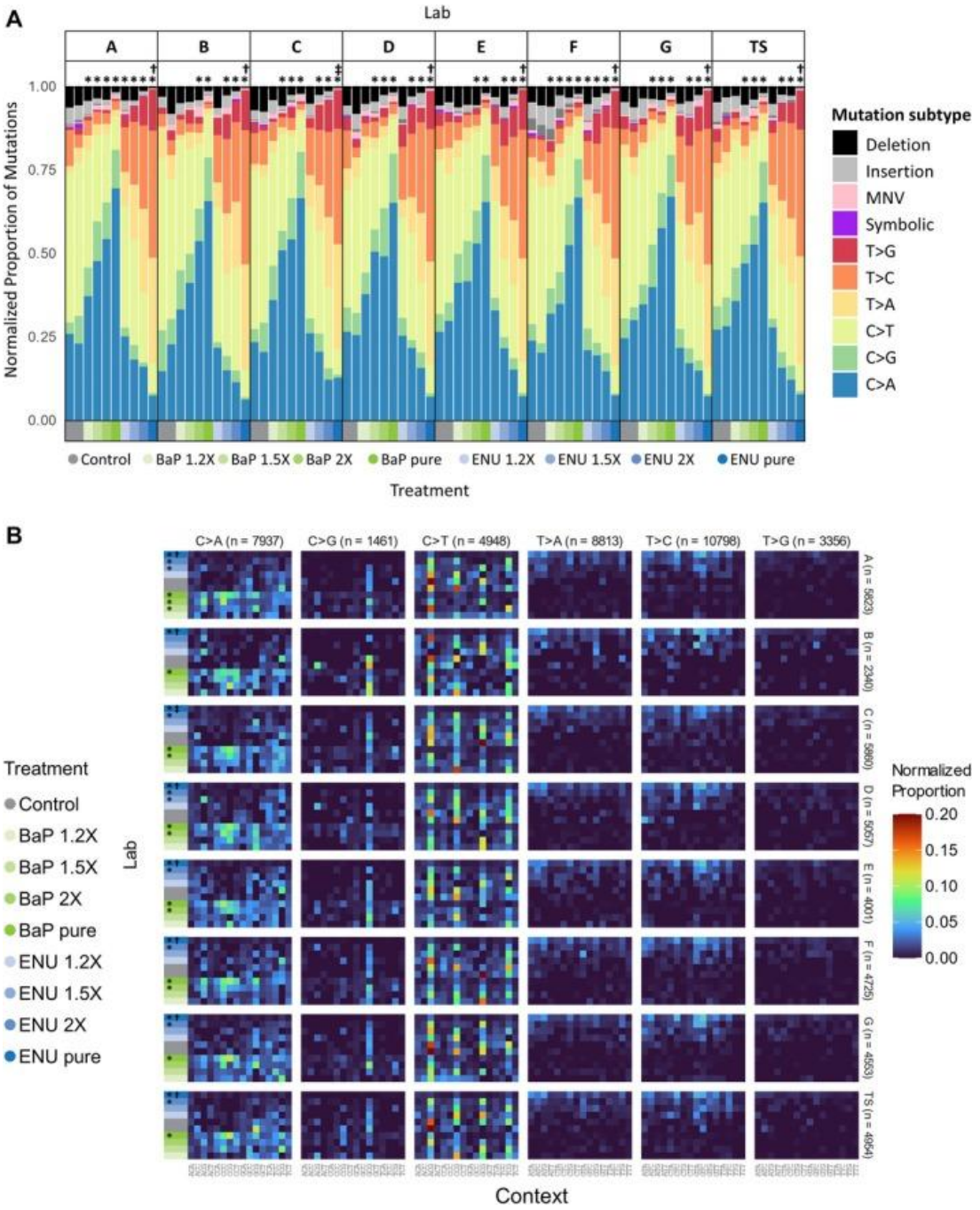
Appendix Figure B5: Group and replicate analysis of mutation frequencies. (A). MFs plotted by sample type, treating libraries prepared by different labs as technical replicates. Each dot represents one library, with horizontal lines representing the group mean. Asterisks indicate FDR-adjusted p-values less than 1x10-2 (*), 1x10-3 (**), and 1x10-4 (***). Replicate libraries of the untreated control and 1.2x mixtures were randomly selected to make smaller groups and statistical testing was performed on the sampled groups. Independent pairwise comparisons were performed for untreated and the B[a]P 1.2x mixture (B) and for untreated and the ENU 1.2x mixture (C). For each group size, indicated by the x-axis, 100 independent samplings were performed and the p-value for each comparison is plotted on the y-axis. Points indicate individual iterations of sampling and "violin" shapes represent the density distribution of the resulting p-values for each group size. Number above each set of points indicates the number of p-values < 0.05 out of the 100 sampling iterations.



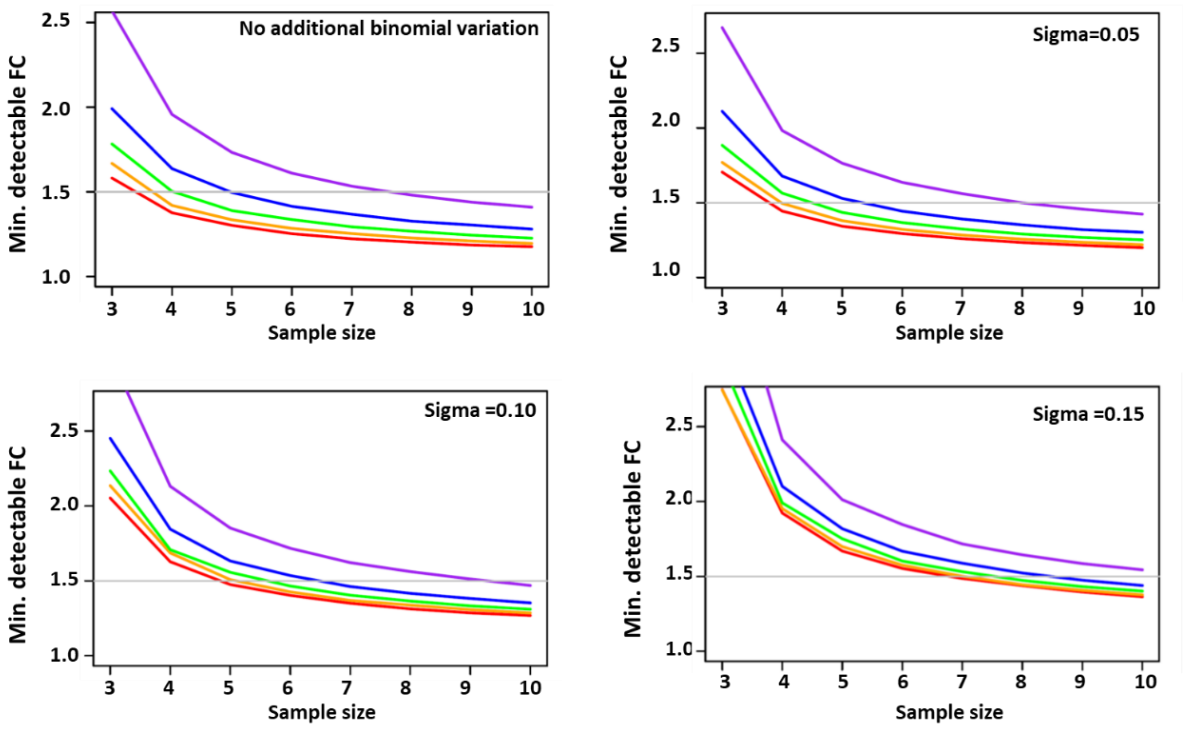
Appendix Figure B6 Unsupervised hierarchical clustering by SBS spectra. For each library, the normalized proportions of simple base substitution types (pyrimidine notation) are plotted as a stacked bar (middle panel). The dendrogram (top) reflects the results of unsupervised hierarchical clustering. The sample identity for each library is indicated along the x-axis by color-coded boxes.



Appendix Figure B7: Replicate analysis of subtype mutation frequencies. Similar to the replicate sampling analysis performed for total MF, replicates were randomly selected into smaller groups and pairwise testing was performed to determine if a significant increase in subtype MF could be detected. For each subtype in each sample pairing, 100 iterations of sampling and statistical testing were performed, represented by individual points. Larger shapes show the density distribution of p-values for each group size and numbers represent the number of iterations that yielded a p-value < 0.05. For each mutagen, the most dominant mutation subtypes are shown C>A (A) and C>G (B) mutations for the B[a]P 1.2× mixture and T>A (C) and T>C (D) mutations for the ENU 1.2× mixture



Appendix Figure B8: Comparison of mutation spectra between treatment groups and across labs. (A) The normalized proportions of simple base substitution (SNV) types (pyrimidine notation) and non-SNV variant types are plotted as a stacked bar for each library. The treatment for each library is indicated along the x-axis by color-coded boxes. Samples are organized by lab, which are denoted at the top of the plot. Asterisks indicate significant differences between a sample and its within-lab control. Daggers represent comparisons of a treatment group between labs; samples with the same dagger are not significantly different from one another. All comparisons were made at a significance level of 0.05. B. The normalized proportions of SNV subtypes (pyrimidine notation) in their trinucleotide context are plotted in a heatmap for each library. The treatment for each library is indicated along the y-axis by color-coded boxes. Samples are faceted by lab (right) and include the total mutation count for all samples per lab. Trinucleotide subtypes are faceted by SNV subtype and include the total mutation count for each subtype across all samples. Asterisks indicate significant differences between a sample and its within-lab control. Daggers represent comparisons of a treatment group between labs; samples with the same dagger are not significantly different from one another. All comparisons were made at a significance level of 0.05.



Appendix Figure B9: Power analysis of minimum detectable fold-change using different sample size. The calculation was based on MF of 6.7x10-8 with total number of informative duplex bases of: 1.71x109 (Red), 1.34x109 (Orange), 1.03x109 (Green), 6.84x108 (Blue), and 3.42x108 (Purple). The grey line represents the minimum detectable fold change of 1.5-fold. Power analysis was conducted assuming no additional binomial variation or three different sample variances (sigma = 0.05, 0.10, and 0.15) with a desired power of 80%.

Appendix C: MutSeqR: An Open-Source R Package for Standardized Analysis of Error-Corrected Next-Generation Sequencing Data in Genetic Toxicology

Modified from Annette E. Dodge, Andrew Williams, Danielle P.M. LeBlanc, David M. Schuster, Elena Esina, Clint C. Valentine, Jesse J. Salk, Alex Maslov, Chris Bradley, Carole L. Yauk, Francesco Marchetti, Matthew Meier. MutSeqR: an open-source R package for standardized analysis of error-corrected next-generation sequencing data in genetic toxicology. Bioinform Adv. 2025 Oct 23;5(1):vbaf265. doi: 10.1093/bioadv/vbaf265

C.1 Abstract

Motivation: Error-corrected next-generation sequencing (ECS) methods are increasingly used to assess mutagenicity and other genetic toxicology endpoints. The lack of open and standardized bioinformatic workflows and tools poses challenges to data reproducibility, comparability, and consistency in interpretation for its application in genetic toxicity assessment.

Results: We present MutSeqR, an open-source R package to analyze ECS mutation data for genetic toxicology studies. MutSeqR offers practical variant filtering, comparative analysis of mutation frequency between experimental conditions, dose-response assessment via benchmark dose calculations, mutation spectrum analysis, and clonality analyses. We demonstrate MutSeqR's application using published datasets on mice treated with benzo[*a*]pyrene or benzo[*b*]fluoranthene, analyzed using Duplex Sequencing and SMM-seq, respectively. MutSeqR's flexible functions enable reproducible analyses across ECS platforms, facilitating research and regulatory applications in mutagenicity testing.

Availability and Implementation: MutSeqR is freely available under an open-source license at https://github.com/EHSRB-BSRSE-Bioinformatics/MutSeqR. Implemented in R (version 3.4.0 or greater), it supports all major operating systems. Sequencing data for Project 1 has been deposited in the Sequence Read Archive under accession number PRJNA803048. Variant call files for Project 2 are available on Mendeley Data (doi:10.17632/65dnysxym8.1).

C.2 Introduction

Assessing xenobiotics for mutagenicity is vital to human health, as somatic mutagenesis contributes to many diseases(Erickson, 2003; Godschalk et al., 2020; Jackson et al., 2018), including cancer, and germline mutations can cause hereditary diseases (Beal, et al., 2017; Marchetti, et al., 2020; Nguengang Wakap et al. 2020). Thus, human and environmental risk assessments require mutagenicity testing of existing and new substances (Krewski et al., 2010). Error-corrected next-generation sequencing (ECS) is an emerging, high-throughput method for direct measurement of mutations in any DNA-based model system (Marchetti et al., 2023b; Menon & Brash, 2023). ECS quantifies spontaneous and genotoxicant-induced mutations and provides in-depth characterization of mutational spectra (Cho et al., 2023; Dodge et al., 2023; LeBlanc et al., 2022; Schuster et al., 2024; Valentine et al., 2020). ECS enables mechanistic insights and accurate quantification of mutation frequency and holds potential to advance the translation of mutagenicity testing on xenobiotics (using cell culture and animal models) to human health risk assessment.

ECS improves sequencing accuracy by grouping sequence reads derived from the same DNA molecule to subtract artifacts introduced during library preparation and sequencing (Salk & Kennedy, 2020). This enables the detection of ultra-low frequency mutations within large populations of non-mutant DNA molecules. ECS technologies are being developed for a broad range of applications, including cancer biology (Abbasi & Alexandrov, 2021; Kennedy et al., 2019; Kostecka et al., 2022; Valentine et al., 2020) and genetic toxicology (Bercu et al., 2023; Cho et al., 2023; Dodge et al., 2023; LeBlanc et al., 2022; Schuster et al., 2024; Smith-Roe et al., 2023). ECS strategies include single-strand (e.g. Safe-SeqS, SiMSen-Seq), tandem-strand (e.g., o2n-Seq, SMM-Seq) and double-strand (e.g. Duplex Sequencing (DS), PECC-Seq, Jade-Seq, PacBio HiFi, CODEC, etc.) consensus sequence methods. Researchers are advocating for the regulatory adoption of ECS for genetic toxicology and cancer risk assessment of xenobiotics (Marchetti et al., 2023b).

The absence of standardized, transparent methods for ECS data analysis remains a barrier to regulatory implementation, as current studies often use proprietary or inconsistent approaches (Marchetti et al., 2023b). Before ECS can be used in regulatory risk assessment, steps must be taken to establish standard guidance for

analyzing ECS data that will unite the many ECS technologies into a single analytical workflow. To address this, we present MutSeqR, an open-source R package for downstream mutation analysis of ECS data (). While MutSeqR does not handle pre-processing of raw sequences or variant calling, it provides robust tools for statistical analysis, integrates tools for characterizing mutational signatures, and generates informative figures. This package aims to enable open-source analyses of ECS data, empowering users with flexibility during exploratory analyses while ensuring compatibility across technologies. We demonstrate its utility on two mouse datasets: 1) mouse bone marrow samples sequenced using TwinStrand's DS, and 2) mouse liver samples using MutagenTech's SMM-Seq (single-molecule mutation sequencing). DS sequenced a panel of 48kb at > 10,000X depth. Alternatively, SMM-Seq randomly sequenced genome-wide, trading read depth for breadth, spanning ~200Mb. Our analyses highlight how MutSeqR enables detailed, reproducible, and platform-independent assessment of mutagenicity.

C.3 Methods

C.3.1 Datasets

Mutation data generated using two ECS technologies were analyzed using MutSeqR to demonstrate its compatibility across platforms. Project 1 included 24 bone marrow DNA samples of MutaMouse males exposed to one of three doses of benzo[*a*]pyrene (BaP) or a vehicle control for 28 days (LeBlanc et al., 2022). DNA samples were sequenced at ≥10,000X depth using DS on the Mouse Mutagenesis Panel (https://github.com/twinstrandbio/twinstrandbio-reference-data/), which consists of twenty 2.4kb target regions (Appendix Figure C2A). Pre-processing of sequencing data was done using TwinStrand's Mutagenesis App (v. 3.20.1) as described in Valentine et al., 2020 to produce tabular mutation data files. Briefly, pre-processing included: raw read alignment, grouping reads by their unique molecular identifiers (UMI) and strand-defining element, quality trimming, error correction of read groups by duplex consensus calling, consensus post-processing, realignment, and variant calling via VarDictJava (Lai et al., 2016).

Project 2 comprised 24 liver DNA samples from MutaMouse males following 28-day exposure to one of five doses of benzo[*b*]fluoranthene (BbF) or a vehicle control(Schuster et al., 2024). Samples were subjected to random genome-wide

sequencing (~200Mb per sample) using SMM-Seq as described (Maslov et al., 2022) (Appendix Figure C2B) resulting in VCF files of single-nucleotide variants. Briefly, raw reads were quality trimmed and aligned to the reference genome. Next, reads were grouped by UMI, then collapsed into a single consensus sequence. Variants were identified from the consensus sequences, and germline mutations were removed.

C.3.2 Data import

Mutation data is imported into R as VCF files or tabular data using the functions *import_vcf_data()* and *import_mut_data(),* respectively. The functions load all files from a user-provided file path and parse them into a unified data frame. VCF files (or block-compressed VCFs) must be single-sample and should conform to the VCF specification (version 4.5; see required fields in Table 1). We recommend providing a record for every sequenced position (which should include the *total_depth* column), to enable site-specific depth frequency calculations (e.g. subtype frequencies). If non-variant sites are unavailable, users may supply a separate data frame to specify precalculated depths for the desired level of granularity.

Upon import, records are categorized as: 1) no_variant; 2) snv: single-nucleotide variant; 3) mnv: multi-nucleotide variant; 4) insertion; 5) deletion; 6) complex: REF and ALT are of different lengths and compositions; 7) sv: structural variants; 8) ambiguous: IUPAC ambiguity; and, 9) uncategorized: the variant does not fall into any of the preceding categories (Supp. Table 1). The trinucleotide context for each site is retrieved using *BSgenome*. This includes the reference base plus its two flanking nucleotides. Additional columns are appended to characterize variants (e.g. variant length, variant allele fraction (VAF), GC content). Lastly, SNV subtypes are stratified by pyrimidine reference (C/T) in conjunction with their trinucleotide context, enabling flexible calculation of subtype frequencies. Mutation data can be output as a data frame or a *GRange* object for downstream analysis (Lawrence et al., 2013).

Project 1: Data for each sample was provided as tabular files containing records for all 48,000 positions on the Mouse Mutagenesis Panel, including *alt_depth* and *total_depth* values. Data was imported using *import_mut_data()*.

Project 2: Data for each sample was provided as compressed VCF files, listing only variant sites. No-variant positions were not included; instead, per-sample total

depths were provided separately for mutation frequency calculations. The data was imported using *import_vcf_data()*.

C.3.3 Variant Filtering

Following import, mutation data is filtered using *filter_mut(),* which flags variants according to user-specified parameters in the *filter_mut* column. Variants flagged as TRUE in *filter_mut* are automatically excluded from mutation counts in downstream analyses (e.g. in *calculate_mf()*), but are retained in the dataset so their *total_depth* values remain available for frequency calculations. Optionally, filtered records can be removed entirely.

Project 1. Filters were applied per TwinStrand's recommendations:

1. *vaf_cutoff = 0.01*. Variants exceeding this VAF were flagged as putative germline variants.
2. *snv_in_germ_mnv = TRUE.* SNVs overlapping germline MNVs were flagged as probable artifacts from variant calling. No-calls in reads supporting the germline MNV will create false minor haplotypes from the original MNV that can appear as sub-clonal SNVs.
3. *custom_filter = filter, custom_filter_val = "*EndRepairFillinArtifact*".* A customizable filter was applied to TwinStrand's column "*filter*". Variants reported as an End-Repair Fill-in Artifact by TwinStrand's pipeline were flagged.
4. *regions = TSpanel_mouse".* Records outside target regions were removed. The regions for TwinStrand's Mouse Mutagenesis Panel are stored in the package files. Upon specification, records that are >1bp must start and end within the target region or be removed.
5. *rm_filtered_mut_from_depth* = TRUE. For flagged variants (other than putative germline), *alt_depth* was subtracted from *total_depth*, treating these as No-calls.

Project 2. Variant filtering was performed by MutagenTech. Variants were filtered against the single-nucleotide polymorphisms database (dbSNP) to remove known mouse variants, and additional germline variants were removed based on matched whole-genome sequencing at 20X depth.

C.3.4 Calculating Mutation Frequencies

Mutation frequency (MF) is calculated by dividing the sum of mutations by the sum of the *total_depth* (mutations/bp) using the function *calculate_mf().* Users can supply grouping variables (e.g. by sample, experimental group, genomic region, subtype), allowing MF to be computed across various categories. To prevent double-counting read depth when multiple variants are called at the sample position in a sample, only the first record retains its *total_depth*, while subsequent records are set to zero. For datasets lacking per-site depth, precalculated depths can be provided via the *precalc_depth_data* parameter.

Two mutation counting methods are implemented (Dodge et al., 2023):

1. The Minimum Independent Mutation Counting Method (Min): Each mutation is counted once, regardless of the number of reads supporting the non-reference allele. This method assumes that multiple instances of the same mutation within a sample are the result of a single mutational event that underwent subsequent clonal expansion.
2. The Maximum Independent Mutation Counting Method (Max): Multiple identical mutations at the same position within a sample are counted as independent mutational events. All reads with the non-reference allele are included in the count.

MutSeqR automatically calculates both Min and Max mutation frequencies (MFmin and MFmax).

C.3.5 Mutation Subtypes: Frequencies and Proportions

calculate_mf() can also group mutations by their subtype at various resolutions. All mutations can be resolved to their variation type, and SNVs can be further resolved into: 1) 6-base: SNV substitution types (subtypes) reported in their pyrimidine reference; 2) 12-base: SNV subtypes; 3) 96-base: 6-base SNV subtypes reported within their trinucleotide context; and, 4) 192-base: 12-base SNV subtypes reported within their trinucleotide context. Mutations for each subtype are summed across groups. The *total_depth* is summed across groups for positions at which the subtype can occur. In the simplest example, for the 6-base SNV subtypes, the two possible reference bases are C or T; hence, the depth is calculated separately for C:G positions

and T:A positions. The function also calculates the proportion of mutations for each subtype, normalized to read depth and coverage of those reads. The normalized proportion of a subtype is calculated by dividing the MF of the given subtype by the sum of all subtype frequencies in a group.

$$P_s = \frac{\frac{M_s}{D_s}}{\sum_s \frac{M_s}{D_s}}$$

P_s Mutation proportion for subtype s, normalized to the read depth. M_s Mutation count for the subtype s. D_s read depth for s context.

If per-site depth is unavailable, users can supply precalculated depth at the chosen subtype resolution. Per our previous example, users would supply precalculated depth for all C:G positions and T:A positions in each group. When this is not possible, (non-normalized) subtype proportions (P'_s) are reported by dividing subtype mutation counts (M_s) by the total number of mutations in the group (M_{total}).

$$P'_s = \frac{M_s}{M_{total}}$$

Project 1: Per-site *total_depth* values were summed to calculate MF and normalized subtype proportions.

Project 2: Precalculated *total_depth* values provided by MutagenTech were used for MF calculation.

C.4 Statistical Approaches

C.4.1 Mutation Frequency Modeling

The *model_mf()* function quantitatively assesses how MF varies with experimental factors. It models the proportion of mutated reads as a function of user-defined fixed and/or random effects and interaction parameters. Depending on the supplied effects, *model_mf()* will automatically choose to fit either a generalized linear model (GLM) or a generalized linear mixed model (GLMM). Mutation counts are treated as binomially distributed, given (a) there is a finite number of sequenced bases, (b) a mutation at any given base is equally probable (acknowledging this does not fully

reflect biological reality), and (c) mutations occur independently of other mutations. To handle over-dispersion, GLM fits use a quasibinomial distribution, unless dispersion <1, in which case a binomial distribution is used; GLMMs employ a binomial distribution throughout.

model_*mf()* provides estimates of the mean for each level of the fixed effects. Furthermore, pairwise comparisons can be performed based on a user-supplied table. Mean estimates and comparisons are conducted using the doBy R library (Halekoh & Højsgaard, 2006). The *esticon()* function is used to generate estimates for specified linear combinations of the model parameters based on the fitted model:

$$g(\mu_i) = logit(\mu_i) = \log\left(\frac{\mu_i}{1-\mu_i}\right) = X_i\beta$$

Where μ_i is the expected value of the response for group *i*, X_i is the ith column of the design matrix for the fixed effects and β is the vector of coefficients associated with the fixed effects.

This formula can be re-written such that:

$$\mu_i = \frac{exp(X_i\beta)}{1+exp(X_i\beta)} \approx exp(X_i\beta) = MF_i$$

Since the mutation frequencies (MF_i) are assumed to be small (~10^{-8}), $1 + exp(X_i\beta) \approx 1$.

The linear combination of the model parameters would then be:

$$\hat{\theta} = c^T\hat{\beta}$$

Where c^Twould represent the contrast vector specifying the weights, for example, if the model includes two treatment groups with estimated coefficients $\hat{\beta}_1$and $\hat{\beta}_2$, the contrast or linear combination to estimate the difference would be:

$$c = (0, 1, -1, 0, \ldots, 0)$$

resulting in

$$\hat{\theta} = c^T \hat{\beta} = \hat{\beta}_1 - \hat{\beta}_2$$

Since the mutation frequencies are small, exponentiating $\hat{\theta}$ would approximate the fold change of group 1 to group 2. In the case of a log-link (common for rare event models like Poisson or binomial GLMs with small probabilities), this provides a direct estimate of the relative risk or the rate ratio between groups.

Estimates of the back-transformed standard errors are approximated using the delta method. The p-values are adjusted for multiple comparisons using the Sidak method. These methods are based on approaches developed for the Transgenic Gene Mutation Reporter (TGR) assay (Fung et al. 1998).

We used *model_mf()* to assess the effect of BaP (Project 1) and BbF (Project 2) dose on MFmin using a GLM with quasibinomial dispersion. For Project 1, we extended this model to a GLMM using effects of dose and sequencing target with the sample as a random effect.

C.4.2 Benchmark Dose (BMD) Modeling

Dose-response models are essential for quantitative risk assessment of mutagenicity, as they provide a framework to evaluate the levels at which exposure to a substance might cause an adverse effect. The BMD is a dose that produces a predetermined change in the measured response, defined as the benchmark response (BMR). The BMD is used as a point of departure to derive human health-based guidance values to inform regulatory risk assessment. Our package provides two functions for BMD modoling, each employing a widely used software designed to be consistent with methods used by regulatory authorities.

1. *bmd_proast()*: Runs a parameterized version of the proast71.1 R package (www.rivm.nl/en/proast). This function will analyze continuous, individual MF data following a log transformation using exponential, Hill, inverse exponential, and log-normal models, selecting the best fit by AIC. Confidence intervals are computed via profile likelihood or bootstrap model averaging (recommended). MutSeqR calculates the model-averaged BMD as the median BMD of all bootstraps. BMR values may be specified as a relative increase or as one standard deviation from the control. Covariates are supported.

2. *bmd_toxicr():* Utilizes the ToxicR package (Wheeler et al., 2023) to analyze continuous/individual or continuous/summary MF data, assuming either a normal (default) or log-normal distribution. This function employs Bayesian estimation using either the Laplace Maximum *a posteriori* approach (default) or Markov chain Monte Carlo (MCMC) simulation. The default model parameters' prior distributions are specified in (Wheeler et al., 2020), but they are user-modifiable. The default models are described in the European Food Safety Authority's 2022 Guidance on the use of the benchmark dose approach in risk assessment (EFSA Scientific Committee et al., 2022). Model averaging may be applied using described methodologies (Wheeler et al., 2020, 2022). Finally, the BMR can be specified as an absolute or relative increase, or as a specified number of standard deviations from control.

We used both functions to perform BMD analysis with model averaging to calculate the BaP/BbF dose that produces a 50% relative increase in MFmin compared to control with 90% confidence intervals (White et al., 2020).

C.4.3 Comparing Mutation Spectra

A mutation spectrum is the pattern of mutation subtypes within a sample or group, which can reveal mutagenic mechanisms associated with endogenous processes or chemical exposure. *spectra_comparison*() constructs an $R*T$ contingency table of mutation counts where R is the number of subtypes and T represents the two groups being compared (Piegorsch & Bailer, 1994). It then tests for differences in subtype proportions between groups using the the G^2 likelihood ratio statistic.

$$G^2 = 2\sum_{i=1}^{R}\sum_{j=1}^{T} Y_{ij}\log(\frac{Y_{ij}}{E_{ij}})$$

where Y_{ij} is the observed mutation count and E_{ij} the *expected* counts under the null hypothesis of no difference. For large sample sizes, G^2 is referred to a χ^2 distribution with $(R-1)(T-1)$ degrees of freedom. However, when N/(R-1) < 20 (where N is the total mutation count across groups), an F-distribution is used for greater reliability, reducing false positives when sample sizes are small.

Spectrum analysis assumes independence among the observations. This is appropriate for mutations from mixed populations, but not for those derived from a single progenitor cell. Accordingly, the MFmin method is used here to ensure counts are independent.

Mutation spectra were compared between BaP/BbF dose groups and controls at multiple subtype resolutions.

C.4.4 Hierarchical Clustering of Mutation Spectra

Hierarchical clustering groups similar data points into nested clusters based on pairwise similarities. *cluster_spectra()* clusters samples (or user-defined groups) according to their mutation spectra, which may be represented as subtype counts, frequencies, or proportions. MutSeqR computes distance using *dist()* from the stats library (default Euclidean). The resulting distance matrix is then passed to *hclust()* to cluster samples using the specified linkage method (default Ward). Optimal leaf ordering is performed with the dendsort library, producing a dendrogram of sample relationships.

For each Project, samples were clustered by the proportions of their 6-base SNV subtypes and non-SNV variation types using default parameters.

C.4.5 Mutational Signature Assignment

Mutational processes generate characteristic patterns of mutations, known as mutational signatures. Distinct mutational signatures have been extracted from various cancer types and normal tissues using data from the Catalogue of Somatic Mutations in Cancer (COSMIC) database and compiled into a known collection (https://cancer.sanger.ac.uk/signatures/). Signatures have also been characterized for specific environmental chemicals (Kucab et al., 2019) (https://signal.mutationalsignatures.com/), germline mutations (Seplyarskiy et al., 2021), and DNA repair deficiencies (Zou et al., 2021). Linking ECS mutational profiles of specific mutagens to existing mutational signatures provides empirical evidence for the contribution of environmental mutagens to the mutations found in human cancers and informs on mutagenic mechanisms.

MutSeqR's *signature_fitting()* utilizes the SigProfiler suite of tools (Díaz-Gay et al., 2023; Khandekar et al., 2023) to assign SBS signatures from the COSMIC database to the 96-base SNV subtypes of a given group by creating a virtual environment to run Python using *reticulate*. Thus, *signature_fitting()* function facilitates interoperability between these tools for users less familiar with Python and assists users by coercing the mutation data to the necessary structure for the SigProfiler tools. SNVs in their 96-base trinucleotide context are summed across groups to create a mutation count matrix by *SigProfilerMatrixGeneratorR()* (Khandekar et al., 2023). *Analyze.cosmic_fit* (SigProfilerAssignment) is then run to assign mutational signatures to each group using refitting methods (Díaz-Gay et al., 2023), which quantifies the contribution of a set of signatures to the group's mutational profile. Cosine similarity values are generated to compare the reconstructed mutational profile to the original mutational profile of the group, with values > 0.9 indicating a robust reconstruction.

Mutations were grouped by BaP/BbF dose group for signature assignment

C.4.6 Data Visualization

MutSeqR has several plotting functions for data visualization.

1. *plot_mf()/plot_mean_mf()* creates a plot of the MF or mean MF from *calculate_mf()* output (Appendix Figure C3). Variable plotting aesthetics (bars, points, lines) are offered. MFmin and MFmax can be displayed individually, side by side, or stacked. Data can be coloured based on a specified variable and labelled with mutation count or frequency. The standard error of the mean may be shown as error bars. Individual samples may be plotted alongside the mean values.
2. *plot_model_mf()* plots the point estimates derived using *model_mf()* (Appendix Figure C4A and Appendix Figure C7A). Data is presented as either bars or points. Users can include the estimated standard error as error bars and add significance labels based on pairwise comparisons.
3. *plot_ci()* creates a plot of BMD confidence intervals for easy comparison between responses (Appendix Figure C4B).
4. *plot_spectra()* creates a transition-transition plot of mutation spectra for user-defined groups at the desired subtype resolution (Appendix Figure C5A). Data can represent subtype mutation count, MF, or proportion. This function can call

upon *cluster_spectra()* to overlay a dendrogram of the clustered groups (Appendix Figure C5B).

5. *plot_trinucleotide()* creates bar plots of the 96-base SNV (trinucleotide) spectrum for all levels of a user-defined group (Appendix Figure C6). Data can represent subtype mutation count, MF, or proportion. Aesthetics are consistent with COSMIC trinucleotide plots.
6. *plot_trinucleotide_heatmap()* creates a heatmap plot of 96-base SNV proportions. Plots can be facetted by a specified grouping variable in the MF data.
7. *plot _radar()* generates a radar/spider plot of MF given a user-defined group (Appendix Figure C7B). Data may be plotted on a single plot or facetted by group.
8. *plot_bubbles()* Creates a bubble plot with circles representing each mutation in a given group (Appendix Figure C8). Plots are facetted by group. Circle size can reflect either *alt_depth* or *VAF.* Circles may be coloured by a specified variable in the mutation data (e.g. subtype/loci).

C.5 Results

C.5.1 Variant Filtering

Project 1: *filter_mut()* flagged or removed 2660 out of 1,152,911 rows of mutation data. 612 variants were identified as putatively germline, and 20 SNVs were flagged because they overlapped with germline MNVs. 2021 variants were flagged by the custom filter. 22 rows were removed because they extended outside of the target regions. 909 rows had their *total_depth* corrected.

C.5.2 Quantifying Mutation Frequencies

MutSeqR's *calculate_mf()* provided MF per sample for both projects, and results were visualized as a bar plot using *plot_mf().* Appendix Figure C3A and B display MF per sample for Projects 1 and 2, respectively. Bars were coloured and ordered by dose. Optional data labels were added to Appendix Figure C3B, displaying the mutation count per sample. These plots offer a useful first exploration of the mutation data. Project 1 shows that BaP causes a dose-dependent increase in MFmin and MFmax, with low inter-sample variability, and a high number of recurrent mutations at high

doses. In Project 2, MFmin increases with BbF dose. As SMM-Seq sequenced each site only once, MFmax could not be assessed.

Mean ± SEM. MF per dose was calculated and visualized using *plot_mean_mf()*. Appendix Figure C3C displays Project 1 mean MFmin and MFmax as a bar plot with bars coloured by MF type. Under the Min assumption, we calculated a mean of 17.4 ± 2.26, 34.3 ± 1.38, 68.4 ± 5.72, and 95.4 ± 2.87 ($x10^{-8}$ mutations/bp) for the control, low, medium, and high doses, respectively. Under MFmax, a mean of 21.7 ± 2.26, 51.5 ± 3.18, 105 ± 5.11, and 263 ± 26.7 ($x10^{-8}$ mutations/bp) was calculated, respectively. Appendix Figure C3D displays Project 2 mean MFmin as a line alongside the individual values for each sample, demonstrating the flexible aesthetics of the plotting function. Mean MFmin ($x10^{-8}$ mutations/bp) was 8.70 ± 0.792, 6.82 ± 0.521, 10.2 ± 0.141, 10.3 ± 1.25, 17.5 ± 1.24, and 23.1 ± 2.11 for the control, low, medium-low, medium, medium-high, and high doses, respectively.

C.5.3 Modeling Dose Effects

The dose-dependent effects of BaP/BbF on MFmin were quantified using two methods. First, a GLM was fit to estimate the dose effect on the MFmin using *model_mf()*. This model assumes that the response variable follows a specified distribution (e.g., an over-dispersed binomial) and uses a link function to relate the mean response to the predictors. To assess model fit, Pearson residuals were visualized using a histogram (Supplementary data C: Figure 1A & C) and a Q–Q plot (Supplementary data C: Figure 1B & D) to check for deviations from the model assumptions. These plots were automatically generated by *model_mf().* BaP induced a significant increase in MFmin for all dose groups, reaching a maximum of a 5.5-fold-increase at the high dose (Appendix Figure C4A). BbF induced a significant increase in MFmin for the top two dose groups, reaching a maximum of a 2.6-fold-increase at the high dose (Supplementary data C: Figure 2A). Model results were plotted using *plot_model_mf(),* which created a bar plot of the model-estimated mean MF per dose. Significance labels were added based on the specified comparisons. The output is a *ggplot* object that is easily modifiable using *ggplot2* functions. In this way, we included the empirical mean MFmin as dark blue lines. Estimated mean values were almost identical to empirical values for all BaP and BbF doses.

Next, BMD models were used to analyze the dose-response (Appendix Figure C4B). MutSeqR provides two methods for BMD modeling: *bmd_toxicr()* and *bmd_proast()*. We used both methods with model averaging to calculate the lower (BMDL) and upper (BMDU) 90% BMD confidence intervals. ToxicR estimated that a BaP dose of 8.06 mg/kg-bw/d (BMDL = 5.58, BMDU = 10.43) induces a 50% relative increase in MFmin from control (Supp. Table 2). For PROAST (Supp. Table 3), the model-averaged BMD_{50} was 9.11 mg/kg-bw/d (BMDL = 7.38, BMDU = 10.9). ToxicR estimated that a BbF dose of 30.9 mg/kg-bw/d (BMDL = 19.0, BMDU = 40.7) induces a 50% relative increase in MFmin from control (Supp. Table 4). For PROAST (Supp. Table 5), the model-averaged BMD_{50} was 28.8 mg/kg-bw/d (BMDL = 17.0, BMDU = 39.4). The BMD_{50} and their confidence intervals were plotted using *plot_ci()* (BaP - Appendix Figure C4B & BbF - Supplementary data C: Figure 2B).

C.5.4 Mutation Spectra

MutSeqR's *calculate_mf()* was used to resolve mutations to their variation types. SNVs were examined at both the 6-base and 96-base resolutions and visualized for each dose using *plot_spectra()*. BaP induced primarily SNVs, followed by small deletions, small insertions, MNVs, and lastly SVs (Appendix Figure C5A). *spectra_comparison()* revealed that the BaP subtype proportions were significantly different from the controls at all doses ($p < 0.001$), indicating that BaP induces a unique pattern of mutation subtypes. Specifically, we observed dose-dependent increases in the proportions of C:G>A:T mutations. This is consistent with the known mode of action of BaP, which forms bulky DNA adducts mostly at the N2 of guanine through its metabolite benzo(a)pyrene-7,8-diol-9,10-epoxide (Beal et al., 2015; LeBlanc et al., 2022; Liamin et al., 2017). We next used *cluster_spectra()* to cluster the samples based on their subtype proportions (Appendix Figure C5B). Samples from the same dose groups largely clustered together. Clear distinctions were made between the high and medium doses versus the control and low doses. Finally, we plotted the proportions of the 96-base spectra for each dose using *plot_trinucleotide()* (Appendix Figure C6). We observed that high proportions of C:G>T:A mutations in the control were specific to CpG sites. These proportions declined with increasing doses of BaP in favour of C:G>A:T mutations at CpG sites.

SMM-Seq recorded only SNV mutations for Project 2. BbF induced primarily C:T>A:G mutations (Supplementary data C: Figure 3A). *spectra_comparison()* revealed that the BbF subtype proportions were significantly different from the controls at the top three doses ($p < 0.001$) at the 6-base resolution. Thus, despite not inducing a significant increase in MFmin, BbF causes significant shifts in the mutation spectrum at the medium dose. Specifically, we observed dose-dependent increases in the proportions of C:G>A:T mutations. Following clustering, distinct sample groups were made between the high, medium-high, and medium dose groups versus the control, low and medium-low dose groups (Supplementary data C: Figure 3B). Finally, trinucleotide plots increases in the C:G>A:T mutations within the CCA, CCC, and CCT context at the medium-high and high-dose groups (Supplementary data C: Figure 4 shows control and high dose results).

C.5.5 Signature Assignment

We next compared the 96-base spectra to the SBS signatures of the COSMIC database using *signature_fitting()*. The cosine similarity between the reconstructed profile and the spectra was high (>0.9) for all dose groups with the exception of BbF medium-low dose (0.855). Within the BaP groups, we observed a dose-dependent increase in the contribution of SBS4, a signature found in lung cancers attributed to tobacco smoking. BaP is prevalent in tobacco smoke (Adesina et al., 2022); thus, we expect its mutation spectra to be associated with tobacco-smoke mutational signatures. Additional signatures with minor association with our BaP groups were SBS1 and SBS5, each associated with aging, and SBS94, whose etiology is unknown. Similarly, SBS 4 was the prevalent signature in the BbF medium-high and high dose groups at 63% and 66% respectively. BbF is also prevalent in tobacco smoke thus this result is aligned with expectations. Additional signatures found included SBS 5 and SBS 1, SBS 58 – a known sequencing artefact, and SBS 3, associated with Azathioprine treatment. Plots depicting the original mutation spectra, the reconstructed profile, the contributing signatures and the solution statistics were automatically generated for each group by *signature_fitting()* for BaP (Supplementary data C: Figure 5) and BbF (Supplementary data C: Figure 6, Medium-high and High dose groups shown).

C.5.6 Region Analysis

Analyzing the individual responses of the 20 DS targets provides clues on inter-locus variability in mutation susceptibility. These regions were supplied to *import_mut_data()*, which annotated sites with their targets based on position. MutSeqR then calculated the per-sample MF for each target with *calculate_mf()*. This was then supplied to *model_mf()*, which ran a GLMM that evaluated the effects of target and dose on the MFmin. Residuals were assessed (Supplementary data C: Figure 7). We specified comparisons between the BaP doses and the control for each target, revealing a significant increase in MFmin in the high-dose group compared to the control for all targets. All targets except for those on chr3 and chr15 were significantly increased at the medium dose, and eight of the 20 targets were significantly increased at the low dose. The model-estimated mean MFmin was plotted using *plot_model_mf()* (Appendix Figure C7A). This function can plot model results with up to two fixed effects. Here, we used target as the x-axis and dose as the colour. The x-axis order was defined as targets ascending by the MFmin value at the high dose. Significance labels were added based on the specified comparisons. Our visualization of the data revealed that MFmin varied extensively between targets, with this effect exacerbated by dose. Within the control, there was a 3-fold difference between the target with the highest MF (chr11) and the one with the lowest (chr19). Within the high-dose group, there was a 5-fold difference in MFmin between the target with the highest MF (chr11) and the one with the lowest MF (chr3). Finally, we plotted the mean MFmin per target for each dose using *plot_radar()*. Targets were organized based on their genic context. This showed a trend for higher MFmin in intergenic targets compared to genic targets at all dose groups (Appendix Figure C7B).

C.5.7 Visualization of Multiplets

In Project 1, we observed considerable inflation of the MFmax compared to the MFmin. However, from an initial analysis, it was unclear whether the increased MFmax was a result of a few highly recurrent mutations or several moderately recurrent mutations. We thus used *plot_bubbles()* to visualize the *alt_depth* of every mutation called for each dose (Appendix Figure C8). Bubble plots allow us to analyze the distribution and density of multiplet mutations. Each mutation is represented by a bubble whose size is scaled on the *alt_depth*. Thus, a large multiplet (i.e. high *alt_depth*) has a large bubble. This plot makes it easy to determine if MFmax is largely

driven by a few large multiplets or several moderate multiplets. Within the control, we observed a slight difference between MFmin and MFmax, driven by 3 relatively large multiplets: two C:G>T:A mutations and one deletion. In contrast, within the BaP dose groups, we observed that multiplets seemed to be evenly distributed, with no single mutation driving up the MFmax.

C.6 Discussion

Herein, we demonstrate the utility of MutSeqR for analyzing ECS data in the assessment of chemical mutagenicity. MutSeqR successfully recapitulates published findings (LeBlanc et al., 2022), replacing previously used *ad hoc* scripts with a version-controlled, function-based workflow. Furthermore, we demonstrate that MutSeqR readily imports data from distinct platforms, illustrating this feature with DS and SMM-seq data.

MutSeqR provides a comprehensive, user-friendly, and reproducible framework for analyzing ECS data in mutagenicity studies, making it a valuable tool to the regulatory community. It offers practical variant filtering methods to ensure data quality, providing a strong foundation for downstream analyses. We provide functionality for statistical testing between experimental groups and BMD derivation for a given mutagen. In a regulatory use-case, these features deliver quantitative data on a substance's mutagenic ability, directly informing risk assessment decisions. The package also supports mutation spectrum analysis and signature assignment, offering insight into a mutagen's mechanism of action and potentially linking mutagen exposure with established human health effects like cancer. Furthermore, data visualization facilitates clear communication and transparent reporting, allowing regulators to interpret complex findings efficiently and justify their decisions to stakeholders with confidence. Finally, to enhance reproducibility, an RMarkdown vignette guides users through standard workflows and serves as a template for consistent and verifiable data analysis. We used preliminary, less formalized versions of this code in our previous ECS studies (Cho et al., 2023; Dodge et al., 2023; Esina et al., 2024; LeBlanc et al., 2022; Schuster et al., 2024). MutSeqR formalizes these workflows, making them accessible, version-controlled, and easily adapted to any ECS project, thereby strengthening the workflow for regulatory genotoxicity evaluations.

Outside of the use cases considered here, it is important to recognize that MutSeqR's utility can extend beyond mutagenicity assessment. ECS methods are rapidly being developed to study a broad range of genetic toxicology endpoints, not just MF and spectra. For instance, applications in carcinogenesis have begun to rely on ECS to track the clonal expansion of cells carrying mutated cancer driver genes. MutSeqR provides users with the ability to analyze clonal mutations using the Maximum Mutation count and the ability to visualize such mutations using *plot_bubbles()*.

Given the increased use of ECS in genetic toxicology (Dodge et al., 2023; LeBlanc et al., 2022; Schuster et al., 2024; Smith-Roe et al., 2023; Valentine et al., 2020), the statistical techniques captured within and simplified by MutSeqR provide a valuable resource for efficient and reproducible analysis of mutation data. Thanks to its compatibility across ECS platforms, MutSeqR will help standardize the bioinformatic workflow involved with characterizing mutations using ECS and act as a guide to help unite the protocols for data processing of the many different ECS strategies. Ultimately, our vision is that this package informs a standardized approach to ECS data processing and analysis that can be applied within a regulatory and research context. Future development will focus on integrating the fastquorum pipeline for upstream variant calling (based on fgbio; https://nf-co.re/fastquorum/) and expanding the available mutational signatures by incorporating databases such as SIGNAL. We also plan to implement additional exploratory analyses, such as transcriptional strand bias tests and screening for extra-chromosomal circular DNA. MutSeqR will remain openly available on GitHub and will continue to evolve through transparent collaboration between maintainers and the broader community to address future challenges in ECS data analysis.

C.7 Acknowledgements

We thank Stephanie Smith-Roe and Xinwen Zhang for their assistance in package testing. We further extend our thanks to Geromino Matteo for reviewing the source code.

C.8 Funding Information

This work was supported by Health Canada's Genomics Research and Development Initiative. CLY is thankful for support from the Canada Research Chairs Program (CRC-2020-00060). CLY and FM are grateful for support from a Burroughs Wellcome Fund Innovation in Regulatory Science Award. ES received stipend support from the Natural Sciences and Engineering Research Council of Canada.

C.9 Conflict of Interest

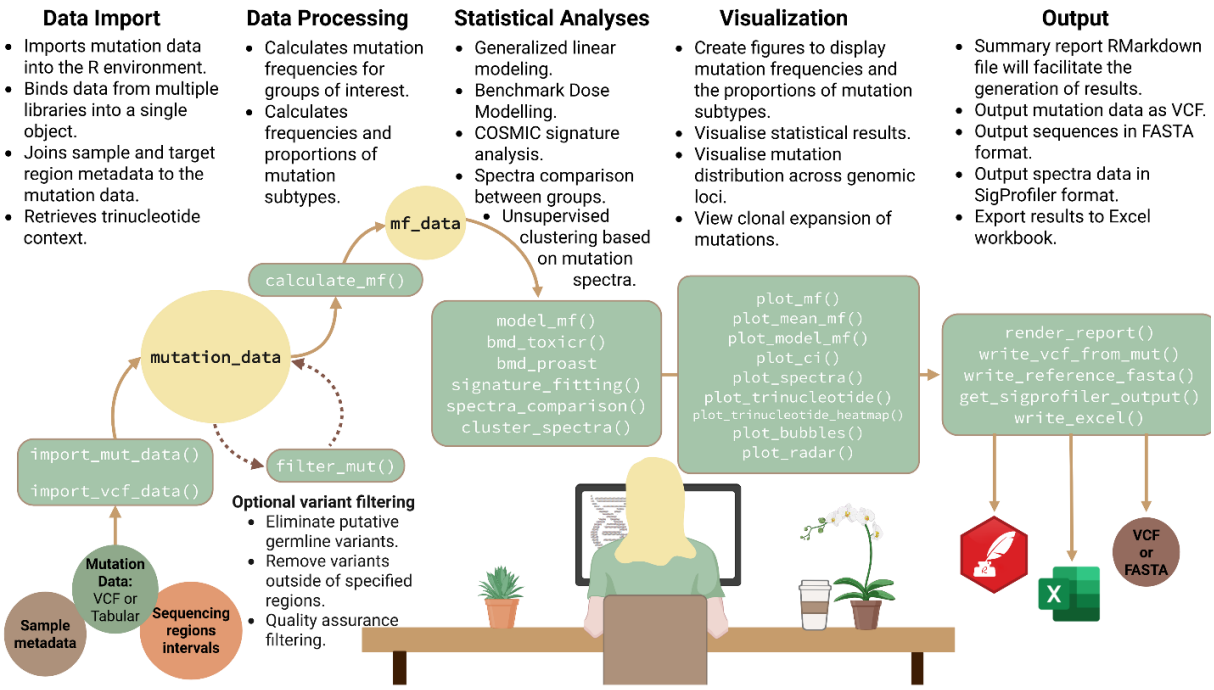
J.J.S. and C.C.V. are shareholders in TwinStrand BioSciences and are inventors on a patent for Duplex Sequencing (US20210269873A1). A.Y.M. is cofounder of MutagenTech and an inventor on a patent pending related to this work filed by Albert Einstein College of Medicine (U.S. provisional application no. 63/277,955, filed 10 November 2021). C.B. is CEO and cofounder of MutagenTech and Matter Bio.

C.10 Figures and Tables

Supplementary data C of Appendix C are available at *Bioinformatics Advances* online.

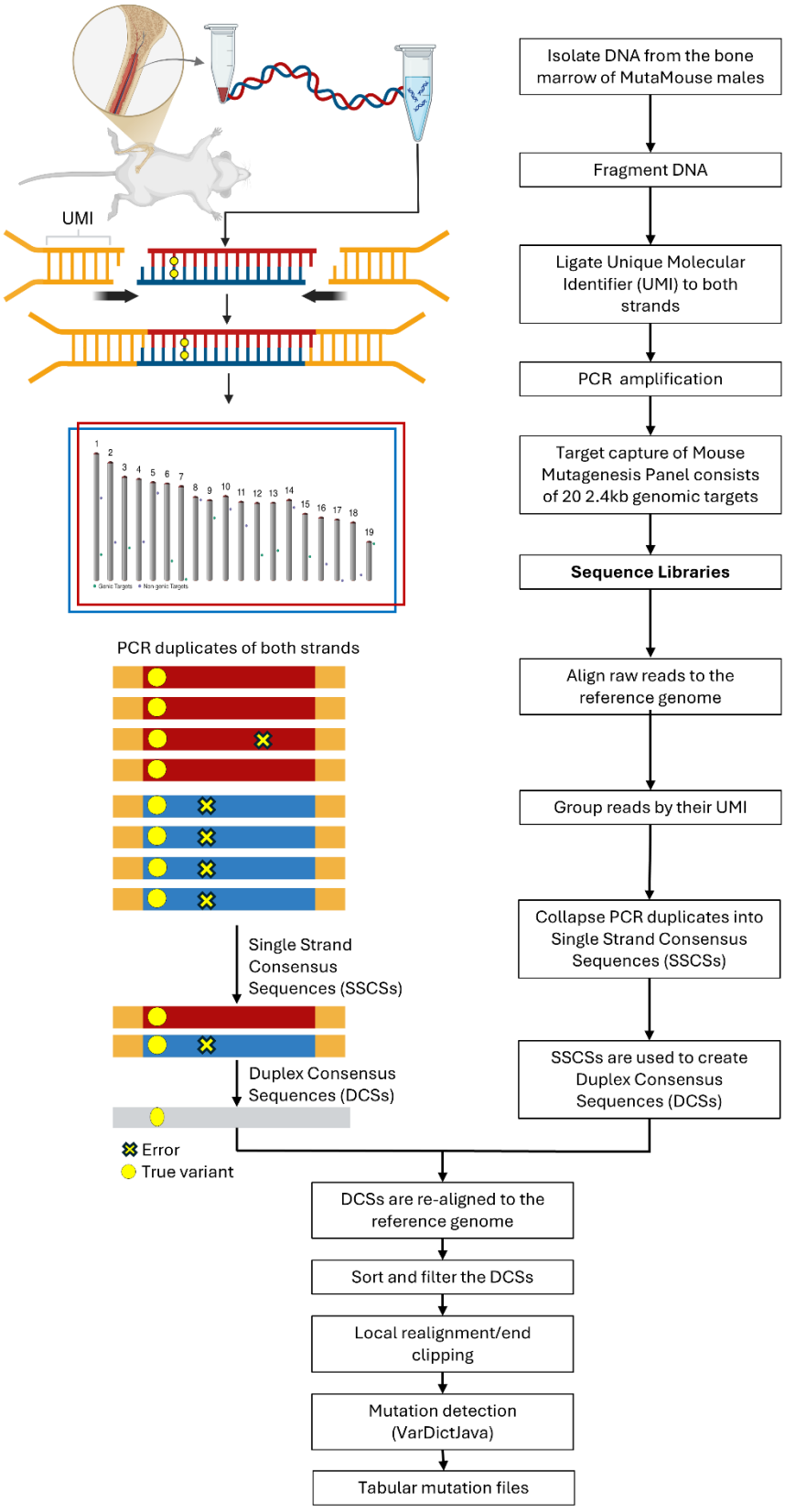
Appendix Table C 1: Required fields for mutation data import

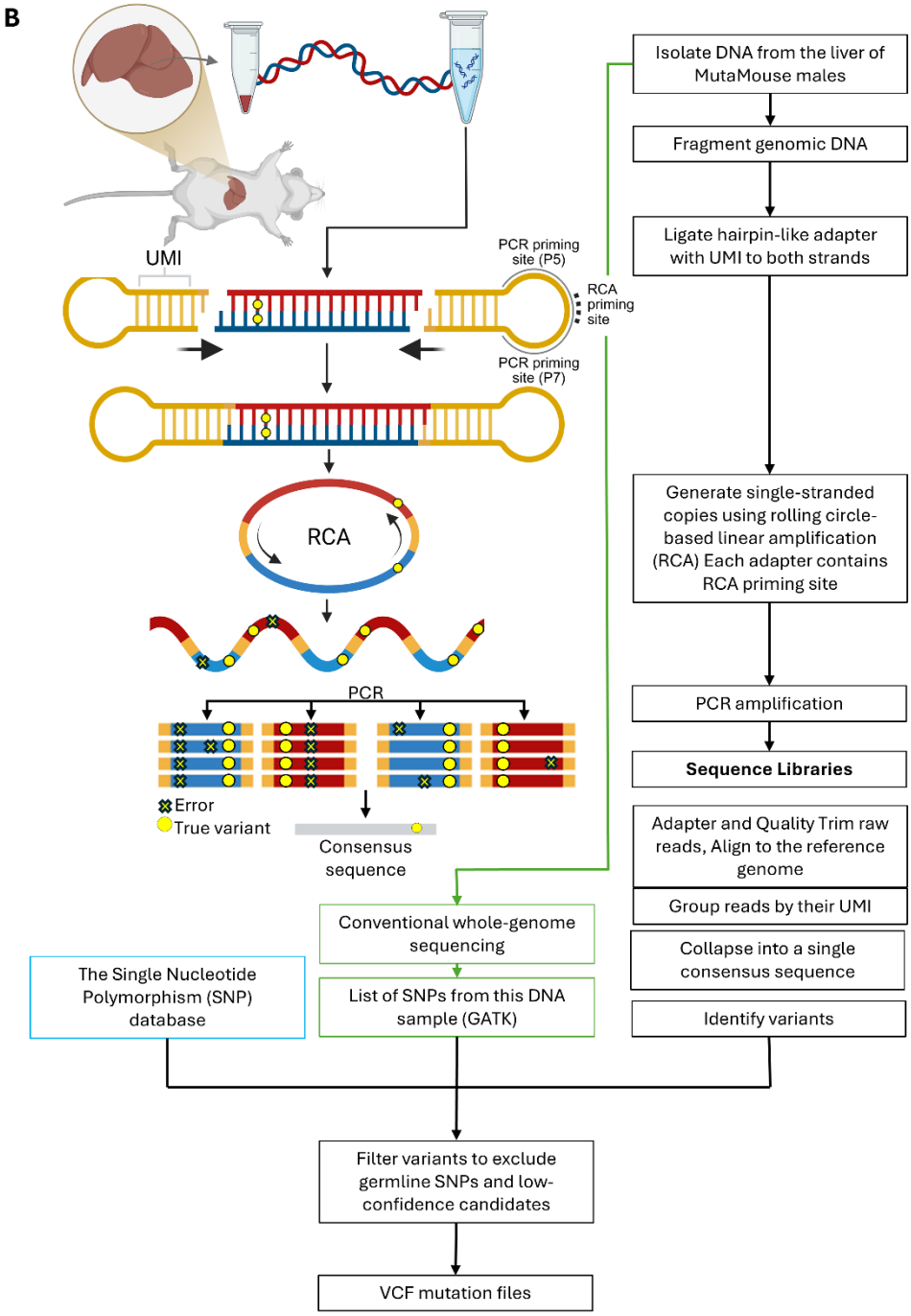
FIELD	VCF Specification	Definition
contig	CHROM	The name of the reference sequence.
start	POS	The start position of the feature. 0-based coordinates are accepted but will be changed to 1-based during import.
end	INFO(END)	The half-open end position of the feature in contig.
ref	REF	The reference allele at this position.
alt	ALT	The left-aligned, normalized, alternate allele at this position.
sample	INFO(sample) | Genotype Header	A unique identifier for the sample library. For VCF files, this field may be provided in either the INFO field or as the header to the GENOTYPE field.
Optional Fields (recommended)		
alt_depth	VD	The read depth that supports the alternate allele. If not included, the function will assume an alt_depth of 1 at variant sites.
total_depth | depth	AD | DP	The total read depth at this position. This field is either “total_depth” which excludes No-calls (Ns), or “depth”, which includes No-calls, if “total_depth” is not available. For VCF files, the total_depth is calculated as the sum of AD.



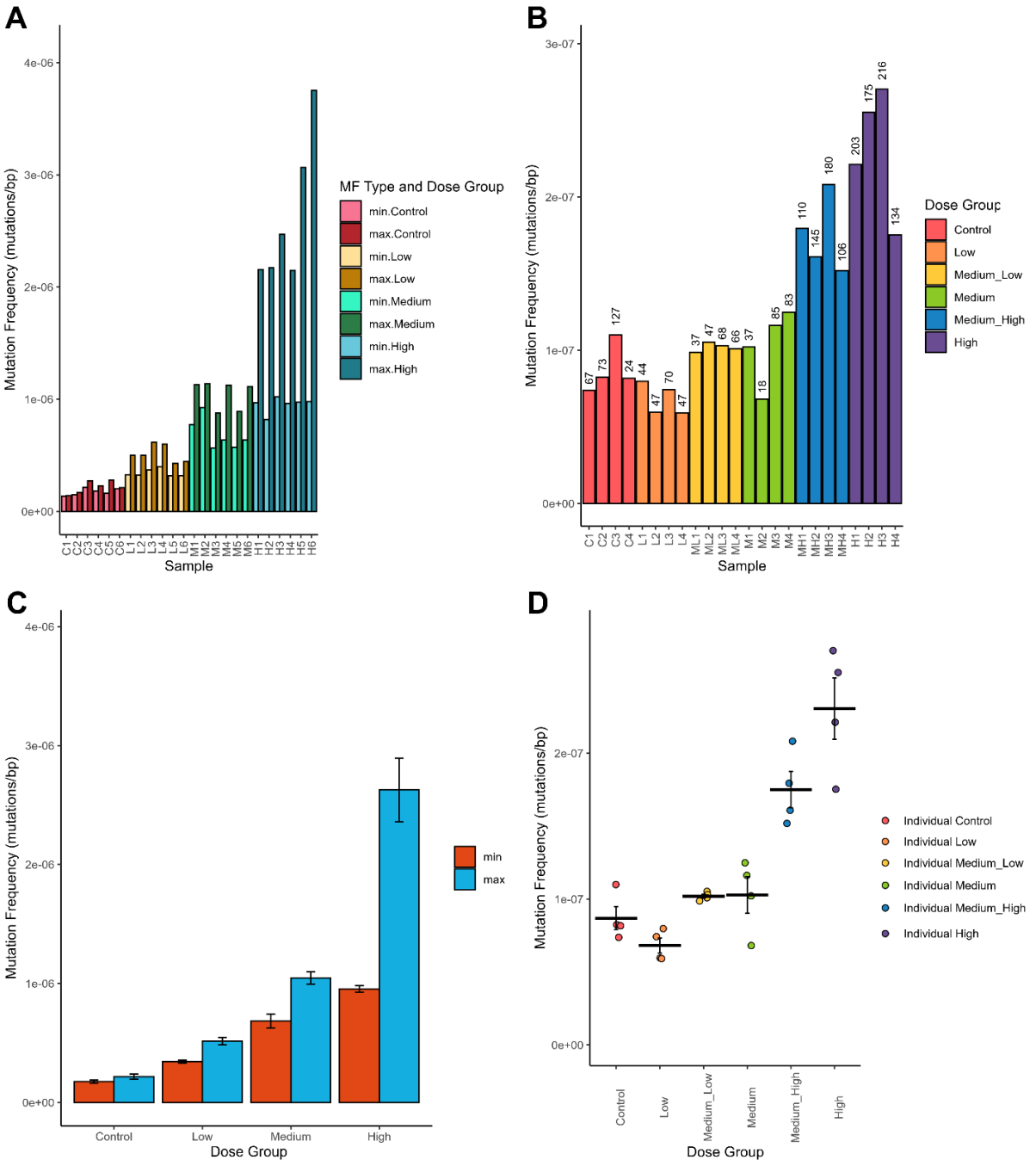
Appendix Figure C1: Overview of MutSeqR utilities. Created using Biorender.

A

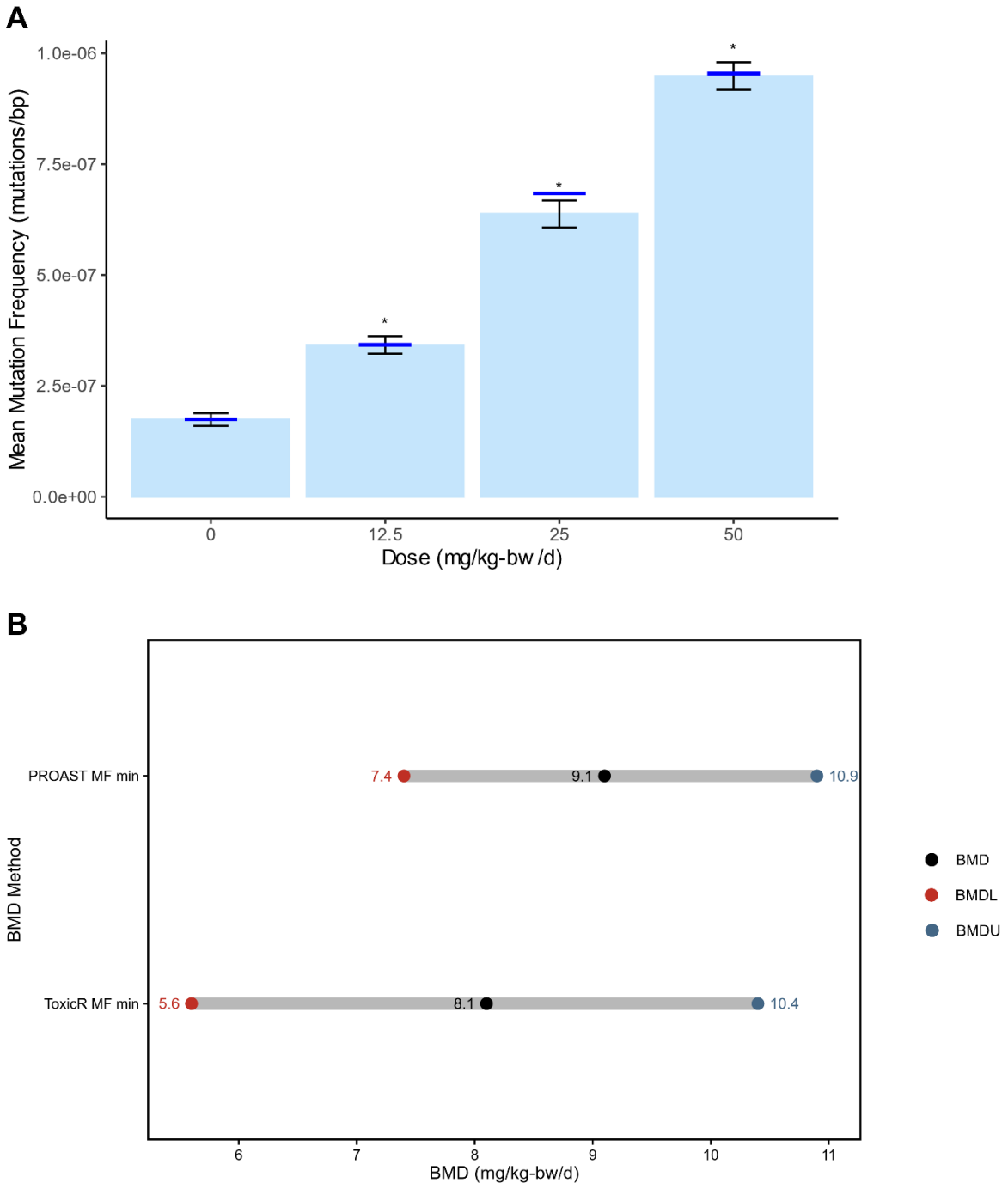

UMI
PCR duplicates of both strands
Single Strand Consensus Sequences (SSCSs)
Duplex Consensus Sequences (DCSs)
Error
True variant
Isolate DNA from the bone marrow of MutaMouse males
Fragment DNA
Ligate Unique Molecular Identifier (UMI) to both strands
PCR amplification
Target capture of Mouse Mutagenesis Panel consists of 20 2.4kb genomic targets
Sequence Libraries
Align raw reads to the reference genome
Group reads by their UMI
Collapse PCR duplicates into Single Strand Consensus Sequences (SSCSs)
SSCSs are used to create Duplex Consensus Sequences (DCSs)
DCSs are re-aligned to the reference genome
Sort and filter the DCSs
Local realignment/end clipping
Mutation detection (VarDictJava)
Tabular mutation files



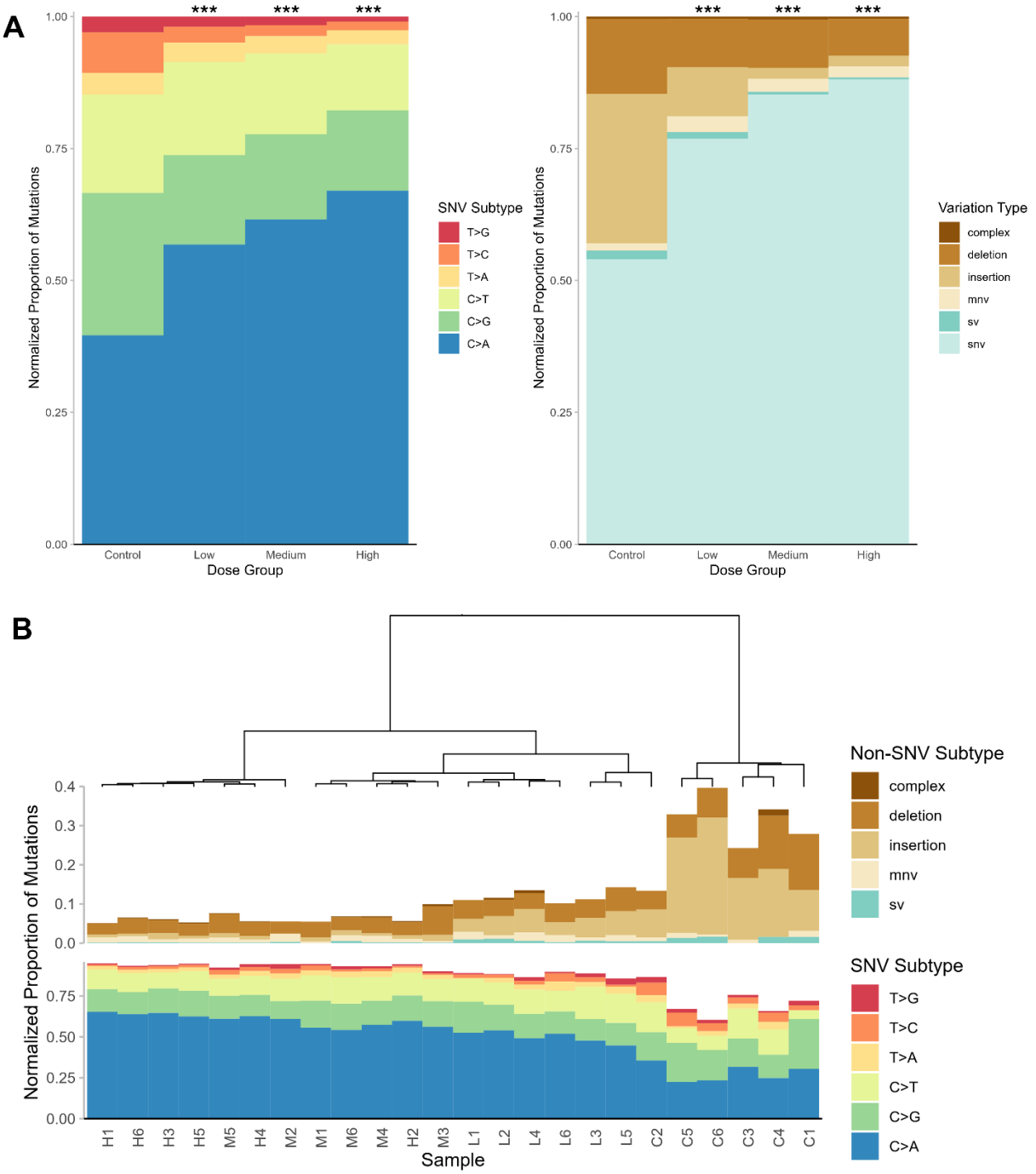
Appendix Figure C2: Overview of Duplex Sequencing, adapted from Kennedy et al. 2014. B) Overview of SMM-seq, adapted from Maslov et al., 2022



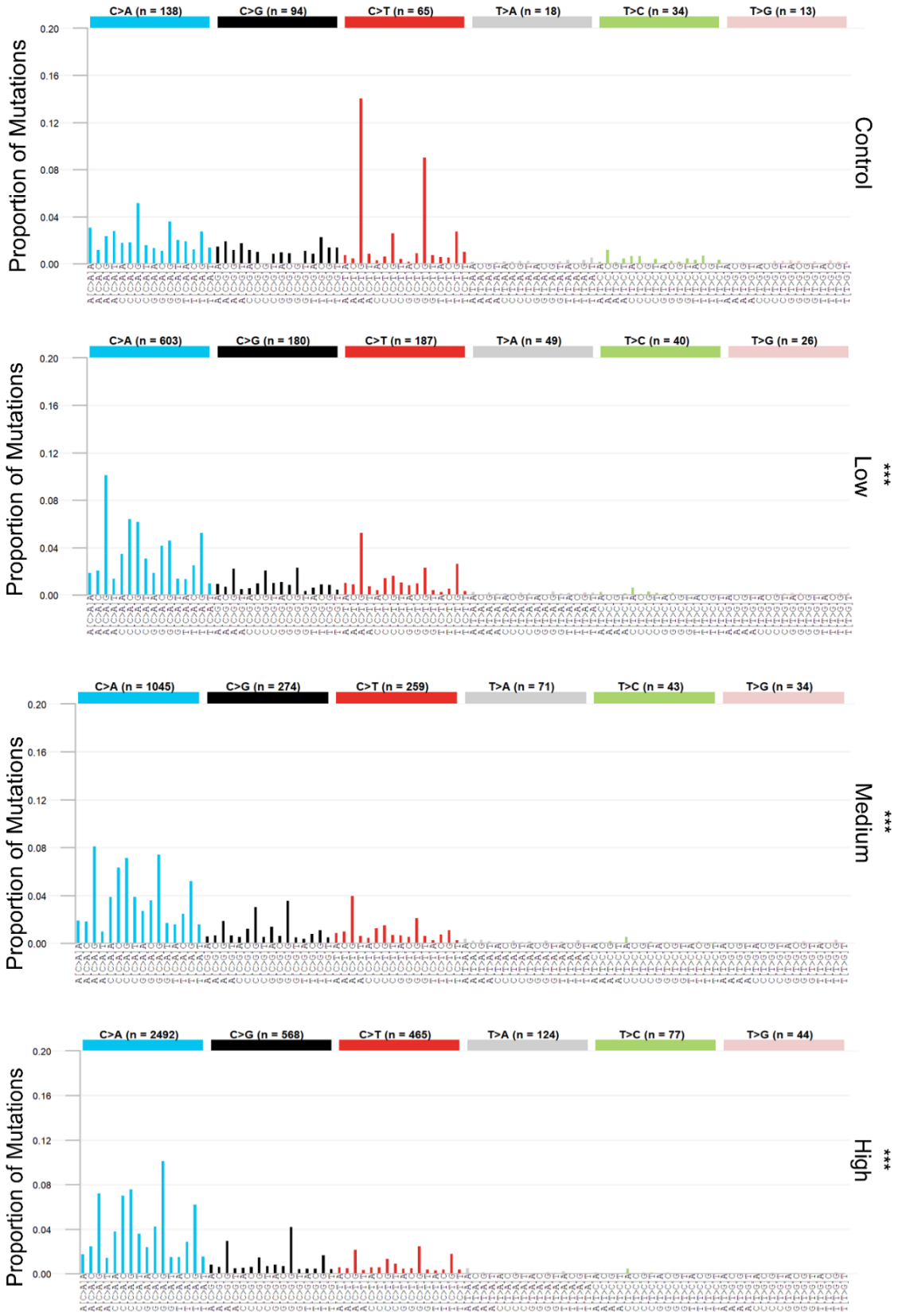
Appendix Figure C3: A) Duplex Sequencing mutation frequencies (MF) in the bone marrow of MutaMouse males at various doses of BaP. Bars represent the MF (mutations per bp) for each animal. Bars are coloured based on BaP dose group and MF type (Min or Max). Plotted using plot_mf(). B) SMM-seq MF in MutaMouse males' liver at various BbF doses. Data labels indicate the total mutation count per animal. C) Mean Duplex Sequencing MF per BaP dose group. Bars are coloured based on MF type. Error bars reflect the SEM. Plotted using plot_mean_mf(). D) Mean SMM-seq MF per BbF dose group. Lines represent the mean MF while the points represent individual MF values per animal and are coloured based on BbF dose group. Error bars represent the SEM. Plotted with plot_mean_mf().



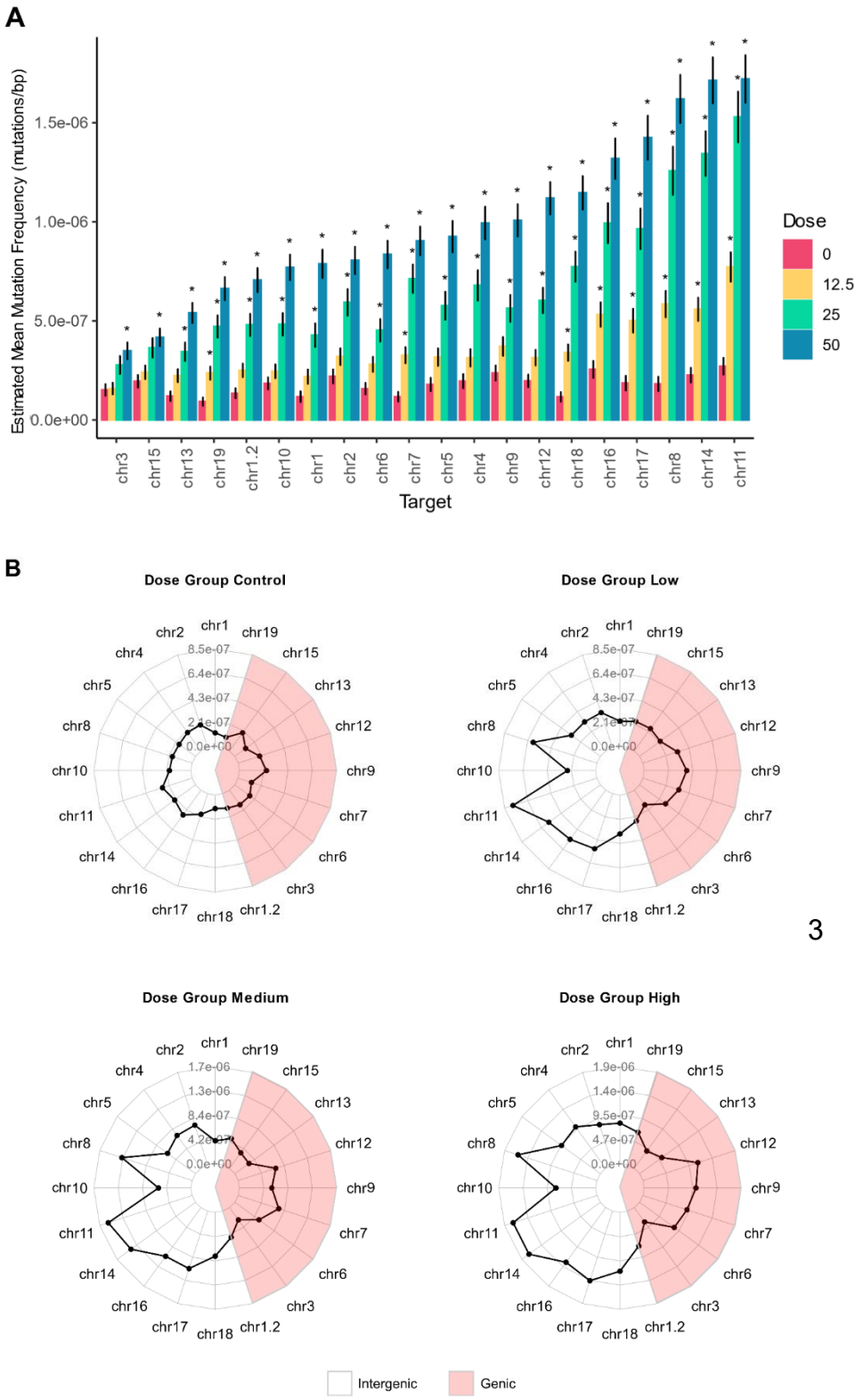
Appendix Figure C4: A) GLM results. Light blue bars represent the model-estimated mean mutation frequency minimum (MFmin) for BaP dose groups. Error bars represent the estimated SEM. Asterisk indicates a significant increase in MFmin relative to control (model_mf() GLM, $p < 0.05$). Plotted using plot_model_mf(). Dark blue bars representing the empirical mean MFmin were added using ggplot2. B) Benchmark dose with 90% confidence intervals representing the dose at which a 50% increase in MFmin from controls occurs for BaP, calculated using bmd_proast() or bmd_toxicr(). Black points represent the BMD, red points represent the BMDL, and blue points represent the BMDU. Plotted using plot_ci().



Appendix Figure C5: A) Mutation spectrum of control and BaP dose groups measured by Duplex Sequencing in the bone marrow of MutaMouse males. Normalized proportions of mutation subtypes are represented by colour within the stacked bar for each BaP dose group. Left is the proportion of SNV subtypes out of total SNVs, right is the proportion of variation types out of total mutations. For each panel, asterisks indicate a significant difference in the mutation spectra compared to the control (spectra_comparison(); modified contingency table, p<0.05). Plotted using plot_spectra(), asterisks added separately. B) Mutation spectrum of individual animals measured by Duplex Sequencing. Animals are clustered into groups based on the Euclidean distance between their subtype proportions using cluster_spectra(). Proportions are SNV subtypes (bottom) and non-SNVs (top) by total mutations. All proportions are normalized to their given reference depth. Sample identifiers (x-axis) denote dose group: C – Control, L – Low, M – Medium, H- High. Plotted with plot_spectra.

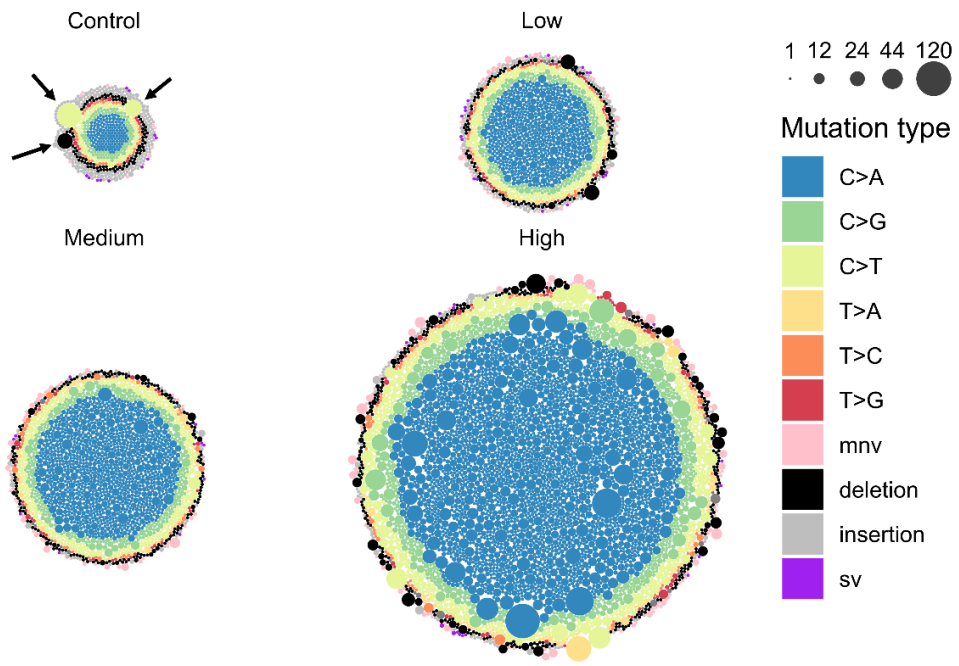


Appendix Figure C6 Normalized proportion of mutation subtypes within their 96-trinucleotide context for control and BaP dose groups, measured by Duplex Sequencing. Bars are coloured based on the normalized SNV subtype. Data labels represent the number of mutations for each normalized SNV subtype within that dose group. Asterisks indicate a significant difference in the trinucleotide mutation spectra relative to the control (spectra_comparison(); modified contingency table, p<0.05). Plotted using plot_trinucleotide(); dose group labels and asterisks were added separately.



3

Appendix Figure C7: A) Spontaneous and BaP-induced MFmin by Duplex Sequencing target, ordered from lowest MFmin in the high BaP dose group to the highest. Data are estimated mean MFmin ± SEM mutation per bp, for each target, separated and coloured by dose group. Asterisks indicate a significant increase in MFmin compared to the control (model_mf(); GLMM, p<0.05). Plotted using plot_model_mf(). B) Mean MFmin per Duplex Sequencing target, facetted by dose group. Duplex Sequencing targets are organized based on genic status; targets within genic regions are highlighted in red, targets within intergenic regions are in white. Plotted with plot_radar(). Red shading and genic status legend added separately.



Appendix Figure C8: Multiplet mutations measured by Duplex Sequencing for control and BaP-treated groups in MutaMouse bone marrow. Each circle represents a mutation, coloured by mutation subtype. The size of the circle is scaled by the mutation's alternative depth. Arrows indicate three multiplets as examples in the Control group. Plotted using plot_bubbles().

Appendix D

Appendix D1: Chapter 2

Appendix D1 Table 1: Peripheral blood micronucleus assay conducted 2 days after 28 days of oral exposure to BbF.

BbF Dose (mg/kg/day)	n	MN-RET[1]/1000 ± SEM	Adj. *p* value[2]	MN-RBC[3]/1000 ± SEM	Adj. *p* value[2]	%RET ± SEM	Adj. *p* value[2]
0	8	2.34 ± 0.1	-	1.63 ± 0.1	-	1.60 ± 0.1	
6.25	8	4.16 ± 0.4	0.0036	2.56 ± 0.2	0.0014	1.81 ± 0.1	0.5763
12.5	8	5.31 ± 0.3	<0.0001	3.00 ± 0.1	<0.0001	1.87 ± 0.1	0.2245
25	8	5.26 ± 0.4	<0.0001	3.00 ± 0.2	<0.0001	2.02 ± 0.1	0.0146
50	8	6.52 ± 0.3	<0.0001	3.90 ± 0.1	<0.0001	2.24 ± 0.1	<0.0001
100	8	8.11 ± 0.4	<0.0001	4.84 ± 0.2	<0.0001	2.17 ± 0.1	0.0005

[1] Micronucleated reticulocytes.
[2] Pairwise comparison to 0 mg/kg/day BbF using the Bonferroni multiple comparison test, significant if $p < 0.05$.
[3] Micronucleated red blood cells.

Appendix D1 Table 2: Technical performance metrics of the libraries built on DNA from BM and liver from animals after exposure to BbF

BbF Dose (mg/kg/day)	Animal Number	Raw Reads Passing Filter	Median Insert Size	Mean Duplex Depth	Peak Tag Family Size	Informative Duplex Bases	Raw Reads Passing Filter	Median Insert Size	Mean Duplex Depth	Peak Tag Family Size	Informative Duplex Bases
		Bone Marrow					Liver				
0	1	3.12^{+08}	299	15,550	14	1.17^{+09}	2.46^{+08}	252	16,764	13	1.04^{+09}
0	2	3.01^{+08}	269	17,271	11	1.34^{+09}	6.81^{+08}	267	17,433	40	1.07^{+09}
0	3	3.25^{+08}	285	15,948	14	1.22^{+09}	2.53^{+08}	276	19,600	11	1.18^{+09}
0	4	3.73^{+08}	249	18,109	13	1.47^{+09}	2.73^{+08}	259	19,867	12	1.23^{+09}
6.25	5	3.42^{+08}	292	13,438	19	1.00^{+09}	5.19^{+08}	268	25,014	20	1.53^{+09}
6.25	6	7.06^{+08}	425	10,718	63	7.06^{+08}	2.10^{+08}	248	16,886	9	1.05^{+09}
6.25	7	3.02^{+08}	311	12,297	19	9.08^{+08}	2.78^{+08}	276	16,524	18	9.60^{+08}
6.25	8	2.40^{+08}	278	14,593	9	1.11^{+09}	2.43^{+08}	261	18,443	10	1.13^{+09}
12.5	9	3.20^{+08}	271	19,131	10	1.47^{+09}	4.99^{+08}	261	26,398	19	1.63^{+09}
12.5	10	5.42^{+08}	308	14,064	29	1.07^{+09}	3.70^{+08}	248	23,489	13	1.49^{+09}
12.5	11	3.48^{+08}	276	18,558	12	1.43^{+09}	4.72^{+08}	261	23,011	20	1.44^{+09}
12.5	12	2.30^{+08}	284	13,785	10	1.04^{+09}	3.55^{+08}	272	21,224	16	1.28^{+09}
25	13	3.56^{+08}	276	19,780	11	1.52^{+09}	3.14^{+08}	268	21,397	13	1.30^{+09}
25	14	2.21^{+08}	279	14,154	9	1.06^{+09}	4.34^{+08}	286	22,006	20	1.33^{+09}
25	15	3.34^{+08}	319	11,071	22	8.31^{+08}	3.34^{+08}	267	17,970	18	1.10^{+09}
25	16	4.57^{+08}	220	19,177	15	1.62^{+09}	2.65^{+08}	202	15,613	9	1.10^{+09}
50	17	2.40^{+08}	273	14,280	11	1.10^{+09}	2.64^{+08}	268	21,146	10	1.28^{+09}
50	18	3.71^{+08}	251	19,331	12	1.55^{+09}	3.12^{+08}	255	22,785	12	1.39^{+09}
50	19	4.03^{+08}	294	13,726	21	1.04^{+09}	3.60^{+08}	254	21,148	16	1.31^{+09}
50	20	3.02^{+08}	268	16,705	12	1.27^{+09}	4.06^{+08}	275	19,002	21	1.17^{+09}
100	21	3.38^{+08}	230	16,505	9	1.39^{+09}	3.40^{+08}	286	19,003	17	1.15^{+09}
100	22	2.38^{+08}	266	13,861	11	1.06^{+09}	3.80^{+08}	253	24,706	14	1.54^{+09}
100	23	4.56^{+08}	282	21,661	14	1.67^{+09}	2.76^{+08}	236	19,811	11	1.26^{+09}
100	24	2.85^{+08}	271	16,844	10	1.30^{+09}	4.81^{+08}	249	21,772	21	1.39^{+09}

Appendix D1 Table 3: Minimum mutation frequencies by duplex sequencing using the mouse mutagenesis panel in bone marrow and liver samples after exposure to increasing doses of BbF.

BbF Dose (mg/kg/day)	n	MFmin[1] ± SEM ($\times 10^{-8}$)	Adj. *p* value[2]	MFmin Fold-change[3]
		Bone Marrow		
0	4	4.13 ± 0.45	-	-
6.25	4	5.65 ± 0.59	0.6271	1.4
12.5	4	7.54 ± 0.58	0.0052	1.8
25	4	7.95 ± 0.60	0.0025	1.9
50	4	11.66 ± 0.72	<0.0001	2.8
100	4	14.78 ± 0.76	<0.0001	3.6
		Liver		
0	4	5.33 ± 1.01	-	-
6.25	4	5.88 ± 1.05	0.9807	1.1
12.5	4	8.06 ± 1.10	0.3567	1.5
25	4	11.12 ± 1.43	0.0378	2.1
50	4	22.23 ± 1.94	<0.0001	4.2
100	4	36.27 ± 2.49	<0.0001	6.8

[1] Minimum mutation frequency considering unique mutations.
[2] Pairwise comparison to 0 mg/kg/day BbF by GLM, significant if $p < 0.05$.
[3] Fold-change of minimum mutation frequency versus 0 mg/kg/day BbF.

Appendix D1 Table 4: Minimum and maximum mutation frequencies in bone marrow and liver samples from individual animals after exposure to increasing doses of BbF.

BbF Dose (mg/kg/day)	Animal	MFmin[1] ($\times 10^{-8}$)	MFmax[2] ($\times 10^{-8}$)	Fold-change[3]	MFmin[1] ($\times 10^{-8}$)	MFmax[2] ($\times 10^{-8}$)	Fold-change[3]
		Bone marrow			Liver		
0	1	4.51	5.24	1.2	6.32	6.75	1.1
0	2	3.87	4.10	1.1	6.61	12.0	1.8
0	3	4.04	4.16	1.0	5.01	5.75	1.1
0	4	4.09	4.51	1.1	3.39	3.39	1.0
6.25	5	4.51	5.36	1.2	6.72	7.87	1.2
6.25	6	9.29	33.3	3.6	5.70	8.26	1.4
6.25	7	4.15	4.30	1.0	4.60	4.75	1.0
6.25	8	4.63	5.69	1.2	6.51	6.51	1.0
12.5	9	6.72	7.53	1.1	8.38	8.65	1.0
12.5	10	9.49	16.8	1.8	8.74	10.9	1.2
12.5	11	7.38	8.20	1.1	7.11	7.53	1.1
12.5	12	6.58	7.14	1.1	8.02	8.92	1.1
25	13	6.38	7.06	1.1	9.97	10.5	1.1
25	14	8.20	9.98	1.2	11.1	12.9	1.2
25	15	8.34	8.85	1.1	11.4	11.9	1.0
25	16	8.88	9.77	1.1	12.0	12.0	1.0
50	17	9.63	11.2	1.2	22.1	22.3	1.0
50	18	12.0	13.3	1.1	29.0	32.0	1.1
50	19	14.4	16.6	1.2	15.4	15.5	1.0
50	20	10.6	11.7	1.1	22.4	23.1	1.0
100	21	16.6	20.5	1.2	31.5	33.9	1.1
100	22	15.9	19.1	1.2	48.5	58.0	1.2
100	23	13.1	15.4	1.2	32.1	33.2	1.0
100	24	13.5	15.5	1.2	33.0	33.9	1.0

[1] Minimum mutation frequency considering unique mutations.
[2] Maximum mutation frequency considering all detected mutations.
[3] Fold-change between MFmin and MFmax

Appendix D1 Table 5: Minimum frequencies per loci detected in bone marrow samples after exposure to increasing doses of BbF.

	0 mg/kg/day	6.25 mg/kg/day			12.5 mg/kg/day			25 mg/kg/day			50 mg/kg/day			100 mg/kg/day		
Location of locus	MFmin[1] ± SEM (x 10^{-8})	MFmin[1] ± SEM (x 10^{-8})	Fold-change[2]	Adj. *p* value[3]	MFmin[1] ± SEM (x 10^{-8})	Fold-change[2]	Adj. *p* value[3]	MFmin[1] ± SEM (x 10^{-8})	Fold-change[2]	Adj. *p* value[3]	MFmin[1] ± SEM (x 10^{-8})	Fold-change[2]	Adj. *p* value[3]	MFmin[1] ± SEM (x 10^{-8})	Fold-change[2]	Adj. *p* value[3]
Chr1	5.91 ± 1.69	15.51 ± 7.44	2.6	0.9986	9.07 ± 4.70	1.5	0.9994	5.14 ± 2.01	0.9	0.7675	7.18 ± 1.94	1.2	0.3633	9.18 ± 1.77	1.6	0.4251
Chr1.2	2.94 ± 1.12	2.31 ± 0.78	0.8	0.1287	2.46 ± 1.19	0.8	0.9299	5.02 ± 1.21	1.7	1.0000	7.46 ± 1.98	2.5	0.9920	6.84 ± 0.83	2.3	0.6902
Chr2	3.77 ± 1.63	5.72 ± 1.86	1.5	0.9564	2.35 ± 1.58	0.6	0.9809	4.16 ± 1.48	1.1	0.9983	3.64 ± 1.50	1.0	1.0000	11.54 ± 1.40	3.1	0.0742
Chr3	2.29 ± 0.98	3.17 ± 1.41	1.4	0.9980	5.80 ± 2.70	2.5	0.4562	7.27 ± 3.49	3.2	0.4201	6.00 ± 1.56	2.6	0.4289	3.80 ± 0.72	1.7	0.9053
Chr4	2.24 ± 0.87	7.07 ± 1.16	3.2	0.3101	4.97 ± 0.45	2.2	0.7340	5.85 ± 1.50	2.6	0.4198	13.75 ± 4.23	6.1	0.0069	13.28 ± 3.32	5.9	0.0090
Chr5	4.53 ± 1.93	4.45 ± 1.85	1.0	1.0000	13.26 ± 4.01	2.9	0.0863	7.37 ± 2.46	1.6	0.7349	13.73 ± 1.97	3.0	0.0421	13.21 ± 2.11	2.9	0.0500
Chr6	1.60 ± 0.98	3.77 ± 1.55	2.4	0.7973	2.85 ± 1.22	1.8	0.9495	4.31 ± 1.68	2.7	0.5109	8.06 ± 3.37	5.0	0.0654	8.29 ± 2.04	5.2	0.0436
Chr7	3.70 ± 1.29	3.22 ± 1.24	0.9	1.0000	1.51 ± 0.90	0.4	0.7279	5.09 ± 0.86	1.4	0.9386	7.80 ± 1.21	2.1	0.3573	8.60 ± 2.30	2.3	0.2542
Chr8	5.91 ± 1.29	10.13 ± 1.20	1.7	0.7182	12.41 ± 2.23	2.1	0.3124	14.86 ± 4.91	2.5	0.1643	28.10 ± 4.90	4.8	0.0002	34.73 ± 10.20	5.9	<0.0001
Chr9	3.92 ± 1.01	5.33 ± 1.84	1.4	0.9921	9.76 ± 4.34	2.5	0.1480	10.51 ± 4.60	2.7	0.2436	16.73 ± 3.36	4.3	0.0041	15.30 ± 0.98	3.9	0.0063
Chr10	1.09 ± 0.63	4.89 ± 1.66	4.5	0.2420	5.63 ± 1.40	5.2	0.1632	3.93 ± 1.46	3.6	0.4414	5.74 ± 1.33	5.3	0.1565	5.81 ± 1.68	5.3	0.1482
Chr11	8.94 ± 2.77	7.46 ± 1.25	0.8	0.9979	16.05 ± 1.77	1.8	0.2784	25.68 ± 5.33	2.9	0.0052	31.80 ± 2.76	3.6	0.0001	35.18 ± 6.74	3.9	<0.0001
Chr12	3.22 ± 1.01	2.12 ± 1.24	0.7	0.9872	4.78 ± 2.07	1.5	0.9858	5.15 ± 1.51	1.6	0.8300	9.50 ± 2.74	3.0	0.0796	7.48 ± 0.89	2.3	0.3355
Chr13	4.31 ± 1.94	5.45 ± 5.45	1.3	1.0000	5.90 ± 1.55	1.4	0.9614	4.61 ± 0.94	1.1	1.0000	4.18 ± 1.72	1.0	1.0000	4.41 ± 1.92	1.0	1.0000
Chr14	4.61 ± 1.34	4.55 ± 0.78	1.0	1.0000	14.40 ± 0.89	3.1	0.0308	17.88 ± 2.20	3.9	0.0047	20.73 ± 4.02	4.5	0.0009	37.50 ± 3.58	8.1	<0.0001
Chr15	1.57 ± 0.53	1.44 ± 0.85	0.9	1.0000	2.88 ± 0.80	1.8	0.9575	2.02 ± 0.76	1.3	0.9957	3.99 ± 1.24	2.5	0.6720	5.12 ± 1.88	3.3	0.3162
Chr16	10.00 ± 1.88	7.31 ± 1.70	0.7	0.9444	14.49 ± 3.40	1.4	0.7512	18.50 ± 1.54	1.9	0.2331	20.18 ± 2.70	2.0	0.1190	30.38 ± 6.63	3.0	0.0020
Chr17	6.23 ± 2.19	12.11 ± 2.72	1.9	0.4845	15.40 ± 2.14	2.5	0.0863	6.47 ± 1.40	1.0	1.0000	14.82 ± 3.87	2.4	0.1516	34.70 ± 3.78	5.6	<0.0001
Chr18	3.87 ± 1.68	5.34 ± 0.99	1.4	0.9899	7.78 ± 1.73	2.0	0.6284	6.34 ± 2.01	1.6	0.8538	5.70 ± 0.70	1.5	0.9361	14.53 ± 1.48	3.8	0.0108
Chr19	3.42 ± 0.70	3.92 ± 1.65	1.1	0.9999	3.16 ± 1.25	0.9	1.0000	3.64 ± 1.31	1.1	1.0000	9.96 ± 4.14	2.9	0.1563	9.08 ± 2.99	2.7	0.2403

[1] Minimum mutation frequency considering unique mutations within a respective locus.
[2] Fold-change of minimum mutation frequency versus 0 mg/kg/day BbF.
[3] Pairwise comparison to 0 mg/kg/day BbF of the respective locus by GLMM, significant if $p < 0.05$.

Appendix D1 Table 6: Minimum frequencies per loci detected in liver samples after exposure to increasing doses of BbF.

	0 mg/kg/day	6.25 mg/kg/day			12.5 mg/kg/day			25 mg/kg/day			50 mg/kg/day			100 mg/kg/day		
Location of locus	MFmin[1] ± SEM (x 10^{-8})	MFmin[1] ± SEM (x 10^{-8})	Fold-change[2]	Adj. *p* value[3]	MFmin[1] ± SEM (x 10^{-8})	Fold-change[2]	Adj. *p* value[3]	MFmin[1] ± SEM (x 10^{-8})	Fold-change[2]	Adj. *p* value[3]	MFmin[1] ± SEM (x 10^{-8})	Fold-change[2]	Adj. *p* value[3]	MFmin[1] ± SEM (x 10^{-8})	Fold-change[2]	Adj. *p* value[3]
Chr1	2.77 ± 1.20	3.96 ± 1.63	1.4	0.9412	8.72 ± 2.22	3.1	0.0679	12.09 ± 2.21	4.4	0.0147	27.15 ± 7.81	9.8	< 0.0001	41.48 ± 4.59	15.0	<0.0001
Chr1.2	9.80 ± 5.52	8.92 ± 1.54	0.9	1.0000	12.93 ± 3.08	1.3	0.7896	23.64 ± 5.74	2.4	0.0097	28.18 ± 5.86	2.9	0.0010	45.28 ± 6.47	4.6	<0.0001
Chr2	2.63 ± 1.03	2.42 ± 0.84	0.9	0.9996	6.87 ± 1.66	2.6	0.3570	4.78 ± 1.63	1.8	0.6473	17.62 ± 4.60	6.7	0.0010	30.70 ± 8.09	11.7	<0.0001
Chr3	3.97 ± 1.79	6.18 ± 1.28	1.6	0.9429	7.04 ± 1.25	1.8	0.9090	8.98 ± 1.89	2.3	0.5024	15.88 ± 2.79	4.0	0.0113	35.23 ± 8.91	8.9	<0.0001
Chr4	5.59 ± 2.76	4.74 ± 2.01	0.8	0.9993	7.01 ± 1.90	1.3	0.9089	15.13 ± 1.96	2.7	0.0282	26.93 ± 2.22	4.8	0.0001	39.43 ± 7.87	7.1	<0.0001
Chr5	6.49 ± 1.59	5.46 ± 1.86	0.8	0.9850	4.82 ± 1.66	0.7	0.9964	9.39 ± 0.42	1.4	0.7252	17.43 ± 3.19	2.7	0.0153	42.63 ± 5.68	6.6	<0.0001
Chr6	5.21 ± 2.90	2.00 ± 1.30	0.4	0.7553	4.84 ± 1.06	0.9	1.0000	7.86 ± 0.52	1.5	0.8318	15.32 ± 4.15	2.9	0.0269	32.38 ± 8.14	6.2	<0.0001
Chr7	2.26 ± 0.90	3.15 ± 1.19	1.4	0.9871	9.21 ± 1.12	4.1	0.0454	11.43 ± 2.15	5.1	0.0125	14.56 ± 4.26	6.4	0.0016	29.85 ± 7.51	13.2	<0.0001
Chr8	4.52 ± 1.51	4.30 ± 2.59	1.0	1.0000	7.09 ± 1.11	1.6	0.9638	7.72 ± 1.45	1.7	0.8537	27.38 ± 3.01	6.1	0.0001	39.25 ± 6.63	8.7	<0.0001
Chr9	6.71 ± 0.73	7.92 ± 3.11	1.2	0.9993	12.33 ± 2.56	1.8	0.2567	11.25 ± 2.00	1.7	0.4410	29.98 ± 4.61	4.5	< 0.0001	32.43 ± 4.93	4.8	<0.0001
Chr10	7.10 ± 1.18	5.30 ± 2.78	0.7	0.9913	3.28 ± 0.70	0.5	0.4860	11.61 ± 3.75	1.6	0.6971	11.09 ± 2.05	1.6	0.7661	24.75 ± 3.71	3.5	0.0014
Chr11	10.46 ± 4.19	8.65 ± 2.48	0.8	1.0000	6.57 ± 0.99	0.6	0.8447	18.65 ± 1.62	1.8	0.1140	29.75 ± 3.68	2.8	0.0005	37.95 ± 7.50	3.6	<0.0001
Chr12	2.12 ± 0.72	6.63 ± 1.52	3.1	0.5879	7.94 ± 1.25	3.7	0.2260	9.33 ± 2.56	4.4	0.0684	23.53 ± 5.03	11.1	0.0001	38.20 ± 4.77	18.0	<0.0001
Chr13	3.12 ± 1.49	3.07 ± 1.84	1.0	0.9999	3.81 ± 1.67	1.2	0.9844	8.84 ± 3.01	2.8	0.1332	9.67 ± 1.18	3.1	0.0789	17.00 ± 1.72	5.4	0.0013
Chr14	10.09 ± 2.96	9.40 ± 3.54	0.9	0.9879	9.59 ± 2.01	1.0	1.0000	14.90 ± 3.13	1.5	0.5698	29.23 ± 4.65	2.9	0.0011	63.35 ± 12.50	6.3	<0.0001
Chr15	3.16 ± 2.46	3.51 ± 2.13	1.1	0.9964	3.77 ± 1.00	1.2	0.9938	7.67 ± 2.81	2.4	0.3084	16.18 ± 1.36	5.1	0.0019	30.10 ± 4.28	9.5	<0.0001
Chr16	2.02 ± 0.68	6.87 ± 3.22	3.4	0.4963	12.29 ± 2.10	6.1	0.0225	12.86 ± 6.65	6.4	0.0496	24.50 ± 4.35	12.1	0.0001	38.85 ± 3.96	19.2	<0.0001
Chr17	4.43 ± 1.85	12.65 ± 2.01	2.9	0.1781	12.27 ± 3.76	2.8	0.1835	5.03 ± 1.85	1.1	0.9972	29.05 ± 4.35	6.6	< 0.0001	42.50 ± 5.09	9.6	<0.0001
Chr18	7.94 ± 3.88	8.18 ± 1.92	1.0	1.0000	11.42 ± 2.33	1.4	0.7965	12.58 ± 1.50	1.6	0.5288	33.00 ± 7.80	4.2	< 0.0001	44.88 ± 4.12	5.7	<0.0001
Chr19	6.78 ± 0.90	5.33 ± 1.17	0.8	0.9893	9.62 ± 2.38	1.4	0.7218	8.28 ± 2.36	1.2	0.9669	18.98 ± 3.27	2.8	0.0123	19.98 ± 1.78	2.9	0.0063

[1] Minimum mutation frequency considering unique mutations within a respective locus.
[2] Fold-change of minimum mutation frequency versus 0 mg/kg/day BbF.
[3] Pairwise comparison to 0 mg/kg/day BbF of the respective locus by GLMM, significant if $p < 0.05$.

Appendix D1 Table 7: Minimum frequencies of single base substitutions detected in bone marrow and liver samples after exposure to increasing doses of BbF.

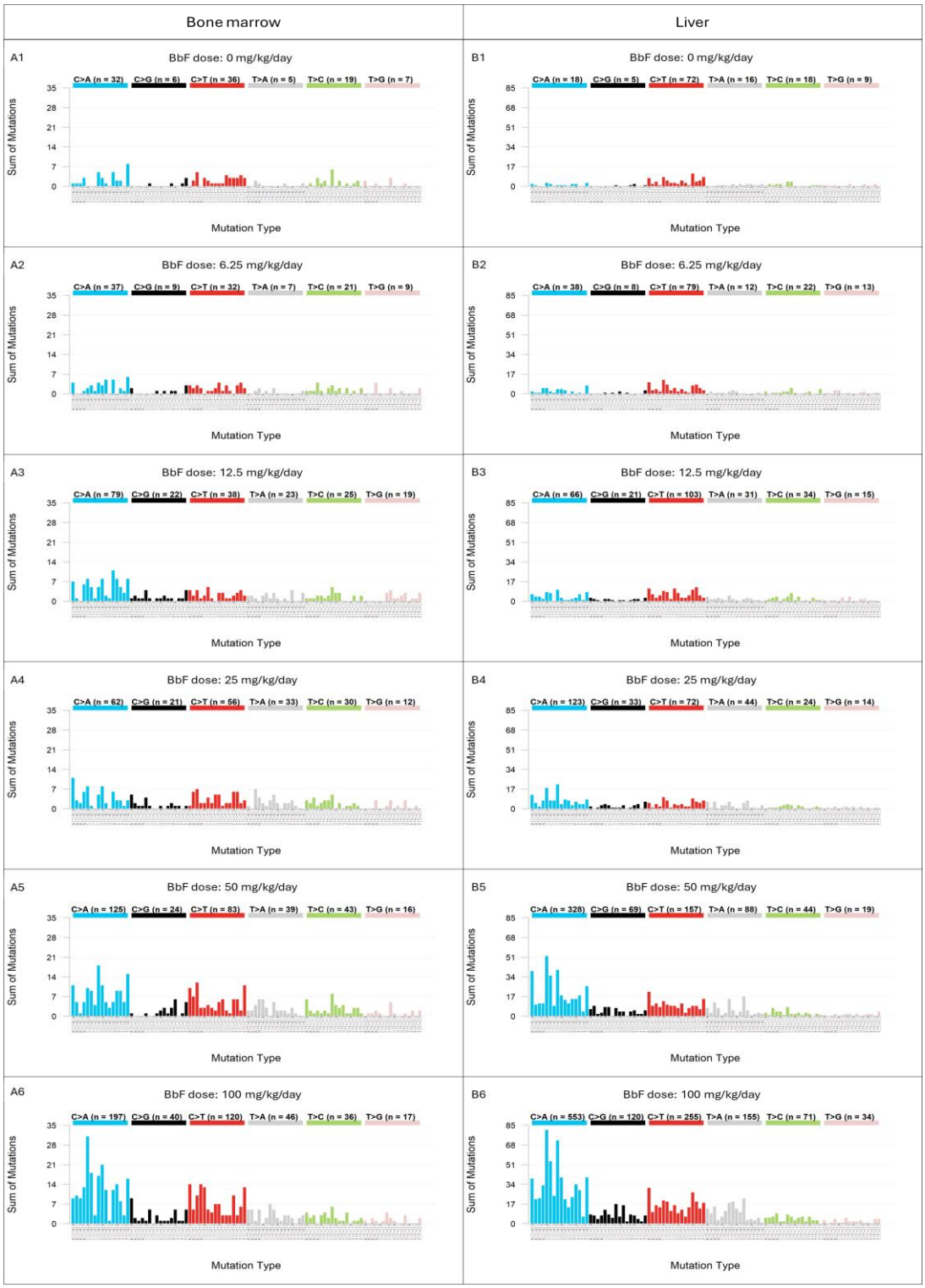
Mutation Subtype	0 mg/kg/day	6.25 mg/kg/day			12.5 mg/kg/day			25 mg/kg/day			50 mg/kg/day			100 mg/kg/day		
	MFmin[1] ± SEM (x 10^{-8})	MFmin[1] ± SEM (x 10^{-8})	Fold-change[2]	Adj. *p* value[3]	MFmin[1] ± SEM (x 10^{-8})	Fold-change[2]	Adj. *p* value[3]	MFmin[1] ± SEM (x 10^{-8})	Fold-change[2]	Adj. *p* value[3]	MFmin[1] ± SEM (x 10^{-8})	Fold-change[2]	Adj. *p* value[3]	MFmin[1] ± SEM (x 10^{-8})	Fold-change[2]	Adj. *p* value[3]
								Bone marrow								
C:G>A:T	2.08 ± 0.40	3.28 ± 0.89	1.6	0.7761	5.48 ± 0.76	2.6	0.0184	4.17 ± 0.32	2.0	0.0912	8.84 ± 1.32	4.3	0.0002	12.78 ± 1.41	6.1	<0.0001
C:G>G:C	0.39 ± 0.23	0.86 ± 0.45	2.2	0.0340	1.45 ± 0.16	3.7	0.0340	1.48 ± 0.16	3.8	0.0523	1.61 ± 0.17	4.1	0.0189	2.50 ± 0.29	6.4	0.0013
C:G>T:A	2.34 ± 0.24	2.79 ± 0.67	1.2	0.9935	2.59 ± 0.64	1.1	0.9998	3.90 ± 0.30	1.7	0.2643	5.63 ± 0.91	2.4	0.0037	7.84 ± 0.81	3.4	0.0001
T:A>A:T	0.26 ± 0.13	0.45 ± 0.10	1.7	0.8744	1.17 ± 0.16	4.5	0.0489	1.76 ± 0.42	6.8	0.0103	2.04 ± 0.16	7.8	0.0044	2.28 ± 0.56	8.8	0.0027
T:A>C:G	0.95 ± 0.15	1.44 ± 0.35	1.5	0.9263	1.24 ± 0.33	1.3	0.8778	1.65 ± 0.24	1.7	0.6433	2.15 ± 0.57	2.3	0.0439	1.75 ± 0.43	1.8	0.2853
T:A>G:C	0.35 ± 0.05	0.61 ± 0.15	1.7	0.5553	0.98 ± 0.11	2.8	0.0052	0.62 ± 0.09	1.8	0.2107	0.81 ± 0.17	2.3	0.0169	0.81 ± 0.11	2.3	0.0168
Insertion	0.12 ± 0.08	0.10 ± 0.06	0.8	1.0000	0.17 ± 0.07	1.4	0.9570	0.24 ± 0.09	2.0	0.9409	0.32 ± 0.08	2.7	0.3450	0.21 ± 0.07	1.8	0.7503
Deletion	0.95 ± 0.17	0.70 ± 0.17	0.7	0.8069	0.99 ± 0.05	1.0	0.9996	1.05 ± 0.08	1.1	0.9949	1.28 ± 0.22	1.3	0.7746	1.03 ± 0.26	1.1	0.9999
MNV[4]	0.06 ± 0.03	0.28 ± 0.14	4.6	0.5715	0.26 ± 0.18	4.3	0.6452	0.19 ± 0.08	3.2	0.5672	0.20 ± 0.05	3.3	0.6186	0.41 ± 0.04	6.8	0.1553
								Liver								
C:G>A:T	1.21 ± 0.29	2.53 ± 0.46	2.1	0.8117	3.85 ± 0.15	3.2	0.4030	9.03 ± 0.76	7.5	0.0319	21.38 ± 4.07	17.7	0.0009	35.48 ± 5.62	29.3	0.0001
C:G>G:C	0.38 ± 0.20	0.44 ± 0.20	1.2	1.0000	1.26 ± 0.33	3.3	0.4026	2.31 ± 0.69	6.1	0.0541	4.56 ± 0.79	12.0	0.0055	7.77 ± 0.94	20.4	0.0007
C:G>T:A	4.75 ± 0.57	4.88 ± 1.29	1.0	0.9910	5.53 ± 0.75	1.2	0.9015	5.34 ± 0.57	1.1	0.9964	10.41 ± 0.85	2.2	0.0043	16.53 ± 1.37	3.5	<0.0001
T:A>A:T	0.87 ± 0.18	0.75 ± 0.33	0.9	0.9617	1.36 ± 0.34	1.6	0.8657	2.47 ± 0.21	2.8	0.1372	4.42 ± 0.55	5.1	0.0034	7.52 ± 1.04	8.6	0.0034
T:A>C:G	1.05 ± 0.30	1.00 ± 0.38	1.0	0.9935	1.54 ± 0.11	1.5	0.6879	1.32 ± 0.09	1.3	0.9461	2.22 ± 0.33	2.1	0.0909	3.47 ± 0.51	3.3	0.0016
T:A>G:C	0.40 ± 0.17	0.65 ± 0.14	1.6	0.8476	0.62 ± 0.08	1.6	0.8725	0.80 ± 0.30	2.0	0.6951	0.96 ± 0.1	2.4	0.3025	1.66 ± 0.32	4.2	0.0137
Insertion	0.24 ± 0.07	0.12 ± 0.05	0.5	0.9145	0.36 ± 0.16	1.5	0.8505	0.23 ± 0.04	1.0	1.0000	0.25 ± 0.1	1.0	0.9986	0.45 ± 0.03	1.9	0.5266
Deletion	0.81 ± 0.27	0.67 ± 0.16	0.8	1.0000	0.76 ± 0.07	0.9	1.0000	0.67 ± 0.23	0.8	0.9999	1.00 ± 0.13	1.2	0.9244	1.65 ± 0.33	2.0	0.0609
MNV[4]	0.14 ± 0.14	0.15 ± 0.01	1.1	0.9975	0.20 ± 0.04	1.4	0.9325	0.39 ± 0.08	2.8	0.2172	0.68 ± 0.09	4.9	0.0298	1.06 ± 0.09	7.6	0.0039

[1] Minimum mutation frequency considering unique mutations.
[2] Fold-change of minimum mutation frequency versus 0 mg/kg/day BbF.
[3] Pairwise comparison to 0 mg/kg/day BbF by GLMM, significant if $p < 0.05$.
[4] Multi-nucleotide variants.

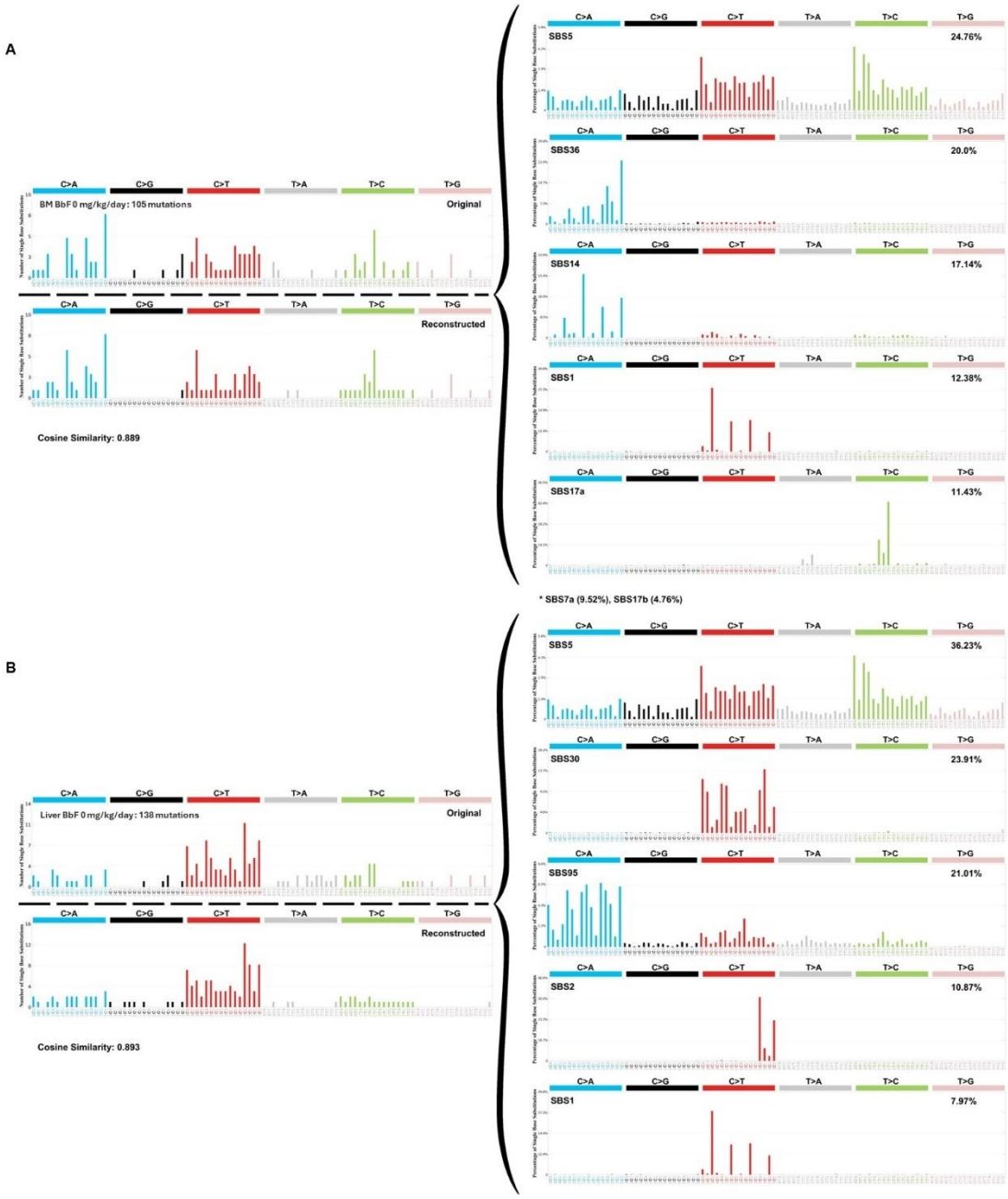
Appendix D1 Table 8: Benchmark doses of measured genotoxic endpoints in blood, bone marrow and liver

Genotoxic endpoint	n	CES[1]	BMDL[2]	BMDU[3]	Log10 BMDL	Log10 BMDU	Bootstrap runs
MN RET	8	0.5	0.408	4.06	-0.3893	0.6085	200
MN RBC	8	0.5	2.23	9.34	0.3483	0.9703	200
MF BM	4	0.5	3.6	12.8	0.5563	1.1072	200
MF Liver	4	0.5	10.1	18.4	1.0043	1.2648	200

[1]Critical effect size used to perform benchmark dose modeling.
[2]Benchmark dose lower confidence limit.
[3]Benchmark dose upper confidence limit.



Appendix D1 Figure 1: Trinucleotide mutation pattern of bone marrow (A) and liver (B) samples after 28-day exposure to the vehicle control (A1, B1) and increasing doses of BbF (A2 = 6.25, A3 = 12.5, A4 = 25, A5 = 50, A6 = 100, B2 = 6.25, B3 = 12.5, B4 = 25, B5 = 50 and B6 = 100 mg/kg/day). The panels represent the sum of mutations of each 96-mutational subtypes for n = 4 samples of each dose group. The sum of all mutation for each single base substitution type; C>A (blue), C>G (black), C>T (red), T>A (gray), T>C (green) and T>G (pink), are indicated.



Appendix D1 Figure 2: Mutation signature analyses for vehicle control in bone marrow (A) and liver (B). The original trinucleotide mutation profile is shown on the top left. SigProfilerAssignment used the single base substitution (SBS) signatures of the Catalogue of Somatic Mutations in Cancer (COSMIC) database to reconstruct the signatures. The SBS signatures and their relative contributions are shown on the right. The reconstructed mutation pattern, number of substitutions used for the assignment and cosine similarity between reconstructed and observed trinucleotide mutation profile are shown on left below the original profile. The total amount of mutations of the original signature are indicated on the left.

Appendix D2: Chapter 3

Appendix D2 Table 1: Body and liver weight

Exposure duration	BbF dose (mg/kg/day)	Initial body weight	Final body weight	%change in body weight	Mean % change per dose	p-value of mean % body weight (Dunnett's test)	Liver weight	Relative liver weigh	Mean liver weight	p-value of mean liver weight (Dunnett's test)	Mean relative liver weight	p-value of mean relative liver weight (Dunnett's test)
28-Days	0	33.4	35.4	5.99			1.40	3.96				
28-Days	0	27.8	30.4	9.35			1.27	4.19				
28-Days	0	27.8	28.6	2.88			1.16	4.04				
28-Days	0	31.2	35.4	13.46	9.56		1.44	4.07	1.38		4.07	
28-Days	0	26.2	27.2	3.82			1.02	3.75				
28-Days	0	28.4	35.0	23.24			1.46	4.18				
28-Days	0	30.0	32.0	6.67			1.45	4.53				
28-Days	0	39.8	44.2	11.06			1.71	3.87				
28-Days	6.25	32.8	34.6	5.49			1.49	4.30				
28-Days	6.25	32.0	36.8	15.00			1.41	3.83				
28-Days	6.25	32.0	36.0	12.50			1.41	3.92				
28-Days	6.25	32.2	35.4	9.94	7.90	0.60	1.44	4.06	1.36	1.00	4.06	0.93
28-Days	6.25	32.8	35.4	7.93			1.49	4.20				
28-Days	6.25	28.6	29.0	1.40			1.18	4.06				
28-Days	6.25	29.4	31.8	8.16			1.33	4.20				
28-Days	6.25	29.0	29.8	2.76			1.18	3.95				
28-Days	12.5	34.2	33.0	-3.51			1.55	4.68				
28-Days	12.5	24.0	26.4	10.00			1.17	4.44				
28-Days	12.5	30.4	31.6	3.95			1.25	3.95				
28-Days	12.5	38.0	38.2	0.53	3.68	0.30	1.72	4.50	1.41	1.00	4.36	0.12
28-Days	12.5	31.0	33.2	7.10			1.41	4.24				
28-Days	12.5	34.8	33.6	-3.45			1.41	4.20				
28-Days	12.5	28.6	29.8	4.20			1.44	4.84				
28-Days	12.5	30.0	33.2	10.67			1.33	4.00				
28-Days	25	30.4	30.4	0.00			1.30	4.28				
28-Days	25	27.4	31.0	13.14			1.36	4.39				
28-Days	25	29.6	32.8	10.81			1.42	4.33				
28-Days	25	27.2	27.4	0.74	5.29	0.60	1.23	4.49	1.37	1.00	4.41	0.02
28-Days	25	29.4	32.2	9.52			1.37	4.25				
28-Days	25	36.4	35.2	-3.30			1.56	4.44				
28-Days	25	28.8	29.4	2.08			1.27	4.31				
28-Days	25	30.0	32.8	9.33			1.58	4.81				
28-Days	50	34.2	35.2	2.92			1.63	4.62				
28-Days	50	26.4	26.6	0.76			1.15	4.31				
28-Days	50	30.0	34.4	14.67			1.49	4.34				
28-Days	50	28.4	29.8	4.93	5.93	0.60	1.34	4.49	1.48	1.00	4.49	<0.01
28-Days	50	29.0	32.0	10.34			1.49	4.66				
28-Days	50	33.6	35.8	6.55			1.69	4.73				
28-Days	50	28.8	28.0	-2.78			1.22	4.34				
28-Days	50	33.8	37.2	10.06			1.65	4.45				
28-Days	100	30.6	30.6	0.00			1.43	4.68				
28-Days	100	26.6	27.8	4.51			1.45	5.21				
28-Days	100	30.2	28.6	-5.30			1.54	5.38				
28-Days	100	28.8	28.0	-2.78	2.58	0.15	1.32	4.70	1.54	0.98	4.94	<0.01
28-Days	100	35.0	36.0	2.86			1.79	4.98				
28-Days	100	36.8	39.0	5.98			1.89	4.85				
28-Days	100	27.2	29.4	8.09			1.41	4.81				
28-Days	100	27.4	29.4	7.30			1.45	4.93				
90-Days	0	30.2	36.4	20.53			1.26	3.47				
90-Days	0	30.2	35.8	18.54			1.37	3.83				
90-Days	0	26.6	31.0	16.54			1.40	4.52				
90-Days	0	32.8	39.6	20.73	18.26		1.62	4.08	1.42		4.02	
90-Days	0	31.2	38.8	24.36			1.45	3.74				
90-Days	0	28.8	35.2	22.22			1.46	4.15				
90-Days	0	33.4	34.4	2.99			1.50	4.37				
90-Days	0	27.8	33.4	20.14			1.32	3.96				
90-Days	3.125	34.2	47.4	38.60			1.89	3.99				
90-Days	3.125	27.2	36.0	32.35			1.44	4.01				
90-Days	3.125	32.0	35.6	11.25			1.63	4.58				
90-Days	3.125	25.0	28.2	12.80	25.01	0.41	1.23	4.37	1.59	1.00	4.11	0.72
90-Days	3.125	31.8	43.8	37.74			1.70	3.88				
90-Days	3.125	29.2	34.0	16.44			1.39	4.09				
90-Days	3.125	37.8	47.6	25.93			1.82	3.83				
90-Days	6.25	26.2	30.0	14.50			1.20	4.00				
90-Days	6.25	32.2	48.6	50.93			1.88	3.87				
90-Days	6.25	32.4	37.8	16.67			1.41	3.74				
90-Days	6.25	34.6	40.4	16.76	24.58	0.41	1.35	3.33	1.40	1.00	3.82	0.74
90-Days	6.25	26.2	32.6	24.43			1.28	3.94				
90-Days	6.25	27.6	31.4	13.77			1.30	4.15				
90-Days	6.25	32.6	44.0	34.97			1.65	3.74				
90-Days	12.5	28.6	29.6	3.50			1.27	4.30				
90-Days	12.5	30.8	38.2	24.03			1.57	4.12				
90-Days	12.5	30.8	34.0	10.39			1.51	4.43				
90-Days	12.5	36.2	37.4	3.31	9.79	0.09	1.60	4.27	1.40	0.99	4.20	0.71
90-Days	12.5	27.4	28.4	3.65			1.31	4.60				
90-Days	12.5	29.6	33.8	14.19			1.34	3.96				
90-Days	12.5	32.6	37.2	14.11			1.42	3.82				
90-Days	12.5	27.4	28.8	5.11			1.19	4.12				
90-Days	25	30.8	33.0	7.14			1.46	4.42				
90-Days	25	29.2	31.0	6.16			1.53	4.93				
90-Days	25	28.0	34.2	22.14			1.53	4.47				
90-Days	25	30.8	31.6	2.60	4.60	0.01	1.48	4.68	1.41	1.00	4.45	0.05
90-Days	25	30.4	29.0	-4.61			1.23	4.24				
90-Days	25	28.4	29.0	2.11			1.22	4.20				
90-Days	25	32.6	32.2	-1.23			1.43	4.45				
90-Days	25	32.0	32.8	2.50			1.38	4.21				
90-Days	50	33.0	32.8	-0.61			1.38	4.22				
90-Days	50	30.8	29.0	-5.84			1.41	4.86				
90-Days	50	28.6	29.6	3.50			1.41	4.76				
90-Days	50	32.4	32.8	1.23			1.66	5.07				
90-Days	50	34.0	36.6	7.65	2.14	0.00	1.72	4.69	1.52	0.98	4.76	<0.01
90-Days	50	30.0	32.6	8.67			1.65	5.05				
90-Days	50	34.0	32.6	-4.12			1.51	4.63				
90-Days	50	27.2	29.0	6.62			1.40	4.82				
180-Days	0	32.2	42.8	32.92			1.76	4.11				
180-Days	0	29.6	46.6	57.43			1.62	3.47				
180-Days	0	31.4	50.2	59.87			2.05	4.08				
180-Days	0	32.0	47.2	47.50	43.77		1.71	3.63	1.71		3.79	
180-Days	0	32.4	46.4	43.21			1.72	3.71				
180-Days	0	34.4	42.8	24.42			1.55	3.61				
180-Days	0	29.8	43.0	44.30			1.73	4.01				
180-Days	0	30.6	43.0	40.52			1.58	3.68				
180-Days	1.56	30.6	40.6	32.68			1.50	3.69				
180-Days	1.56	30.2	47.0	55.63			1.79	3.80				
180-Days	1.56	27.6	40.8	47.83			1.58	3.88				
180-Days	1.56	29.6	49.8	68.24	48.04	1.00	1.78	3.57	1.77	0.98	4.06	0.54
180-Days	1.56	30.2	46.0	52.32			1.88	4.09				
180-Days	1.56	29.4	38.4	30.61			1.59	4.13				
180-Days	1.56	29.0	47.2	62.76			2.40	5.09				
180-Days	1.56	29.8	40.0	34.23			1.68	4.21				
180-Days	3.125	34.2	44.0	28.66			1.73	3.94				
180-Days	3.125	30.0	44.2	47.33			1.58	3.56				
180-Days	3.125	30.4	40.4	32.89			1.57	3.89				
180-Days	3.125	32.6	49.0	50.31	43.7	1.00	1.89	3.85	1.79	0.96	3.88	0.69
180-Days	3.125	33.6	51.0	51.79			2.00	3.92				
180-Days	3.125	32.2	48.0	49.07			2.18	4.54				
180-Days	3.125	31.4	45.8	45.86			1.58	3.45				
180-Days	6.26	31.4	43.0	36.94			1.79	4.15				
180-Days	6.26	29.8	50.0	67.79			2.33	4.66				
180-Days	6.26	29.4	40.6	38.10			1.51	3.71				
180-Days	6.26	31.6	42.2	33.54	41.93	1.00	1.43	3.38	1.73	0.99	3.94	0.68
180-Days	6.26	29.6	48.0	62.16			1.86	3.88				
180-Days	6.26	34.2	42.8	25.15			1.60	3.73				
180-Days	6.26	28.8	38.2	32.64			1.54	4.02				
180-Days	6.26	32.2	44.8	39.13			1.80	4.01				
180-Days	12.5	28.4	37.6	32.39			1.62	4.32				
180-Days	12.5	27.0	33.6	24.44			1.33	3.95				
180-Days	12.5	32.2	44.8	39.13			1.66	3.70				
180-Days	12.5	30.0	35.8	19.33	33.78	0.56	1.52	4.26	1.58	0.85	4.07	0.23
180-Days	12.5	30.2	47.4	56.95			1.87	3.95				
180-Days	12.5	29.8	44.4	48.99			1.65	3.72				
180-Days	12.5	28.6	36.8	28.67			1.68	4.56				
180-Days	12.5	25.6	30.8	20.31			1.27	4.14				
180-Days	25	31.8	39.0	22.64			1.67	4.29				
180-Days	25	32.8	46.0	40.24			2.08	4.52				
180-Days	25	27.8	34.0	22.30			1.57	4.61				
180-Days	25	29.2	38.4	31.51	26.48	0.03	1.51	3.92	1.61	0.92	4.30	<0.01
180-Days	25	28.6	35.4	23.78			1.51	4.27				
180-Days	25	29.0	34.8	20.00			1.35	3.89				
180-Days	25	28.0	38.6	37.86			1.64	4.25				
180-Days	25	29.6	33.6	13.51			1.56	4.64				

Appendix D2 Table 2: Histopathology of MutaMouse males after exposure to increasing dose of BbF.

BbF dose (mg/kg/day)	Exposure duration	Observation: focus of cellular alteration	Observation: necrosis, focal	Observation: rarefaction, cytoplasmic or glycogen accumulation increased
0	28-Days	WNL	WNL	WNL
0	28-Days	WNL	WNL	WNL
0	28-Days	WNL	WNL	WNL
0	28-Days	WNL	WNL	WNL
0	28-Days	WNL	WNL	WNL
0	28-Days	0	1	0
0	28-Days	WNL	WNL	WNL
0	28-Days	0	0	2, diffuse
6.25	28-Days	WNL	WNL	WNL
6.25	28-Days	0	1, multifocal with mixed cell infiltration	0
6.25	28-Days	WNL	WNL	WNL
6.25	28-Days	0	0	2, diffuse
6.25	28-Days	WNL	WNL	WNL
6.25	28-Days	WNL	WNL	WNL
6.25	28-Days	0	1	0
6.25	28-Days	WNL	WNL	WNL
12.5	28-Days	WNL	WNL	WNL
12.5	28-Days	WNL	WNL	WNL
12.5	28-Days	0	1, with mixed cell infiltration	0
12.5	28-Days	WNL	WNL	WNL
12.5	28-Days	WNL	WNL	WNL
12.5	28-Days	WNL	WNL	WNL
12.5	28-Days	WNL	WNL	WNL
12.5	28-Days	WNL	WNL	WNL
25	28-Days	WNL	WNL	WNL
25	28-Days	0	0	2, periportal
25	28-Days	WNL	WNL	WNL
25	28-Days	WNL	WNL	WNL
25	28-Days	WNL	WNL	WNL
25	28-Days	WNL	WNL	WNL
25	28-Days	WNL	WNL	WNL
25	28-Days	0	0	2, periportal
50	28-Days	WNL	WNL	WNL
50	28-Days	WNL	WNL	WNL
50	28-Days	0	0	0
50	28-Days	WNL	WNL	WNL
50	28-Days	WNL	WNL	WNL
50	28-Days	WNL	WNL	WNL
50	28-Days	WNL	WNL	WNL
50	28-Days	WNL	WNL	WNL
100	28-Days	0	1, focal	0
100	28-Days	WNL	WNL	WNL
100	28-Days	WNL	WNL	WNL
100	28-Days	WNL	WNL	WNL
100	28-Days	0	0	2, periportal
100	28-Days	WNL	WNL	WNL
100	28-Days	WNL	WNL	WNL
100	28-Days	WNL	WNL	WNL
0	90-Days	WNL	WNL	WNL
0	90-Days	WNL	WNL	WNL
0	90-Days	WNL	WNL	WNL
0	90-Days	WNL	WNL	WNL
0	90-Days	WNL	WNL	WNL
0	90-Days	WNL	WNL	WNL
0	90-Days	WNL	WNL	WNL
0	90-Days	WNL	WNL	WNL
3.125	90-Days	WNL	WNL	WNL
3.125	90-Days	WNL	WNL	WNL
3.125	90-Days	WNL	WNL	WNL
3.125	90-Days	WNL	WNL	WNL
3.125	90-Days	WNL	WNL	WNL
3.125	90-Days	WNL	WNL	WNL
3.125	90-Days	0	0	2, centrilobular
6.25	90-Days	WNL	WNL	WNL
6.25	90-Days	WNL	WNL	WNL
6.25	90-Days	WNL	WNL	WNL
6.25	90-Days	WNL	WNL	WNL
6.25	90-Days	WNL	WNL	WNL
6.25	90-Days	WNL	WNL	WNL
6.25	90-Days	WNL	WNL	WNL
12.5	90-Days	0	1, mf, two, with inflammtion, neutrophils	0
12.5	90-Days	WNL	WNL	WNL
12.5	90-Days	WNL	WNL	WNL
12.5	90-Days	0	0	2, centrilobular
12.5	90-Days	0	0	0
12.5	90-Days	WNL	WNL	WNL
12.5	90-Days	WNL	WNL	WNL
12.5	90-Days	WNL	WNL	WNL
25	90-Days	WNL	WNL	WNL
25	90-Days	WNL	WNL	WNL
25	90-Days	WNL	WNL	WNL
25	90-Days	WNL	WNL	WNL
25	90-Days	WNL	WNL	WNL
25	90-Days	WNL	WNL	WNL
25	90-Days	WNL	WNL	WNL
25	90-Days	WNL	WNL	WNL
50	90-Days	WNL	WNL	WNL
50	90-Days	WNL	WNL	WNL
50	90-Days	WNL	WNL	WNL
50	90-Days	0	1, mf, with inflammtion, neutrophils	0
50	90-Days	WNL	WNL	WNL
50	90-Days	WNL	WNL	WNL
50	90-Days	WNL	WNL	WNL
50	90-Days	WNL	WNL	WNL
0	180-Days	1, single, eosinophilic	0	0
0	180-Days	WNL	WNL	WNL
0	180-Days	WNL	WNL	WNL
0	180-Days	WNL	WNL	WNL
0	180-Days	WNL	WNL	WNL
0	180-Days	WNL	WNL	WNL
0	180-Days	WNL	WNL	WNL
0	180-Days	WNL	WNL	WNL
1.56	180-Days	WNL	WNL	WNL
1.56	180-Days	WNL	WNL	WNL
1.56	180-Days	WNL	WNL	WNL
1.56	180-Days	WNL	WNL	WNL
1.56	180-Days	WNL	WNL	WNL
1.56	180-Days	WNL	WNL	WNL
1.56	180-Days	0	0	2, centrilobular
1.56	180-Days	WNL	WNL	WNL
3.125	180-Days	1, two, eosinophilic	0	0
3.125	180-Days	WNL	WNL	WNL
3.125	180-Days	WNL	WNL	WNL
3.125	180-Days	WNL	WNL	WNL
3.125	180-Days	0	0	2, centrilobular
3.125	180-Days	0	0	3, centrilobular
3.125	180-Days	WNL	WNL	WNL
6.25	180-Days	WNL	WNL	WNL
6.25	180-Days	0	0	3, centrilobular
6.25	180-Days	WNL	WNL	WNL
6.25	180-Days	WNL	WNL	WNL
6.25	180-Days	WNL	WNL	WNL
6.25	180-Days	WNL	WNL	WNL
6.25	180-Days	WNL	WNL	WNL
6.25	180-Days	WNL	WNL	WNL
12.5	180-Days	WNL	WNL	WNL
12.5	180-Days	1, single, eosinophilic	0	0
12.5	180-Days	WNL	WNL	WNL
12.5	180-Days	0	1	0
12.5	180-Days	WNL	WNL	WNL
12.5	180-Days	WNL	WNL	WNL
12.5	180-Days	WNL	WNL	WNL
12.5	180-Days	WNL	WNL	WNL
25	180-Days	WNL	WNL	WNL
25	180-Days	WNL	WNL	WNL
25	180-Days	WNL	WNL	WNL
25	180-Days	WNL	WNL	WNL
25	180-Days	WNL	WNL	WNL
25	180-Days	WNL	WNL	WNL
25	180-Days	WNL	WNL	WNL

Appendix D2 Table 3: DNA Sequencing Metrics

Sample	Tissue	Exposure Duration	BbF dose (mg/kg/day)	Raw Reads	% selected from raw reads	Median Insert Size	SD Insert Size	Mean On-Target Duplex Depth	Maximum On-Target Duplex Depth	Peak Tag Family Size	Informative Duplex Bases
25	Bone Marrow	90-Days	0	409,222,614	78.62%	224	126	11,811	16,002	25	759856338
26	Bone Marrow	90-Days	0	400,402,998	78.81%	194	111	15,323	20,299	17	1046000000
27	Bone Marrow	90-Days	0	420,790,610	78.62%	209	119	14,602	20,322	21	953211698
28	Bone Marrow	90-Days	0	421,243,686	78.55%	204	113	14,464	19,441	20	973783759
29	Bone Marrow	90-Days	3.125	546,765,928	78.78%	211	120	17,818	23,598	22	1194000000
30	Bone Marrow	90-Days	3.125	478,088,010	78.62%	197	109	16,488	22,201	19	1135000000
31	Bone Marrow	90-Days	3.125	422,638,398	79.06%	206	113	10,238	13,323	30	667931725
32	Bone Marrow	90-Days	3.125	468,287,390	78.42%	208	121	21,676	28,987	14	1442000000
33	Bone Marrow	90-Days	6.25	409,737,270	78.56%	206	118	20,502	27,564	13	1385000000
34	Bone Marrow	90-Days	6.25	552,581,248	78.51%	210	119	25,900	35,146	14	1697000000
35	Bone Marrow	90-Days	6.25	449,524,598	78.69%	195	110	16,948	23,011	17	1175000000
36	Bone Marrow	90-Days	6.25	372,927,830	78.80%	214	114	12,238	16,005	23	804685419
37	Bone Marrow	90-Days	12.5	389,057,852	78.58%	201	116	15,752	21,037	16	1065000000
38	Bone Marrow	90-Days	12.5	356,289,346	78.55%	184	99	16,546	23,786	12	1173000000
39	Bone Marrow	90-Days	12.5	551,392,676	79.08%	196	96	16,318	22,171	23	1106000000
40	Bone Marrow	90-Days	12.5	424,511,050	78.66%	195	108	17,041	23,307	16	1168000000
41	Bone Marrow	90-Days	25	363,854,086	79.06%	214	113	16,249	21,723	16	1059000000
42	Bone Marrow	90-Days	25	388,973,088	78.40%	208	124	20,040	26,922	12	1338000000
43	Bone Marrow	90-Days	25	413,097,582	78.38%	210	122	20,542	27,536	13	1359000000
44	Bone Marrow	90-Days	25	582,767,840	78.63%	214	121	12,736	16,423	33	844992742
45	Bone Marrow	90-Days	50	681,259,794	78.48%	219	124	16,793	22,005	29	1097000000
46	Bone Marrow	90-Days	50	372,100,086	78.89%	192	106	13,795	19,250	17	976868829
47	Bone Marrow	90-Days	50	424,757,280	78.42%	212	126	20,253	26,896	14	1347000000
48	Bone Marrow	90-Days	50	445,421,856	78.61%	207	118	17,224	22,734	17	1152000000
25	Liver	90-Days	0	165,211,796	99.30%	228	132	15,620	22,028	7	982894433
26	Liver	90-Days	0	332,408,202	99.35%	233	135	15,056	20,102	20	960642901
27	Liver	90-Days	0	988,209,776	99.54%	222	118	13,073	18,034	69	887912598
28	Liver	90-Days	0	565,787,264	99.46%	227	124	14,132	18,624	39	920244545
29	Liver	90-Days	3.125	562,530,562	99.49%	227	124	15,503	21,172	34	1038000000
30	Liver	90-Days	3.125	451,020,442	99.42%	224	130	14,694	19,634	29	939268813
31	Liver	90-Days	3.125	304,416,792	99.53%	213	119	17,715	23,101	15	1140000000
32	Liver	90-Days	3.125	262,131,980	99.29%	246	141	13,187	17,735	19	831666815
33	Liver	90-Days	6.25	453,566,682	99.51%	214	122	14,465	19,191	30	935327508
34	Liver	90-Days	6.25	494,580,594	99.53%	220	123	13,895	18,294	34	912078389
35	Liver	90-Days	6.25	533,611,424	98.78%	282	149	10,789	14,578	52	666576459
36	Liver	90-Days	6.25	213,529,158	99.59%	205	108	15,403	20,574	12	999512105
37	Liver	90-Days	12.5	347,810,504	99.50%	210	119	14,099	18,558	23	908981483
38	Liver	90-Days	12.5	439,046,956	99.30%	243	135	13,066	17,119	30	851367935
39	Liver	90-Days	12.5	386,584,000	99.36%	215	123	21,465	27,282	16	1378000000
40	Liver	90-Days	12.5	290,469,592	99.51%	206	119	15,251	20,235	17	992949684
41	Liver	90-Days	25	548,620,496	99.45%	216	127	14,010	19,075	34	952909718
42	Liver	90-Days	25	533,463,110	99.51%	223	122	13,343	17,335	39	851071507
43	Liver	90-Days	25	458,087,918	99.36%	229	136	13,883	19,024	31	884589593
44	Liver	90-Days	25	629,416,404	99.55%	227	122	13,216	18,224	45	886882874
45	Liver	90-Days	50	232,749,574	99.49%	216	119	14,402	19,604	14	917636266
46	Liver	90-Days	50	651,101,106	99.41%	223	126	12,416	16,622	49	822106477
47	Liver	90-Days	50	577,351,704	99.40%	241	135	14,154	18,474	42	912842834
48	Liver	90-Days	50	324,100,880	99.38%	238	135	12,631	17,388	24	787385149
49	Bone Marrow	180-Days	0	323,153,452	78.82%	221	130	21,735	30,763	8	1396000000
50	Bone Marrow	180-Days	0	466,735,022	79.00%	222	131	25,179	33,071	11	1645000000
51	Bone Marrow	180-Days	0	326,417,210	78.78%	233	144	19,690	26,542	10	1254000000
52	Bone Marrow	180-Days	0	418,176,396	78.28%	251	155	24,284	32,228	10	1519000000
53	Bone Marrow	180-Days	1.56	479,469,258	79.08%	204	120	24,762	33,429	11	1679000000
54	Bone Marrow	180-Days	1.56	510,565,138	78.93%	228	136	27,373	35,862	11	1768000000
55	Bone Marrow	180-Days	1.56	474,732,908	78.82%	226	129	24,594	32,283	12	1602000000
56	Bone Marrow	180-Days	1.56	466,028,754	79.05%	209	124	25,335	33,798	11	1687000000
57	Bone Marrow	180-Days	3.125	485,097,690	78.87%	214	125	27,815	37,291	10	1824000000
58	Bone Marrow	180-Days	3.125	368,960,226	78.96%	247	141	25,255	36,010	8	1564000000
59	Bone Marrow	180-Days	3.125	470,190,918	78.73%	228	138	27,794	37,616	9	1800000000
60	Bone Marrow	180-Days	3.125	508,265,514	78.67%	271	159	26,873	34,950	12	1667000000
61	Bone Marrow	180-Days	6.25	410,580,356	78.76%	250	147	26,129	35,518	9	1631000000
62	Bone Marrow	180-Days	6.25	454,155,952	78.75%	205	129	23,185	30,354	11	1545000000
63	Bone Marrow	180-Days	6.25	413,570,910	78.95%	224	135	23,571	30,848	11	1520000000
64	Bone Marrow	180-Days	6.25	529,849,678	78.75%	244	144	29,992	39,839	10	1921000000
65	Bone Marrow	180-Days	12.5	386,066,926	78.95%	230	136	23,218	31,294	9	1493000000
66	Bone Marrow	180-Days	12.5	457,865,142	79.04%	210	124	25,026	33,388	11	1664000000
67	Bone Marrow	180-Days	12.5	413,598,060	78.93%	186	120	23,387	32,597	9	1597000000
68	Bone Marrow	180-Days	12.5	399,282,020	78.86%	219	132	25,604	36,974	8	1666000000
69	Bone Marrow	180-Days	25	412,788,772	78.87%	226	137	24,713	33,465	10	1592000000
70	Bone Marrow	180-Days	25	597,846,306	78.39%	249	147	26,336	34,101	15	1661000000
71	Bone Marrow	180-Days	25	421,850,592	78.67%	236	140	25,500	34,010	10	1627000000
72	Bone Marrow	180-Days	25	404,924,326	78.60%	252	149	25,701	35,181	9	1594000000
49	Liver	180-Days	0	568,537,292	99.54%	195	125	27,901	36,292	17	1845000000
50	Liver	180-Days	0	347,602,562	99.35%	220	132	17,043	21,387	17	1149000000
51	Liver	180-Days	0	484,774,308	99.41%	217	140	20,522	26,043	20	1359000000
52	Liver	180-Days	0	444,578,592	99.47%	208	126	20,933	26,767	16	1442000000
53	Liver	180-Days	1.56	681,366,774	99.35%	226	145	21,367	27,318	28	1408000000
54	Liver	180-Days	1.56	476,578,924	99.42%	222	136	20,558	25,791	20	1379000000
55	Liver	180-Days	1.56	311,903,974	99.50%	200	128	19,259	24,814	12	1319000000
56	Liver	180-Days	1.56	506,777,784	99.45%	215	133	22,889	28,876	18	1545000000
57	Liver	180-Days	3.125	602,854,626	99.64%	196	119	25,193	32,574	21	1660000000
58	Liver	180-Days	3.125	416,422,206	99.48%	294	131	20,388	25,948	16	1399000000
59	Liver	180-Days	3.125	613,332,696	99.42%	199	127	22,264	28,032	22	1520000000
60	Liver	180-Days	3.125	402,583,334	99.36%	220	140	20,485	26,276	16	1365000000
61	Liver	180-Days	6.25	456,990,126	99.44%	206	132	18,924	23,927	19	1294000000
62	Liver	180-Days	6.25	517,981,434	99.39%	215	137	18,408	23,295	25	1211000000
63	Liver	180-Days	6.25	404,336,322	99.28%	227	142	19,947	25,294	17	1302000000
64	Liver	180-Days	6.25	353,889,678	99.50%	197	123	20,396	25,991	13	1395000000
65	Liver	180-Days	12.5	343,950,724	99.29%	219	138	19,167	24,077	15	1274000000
66	Liver	180-Days	12.5	551,044,450	99.49%	208	124	23,964	30,186	20	1623000000
67	Liver	180-Days	12.5	525,218,476	99.27%	231	144	19,659	24,974	24	1293000000
68	Liver	180-Days	12.5	471,398,910	99.14%	251	153	18,794	24,246	22	1178000000
69	Liver	180-Days	25	195,122,678	99.31%	210	126	7,763	10,444	13	514337139
70	Liver	180-Days	25	543,278,296	99.51%	212	130	23,273	29,422	20	1546000000
71	Liver	180-Days	25	285,433,610	99.01%	246	155	16,506	21,957	15	1028000000
72	Liver	180-Days	25	433,148,252	99.50%	212	130	24,663	31,685	15	1626000000

Appendix D2 Table 4: Mutagenicity data of MutaMouse males post exposure to BbF for 28-days

Animal ID	Sample tissue	Exposure duration	BbF dose (mg/kg/day)	Mutation count unique	Mutation count all	Sample depth	Mutation frequency minimum	Mutation Frequency maximum
1	Bone Marrow	28-Days	0	37	43	820319269	4.51E-08	5.24E-08
2	Bone Marrow	28-Days	0	35	37	903268416	3.87E-08	4.10E-08
3	Bone Marrow	28-Days	0	34	35	841007628	4.04E-08	4.16E-08
4	Bone Marrow	28-days	0	39	43	953865140	4.09E-08	4.51E-08
5	Bone Marrow	28-Days	6.25	32	38	709601657	4.51E-08	5.36E-08
6	Bone Marrow	28-Days	6.25	53	190	570223160	9.29E-08	3.33E-07
7	Bone Marrow	28-Days	6.25	27	28	650976772	4.15E-08	4.30E-08
8	Bone Marrow	28-days	6.25	35	43	755934082	4.63E-08	5.69E-08
9	Bone Marrow	28-Days	12.5	67	75	996552403	6.72E-08	7.53E-08
10	Bone Marrow	28-Days	12.5	71	126	748281155	9.49E-08	1.68E-07
11	Bone Marrow	28-Days	12.5	72	80	975415700	7.38E-08	8.20E-08
12	Bone Marrow	28-days	12.5	47	51	714621726	6.58E-08	7.14E-08
13	Bone Marrow	28-Days	25	66	73	1034555738	6.38E-08	7.06E-08
14	Bone Marrow	28-Days	25	60	73	731411483	8.20E-08	9.98E-08
15	Bone Marrow	28-Days	25	49	52	587263813	8.34E-08	8.85E-08
16	Bone Marrow	28-days	25	90	99	1013524373	8.88E-08	9.77E-08
17	Bone Marrow	28-Days	50	72	84	747503560	9.63E-08	1.12E-07
18	Bone Marrow	28-Days	50	122	135	1015314151	1.20E-07	1.33E-07
19	Bone Marrow	28-Days	50	105	121	727894703	1.44E-07	1.66E-07
20	Bone Marrow	28-days	50	93	103	876901612	1.06E-07	1.17E-07
21	Bone Marrow	28-Days	100	143	176	859890285	1.66E-07	2.05E-07
22	Bone Marrow	28-Days	100	115	138	723027618	1.59E-07	1.91E-07
23	Bone Marrow	28-Days	100	149	176	1141474460	1.31E-07	1.54E-07
24	Bone Marrow	28-days	100	118	136	876503251	1.35E-07	1.55E-07
1	Liver	28-Days	0	44	47	749609387	6.32E-08	6.75E-08
2	Liver	28-Days	0	48	87	722340262	6.61E-08	1.20E-07
3	Liver	28-Days	0	41	47	627025630	5.01E-08	5.75E-08
4	Liver	28-days	0	28	28	677732398	3.39E-08	3.39E-08
5	Liver	28-Days	6.25	70	82	743455407	6.72E-08	7.87E-08
6	Liver	28-Days	6.25	40	58	704928576	5.70E-08	8.26E-08
7	Liver	28-Days	6.25	32	33	849818048	4.60E-08	4.75E-08
8	Liver	28-days	6.25	50	50	632701328	6.51E-08	6.51E-08
9	Liver	28-Days	12.5	92	95	693759043	8.38E-08	8.65E-08
10	Liver	28-Days	12.5	85	106	666531230	8.74E-08	1.09E-07
11	Liver	28-Days	12.5	68	72	517474384	7.11E-08	7.53E-08
12	Liver	28-days	12.5	71	79	738972435	8.02E-08	8.92E-08
13	Liver	28-Days	25	89	94	676238056	9.97E-08	1.05E-07
14	Liver	28-Days	25	102	119	626642186	1.11E-07	1.29E-07
15	Liver	28-Days	25	85	89	1029731981	1.14E-07	1.19E-07
16	Liver	28-days	25	76	76	731701362	1.20E-07	1.20E-07
17	Liver	28-Days	50	195	197	671817466	2.21E-07	2.23E-07
18	Liver	28-Days	50	275	304	639888577	2.90E-07	3.20E-07
19	Liver	28-Days	50	135	136	665939020	1.54E-07	1.55E-07
20	Liver	28-days	50	177	183	633802424	2.24E-07	2.31E-07
21	Liver	28-Days	100	250	269	690876985	3.15E-07	3.39E-07
22	Liver	28-Days	100	498	595	595451784	4.85E-07	5.80E-07
23	Liver	28-Days	100	263	272	678943393	3.21E-07	3.32E-07
24	Liver	28-days	100	297	305	605952416	3.30E-07	3.39E-07

Appendix D2 Table 5: Mutagenicity data of MutaMouse males post exposure to BbF for 90-days

Animal ID	Sample tissue	Exposure duration	BbF dose (mg/kg/day)	Mutation count unique	Mutation count all	Sample depth	Mutation frequency minimum	Mutation Frequency maximum
25	Bone Marrow	90-Days	0	16	16	566484337	2.82E-08	2.82E-08
26	Bone Marrow	90-Days	0	42	49	735141570	5.71E-08	6.67E-08
27	Bone Marrow	90-Days	0	18	23	700385118	2.57E-08	3.28E-08
28	Bone Marrow	90-Days	0	30	30	693878407	4.32E-08	4.32E-08
29	Bone Marrow	90-Days	3.125	55	72	854588642	6.44E-08	8.43E-08
30	Bone Marrow	90-Days	3.125	49	62	791006370	6.19E-08	7.84E-08
31	Bone Marrow	90-Days	3.125	21	21	491063562	4.28E-08	4.28E-08
32	Bone Marrow	90-Days	3.125	39	49	1040003431	3.75E-08	4.71E-08
33	Bone Marrow	90-Days	6.25	68	82	983841168	6.91E-08	8.33E-08
34	Bone Marrow	90-Days	6.25	68	89	1242699186	5.47E-08	7.16E-08
35	Bone Marrow	90-Days	6.25	82	119	813103920	1.01E-07	1.46E-07
36	Bone Marrow	90-Days	6.25	43	46	586972650	7.33E-08	7.84E-08
37	Bone Marrow	90-Days	12.5	72	93	755751466	9.53E-08	1.23E-07
38	Bone Marrow	90-Days	12.5	113	214	793953742	1.42E-07	2.70E-07
39	Bone Marrow	90-Days	12.5	78	88	782605930	9.97E-08	1.12E-07
40	Bone Marrow	90-Days	12.5	116	143	817546833	1.42E-07	1.75E-07
41	Bone Marrow	90-Days	25	103	165	779512147	1.32E-07	2.12E-07
42	Bone Marrow	90-Days	25	137	198	961611346	1.42E-07	2.06E-07
43	Bone Marrow	90-Days	25	141	209	985667893	1.43E-07	2.12E-07
44	Bone Marrow	90-Days	25	88	143	610801986	1.44E-07	2.34E-07
45	Bone Marrow	90-Days	50	172	232	805599949	2.14E-07	2.88E-07
46	Bone Marrow	90-Days	50	109	143	661835407	1.65E-07	2.16E-07
47	Bone Marrow	90-Days	50	187	340	971821068	1.92E-07	3.50E-07
48	Bone Marrow	90-Days	50	136	222	826366527	1.65E-07	2.69E-07
25	Liver	90-Days	0	22	25	749609387	2.93E-08	3.34E-08
26	Liver	90-Days	0	41	45	722340262	5.68E-08	6.23E-08
27	Liver	90-Days	0	39	68	627025630	6.22E-08	1.08E-07
28	Liver	90-Days	0	35	48	677732398	5.16E-08	7.08E-08
29	Liver	90-Days	3.125	63	88	743455407	8.47E-08	1.18E-07
30	Liver	90-Days	3.125	35	35	704928576	4.97E-08	4.97E-08
31	Liver	90-Days	3.125	46	49	849818048	5.41E-08	5.77E-08
32	Liver	90-Days	3.125	36	37	632701328	5.69E-08	5.85E-08
33	Liver	90-Days	6.25	58	64	693759043	8.36E-08	9.23E-08
34	Liver	90-Days	6.25	69	75	666531230	1.04E-07	1.13E-07
35	Liver	90-Days	6.25	72	161	517474384	1.39E-07	3.11E-07
36	Liver	90-Days	6.25	56	58	738972435	7.58E-08	7.85E-08
37	Liver	90-Days	12.5	96	195	676238056	1.42E-07	2.88E-07
38	Liver	90-Days	12.5	110	130	626642186	1.76E-07	2.07E-07
39	Liver	90-Days	12.5	130	132	1029731981	1.26E-07	1.28E-07
40	Liver	90-Days	12.5	129	130	731701362	1.76E-07	1.78E-07
41	Liver	90-Days	25	244	260	671817466	3.63E-07	3.87E-07
42	Liver	90-Days	25	238	267	639888577	3.72E-07	4.17E-07
43	Liver	90-Days	25	279	281	665939020	4.19E-07	4.22E-07
44	Liver	90-Days	25	333	360	633802424	5.25E-07	5.68E-07
45	Liver	90-Days	50	844	862	690876985	1.22E-06	1.25E-06
46	Liver	90-Days	50	462	487	595451784	7.76E-07	8.18E-07
47	Liver	90-Days	50	568	581	678943393	8.37E-07	8.56E-07
48	Liver	90-Days	50	401	407	605952416	6.62E-07	6.72E-07

Appendix D2 Table 6: Mutagenicity data of MutaMouse males post exposure to BbF for 180-days

Animal ID	Sample tissue	Exposure duration	BbF dose (mg/kg/day)	Mutation count unique	Mutation count all	Sample depth	Mutation frequency minimum	Mutation Frequency maximum
49	Bone Marrow	180-Days	0	43	47	1043357842	4.12E-08	4.50E-08
50	Bone Marrow	180-Days	0	56	71	1208360081	4.63E-08	5.88E-08
51	Bone Marrow	180-Days	0	31	39	944993138	3.28E-08	4.13E-08
52	Bone Marrow	180-Days	0	54	68	1165443482	4.63E-08	5.83E-08
53	Bone Marrow	180-Days	1.56	59	85	1188414445	4.96E-08	7.15E-08
54	Bone Marrow	180-Days	1.56	53	58	1313694711	4.03E-08	4.42E-08
55	Bone Marrow	180-Days	1.56	63	70	1180341418	5.34E-08	5.93E-08
56	Bone Marrow	180-Days	1.56	64	68	1215978919	5.26E-08	5.59E-08
57	Bone Marrow	180-Days	3.125	84	143	1335013466	6.29E-08	1.07E-07
58	Bone Marrow	180-Days	3.125	71	81	1212318653	5.86E-08	6.68E-08
59	Bone Marrow	180-Days	3.125	99	115	1334128426	7.42E-08	8.62E-08
60	Bone Marrow	180-Days	3.125	78	92	1289621996	6.05E-08	7.13E-08
61	Bone Marrow	180-Days	6.25	104	162	1254094598	8.29E-08	1.29E-07
62	Bone Marrow	180-Days	6.25	114	148	1112669546	1.02E-07	1.33E-07
63	Bone Marrow	180-Days	6.25	103	153	1131192554	9.11E-08	1.35E-07
64	Bone Marrow	180-Days	6.25	156	266	1439588653	1.08E-07	1.85E-07
65	Bone Marrow	180-Days	12.5	165	281	1114445535	1.48E-07	2.52E-07
66	Bone Marrow	180-Days	12.5	159	275	1201066020	1.32E-07	2.29E-07
67	Bone Marrow	180-Days	12.5	184	230	1122331889	1.64E-07	2.05E-07
68	Bone Marrow	180-Days	12.5	210	280	1229131410	1.71E-07	2.28E-07
69	Bone Marrow	180-Days	25	248	548	1186149601	2.09E-07	4.62E-07
70	Bone Marrow	180-Days	25	273	491	1263726420	2.16E-07	3.89E-07
71	Bone Marrow	180-Days	25	210	536	1223966951	1.72E-07	4.38E-07
72	Bone Marrow	180-Days	25	224	395	1233512374	1.82E-07	3.20E-07
49	Liver	180-Days	0	81	87	1338528345	6.05E-08	6.50E-08
50	Liver	180-Days	0	39	45	817844502	4.77E-08	5.50E-08
51	Liver	180-Days	0	50	60	984471002	5.08E-08	6.09E-08
52	Liver	180-Days	0	79	83	1004542293	7.86E-08	8.26E-08
53	Liver	180-Days	1.56	76	84	1025217179	7.41E-08	8.19E-08
54	Liver	180-Days	1.56	57	59	986447253	5.78E-08	5.98E-08
55	Liver	180-Days	1.56	64	69	924442350	6.92E-08	7.46E-08
56	Liver	180-Days	1.56	62	66	1098103567	5.65E-08	6.01E-08
57	Liver	180-Days	3.125	73	76	1208627273	6.04E-08	6.29E-08
58	Liver	180-Days	3.125	68	87	978382774	6.95E-08	8.89E-08
59	Liver	180-Days	3.125	85	119	1068256281	7.96E-08	1.11E-07
60	Liver	180-Days	3.125	63	64	983049722	6.41E-08	6.51E-08
61	Liver	180-Days	6.25	135	135	908053521	1.49E-07	1.49E-07
62	Liver	180-Days	6.25	110	129	883228643	1.25E-07	1.46E-07
63	Liver	180-Days	6.25	153	173	957144595	1.60E-07	1.81E-07
64	Liver	180-Days	6.25	137	141	978861139	1.40E-07	1.44E-07
65	Liver	180-Days	12.5	367	375	919815508	3.99E-07	4.08E-07
66	Liver	180-Days	12.5	409	417	1149746878	3.56E-07	3.63E-07
67	Liver	180-Days	12.5	472	493	943037743	5.01E-07	5.23E-07
68	Liver	180-Days	12.5	332	334	901595376	3.68E-07	3.70E-07
69	Liver	180-Days	25	404	413	372700610	1.08E-06	1.11E-06
70	Liver	180-Days	25	1200	1252	1116483211	1.07E-06	1.12E-06
71	Liver	180-Days	25	635	650	792068914	8.02E-07	8.21E-07
72	Liver	180-Days	25	1040	1109	1183336314	8.79E-07	9.37E-07

Appendix D2 Table 7: Subtype specific mutagenicity data of MutaMouse males post exposure to BbF for 28-days

Animal ID	Sample Tissue	Exposure Duration	BbF dose (mg/kg/day)	C:G>A:T Count	C:G>A:T Count All	C:G>A:T MFmin	C:G>A:T MFmax	C:G>G:C Count	C:G>G:C Count All	C:G>G:C MFmin	C:G>G:C MFmax	C:G>T:A Count	C:G>T: A	C:G>T:A MFmin	C:G>T:A MFmax	T:A>A:T Count	T:A>A: T	T:A>A:T MFmin	T:A>A:T MFmax	T:A>C:G Count	T:A>C: G	T:A>C:G MFmin	T:A>C:G MFmax	T:A>G:C Count	T:A>G: C	T:A>G:C MFmin	T:A>G:C MFmax	Insertion Count Unique	Insertion Count All	Insertion MFmin	Insertion MFmax	Deletion Count	Deletio n	Deletion MFmin	Deletion MFmax
1	Bone Marrow	28-Days	0	9	10	2.53E-08	2.82E-08	3	4	8.45E-09	1.13E-08	7	7	1.97E-08	1.97E-08	1	1	2.15E-09	2.15E-09	4	4	8.60E-09	8.60E-09	2	5	4.30E-09	1.07E-08	0	0	0	0	11	12	1.34E-08	1.46E-08
2	Bone Marrow	28-Days	0	7	7	1.78E-08	1.78E-08	0	0	0	0	12	13	3.05E-08	3.31E-08	1	1	1.96E-09	1.96E-09	3	3	5.88E-09	5.88E-09	2	2	3.92E-09	3.92E-09	1	2	1.11E-09	2.21E-09	9	9	9.96E-09	9.96E-09
3	Bone Marrow	28-Days	0	4	4	1.10E-08	1.10E-08	0	0	0	0	8	8	2.20E-08	2.20E-08	3	3	6.29E-09	6.29E-09	6	6	1.26E-08	1.26E-08	1	1	2.10E-09	2.10E-09	3	4	3.57E-09	4.76E-09	8	8	9.51E-09	9.51E-09
4	Bone Marrow	28-Days	0	12	12	2.90E-08	2.90E-08	3	3	7.25E-09	7.25E-09	9	9	2.17E-08	2.17E-08	0	0	0	0	6	6	1.11E-08	1.11E-08	2	2	3.70E-09	3.70E-09	0	0	0	0	5	8	5.24E-09	8.39E-09
5	Bone Marrow	28-Days	6.25	9	9	2.94E-08	2.94E-08	1	3	3.27E-09	9.81E-09	11	11	3.60E-08	3.60E-08	2	2	4.95E-09	4.95E-09	4	4	9.91E-09	9.91E-09	1	2	2.48E-09	4.95E-09	0	0	0	0	3	3	4.23E-09	4.23E-09
6	Bone Marrow	28-Days	6.25	14	14	5.71E-08	5.71E-08	5	30	2.04E-08	1.22E-07	10	50	4.08E-08	2.04E-07	1	1	3.08E-09	3.08E-09	8	14	2.46E-08	4.31E-08	3	3	9.23E-09	9.23E-09	0	0	0	0	6	20	1.05E-08	3.51E-08
7	Bone Marrow	28-Days	6.25	4	5	1.43E-08	1.79E-08	3	3	1.07E-08	1.07E-08	3	3	1.07E-08	1.07E-08	1	1	2.70E-09	2.70E-09	5	5	1.35E-08	1.35E-08	3	3	8.09E-09	8.09E-09	1	1	1.54E-09	1.54E-09	6	6	9.22E-09	9.22E-09
8	Bone Marrow	28-Days	6.25	10	10	3.02E-08	3.02E-08	0	0	0	0	8	9	2.41E-08	2.71E-08	3	8	7.07E-09	1.89E-08	4	4	9.43E-09	9.43E-09	2	2	4.71E-09	4.71E-09	2	2	2.65E-09	2.65E-09	3	3	3.97E-09	3.97E-09
9	Bone Marrow	28-Days	12.5	15	17	3.44E-08	3.90E-08	7	7	1.61E-08	1.61E-08	14	16	3.21E-08	3.67E-08	6	6	1.07E-08	1.07E-08	8	8	1.43E-08	1.43E-08	4	6	7.13E-09	1.07E-08	1	1	1.00E-09	1.00E-09	10	11	1.00E-08	1.10E-08
10	Bone Marrow	28-Days	12.5	21	32	6.54E-08	9.97E-08	5	15	1.56E-08	4.67E-08	13	25	4.05E-08	7.79E-08	6	6	1.40E-08	1.40E-08	7	9	1.64E-08	2.11E-08	4	4	9.36E-09	9.36E-09	2	2	2.67E-09	2.67E-09	7	7	9.35E-09	9.35E-09
11	Bone Marrow	28-Days	12.5	22	25	5.19E-08	5.90E-08	7	7	1.65E-08	1.65E-08	5	7	1.18E-08	1.65E-08	8	9	1.45E-08	1.63E-08	9	10	1.63E-08	1.81E-08	7	7	1.27E-08	1.27E-08	3	4	3.08E-09	4.10E-09	9	9	9.23E-09	9.23E-09
12	Bone Marrow	28-Days	12.5	21	23	6.73E-08	7.37E-08	3	3	9.62E-09	9.62E-09	6	7	1.92E-08	2.24E-08	3	4	7.45E-09	9.93E-09	1	1	2.48E-09	2.48E-09	4	4	9.93E-09	9.93E-09	0	0	0	0	8	8	1.12E-08	1.12E-08
13	Bone Marrow	28-Days	25	17	17	3.77E-08	3.77E-08	6	6	1.33E-08	1.33E-08	17	22	3.77E-08	4.88E-08	4	5	6.86E-09	8.57E-09	6	6	1.03E-08	1.03E-08	3	3	5.14E-09	5.14E-09	1	1	9.67E-10	9.67E-10	9	10	8.70E-09	9.67E-09
14	Bone Marrow	28-Days	25	14	19	4.36E-08	5.92E-08	6	6	1.87E-08	1.87E-08	11	16	3.43E-08	4.98E-08	10	12	2.44E-08	2.92E-08	9	9	2.19E-08	2.19E-08	2	2	4.87E-09	4.87E-09	1	2	1.37E-09	2.73E-09	7	7	9.57E-09	9.57E-09
15	Bone Marrow	28-Days	25	9	10	3.57E-08	3.96E-08	4	4	1.59E-08	1.59E-08	12	12	4.76E-08	4.76E-08	5	5	1.49E-08	1.49E-08	6	7	1.79E-08	2.09E-08	2	2	5.97E-09	5.97E-09	3	4	5.11E-09	6.81E-09	7	7	1.19E-08	1.19E-08
16	Bone Marrow	28-Days	25	22	23	4.99E-08	5.22E-08	5	9	1.14E-08	2.04E-08	16	17	3.63E-08	3.86E-08	14	15	2.44E-08	2.62E-08	9	10	1.57E-08	1.75E-08	5	5	8.73E-09	8.73E-09	2	2	1.97E-09	1.97E-09	12	12	1.18E-08	1.18E-08
17	Bone Marrow	28-Days	50	32	42	9.84E-08	1.29E-07	4	5	1.23E-08	1.54E-08	10	11	3.07E-08	3.38E-08	7	7	1.66E-08	1.66E-08	3	3	7.11E-09	7.11E-09	4	4	9.47E-09	9.47E-09	2	2	2.68E-09	2.68E-09	8	8	1.07E-08	1.07E-08
18	Bone Marrow	28-Days	50	28	28	6.34E-08	6.34E-08	9	10	2.04E-08	2.26E-08	28	35	6.34E-08	7.92E-08	11	13	1.92E-08	2.27E-08	20	21	3.49E-08	3.66E-08	7	8	1.22E-08	1.39E-08	5	5	4.92E-09	4.92E-09	10	11	9.85E-09	1.08E-08
19	Bone Marrow	28-Days	50	38	46	1.21E-07	1.47E-07	5	5	1.60E-08	1.60E-08	23	26	7.35E-08	8.31E-08	9	10	2.17E-08	2.41E-08	9	9	2.17E-08	2.17E-08	2	2	4.82E-09	4.82E-09	3	4	4.12E-09	5.50E-09	14	17	1.92E-08	2.34E-08
20	Bone Marrow	28-Days	50	27	28	7.08E-08	7.35E-08	6	6	1.57E-08	1.57E-08	22	27	5.77E-08	7.08E-08	12	16	2.42E-08	3.23E-08	11	11	2.22E-08	2.22E-08	3	3	6.05E-09	6.05E-09	1	1	1.14E-09	1.14E-09	10	10	1.14E-08	1.14E-08
21	Bone Marrow	28-Days	100	43	57	1.14E-07	1.52E-07	11	14	2.93E-08	3.73E-08	32	35	8.52E-08	9.32E-08	18	20	3.72E-08	4.13E-08	9	9	1.86E-08	1.86E-08	4	9	8.26E-09	1.86E-08	3	4	3.49E-09	4.65E-09	14	15	1.63E-08	1.74E-08
22	Bone Marrow	28-Days	100	54	58	1.70E-07	1.83E-07	7	7	2.21E-08	2.21E-08	31	43	9.79E-08	1.36E-07	9	10	2.22E-08	2.46E-08	3	3	7.38E-09	7.38E-09	2	2	4.92E-09	4.92E-09	0	0	0	0	5	6	6.92E-09	8.30E-09
23	Bone Marrow	28-Days	100	57	71	1.15E-07	1.43E-07	15	17	3.02E-08	3.43E-08	31	33	6.25E-08	6.65E-08	14	16	2.17E-08	2.48E-08	10	14	1.55E-08	2.17E-08	6	7	9.29E-09	1.08E-08	3	3	2.63E-09	2.63E-09	6	7	5.26E-09	6.13E-09
24	Bone Marrow	28-Days	100	43	49	1.12E-07	1.28E-07	7	7	1.83E-08	1.83E-08	26	33	6.79E-08	8.61E-08	5	5	1.01E-08	1.01E-08	14	18	2.84E-08	3.65E-08	5	5	1.01E-08	1.01E-08	2	2	2.28E-09	2.28E-09	11	11	1.26E-08	1.26E-08
1	Liver	28-Days	0	2	2	6.64E-09	6.64E-09	0	0	0	0	19	19	6.31E-08	6.31E-08	5	5	1.26E-08	1.26E-08	4	6	1.01E-08	1.52E-08	0	0	0	0	3	3	4.31E-09	4.31E-09	10	11	1.44E-08	1.58E-08
2	Liver	28-Days	0	6	6	1.93E-08	1.93E-08	3	7	9.63E-09	2.25E-08	15	35	4.81E-08	1.12E-07	2	2	4.83E-09	4.83E-09	6	6	1.45E-08	1.45E-08	3	3	7.24E-09	7.24E-09	2	2	2.76E-09	2.76E-09	7	7	9.64E-09	9.64E-09
3	Liver	28-Days	0	3	3	8.45E-09	8.45E-09	1	1	2.82E-09	2.82E-09	15	21	4.23E-08	5.92E-08	5	5	1.08E-08	1.08E-08	7	7	1.51E-08	1.51E-08	3	3	6.48E-09	6.48E-09	1	1	1.22E-09	1.22E-09	6	6	7.34E-09	7.34E-09
4	Liver	28-Days	0	5	5	1.40E-08	1.40E-08	1	1	2.79E-09	2.79E-09	13	13	3.63E-08	3.63E-08	3	3	6.42E-09	6.42E-09	1	1	2.14E-09	2.14E-09	1	1	2.14E-09	2.14E-09	1	1	1.21E-09	1.21E-09	1	1	1.21E-09	1.21E-09
5	Liver	28-Days	6.25	15	16	3.35E-08	3.57E-08	2	2	4.47E-09	4.47E-09	26	30	5.81E-08	6.70E-08	0	0	0	0	11	12	1.85E-08	2.02E-08	4	4	6.73E-09	6.73E-09	1	1	9.60E-10	9.60E-10	9	12	8.64E-09	1.15E-08
6	Liver	28-Days	6.25	10	19	3.27E-08	6.21E-08	3	3	9.81E-09	9.81E-09	9	15	2.94E-08	4.90E-08	4	4	1.01E-08	1.01E-08	4	4	1.01E-08	1.01E-08	1	1	2.53E-09	2.53E-09	1	2	1.42E-09	2.85E-09	2	2	2.85E-09	2.85E-09
7	Liver	28-Days	6.25	6	6	2.00E-08	2.00E-08	1	1	3.34E-09	3.34E-09	8	8	2.67E-08	2.67E-08	6	6	1.52E-08	1.52E-08	0	0	0	0	3	4	7.58E-09	1.01E-08	0	0	0	0	7	7	1.01E-08	1.01E-08
8	Liver	28-Days	6.25	5	5	1.50E-08	1.50E-08	0	0	0	0	27	27	8.10E-08	8.10E-08	2	2	4.61E-09	4.61E-09	5	5	1.15E-08	1.15E-08	4	4	9.22E-09	9.22E-09	2	2	2.61E-09	2.61E-09	4	4	5.21E-09	5.21E-09
9	Liver	28-Days	12.5	19	19	4.02E-08	4.02E-08	6	7	1.27E-08	1.48E-08	35	36	7.40E-08	7.61E-08	5	5	8.00E-09	8.00E-09	8	8	1.28E-08	1.28E-08	4	4	6.40E-09	6.40E-09	2	2	1.82E-09	1.82E-09	9	9	8.19E-09	8.19E-09
10	Liver	28-Days	12.5	17	17	4.05E-08	4.05E-08	5	5	1.19E-08	1.19E-08	17	19	4.05E-08	4.53E-08	13	13	2.35E-08	2.35E-08	10	26	1.81E-08	4.71E-08	4	4	7.24E-09	7.24E-09	6	8	6.17E-09	8.23E-09	8	8	8.23E-09	8.23E-09
11	Liver	28-Days	12.5	14	14	3.41E-08	3.41E-08	2	2	4.88E-09	4.88E-09	19	19	4.63E-08	4.63E-08	6	6	1.10E-08	1.10E-08	8	8	1.47E-08	1.47E-08	4	4	7.33E-09	7.33E-09	6	6	6.28E-09	6.28E-09	8	8	8.37E-09	8.37E-09
12	Liver	28-Days	12.5	15	15	3.92E-08	3.92E-08	8	8	2.09E-08	2.09E-08	23	31	6.02E-08	8.11E-08	6	6	1.19E-08	1.19E-08	8	8	1.59E-08	1.59E-08	2	2	3.97E-09	3.97E-09	0	0	0	0	5	5	5.65E-09	5.65E-09
13	Liver	28-Days	25	32	32	8.30E-08	8.30E-08	5	6	1.30E-08	1.56E-08	15	17	3.89E-08	4.41E-08	11	11	2.17E-08	2.17E-08	6	6	1.18E-08	1.18E-08	7	7	1.38E-08	1.38E-08	1	1	1.12E-09	1.12E-09	8	8	8.96E-09	8.96E-09
14	Liver	28-Days	25	35	35	8.84E-08	8.84E-08	16	19	4.04E-08	4.80E-08	20	28	5.05E-08	7.07E-08	12	12	2.29E-08	2.29E-08	8	10	1.53E-08	1.91E-08	0	0	0	0	2	2	2.18E-09	2.18E-09	4	4	4.35E-09	4.35E-09
15	Liver	28-Days	25	25	27	7.77E-08	8.39E-08	9	9	2.80E-08	2.80E-08	19	21	5.91E-08	6.53E-08	10	10	2.34E-08	2.34E-08	5	5	1.17E-08	1.17E-08	3	3	7.02E-09	7.02E-09	2	2	2.67E-09	2.67E-09	9	9	1.20E-08	1.20E-08
16	Liver	28-Days	25	31	31	1.12E-07	1.12E-07	3	3	1.09E-08	1.09E-08	18	18	6.52E-08	6.52E-08	11	11	3.09E-08	3.09E-08	5	5	1.40E-08	1.40E-08	4	4	1.12E-08	1.12E-08	2	2	3.16E-09	3.16E-09	1	1	1.58E-09	1.58E-09
17	Liver	28-Days	50	75	76	1.96E-07	1.98E-07	17	17	4.43E-08	4.43E-08	42	42	1.10E-07	1.10E-07	21	21	4.22E-08	4.22E-08	10	10	2.01E-08	2.01E-08	5	5	1.00E-08	1.00E-08	1	1	1.13E-09	1.13E-09	12	12	1.36E-08	1.36E-08
18	Liver	28-Days	50	136	136	3.31E-07	3.31E-07	24	24	5.84E-08	5.84E-08	45	52	1.09E-07	1.27E-07	30	30	5.58E-08	5.58E-08	14	16	2.60E-08	2.98E-08	4	4	7.44E-09	7.44E-09	5	6	5.27E-09	6.33E-09	7	25	7.38E-09	2.64E-08
19	Liver	28-Days	50	54	54	1.43E-07	1.43E-07	9	9	2.38E-08	2.38E-08	30	30	7.94E-08	7.94E-08	15	15	3.00E-08	3.00E-08	7	7	1.40E-08	1.40E-08	6	7	1.20E-08	1.40E-08	1	1	1.14E-09	1.14E-09	9	9	1.03E-08	1.03E-08
20	Liver	28-Days	50	63	63	1.85E-07	1.85E-07	19	20	5.59E-08	5.89E-08	40	44	1.18E-07	1.30E-07	22	22	4.87E-08	4.87E-08	13	14	2.88E-08	3.10E-08	4	4	8.86E-09	8.86E-09	2	2	2.53E-09	2.53E-09	7	7	8.85E-09	8.85E-09
21	Liver	28-Days	100	107	109	3.13E-07	3.19E-07	28	33	8.20E-08	9.67E-08	56	63	1.64E-07	1.85E-07	28	28	6.19E-08	6.19E-08	9	10	1.99E-08	2.21E-08	5	5	1.11E-08	1.11E-08	3	3	3.78E-09	3.78E-09	6	6	7.56E-09	7.56E-09
22	Liver	28-Days	100	229	232	5.17E-07	5.24E-07	45	45	1.02E-07	1.02E-07	88	181	1.99E-07	4.09E-07	62	63	1.06E-07	1.08E-07	23	23	3.94E-08	3.94E-08	10	10	1.71E-08	1.71E-08	5	5	4.87E-09	4.87E-09	21	21	2.05E-08	2.05E-08
23	Liver	28-Days	100	117	120	3.30E-07	3.38E-07	21	21	5.92E-08	5.92E-08	47	48	1.32E-07	1.35E-07	30	30	6.47E-08	6.47E-08	17	17	3.66E-08	3.66E-08	6	7	1.29E-08	1.51E-08	4	6	4.89E-09	7.33E-09	13	14	1.59E-08	1.71E-08
24	Liver	28-Days	100	100	101	2.59E-07	2.62E-07	26	26	6.74E-08	6.74E-08	64	66	1.66E-07	1.71E-07	35	35	6.81E-08	6.81E-08	22	22	4.28E-08	4.28E-08	13	13	2.53E-08	2.53E-08	4	4	4.45E-09	4.45E-09	20	21	2.22E-08	2.33E-08

Appendix D2 Table 8: Subtype specific mutagenicity data of MutaMouse males post exposure to BbF for 90-days

Animal ID	Sample Tissue	Exposure Duration	BbF dose (mg/kg/day)	C:G>A:T Count	C:G>A:T	C:G>A:T MFmin	C:G>A:T MFmax	C:G>G:C Count	C:G>G:C	C:G>G:C MFmin	C:G>G:C MFmax	C:G>T:A Count	C:G>T:A	C:G>T:A MFmin	C:G>T:A MFmax	T:A>A:T Count	T:A>A:T Count	T:A>A:T MFmin	T:A>A:T MFmax	T:A>C:G Count	T:A>C:G	T:A>C:G MFmin	T:A>C:G MFmax	T:A>G:C Count	T:A>G:C	T:A>G:C MFmin	T:A>G:C MFmax	Insertion Count	Insertion Count	Insertion MFmin	Insertion MFmax	Deletion Count	Deletion Count	Deletion MFmin	Deletion MFmax
25	Bone Marrow	90-Days	0	2	2	8.28E-09	8.28E-09	1	1	4.14E-09	4.14E-09	3	3	1.24E-08	1.24E-08	0	0	0	0	1	1	3.08E-09	3.08E-09	1	1	3.08E-09	3.08E-09	1	1	1.77E-09	1.77E-09	7	7	1.24E-08	1.24E-08
26	Bone Marrow	90-Days	0	8	8	2.53E-08	2.53E-08	4	4	1.26E-08	1.26E-08	13	20	4.11E-08	6.32E-08	3	3	7.16E-09	7.16E-09	7	7	1.67E-08	1.67E-08	2	2	4.78E-09	4.78E-09	1	1	1.36E-09	1.36E-09	3	3	4.08E-09	4.08E-09
27	Bone Marrow	90-Days	0	3	3	1.00E-08	1.00E-08	0	0	0	0	4	9	1.34E-08	3.00E-08	2	2	4.99E-09	4.99E-09	2	2	4.99E-09	4.99E-09	2	2	4.99E-09	4.99E-09	1	1	1.43E-09	1.43E-09	2	2	2.86E-09	2.86E-09
28	Bone Marrow	90-Days	0	2	2	6.71E-09	6.71E-09	0	0	0	0	14	14	4.70E-08	4.70E-08	2	2	5.05E-09	5.05E-09	3	3	7.58E-09	7.58E-09	4	4	1.01E-08	1.01E-08	1	1	1.44E-09	1.44E-09	4	4	5.76E-09	5.76E-09
29	Bone Marrow	90-Days	3.125	15	16	4.11E-08	4.39E-08	4	12	1.10E-08	3.29E-08	11	17	3.01E-08	4.66E-08	4	4	8.17E-09	8.17E-09	8	8	1.63E-08	1.63E-08	3	3	6.13E-09	6.13E-09	2	4	2.34E-09	4.68E-09	5	5	5.85E-09	5.85E-09
30	Bone Marrow	90-Days	3.125	8	9	2.35E-08	2.64E-08	2	2	5.87E-09	5.87E-09	12	21	3.52E-08	6.17E-08	6	6	1.33E-08	1.33E-08	11	12	2.44E-08	2.66E-08	4	4	8.88E-09	8.88E-09	0	0	0	0	2	2	2.53E-09	2.53E-09
31	Bone Marrow	90-Days	3.125	6	6	2.87E-08	2.87E-08	1	1	4.78E-09	4.78E-09	3	3	1.43E-08	1.43E-08	0	0	0	0	2	2	7.10E-09	7.10E-09	6	6	2.13E-08	2.13E-08	1	1	2.04E-09	2.04E-09	1	1	2.04E-09	2.04E-09
32	Bone Marrow	90-Days	3.125	12	17	2.68E-08	3.79E-08	2	2	4.46E-09	4.46E-09	9	11	2.01E-08	2.45E-08	3	3	5.07E-09	5.07E-09	6	7	1.01E-08	1.18E-08	3	3	5.07E-09	5.07E-09	1	2	9.62E-10	1.92E-09	2	2	1.92E-09	1.92E-09
33	Bone Marrow	90-Days	6.25	26	28	6.11E-08	6.58E-08	4	4	9.41E-09	9.41E-09	15	23	3.53E-08	5.41E-08	4	4	7.16E-09	7.16E-09	5	5	8.95E-09	8.95E-09	2	2	3.58E-09	3.58E-09	3	3	3.05E-09	3.05E-09	5	5	5.08E-09	5.08E-09
34	Bone Marrow	90-Days	6.25	27	30	5.04E-08	5.60E-08	4	4	7.47E-09	7.47E-09	14	27	2.61E-08	5.04E-08	4	4	5.66E-09	5.66E-09	3	3	4.24E-09	4.24E-09	7	7	9.90E-09	9.90E-09	1	1	8.05E-10	8.05E-10	5	7	4.02E-09	5.63E-09
35	Bone Marrow	90-Days	6.25	33	46	9.42E-08	1.31E-07	3	4	8.57E-09	1.14E-08	13	14	3.71E-08	4.00E-08	12	17	2.59E-08	3.67E-08	9	15	1.94E-08	3.24E-08	4	4	8.64E-09	8.64E-09	2	3	2.46E-09	3.69E-09	1	1	1.23E-09	1.23E-09
36	Bone Marrow	90-Days	6.25	13	14	5.19E-08	5.59E-08	2	2	7.98E-09	7.98E-09	9	9	3.59E-08	3.59E-08	4	4	1.19E-08	1.19E-08	5	6	1.49E-08	1.78E-08	3	3	8.92E-09	8.92E-09	2	3	3.41E-09	5.11E-09	4	4	6.81E-09	6.81E-09
37	Bone Marrow	90-Days	12.5	22	28	6.76E-08	8.60E-08	5	5	1.54E-08	1.54E-08	19	28	5.84E-08	8.60E-08	8	9	1.86E-08	2.09E-08	9	14	2.09E-08	3.25E-08	3	3	6.97E-09	6.97E-09	1	1	1.32E-09	1.32E-09	2	2	2.65E-09	2.65E-09
38	Bone Marrow	90-Days	12.5	41	112	1.20E-07	3.27E-07	11	12	3.21E-08	3.50E-08	21	43	6.13E-08	1.25E-07	11	11	2.44E-08	2.44E-08	8	10	1.77E-08	2.22E-08	5	5	1.11E-08	1.11E-08	4	4	5.04E-09	5.04E-09	9	14	1.13E-08	1.76E-08
39	Bone Marrow	90-Days	12.5	29	29	8.68E-08	8.68E-08	5	5	1.50E-08	1.50E-08	15	21	4.49E-08	6.29E-08	10	12	2.23E-08	2.68E-08	3	3	6.69E-09	6.69E-09	1	1	2.23E-09	2.23E-09	8	9	1.02E-08	1.15E-08	5	5	6.39E-09	6.39E-09
40	Bone Marrow	90-Days	12.5	47	57	1.33E-07	1.62E-07	10	11	2.84E-08	3.12E-08	24	33	6.82E-08	9.37E-08	11	15	2.36E-08	3.22E-08	8	8	1.72E-08	1.72E-08	7	7	1.50E-08	1.50E-08	3	5	3.67E-09	6.12E-09	2	2	2.45E-09	2.45E-09
41	Bone Marrow	90-Days	25	37	59	1.11E-07	1.76E-07	7	7	2.09E-08	2.09E-08	28	56	8.37E-08	1.67E-07	8	19	1.80E-08	4.27E-08	12	13	2.70E-08	2.92E-08	4	4	8.99E-09	8.99E-09	5	5	6.41E-09	6.41E-09	1	1	1.28E-09	1.28E-09
42	Bone Marrow	90-Days	25	56	89	1.35E-07	2.15E-07	12	22	2.89E-08	5.30E-08	30	35	7.23E-08	8.44E-08	13	13	2.38E-08	2.38E-08	11	20	2.01E-08	3.66E-08	0	0	0	0	0	0	0	0	9	13	9.36E-09	1.35E-08
43	Bone Marrow	90-Days	25	48	66	1.13E-07	1.55E-07	17	51	4.00E-08	1.20E-07	27	28	6.35E-08	6.59E-08	14	24	2.50E-08	4.28E-08	12	14	2.14E-08	2.50E-08	4	4	7.14E-09	7.14E-09	3	3	3.04E-09	3.04E-09	13	14	1.32E-08	1.42E-08
44	Bone Marrow	90-Days	25	35	67	1.35E-07	2.58E-07	5	5	1.92E-08	1.92E-08	18	23	6.92E-08	8.85E-08	9	16	2.57E-08	4.56E-08	8	12	2.28E-08	3.42E-08	3	3	8.55E-09	8.55E-09	1	1	1.64E-09	1.64E-09	8	15	1.31E-08	2.46E-08
45	Bone Marrow	90-Days	50	69	97	2.00E-07	2.81E-07	17	19	4.92E-08	5.50E-08	29	39	8.40E-08	1.13E-07	14	14	3.04E-08	3.04E-08	14	18	3.04E-08	3.91E-08	10	10	2.17E-08	2.17E-08	2	2	2.48E-09	2.48E-09	9	20	1.12E-08	2.48E-08
46	Bone Marrow	90-Days	50	37	46	1.30E-07	1.62E-07	18	19	6.32E-08	6.67E-08	20	22	7.02E-08	7.72E-08	11	25	2.92E-08	6.63E-08	8	10	2.12E-08	2.65E-08	5	5	1.33E-08	1.33E-08	2	2	3.02E-09	3.02E-09	7	13	1.06E-08	1.96E-08
47	Bone Marrow	90-Days	50	65	138	1.55E-07	3.30E-07	20	32	4.78E-08	7.64E-08	31	70	7.40E-08	1.67E-07	29	52	5.24E-08	9.40E-08	17	20	3.07E-08	3.62E-08	7	7	1.27E-08	1.27E-08	4	4	4.12E-09	4.12E-09	7	10	7.20E-09	1.03E-08
48	Bone Marrow	90-Days	50	57	71	1.60E-07	2.00E-07	9	11	2.53E-08	3.10E-08	22	64	6.19E-08	1.80E-07	15	31	3.18E-08	6.58E-08	13	16	2.76E-08	3.40E-08	8	9	1.70E-08	1.91E-08	2	3	2.42E-09	3.63E-09	3	7	3.63E-09	8.47E-09
25	Liver	90-Days	0	4	5	1.23E-08	1.53E-08	1	1	3.07E-09	3.07E-09	6	6	1.84E-08	1.84E-08	3	3	7.08E-09	7.08E-09	3	3	7.08E-09	7.08E-09	0	0	0	0	0	0	0	0	3	4	4.00E-09	5.34E-09
26	Liver	90-Days	0	4	4	1.30E-08	1.30E-08	3	3	9.74E-09	9.74E-09	10	13	3.25E-08	4.22E-08	5	5	1.21E-08	1.21E-08	10	10	2.41E-08	2.41E-08	1	1	2.41E-09	2.41E-09	2	2	2.77E-09	2.77E-09	4	4	5.54E-09	5.54E-09
27	Liver	90-Days	0	3	3	1.12E-08	1.12E-08	1	1	3.74E-09	3.74E-09	7	11	2.62E-08	4.12E-08	5	5	1.39E-08	1.39E-08	3	3	8.34E-09	8.34E-09	4	10	1.11E-08	2.78E-08	2	2	3.19E-09	3.19E-09	10	25	1.59E-08	3.99E-08
28	Liver	90-Days	0	5	5	1.73E-08	1.73E-08	1	2	3.47E-09	6.93E-09	7	7	2.43E-08	2.43E-08	3	3	7.71E-09	7.71E-09	9	9	2.31E-08	2.31E-08	2	2	5.14E-09	5.14E-09	1	1	1.48E-09	1.48E-09	5	14	7.38E-09	2.07E-08
29	Liver	90-Days	3.125	12	12	3.78E-08	3.78E-08	3	3	9.46E-09	9.46E-09	13	24	4.10E-08	7.57E-08	10	10	2.35E-08	2.35E-08	9	9	2.11E-08	2.11E-08	4	4	9.38E-09	9.38E-09	4	4	5.38E-09	5.38E-09	5	19	6.73E-09	2.56E-08
30	Liver	90-Days	3.125	5	5	1.66E-08	1.66E-08	2	2	6.65E-09	6.65E-09	13	13	4.32E-08	4.32E-08	4	4	9.90E-09	9.90E-09	4	4	9.90E-09	9.90E-09	1	1	2.47E-09	2.47E-09	0	0	0	0	2	2	2.84E-09	2.84E-09
31	Liver	90-Days	3.125	8	8	2.19E-08	2.19E-08	4	4	1.10E-08	1.10E-08	13	13	3.56E-08	3.56E-08	3	3	6.19E-09	6.19E-09	6	7	1.24E-08	1.44E-08	5	6	1.03E-08	1.24E-08	3	4	3.53E-09	4.71E-09	0	0	0	0
32	Liver	90-Days	3.125	6	6	2.22E-08	2.22E-08	3	3	1.11E-08	1.11E-08	7	7	2.59E-08	2.59E-08	5	5	1.38E-08	1.38E-08	5	5	1.38E-08	1.38E-08	5	5	1.38E-08	1.38E-08	0	0	0	0	4	4	6.32E-09	6.32E-09
33	Liver	90-Days	6.25	26	26	8.78E-08	8.78E-08	4	4	1.35E-08	1.35E-08	10	13	3.38E-08	4.39E-08	6	6	1.51E-08	1.51E-08	6	6	1.51E-08	1.51E-08	1	1	2.51E-09	2.51E-09	1	1	1.44E-09	1.44E-09	2	3	2.88E-09	4.32E-09
34	Liver	90-Days	6.25	21	21	7.37E-08	7.37E-08	5	5	1.76E-08	1.76E-08	11	11	3.86E-08	3.86E-08	7	7	1.83E-08	1.83E-08	7	7	1.83E-08	1.83E-08	6	6	1.57E-08	1.57E-08	2	2	3.00E-09	3.00E-09	7	13	1.05E-08	1.95E-08
35	Liver	90-Days	6.25	21	21	9.56E-08	9.56E-08	5	8	2.28E-08	3.64E-08	18	20	8.19E-08	9.10E-08	5	5	1.68E-08	1.68E-08	8	8	2.69E-08	2.69E-08	4	4	1.34E-08	1.34E-08	1	1	1.93E-09	1.93E-09	9	93	1.74E-08	1.80E-07
36	Liver	90-Days	6.25	17	17	5.30E-08	5.30E-08	3	3	9.36E-09	9.36E-09	15	16	4.68E-08	4.99E-08	4	4	9.56E-09	9.56E-09	5	5	1.20E-08	1.20E-08	5	6	1.20E-08	1.43E-08	0	0	0	0	7	7	9.47E-09	9.47E-09
37	Liver	90-Days	12.5	39	39	1.35E-07	1.35E-07	9	9	3.11E-08	3.11E-08	18	43	6.22E-08	1.49E-07	8	8	2.07E-08	2.07E-08	9	82	2.33E-08	2.12E-07	6	6	1.55E-08	1.55E-08	0	0	0	0	3	4	4.44E-09	5.92E-09
38	Liver	90-Days	12.5	42	42	1.57E-07	1.57E-07	7	7	2.62E-08	2.62E-08	20	20	7.50E-08	7.50E-08	14	14	3.89E-08	3.89E-08	10	10	2.78E-08	2.78E-08	3	3	8.34E-09	8.34E-09	2	2	3.19E-09	3.19E-09	8	25	1.28E-08	3.99E-08
39	Liver	90-Days	12.5	50	50	1.13E-07	1.13E-07	12	12	2.70E-08	2.70E-08	31	32	6.99E-08	7.21E-08	14	14	2.39E-08	2.39E-08	12	13	2.05E-08	2.22E-08	3	3	5.12E-09	5.12E-09	1	1	9.71E-10	9.71E-10	4	4	3.88E-09	3.88E-09
40	Liver	90-Days	12.5	41	42	1.31E-07	1.34E-07	17	17	5.43E-08	5.43E-08	27	27	8.62E-08	8.62E-08	20	20	4.78E-08	4.78E-08	6	6	1.43E-08	1.43E-08	4	4	9.56E-09	9.56E-09	1	1	1.37E-09	1.37E-09	10	10	1.37E-08	1.37E-08
41	Liver	90-Days	25	119	120	4.12E-07	4.16E-07	21	21	7.28E-08	7.28E-08	41	43	1.42E-07	1.49E-07	22	22	5.74E-08	5.74E-08	14	14	3.65E-08	3.65E-08	4	4	1.04E-08	1.04E-08	4	4	5.95E-09	5.95E-09	13	20	1.94E-08	2.98E-08
42	Liver	90-Days	25	113	134	4.14E-07	4.91E-07	22	22	8.06E-08	8.06E-08	42	43	1.54E-07	1.58E-07	33	33	8.99E-08	8.99E-08	12	12	3.27E-08	3.27E-08	6	6	1.64E-08	1.64E-08	2	2	3.13E-09	3.13E-09	5	11	7.81E-09	1.72E-08
43	Liver	90-Days	25	124	125	4.34E-07	4.37E-07	36	36	1.26E-07	1.26E-07	43	44	1.50E-07	1.54E-07	39	39	1.03E-07	1.03E-07	17	17	4.47E-08	4.47E-08	2	2	5.26E-09	5.26E-09	1	1	1.50E-09	1.50E-09	7	7	1.05E-08	1.05E-08
44	Liver	90-Days	25	165	166	6.05E-07	6.09E-07	34	35	1.25E-07	1.28E-07	49	50	1.80E-07	1.83E-07	32	32	8.86E-08	8.86E-08	11	11	3.05E-08	3.05E-08	12	12	3.32E-08	3.32E-08	3	4	4.73E-09	6.31E-09	11	31	1.74E-08	4.89E-08
45	Liver	90-Days	50	428	443	1.44E-06	1.49E-06	97	99	3.26E-07	3.33E-07	135	136	4.54E-07	4.57E-07	82	82	2.09E-07	2.09E-07	27	27	6.87E-08	6.87E-08	10	10	2.54E-08	2.54E-08	5	5	7.24E-09	7.24E-09	23	23	3.33E-08	3.33E-08
46	Liver	90-Days	50	224	227	8.82E-07	8.94E-07	41	41	1.61E-07	1.61E-07	74	75	2.91E-07	2.95E-07	53	55	1.55E-07	1.61E-07	24	27	7.03E-08	7.91E-08	10	10	2.93E-08	2.93E-08	1	2	1.68E-09	3.36E-09	25	36	4.20E-08	6.05E-08
47	Liver	90-Days	50	287	297	9.87E-07	1.02E-06	57	57	1.96E-07	1.96E-07	80	80	2.75E-07	2.75E-07	72	73	1.85E-07	1.88E-07	26	27	6.70E-08	6.96E-08	8	8	2.06E-08	2.06E-08	2	2	2.95E-09	2.95E-09	16	17	2.36E-08	2.50E-08
48	Liver	90-Days	50	198	202	7.60E-07	7.76E-07	35	35	1.34E-07	1.34E-07	60	60	2.30E-07	2.30E-07	39	40	1.13E-07	1.16E-07	25	25	7.23E-08	7.23E-08	14	14	4.05E-08	4.05E-08	1	1	1.65E-09	1.65E-09	17	18	2.81E-08	2.97E-08

Appendix D2 Table 9: Subtype specific mutagenicity data of MutaMouse males post exposure to BbF for 180-days

Animal ID	Sample Tissue	Exposure Duration	BbF dose (mg/kg/day)	C:G>A:T Count	C:G>A: T	C:G>A:T MFmin	C:G>A:T MFmax	C:G>G:C Count	C:G>G: C	C:G>G:C MFmin	C:G>G:C MFmax	C:G>T:A Count	C:G>T: A	C:G>T:A MFmin	C:G>T:A MFmax	T:A>A:T Count	T:A>A: T	T:A>A:T MFmin	T:A>A:T MFmax	T:A>C:G Count	T:A>C: G	T:A>C:G MFmin	T:A>C:G MFmax	T:A>G:C Count	T:A>G: C	T:A>G:C MFmin	T:A>G:C MFmax	Insertion Count	Insertio n	Insertion MFmin	Insertion MFmax	Deletion Count	Deletio n	Deletion MFmin	Deletion MFmax
49	Bone Marrow	180-Days	0	6	9	1.32E-08	1.98E-08	1	1	2.20E-09	2.20E-09	14	15	3.08E-08	3.30E-08	1	1	1.70E-09	1.70E-09	12	12	2.04E-08	2.04E-08	4	4	6.80E-09	6.80E-09	2	2	1.92E-09	1.92E-09	3	3	2.88E-09	2.88E-09
50	Bone Marrow	180-Days	0	12	12	2.29E-08	2.29E-08	3	3	5.72E-09	5.72E-09	15	18	2.86E-08	3.43E-08	4	4	5.85E-09	5.85E-09	9	9	1.32E-08	1.32E-08	3	5	4.39E-09	7.31E-09	2	2	1.66E-09	1.66E-09	5	11	4.14E-09	9.10E-09
51	Bone Marrow	180-Days	0	3	3	7.30E-09	7.30E-09	2	2	4.87E-09	4.87E-09	7	12	1.70E-08	2.92E-08	7	7	1.31E-08	1.31E-08	5	8	9.36E-09	1.50E-08	3	3	5.62E-09	5.62E-09	1	1	1.06E-09	1.06E-09	1	1	1.06E-09	1.06E-09
52	Bone Marrow	180-Days	0	10	11	1.98E-08	2.18E-08	1	1	1.98E-09	1.98E-09	8	8	1.58E-08	1.58E-08	5	6	7.57E-09	9.09E-09	14	22	2.12E-08	3.33E-08	5	6	7.57E-09	9.09E-09	4	4	3.43E-09	3.43E-09	5	5	4.29E-09	4.29E-09
53	Bone Marrow	180-Days	1.56	13	13	2.51E-08	2.51E-08	0	0	0	0	13	32	2.51E-08	6.19E-08	5	5	7.45E-09	7.45E-09	3	3	4.47E-09	4.47E-09	6	8	8.94E-09	1.19E-08	4	5	3.37E-09	4.21E-09	10	10	8.41E-09	8.41E-09
54	Bone Marrow	180-Days	1.56	9	9	1.58E-08	1.58E-08	3	3	5.26E-09	5.26E-09	11	11	1.93E-08	1.93E-08	4	6	5.38E-09	8.07E-09	10	12	1.34E-08	1.61E-08	7	7	9.41E-09	9.41E-09	3	3	2.28E-09	2.28E-09	5	5	3.81E-09	3.81E-09
55	Bone Marrow	180-Days	1.56	13	13	2.54E-08	2.54E-08	0	0	0	0	16	18	3.13E-08	3.52E-08	6	6	8.96E-09	8.96E-09	14	15	2.09E-08	2.24E-08	1	1	1.49E-09	1.49E-09	3	4	2.54E-09	3.39E-09	9	11	7.62E-09	9.32E-09
56	Bone Marrow	180-Days	1.56	14	14	2.65E-08	2.65E-08	1	1	1.89E-09	1.89E-09	12	12	2.27E-08	2.27E-08	5	5	7.28E-09	7.28E-09	16	16	2.33E-08	2.33E-08	9	9	1.31E-08	1.31E-08	2	2	1.64E-09	1.64E-09	4	7	3.29E-09	5.76E-09
57	Bone Marrow	180-Days	3.125	25	27	4.30E-08	4.65E-08	3	12	5.17E-09	2.07E-08	18	37	3.10E-08	6.37E-08	10	13	1.33E-08	1.72E-08	10	14	1.33E-08	1.86E-08	10	32	1.33E-08	4.24E-08	2	2	1.50E-09	1.50E-09	5	5	3.75E-09	3.75E-09
58	Bone Marrow	180-Days	3.125	18	22	3.40E-08	4.15E-08	2	3	3.78E-09	5.66E-09	17	18	3.21E-08	3.40E-08	9	9	1.32E-08	1.32E-08	7	10	1.03E-08	1.47E-08	6	6	8.79E-09	8.79E-09	1	1	8.25E-10	8.25E-10	7	8	5.77E-09	6.60E-09
59	Bone Marrow	180-Days	3.125	24	27	4.13E-08	4.65E-08	3	3	5.16E-09	5.16E-09	22	29	3.79E-08	4.99E-08	14	16	1.86E-08	2.12E-08	16	17	2.12E-08	2.26E-08	11	11	1.46E-08	1.46E-08	1	1	7.50E-10	7.50E-10	5	5	3.75E-09	3.75E-09
60	Bone Marrow	180-Days	3.125	16	20	2.87E-08	3.58E-08	4	4	7.17E-09	7.17E-09	18	26	3.23E-08	4.66E-08	4	4	5.47E-09	5.47E-09	16	18	2.19E-08	2.46E-08	8	8	1.09E-08	1.09E-08	4	4	3.10E-09	3.10E-09	6	6	4.65E-09	4.65E-09
61	Bone Marrow	180-Days	6.25	39	44	7.15E-08	8.07E-08	8	9	1.47E-08	1.65E-08	18	19	3.30E-08	3.48E-08	10	61	1.41E-08	8.61E-08	12	12	1.69E-08	1.69E-08	5	5	7.06E-09	7.06E-09	0	0	0	0	10	10	7.97E-09	7.97E-09
62	Bone Marrow	180-Days	6.25	47	72	9.73E-08	1.49E-07	8	8	1.66E-08	1.66E-08	20	26	4.14E-08	5.38E-08	16	17	2.54E-08	2.70E-08	6	8	9.53E-09	1.27E-08	5	5	7.94E-09	7.94E-09	4	4	3.59E-09	3.59E-09	7	7	6.29E-09	6.29E-09
63	Bone Marrow	180-Days	6.25	46	68	9.37E-08	1.39E-07	6	9	1.22E-08	1.83E-08	19	24	3.87E-08	4.89E-08	10	16	1.56E-08	2.50E-08	10	20	1.56E-08	3.12E-08	4	5	6.25E-09	7.81E-09	3	5	2.65E-09	4.42E-09	3	3	2.65E-09	2.65E-09
64	Bone Marrow	180-Days	6.25	52	95	8.31E-08	1.52E-07	16	16	2.56E-08	2.56E-08	27	59	4.32E-08	9.43E-08	14	15	1.72E-08	1.84E-08	16	21	1.97E-08	2.58E-08	14	15	1.72E-08	1.84E-08	4	4	2.78E-09	2.78E-09	9	34	6.25E-09	2.36E-08
65	Bone Marrow	180-Days	12.5	70	156	1.45E-07	3.22E-07	8	8	1.65E-08	1.65E-08	27	46	5.57E-08	9.50E-08	21	28	3.33E-08	4.44E-08	16	17	2.54E-08	2.70E-08	11	13	1.75E-08	2.06E-08	2	2	1.79E-09	1.79E-09	8	8	7.18E-09	7.18E-09
66	Bone Marrow	180-Days	12.5	60	139	1.15E-07	2.66E-07	8	11	1.53E-08	2.11E-08	33	55	6.32E-08	1.05E-07	17	20	2.50E-08	2.95E-08	17	22	2.50E-08	3.24E-08	4	4	5.89E-09	5.89E-09	4	6	3.33E-09	5.00E-09	11	11	9.16E-09	9.16E-09
67	Bone Marrow	180-Days	12.5	64	85	1.31E-07	1.74E-07	13	14	2.65E-08	2.86E-08	41	51	8.37E-08	1.04E-07	22	32	3.48E-08	5.06E-08	19	19	3.00E-08	3.00E-08	7	7	1.11E-08	1.11E-08	6	7	5.35E-09	6.24E-09	5	5	4.46E-09	4.46E-09
68	Bone Marrow	180-Days	12.5	88	128	1.64E-07	2.38E-07	18	20	3.35E-08	3.72E-08	40	49	7.45E-08	9.12E-08	16	17	2.31E-08	2.46E-08	20	32	2.89E-08	4.62E-08	9	10	1.30E-08	1.44E-08	2	3	1.63E-09	2.44E-09	11	13	8.95E-09	1.06E-08
69	Bone Marrow	180-Days	25	105	221	2.03E-07	4.28E-07	15	75	2.90E-08	1.45E-07	39	95	7.55E-08	1.84E-07	26	85	3.88E-08	1.27E-07	17	19	2.54E-08	2.84E-08	21	21	3.14E-08	3.14E-08	1	1	8.43E-10	8.43E-10	16	21	1.35E-08	1.77E-08
70	Bone Marrow	180-Days	25	108	197	1.98E-07	3.61E-07	25	68	4.58E-08	1.25E-07	50	65	9.17E-08	1.19E-07	25	37	3.48E-08	5.15E-08	26	52	3.62E-08	7.24E-08	8	9	1.11E-08	1.25E-08	5	5	3.96E-09	3.96E-09	20	31	1.58E-08	2.45E-08
71	Bone Marrow	180-Days	25	78	255	1.47E-07	4.79E-07	19	60	3.57E-08	1.13E-07	40	104	7.52E-08	1.95E-07	28	47	4.05E-08	6.79E-08	18	21	2.60E-08	3.04E-08	9	12	1.30E-08	1.73E-08	1	1	8.17E-10	8.17E-10	11	29	8.99E-09	2.37E-08
72	Bone Marrow	180-Days	25	102	242	1.90E-07	4.51E-07	22	28	4.10E-08	5.21E-08	34	45	6.33E-08	8.38E-08	27	32	3.88E-08	4.59E-08	17	19	2.44E-08	2.73E-08	3	6	4.31E-09	8.61E-09	3	3	2.43E-09	2.43E-09	10	11	8.11E-09	8.92E-09
49	Liver	180-Days	0	13	13	2.26E-08	2.26E-08	3	3	5.21E-09	5.21E-09	27	29	4.69E-08	5.04E-08	6	6	7.86E-09	7.86E-09	11	11	1.44E-08	1.44E-08	7	7	9.17E-09	9.17E-09	4	4	2.99E-09	2.99E-09	5	5	3.74E-09	3.74E-09
50	Liver	180-Days	0	7	7	1.99E-08	1.99E-08	1	1	2.85E-09	2.85E-09	12	12	3.42E-08	3.42E-08	5	5	1.07E-08	1.07E-08	5	5	1.07E-08	1.07E-08	3	7	6.43E-09	1.50E-08	2	2	2.45E-09	2.45E-09	3	4	3.67E-09	4.89E-09
51	Liver	180-Days	0	10	10	2.37E-08	2.37E-08	2	2	4.74E-09	4.74E-09	8	9	1.89E-08	2.13E-08	9	9	1.60E-08	1.60E-08	9	10	1.60E-08	1.78E-08	3	9	5.34E-09	1.60E-08	2	2	2.03E-09	2.03E-09	4	5	4.06E-09	5.08E-09
52	Liver	180-Days	0	19	20	4.40E-08	4.63E-08	2	2	4.63E-09	4.63E-09	18	18	4.17E-08	4.17E-08	9	9	1.57E-08	1.57E-08	14	15	2.44E-08	2.62E-08	4	5	6.98E-09	8.73E-09	4	5	3.98E-09	4.98E-09	6	6	5.97E-09	5.97E-09
53	Liver	180-Days	1.56	10	13	2.28E-08	2.96E-08	4	4	9.12E-09	9.12E-09	24	26	5.47E-08	5.93E-08	7	8	1.19E-08	1.36E-08	7	7	1.19E-08	1.19E-08	8	8	1.36E-08	1.36E-08	5	6	4.88E-09	5.85E-09	8	8	7.80E-09	7.80E-09
54	Liver	180-Days	1.56	7	7	1.65E-08	1.65E-08	2	2	4.73E-09	4.73E-09	18	18	4.25E-08	4.25E-08	10	10	1.77E-08	1.77E-08	7	7	1.24E-08	1.24E-08	3	3	5.32E-09	5.32E-09	3	4	3.04E-09	4.05E-09	5	5	5.07E-09	5.07E-09
55	Liver	180-Days	1.56	14	14	3.51E-08	3.51E-08	3	3	7.53E-09	7.53E-09	23	24	5.77E-08	6.02E-08	8	8	1.52E-08	1.52E-08	9	12	1.71E-08	2.28E-08	2	2	3.80E-09	3.80E-09	1	1	1.08E-09	1.08E-09	2	2	2.16E-09	2.16E-09
56	Liver	180-Days	1.56	16	16	3.39E-08	3.39E-08	4	4	8.48E-09	8.48E-09	10	10	2.12E-08	2.12E-08	7	7	1.12E-08	1.12E-08	9	11	1.44E-08	1.76E-08	2	2	3.19E-09	3.19E-09	4	4	3.64E-09	3.64E-09	8	10	7.29E-09	9.11E-09
57	Liver	180-Days	3.125	12	12	2.31E-08	2.31E-08	4	4	7.70E-09	7.70E-09	24	24	4.62E-08	4.62E-08	10	10	1.45E-08	1.45E-08	9	9	1.31E-08	1.31E-08	4	4	5.81E-09	5.81E-09	3	3	2.48E-09	2.48E-09	3	4	2.48E-09	3.31E-09
58	Liver	180-Days	3.125	14	19	3.33E-08	4.51E-08	8	8	1.90E-08	1.90E-08	11	11	2.61E-08	2.61E-08	10	10	1.79E-08	1.79E-08	7	20	1.26E-08	3.59E-08	6	6	1.08E-08	1.08E-08	3	3	3.07E-09	3.07E-09	7	8	7.15E-09	8.18E-09
59	Liver	180-Days	3.125	11	42	2.40E-08	9.17E-08	4	4	8.73E-09	8.73E-09	23	23	5.02E-08	5.02E-08	13	13	2.13E-08	2.13E-08	16	18	2.62E-08	2.95E-08	4	4	6.56E-09	6.56E-09	1	1	9.36E-10	9.36E-10	7	7	6.55E-09	6.55E-09
60	Liver	180-Days	3.125	15	15	3.55E-08	3.55E-08	4	4	9.48E-09	9.48E-09	17	17	4.03E-08	4.03E-08	8	8	1.43E-08	1.43E-08	10	10	1.78E-08	1.78E-08	2	2	3.57E-09	3.57E-09	2	2	2.03E-09	2.03E-09	2	2	2.03E-09	2.03E-09
61	Liver	180-Days	6.25	46	46	1.18E-07	1.18E-07	12	12	3.08E-08	3.08E-08	24	24	6.16E-08	6.16E-08	14	14	2.70E-08	2.70E-08	20	20	3.86E-08	3.86E-08	5	5	9.65E-09	9.65E-09	3	3	3.30E-09	3.30E-09	6	6	6.61E-09	6.61E-09
62	Liver	180-Days	6.25	35	35	9.25E-08	9.25E-08	7	7	1.85E-08	1.85E-08	21	23	5.55E-08	6.08E-08	18	18	3.56E-08	3.56E-08	13	13	2.57E-08	2.57E-08	2	2	3.96E-09	3.96E-09	4	5	4.53E-09	5.66E-09	3	4	3.40E-09	4.53E-09
63	Liver	180-Days	6.25	58	58	1.41E-07	1.41E-07	8	8	1.95E-08	1.95E-08	37	56	9.01E-08	1.36E-07	13	13	2.38E-08	2.38E-08	13	13	2.38E-08	2.38E-08	7	7	1.28E-08	1.28E-08	1	1	1.04E-09	1.04E-09	11	11	1.15E-08	1.15E-08
64	Liver	180-Days	6.25	61	61	1.45E-07	1.45E-07	11	11	2.61E-08	2.61E-08	25	25	5.92E-08	5.92E-08	11	13	1.98E-08	2.33E-08	12	12	2.15E-08	2.15E-08	1	1	1.80E-09	1.80E-09	6	6	6.13E-09	6.13E-09	7	8	7.15E-09	8.17E-09
65	Liver	180-Days	12.5	167	169	4.22E-07	4.28E-07	35	36	8.85E-08	9.11E-08	62	62	1.57E-07	1.57E-07	44	45	8.39E-08	8.58E-08	18	21	3.43E-08	4.00E-08	9	9	1.72E-08	1.72E-08	6	6	6.52E-09	6.52E-09	15	15	1.63E-08	1.63E-08
66	Liver	180-Days	12.5	163	166	3.30E-07	3.36E-07	38	38	7.70E-08	7.70E-08	74	74	1.50E-07	1.50E-07	63	64	9.60E-08	9.76E-08	28	28	4.27E-08	4.27E-08	9	9	1.37E-08	1.37E-08	6	7	5.22E-09	6.09E-09	16	16	1.39E-08	1.39E-08
67	Liver	180-Days	12.5	224	232	5.54E-07	5.74E-07	44	45	1.09E-07	1.11E-07	77	79	1.91E-07	1.96E-07	66	69	1.22E-07	1.28E-07	18	19	3.34E-08	3.52E-08	9	10	1.67E-08	1.86E-08	2	2	2.12E-09	2.12E-09	22	23	2.33E-08	2.44E-08
68	Liver	180-Days	12.5	144	146	3.73E-07	3.79E-07	29	29	7.52E-08	7.52E-08	58	58	1.50E-07	1.50E-07	53	53	1.03E-07	1.03E-07	13	13	2.52E-08	2.52E-08	5	5	9.69E-09	9.69E-09	5	5	5.55E-09	5.55E-09	13	13	1.44E-08	1.44E-08
69	Liver	180-Days	25	197	200	1.22E-06	1.24E-06	40	40	2.48E-07	2.48E-07	62	63	3.84E-07	3.91E-07	54	54	2.55E-07	2.55E-07	17	18	8.04E-08	8.52E-08	6	10	2.84E-08	4.73E-08	4	4	1.07E-08	1.07E-08	12	12	3.22E-08	3.22E-08
70	Liver	180-Days	25	603	644	1.26E-06	1.34E-06	124	125	2.59E-07	2.61E-07	177	181	3.69E-07	3.77E-07	135	138	2.12E-07	2.17E-07	52	52	8.17E-08	8.17E-08	19	20	2.98E-08	3.14E-08	6	6	5.37E-09	5.37E-09	48	49	4.30E-08	4.39E-08
71	Liver	180-Days	25	327	332	9.63E-07	9.78E-07	45	48	1.33E-07	1.41E-07	98	100	2.89E-07	2.95E-07	79	80	1.75E-07	1.77E-07	26	26	5.74E-08	5.74E-08	9	9	1.99E-08	1.99E-08	4	4	5.05E-09	5.05E-09	30	33	3.79E-08	4.17E-08
72	Liver	180-Days	25	521	584	1.02E-06	1.15E-06	94	94	1.85E-07	1.85E-07	141	142	2.77E-07	2.79E-07	119	122	1.76E-07	1.81E-07	49	49	7.26E-08	7.26E-08	27	27	4.00E-08	4.00E-08	7	7	5.92E-09	5.92E-09	42	42	3.55E-08	3.55E-08

Appendix D2 Table 10: Locus specific mutagenicity data of MutaMouse males post exposure to BbF for 28-days

Animal ID	Sample tissue	Exposure duration	BbF dose (mg/kg/day)	Chr. 1 MFmin	Chr. 1 MFmax	Chr. 1.2 MFmin	Chr. 1.2 MFmax	Chr. 2 MFmin	Chr. 2 MFmax	Chr. 3 MFmin	Chr. 3 MFmax	Chr. 4 MFmin	Chr. 4 MFmax	Chr. 5 MFmin	Chr. 5 MFmax	Chr. 6 MFmin	Chr. 6 MFmax	Chr. 7 MFmin	Chr. 7 MFmax	Chr. 8 MFmin	Chr. 8 MFmax	Chr. 9 MFmin	Chr. 9 MFmax	Chr. 10 MFmin	Chr. 10 MFmax	Chr. 11 MFmin	Chr. 11 MFmax	Chr. 12 MFmin	Chr. 12 MFmax	Chr. 13 MFmin	Chr. 13 MFmax	Chr. 14 MFmin	Chr. 14 MFmax	Chr. 15 MFmin	Chr. 15 MFmax	Chr. 16 MFmin	Chr. 16 MFmax	Chr. 17 MFmin	Chr. 17 MFmax	Chr. 18 MFmin	Chr. 18 MFmax	Chr. 19 MFmin	Chr. 19 MFmax
1	Bone Marrow	28-Days	0	7.11E-08	9.48E-08	4.91435E-08	7.37152E-08	5.09738E-08	7.64607E-08	4.80743E-08	4.80743E-08	0	0	0.00E+00	0.00E+00	2.39E-08	2.39E-08	4.40581E-08	4.40581E-08	2.82E-08	2.82E-08	4.83433E-08	4.83433E-08	2.33523E-08	2.33523E-08	7.50E-08	7.50E-08	2.35542E-08	2.35542E-08	7.09783E-08	7.09783E-08	7.43E-08	9.91E-08	0.00E+00	0.00E+00	7.85E-08	7.85E-08	7.8818E-08	1.31363E-07	7.14237E-08	7.14237E-08	4.85E-08	4.85E-08
2	Bone Marrow	28-Days	0	2.04E-08	4.08E-08	4.47E-08	4.47E-08	2.42E-08	2.42E-08	2.23243E-08	2.23243E-08	2.28E-08	2.28E-08	4.50E-08	4.50E-08	0.00E+00	0.00E+00	6.05E-08	6.05E-08	7.86E-08	7.86E-08	2.17E-08	2.17E-08	2.02302E-08	2.02302E-08	6.87E-08	6.87E-08	6.24E-08	6.24E-08	2.09E-08	2.09E-08	2.21E-08	2.21E-08	2.10E-08	4.20E-08	7.31E-08	7.31E-08	1.02E-07	1.02E-07	2.15E-08	2.15E-08	4.37E-08	4.37E-08
3	Bone Marrow	28-Days	0	4.58E-08	4.58E-08	2.38E-08	2.38E-08	7.54089E-08	7.54089E-08	0	0	2.43E-08	2.43E-08	9.42E-08	9.42E-08	0.00E+00	0.00E+00	4.31E-08	4.31E-08	8.20668E-08	1.09422E-07	2.3351E-08	2.3351E-08	0	0	1.70E-07	1.70E-07	2.30E-08	2.30E-08	0.00E+00	0.00E+00	2.42098E-08	2.42098E-08	2.22E-08	2.22E-08	1.55E-07	1.55E-07	0	0	0.00E+00	0.00E+00	2.37937E-08	2.37937E-08
4	Bone Marrow	28-Days	0	9.90707E-08	1.78327E-07	0.00E+00	0.00E+00	0	0	2.1152E-08	2.1152E-08	4.26E-08	4.26E-08	4.21496E-08	4.21496E-08	3.99E-08	3.99E-08	0.00E+00	0.00E+00	4.75E-08	4.75E-08	6.33E-08	6.33E-08	0.00E+00	0.00E+00	4.40E-08	4.40E-08	1.98E-08	1.98E-08	8.05E-08	8.05E-08	6.38E-08	6.38E-08	1.96E-08	1.96E-08	9.32E-08	9.32E-08	6.85E-08	6.85E-08	6.19727E-08	6.19727E-08	2.07552E-08	2.07552E-08
5	Bone Marrow	28-Days	6.25	1.43E-07	2.86E-07	2.78E-08	2.78E-08	2.89E-08	2.89E-08	2.75E-08	2.75E-08	5.88E-08	5.88E-08	0.00E+00	0.00E+00	5.64E-08	5.64E-08	2.57E-08	2.57E-08	1.27E-07	1.27E-07	8.33E-08	8.33E-08	5.59274E-08	5.59274E-08	5.62E-08	8.43E-08	0.00E+00	0.00E+00	0.00E+00	0.00E+00	2.90E-08	2.90E-08	0	0	2.96E-08	2.96E-08	5.70E-08	5.70E-08	5.59E-08	5.59E-08	5.73003E-08	5.73003E-08
6	Bone Marrow	28-Days	6.25	3.69407E-07	5.20864E-06	3.38E-08	3.38E-08	1.09E-07	1.09E-07	6.87275E-08	6.87275E-08	1.05E-07	1.05E-07	6.49E-08	6.49E-08	6.86E-08	6.86E-08	0.00E+00	0.00E+00	1.10E-07	1.10E-07	6.94E-08	6.94E-08	6.88E-08	6.88E-08	7.00E-08	2.45E-07	3.68E-08	3.68E-08	2.18E-07	2.18E-07	3.65E-08	3.65E-08	3.25973E-08	3.25973E-08	1.09E-07	1.09E-07	1.73E-07	2.07E-07	7.12E-08	7.12E-08	7.39E-08	7.39E-08
7	Bone Marrow	28-Days	6.25	6.16575E-08	6.16575E-08	3.07E-08	3.07E-08	3.16E-08	3.16E-08	3.04E-08	3.04E-08	6.37326E-08	6.37326E-08	2.96E-08	2.96E-08	0.00E+00	0.00E+00	5.52E-08	5.52E-08	7.00E-08	7.00E-08	6.02529E-08	6.02529E-08	0	0	6.11945E-08	6.11945E-08	0	0	0.00E+00	0.00E+00	6.36E-08	6.36E-08	0.00E+00	0.00E+00	6.59E-08	6.59E-08	1.59E-07	1.91E-07	6.11414E-08	6.11414E-08	0.00E+00	0.00E+00
8	Bone Marrow	28-Days	6.25	4.64512E-08	6.96768E-08	0.00E+00	0.00E+00	5.91E-08	5.91E-08	0	0	5.52E-08	5.52E-08	8.33E-08	8.33E-08	2.57E-08	2.57E-08	4.77E-08	4.77E-08	9.83E-08	1.31E-07	0.00E+00	0.00E+00	7.07E-08	7.07E-08	1.11E-07	1.11E-07	4.79E-08	4.79E-08	0.00E+00	0.00E+00	5.29E-08	1.85E-07	2.50E-08	2.50E-08	8.80E-08	8.80E-08	9.5504E-08	9.5504E-08	2.5311E-08	2.5311E-08	2.56E-08	5.12E-08
9	Bone Marrow	28-Days	12.5	9.17E-08	1.10E-07	2.05E-08	2.05E-08	0.00E+00	0.00E+00	4.04E-08	4.04E-08	4.22E-08	4.22E-08	4.14E-08	4.14E-08	0.00E+00	0.00E+00	3.61E-08	5.41E-08	1.72E-07	1.72E-07	5.99704E-08	5.99704E-08	1.85E-08	1.85E-08	1.48E-07	1.90E-07	3.69E-08	5.53E-08	9.28E-08	9.28E-08	1.42E-07	1.63E-07	1.88E-08	1.88E-08	1.33E-07	1.33E-07	2.09E-07	2.55E-07	5.73E-08	5.73E-08	3.86E-08	3.86E-08
10	Bone Marrow	28-Days	12.5	2.20E-07	1.24E-06	0.00E+00	0.00E+00	2.72E-08	2.72E-08	1.29E-07	1.29E-07	5.50E-08	5.50E-08	1.54E-07	1.54E-07	2.65E-08	2.65E-08	2.41E-08	2.41E-08	1.48E-07	1.48E-07	7.89E-08	7.89E-08	7.93E-08	7.93E-08	1.87E-07	2.40E-07	1.09E-07	1.09E-07	2.6963E-08	2.6963E-08	1.68E-07	2.51E-07	2.47E-08	2.47E-08	2.26E-07	3.67E-07	1.07E-07	2.93E-07	1.08E-07	1.35E-07	2.73E-08	2.73E-08
11	Bone Marrow	28-Days	12.5	0.00E+00	0.00E+00	2.06E-08	2.06E-08	6.68E-08	8.90E-08	6.24E-08	6.24E-08	4.21E-08	4.21E-08	1.04E-07	1.04E-07	5.96E-08	5.96E-08	0.00E+00	0.00E+00	7.23E-08	7.23E-08	2.24E-07	2.85E-07	7.54511E-08	9.43139E-08	1.91E-07	2.12E-07	1.93E-08	1.93E-08	3.90E-08	3.90E-08	1.25E-07	1.46E-07	1.91E-08	1.91E-08	1.59E-07	1.82E-07	1.39E-07	1.39E-07	3.98E-08	3.98E-08	6.03E-08	6.03E-08
12	Bone Marrow	28-Days	12.5	5.11E-08	7.67E-08	5.71E-08	5.71E-08	0	0	0.00E+00	0.00E+00	5.96E-08	5.96E-08	2.31E-07	2.31E-07	2.77728E-08	2.77728E-08	0.00E+00	0.00E+00	1.04E-07	1.39E-07	2.75E-08	2.75E-08	5.19E-08	5.19E-08	1.16E-07	1.74E-07	2.59E-08	2.59E-08	7.73E-08	7.73E-08	1.40639E-07	1.40639E-07	5.25E-08	5.25E-08	6.14E-08	6.14E-08	1.61E-07	1.61E-07	1.06E-07	1.06E-07	0.00E+00	0.00E+00
13	Bone Marrow	28-Days	25	3.56E-08	3.56E-08	5.92E-08	5.92E-08	4.12E-08	4.12E-08	1.94826E-08	1.94826E-08	1.01E-07	1.01E-07	9.85E-08	9.85E-08	5.61644E-08	5.61644E-08	5.24E-08	5.24E-08	4.63E-08	4.63E-08	9.71E-08	1.55E-07	5.33951E-08	5.33951E-08	1.43E-07	1.43E-07	5.37E-08	7.16E-08	5.38E-08	8.97E-08	1.18E-07	1.18E-07	3.64E-08	3.64E-08	1.50E-07	1.50E-07	4.39E-08	6.58E-08	1.86E-08	1.86E-08	0.00E+00	0.00E+00
14	Bone Marrow	28-Days	25	2.37E-08	4.75E-08	2.79E-08	2.79E-08	6.07E-08	6.07E-08	8.42417E-08	8.42417E-08	5.75E-08	5.75E-08	0.00E+00	0.00E+00	7.94E-08	7.94E-08	4.85E-08	9.70E-08	2.80E-07	2.80E-07	2.66E-08	2.66E-08	0.00E+00	0.00E+00	4.06E-07	4.64E-07	2.45E-08	4.90E-08	2.50E-08	2.50E-08	2.19E-07	3.84E-07	2.56E-08	2.56E-08	1.86E-07	2.16E-07	6.80E-08	6.80E-08	5.18E-08	5.18E-08	5.26E-08	5.26E-08
15	Bone Marrow	28-Days	25	3.51E-08	3.51E-08	3.34E-08	3.34E-08	0.00E+00	0.00E+00	1.67E-07	1.67E-07	3.55E-08	3.55E-08	9.71E-08	9.71E-08	0.00E+00	0.00E+00	3.05E-08	3.05E-08	1.15E-07	1.54E-07	2.36E-07	2.36E-07	6.79E-08	6.79E-08	2.03E-07	2.03E-07	3.48E-08	6.96E-08	6.81E-08	6.81E-08	1.77E-07	1.77E-07	0.00E+00	0.00E+00	1.79E-07	1.79E-07	1.03E-07	1.03E-07	6.83E-08	1.02E-07	3.49E-08	3.49E-08
16	Bone Marrow	28-Days	25	1.11E-07	1.11E-07	8.03E-08	8.03E-08	6.45E-08	6.45E-08	2.00E-08	2.00E-08	3.98E-08	3.98E-08	9.92E-08	1.19E-07	3.68E-08	3.68E-08	7.22E-08	7.22E-08	1.53E-07	1.74E-07	6.05E-08	8.07E-08	3.59E-08	3.59E-08	2.56E-07	2.56E-07	9.29E-08	9.29E-08	3.73E-08	3.73E-08	2.01E-07	2.21E-07	1.86E-08	1.86E-08	2.25E-07	3.16E-07	4.39E-08	6.58E-08	1.15E-07	1.15E-07	5.80E-08	5.80E-08
17	Bone Marrow	28-Days	50	1.22E-07	1.22E-07	8.19E-08	1.09E-07	0.00E+00	0.00E+00	2.71E-08	2.71E-08	8.28E-08	8.28E-08	1.10E-07	1.10E-07	0.00E+00	0.00E+00	4.82E-08	4.82E-08	1.92E-07	2.24E-07	2.63E-07	2.63E-07	7.34E-08	7.34E-08	3.07E-07	5.03E-07	5.00E-08	5.00E-08	2.53E-08	2.53E-08	1.08E-07	1.08E-07	2.51E-08	2.51E-08	2.07E-07	2.07E-07	1.53E-07	1.53E-07	5.15E-08	7.73E-08	5.30E-08	1.06E-07
18	Bone Marrow	28-Days	50	7.28E-08	1.09E-07	1.99E-08	1.99E-08	6.52E-08	6.52E-08	4.00E-08	4.00E-08	1.40E-07	1.40E-07	9.92E-08	9.92E-08	9.29E-08	9.29E-08	1.07E-07	1.25E-07	3.15E-07	3.60E-07	1.39E-07	1.39E-07	5.38E-08	5.38E-08	3.35E-07	3.35E-07	1.67E-07	1.85E-07	9.32E-08	9.32E-08	1.79E-07	2.59E-07	3.74E-08	3.74E-08	2.43E-07	3.09E-07	1.10E-07	1.10E-07	7.68E-08	7.68E-08	7.80E-08	7.80E-08
19	Bone Marrow	28-Days	50	2.77E-08	2.77E-08	8.15E-08	1.63E-07	5.61E-08	5.61E-08	8.15E-08	8.15E-08	2.56E-07	3.41E-07	1.57E-07	1.57E-07	1.63E-07	1.63E-07	7.43E-08	7.43E-08	4.04E-07	5.60E-07	1.08E-07	1.35E-07	8.08E-08	8.08E-08	2.49E-07	3.04E-07	5.50E-08	5.50E-08	2.72E-08	2.72E-08	2.88E-07	2.88E-07	7.57E-08	1.01E-07	2.33E-07	2.33E-07	2.54E-07	2.82E-07	5.51E-08	5.51E-08	2.22E-07	2.22E-07
20	Bone Marrow	28-Days	50	6.46E-08	6.46E-08	1.15E-07	1.15E-07	2.41E-08	2.41E-08	9.13E-08	9.13E-08	7.11E-08	7.11E-08	1.83E-07	1.83E-07	6.63E-08	6.63E-08	8.23E-08	1.03E-07	2.13E-07	2.93E-07	1.59E-07	1.59E-07	2.13644E-08	2.13644E-08	3.81E-07	4.29E-07	1.08372E-07	1.30046E-07	2.16E-08	2.16E-08	2.54E-07	3.01E-07	2.11873E-08	2.11873E-08	1.24E-07	1.49E-07	7.57E-08	7.57E-08	4.44E-08	4.44E-08	4.55E-08	4.55E-08
21	Bone Marrow	28-Days	100	8.19E-08	8.19E-08	7.06E-08	7.06E-08	7.97E-08	1.06E-07	4.82E-08	7.24E-08	2.15E-07	2.63E-07	1.70E-07	1.70E-07	1.31E-07	1.31E-07	1.50E-07	1.71E-07	6.42E-07	9.09E-07	1.64E-07	1.88E-07	2.01E-08	2.01E-08	2.78E-07	2.78E-07	6.26E-08	1.04E-07	2.13E-08	1.28E-07	4.01E-07	4.96E-07	8.88E-08	1.11E-07	3.44E-07	3.44E-07	2.47E-07	3.02E-07	1.13E-07	1.81E-07	1.60E-07	1.60E-07
22	Bone Marrow	28-Days	100	4.94E-08	4.94E-08	8.63E-08	8.63E-08	1.46E-07	1.75E-07	2.83E-08	2.83E-08	1.16E-07	1.16E-07	8.72E-08	8.72E-08	5.46E-08	5.46E-08	7.37E-08	7.37E-08	3.25E-07	3.25E-07	1.37E-07	1.37E-07	1.01E-07	1.01E-07	5.26E-07	9.06E-07	7.48E-08	1.99E-07	1.01E-07	1.01E-07	2.79E-07	3.07E-07	5.11E-08	5.11E-08	4.64E-07	5.25E-07	4.04E-07	4.37E-07	1.82E-07	1.82E-07	7.72E-08	7.72E-08
23	Bone Marrow	28-Days	100	1.34E-07	1.51E-07	7.08E-08	7.08E-08	1.10E-07	1.10E-07	5.23E-08	5.23E-08	5.50E-08	5.50E-08	1.05E-07	1.41E-07	1.02E-07	1.19E-07	7.97E-08	7.97E-08	2.25E-07	3.07E-07	1.75E-07	1.93E-07	4.95E-08	4.95E-08	2.19E-07	3.47E-07	1.00E-07	1.33E-07	3.31E-08	3.31E-08	4.49E-07	5.39E-07	6.48E-08	6.48E-08	1.52E-07	1.90E-07	3.30E-07	3.30E-07	1.54E-07	1.88E-07	1.70E-08	1.70E-08
24	Bone Marrow	28-Days	100	1.02E-07	1.02E-07	4.60E-08	4.60E-08	1.26E-07	2.78E-07	2.32E-08	2.32E-08	1.45E-07	1.45E-07	1.66E-07	1.66E-07	4.41E-08	6.62E-08	4.07E-08	4.07E-08	1.97E-07	3.09E-07	1.36E-07	1.36E-07	6.19E-08	6.19E-08	3.84E-07	4.32E-07	6.19E-08	6.19E-08	2.11E-08	2.11E-08	3.71E-07	4.17E-07	0.00E+00	0.00E+00	2.55E-07	2.55E-07	4.07E-07	4.89E-07	1.32E-07	1.32E-07	1.09E-07	1.09E-07
1	Liver	28-Days	0	5.84E-08	5.84E-08	8.79E-08	8.79E-08	3.01E-08	3.01E-08	8.27E-08	8.27E-08	1.23851E-07	1.23851E-07	5.72E-08	5.72E-08	2.68807E-08	2.68807E-08	0	0	6.41E-08	6.41E-08	7.91391E-08	7.91391E-08	7.85E-08	7.85E-08	1.46E-07	1.76E-07	3.06E-08	3.06E-08	2.65E-08	2.65E-08	9.30499E-08	9.30499E-08	1.03E-07	1.03E-07	0.00E+00	0.00E+00	6.41E-08	1.28E-07	2.83E-08	2.83E-08	8.79E-08	8.79E-08
2	Liver	28-Days	0	2.75E-08	2.75E-08	2.57E-07	1.37E-06	0.00E+00	0.00E+00	5.25E-08	5.25E-08	0.00E+00	0.00E+00	8.02993E-08	8.02993E-08	1.34E-07	1.34E-07	2.46E-08	2.46E-08	5.98E-08	5.98E-08	7.48E-08	7.48E-08	5.49E-08	5.49E-08	7.95E-08	7.95E-08	2.97E-08	2.97E-08	2.67E-08	2.67E-08	1.20E-07	1.20E-07	0.00E+00	0.00E+00	2.94E-08	2.94E-08	2.95E-08	2.95E-08	1.95E-07	1.95E-07	5.50E-08	5.50E-08
3	Liver	28-Days	0	2.49E-08	2.49E-08	2.37E-08	2.37E-08	5.00E-08	5.00E-08	2.37E-08	2.37E-08	7.51E-08	7.51E-08	9.79E-08	9.79E-08	0	0	2.17E-08	2.17E-08	5.70E-08	2.28E-07	6.80E-08	6.80E-08	1.00588E-07	1.00588E-07	1.92857E-07	1.92857E-07	2.45E-08	2.45E-08	0.00E+00	0.00E+00	2.44E-08	2.44E-08	0.00E+00	0.00E+00	2.59E-08	2.59E-08	8.35E-08	8.35E-08	4.69E-08	4.69E-08	7.77E-08	7.77E-08
4	Liver	28-Days	0	0.00E+00	0.00E+00	2.32E-08	2.32E-08	2.49E-08	2.49E-08	0.00E+00	0.00E+00	2.43187E-08	2.43187E-08	2.42876E-08	2.42876E-08	4.74E-08	4.74E-08	4.38537E-08	4.38537E-08	0.00E+00	0.00E+00	4.63E-08	4.63E-08	4.94E-08	4.94E-08	0.00E+00	0.00E+00	0.00E+00	0.00E+00	7.16E-08	7.16E-08	1.66E-07	1.66E-07	2.28054E-08	2.28054E-08	2.54E-08	2.54E-08	0.00E+00	0.00E+00	2.36E-08	2.36E-08	5.06E-08	5.06E-08
5	Liver	28-Days	6.25	7.66E-08	7.66E-08	1.24E-07	2.48E-07	3.88083E-08	3.88083E-08	5.48E-08	5.48E-08	8.29E-08	1.04E-07	1.85E-08	1.85E-08	5.48E-08	5.48E-08	5.21E-08	5.21E-08	2.11E-08	2.11E-08	6.99E-08	6.99E-08	1.30E-07	1.30E-07	1.50E-07	2.07E-07	2.11E-08	2.11E-08	7.32E-08	9.15E-08	4.16E-08	6.24E-08	1.71E-08	1.71E-08	1.21E-07	1.21E-07	1.01E-07	1.01E-07	5.73E-08	5.73E-08	4.08109E-08	4.08109E-08
6	Liver	28-Days	6.25	0.00E+00	0.00E+00	1.04E-07	1.29E-07	0.00E+00	0.00E+00	2.89179E-08	2.89179E-08	0.00E+00	0.00E+00	8.8228E-08	8.8228E-08	0	0	0.00E+00	0.00E+00	3.28E-08	3.28E-08	5.42E-08	5.42E-08	0	0	8.79E-08	8.79E-08	8.07E-08	8.07E-08	0.00E+00	0.00E+00	1.96E-07	2.24E-07	2.62E-08	2.62E-08	1.26E-07	5.34E-07	1.34E-07	1.34E-07	1.38E-07	2.21E-07	5.96E-08	5.96E-08
7	Liver	28-Days	6.25	2.95E-08	2.95E-08	5.48E-08	5.48E-08	3.05019E-08	3.05019E-08	8.52E-08	8.52E-08	2.83E-08	2.83E-08	8.53E-08	8.53E-08	0	0	2.65E-08	2.65E-08	0.00E+00	0.00E+00	1.68E-07	1.68E-07	3.02765E-08	3.02765E-08	2.89E-08	2.89E-08	8.63E-08	8.63E-08	0	0	8.72E-08	1.16E-07	0.00E+00	0.00E+00	0.00E+00	0.00E+00	9.09E-08	9.09E-08	5.64E-08	5.64E-08	2.96E-08	2.96E-08
8	Liver	28-Days	6.25	5.24E-08	5.24E-08	7.40E-08	7.40E-08	2.76405E-08	2.76405E-08	7.84E-08	7.84E-08	7.83441E-08	7.83441E-08	2.64E-08	2.64E-08	2.52E-08	2.52E-08	4.72E-08	4.72E-08	1.18E-07	1.18E-07	2.46E-08	2.46E-08	5.15E-08	5.15E-08	7.90E-08	7.90E-08	7.70312E-08	7.70312E-08	4.97E-08	4.97E-08	5.13E-08	5.13E-08	9.70E-08	9.70E-08	2.78E-08	2.78E-08	1.80E-07	1.80E-07	7.55E-08	7.55E-08	8.32E-08	8.32E-08
9	Liver	28-Days	12.5	9.20141E-08	9.20141E-08	1.26E-07	1.61E-07	7.52E-08	7.52E-08	7.15E-08	7.15E-08	1.11E-07	1.11E-07	8.92E-08	8.92E-08	7.10E-08	7.10E-08	1.15E-07	1.15E-07	3.85E-08	3.85E-08	1.04E-07	1.04E-07	1.88852E-08	1.88852E-08	7.18E-08	7.18E-08	5.59613E-08	5.59613E-08	0.00E+00	0.00E+00	9.31E-08	9.31E-08	5.12E-08	5.12E-08	1.54E-07	1.54E-07	5.79E-08	5.79E-08	1.44E-07	1.44E-07	1.31E-07	1.50E-07
10	Liver	28-Days	12.5	4.15E-08	4.15E-08	1.58E-07	1.58E-07	4.34E-08	4.34E-08	1.02E-07	1.43E-07	2.02E-08	2.02E-08	6.10E-08	6.10E-08	5.91E-08	3.75E-07	9.49E-08	9.49E-08	8.59E-08	1.07E-07	1.57E-07	1.57E-07	4.26E-08	4.26E-08	6.18E-08	6.18E-08	6.09E-08	6.09E-08	2.04E-08	4.09E-08	1.42E-07	1.63E-07	5.85E-08	5.85E-08	1.55E-07	1.55E-07	1.56E-07	1.56E-07	1.63E-07	1.63E-07	6.65E-08	6.65E-08
11	Liver	28-Days	12.5	1.46E-07	1.46E-07	1.88E-07	2.72E-07	4.30E-08	4.30E-08	4.10E-08	4.10E-08	8.21E-08	8.21E-08	2.02E-08	2.02E-08	4.10E-08	4.10E-08	9.70E-08	9.70E-08	8.49E-08	8.49E-08	6.02E-08	6.02E-08	2.27E-08	2.27E-08	4.09E-08	4.09E-08	1.08E-07	1.08E-07	6.38E-08	6.38E-08	1.04E-07	1.04E-07	1.99E-08	1.99E-08	6.57E-08	6.57E-08	6.38E-08	6.38E-08	8.34E-08	8.34E-08	4.52947E-08	4.52947E-08
12	Liver	28-Days	12.5	6.93E-08	6.93E-08	4.51363E-08	4.51363E-08	1.13E-07	1.13E-07	6.69E-08	6.69E-08	6.70E-08	6.70E-08	2.24E-08	2.24E-08	2.26E-08	2.26E-08	6.13E-08	6.13E-08	7.44E-08	7.44E-08	1.72E-07	1.72E-07	4.70E-08	4.70E-08	8.81E-08	8.81E-08	9.25E-08	9.25E-08	6.82E-08	6.82E-08	4.45E-08	4.45E-08	2.13E-08	2.13E-08	1.17E-07	1.41E-07	2.13E-07	3.79E-07	6.65E-08	6.65E-08	1.42E-07	1.42E-07
13	Liver	28-Days	25	9.09E-08	1.14E-07	2.64E-07	3.52E-07	7.03E-08	7.03E-08	4.44E-08	4.44E-08	1.78E-07	1.78E-07	8.88E-08	8.88E-08	8.11E-08	8.11E-08	1.03E-07	1.03E-07	9.63E-08	9.63E-08	6.50E-08	6.50E-08	4.67E-08	4.67E-08	1.55E-07	1.55E-07	4.55E-08	4.55E-08	2.23E-08	2.23E-08	1.78E-07	1.78E-07	1.49E-07	1.49E-07	7.15E-08	7.15E-08	7.21E-08	7.21E-08	1.32E-07	1.32E-07	2.38E-08	2.38E-08
14	Liver	28-Days	25	1.10E-07	1.10E-07	2.40E-07	6.10E-07	6.67E-08	6.67E-08	1.06E-07	1.06E-07	1.11E-07	1.11E-07	8.48E-08	8.48E-08	6.42E-08	6.42E-08	7.91E-08	7.91E-08	9.26E-08	9.26E-08	1.01E-07	1.01E-07	1.84E-07	1.84E-07	1.90E-07	1.90E-07	1.57E-07	1.57E-07	1.30E-07	1.30E-07	1.80E-07	1.80E-07	4.04E-08	4.04E-08	9.13E-08	9.13E-08	4.58E-08	4.58E-08	8.81E-08	8.81E-08	6.63E-08	6.63E-08
15	Liver	28-Days	25	1.86E-07	1.86E-07	3.59E-07	4.15E-07	5.43E-08	5.43E-08	7.68E-08	7.68E-08	1.91E-07	2.18E-07	1.02E-07	1.02E-07	8.02E-08	8.02E-08	9.79E-08	9.79E-08	8.57E-08	8.57E-08	1.24E-07	1.24E-07	5.57E-08	5.57E-08	2.30E-07	2.30E-07	1.12E-07	1.12E-07	5.32E-08	5.32E-08	5.51E-08	5.51E-08	2.48E-08	2.48E-08	2.77E-08	2.77E-08	8.34E-08	8.34E-08	1.61E-07	1.61E-07	1.13E-07	1.42E-07
16	Liver	28-Days	25	9.66E-08	9.66E-08	8.24E-08	8.24E-08	0.00E+00	0.00E+00	1.32E-07	1.32E-07	1.25E-07	1.25E-07	1.00E-07	1.00E-07	8.89E-08	8.89E-08	1.77E-07	1.77E-07	3.41E-08	3.41E-08	1.60E-07	1.60E-07	1.78E-07	1.78E-07	1.71E-07	1.71E-07	5.88E-08	5.88E-08	1.48E-07	1.48E-07	1.83E-07	1.83E-07	9.24E-08	9.24E-08	3.24E-07	3.24E-07	0.00E+00	0.00E+00	1.22E-07	1.22E-07	1.28E-07	1.28E-07
17	Liver	28-Days	50	2.54E-07	2.54E-07	1.70E-07	1.70E-07	1.92E-07	1.92E-07	1.13E-07	1.13E-07	2.78E-07	2.78E-07	1.37E-07	1.37E-07	1.33E-07	1.33E-07	8.22E-08	8.22E-08	2.89E-07	2.89E-07	2.99E-07	3.20E-07	1.14E-07	1.14E-07	2.93E-07	2.93E-07	3.71E-07	3.93E-07	1.09E-07	1.09E-07	4.05E-07	4.05E-07	1.93E-07	1.93E-07	2.91E-07	2.91E-07	2.33E-07	2.33E-07	3.73E-07	3.73E-07	1.17E-07	1.17E-07
18	Liver	28-Days	50	4.97E-07	4.97E-07	4.43E-07	4.43E-07	2.64E-07	2.86E-07	1.88E-07	1.88E-07	2.96E-07	2.96E-07	2.53E-07	2.53E-07	2.07E-07	2.07E-07	2.30E-07	2.30E-07	3.52E-07	3.52E-07	3.42E-07	4.03E-07	1.52E-07	1.52E-07	3.78E-07	7.56E-07	2.46E-07	2.46E-07	6.19E-08	6.19E-08	3.31E-07	3.31E-07	1.60E-07	1.60E-07	3.38E-07	3.38E-07	4.15E-07	5.76E-07	5.10E-07	5.10E-07	1.52E-07	1.52E-07
19	Liver	28-Days	50	1.82E-07	1.82E-07	2.30E-07	2.30E-07	4.66E-08	4.66E-08	1.11E-07	1.11E-07	2.04E-07	2.04E-07	1.10E-07	1.10E-07	4.49E-08	4.49E-08	6.30E-08	6.30E-08	2.18E-07	2.18E-07	1.72E-07	1.72E-07	1.23E-07	1.23E-07	2.01E-07	2.01E-07	1.39E-07	1.39E-07	1.13E-07	1.13E-07	2.05E-07	2.05E-07	1.27E-07	1.49E-07	1.42E-07	1.42E-07	2.86E-07	2.86E-07	1.35E-07	1.35E-07	2.46E-07	2.46E-07
20	Liver	28-Days	50	1.53E-07	1.53E-07	2.84E-07	4.39E-07	2.02E-07	2.02E-07	2.23E-07	2.23E-07	2.99E-07	2.99E-07	1.97E-07	1.97E-07	2.28E-07	2.28E-07	2.07E-07	2.07E-07	2.36E-07	2.36E-07	3.86E-07	3.86E-07	5.44E-08	5.44E-08	3.18E-07	3.18E-07	1.85E-07	1.85E-07	1.03E-07	1.03E-07	2.28E-07	2.28E-07	1.67E-07	1.67E-07	2.09E-07	2.09E-07	2.28E-07	2.28E-07	3.02E-07	3.02E-07	2.44E-07	2.44E-07
21	Liver	28-Days	100	3.52E-07	3.52E-07	4.90E-07	9.28E-07	3.84E-07	3.84E-07	2.68E-07	2.68E-07	3.01E-07	3.01E-07	3.63E-07	3.63E-07	1.25E-07	1.25E-07	1.38E-07	1.38E-07	4.51E-07	4.51E-07	2.61E-07	2.61E-07	1.62E-07	1.62E-07	1.95E-07	1.95E-07	2.87E-07	2.87E-07	1.52E-07	1.52E-07	8.13E-07	8.38E-07	2.60E-07	2.60E-07	3.12E-07	3.12E-07	3.91E-07	3.91E-07	4.04E-07	4.29E-07	2.19E-07	2.19E-07
22	Liver	28-Days	100	5.32E-07	5.51E-07	6.05E-07	6.24E-07	4.66E-07	4.66E-07	6.19E-07	6.19E-07	4.43E-07	4.43E-07	5.92E-07	6.12E-07	4.37E-07	4.37E-07	5.00E-07	5.00E-07	5.42E-07	5.42E-07	4.68E-07	4.87E-07	2.67E-07	2.67E-07	5.61E-07	2.34E-06	5.05E-07	5.05E-07	1.35E-07	1.35E-07	8.29E-07	8.29E-07	4.06E-07	4.25E-07	4.96E-07	4.96E-07	5.21E-07	5.21E-07	5.35E-07	5.35E-07	2.27E-07	2.27E-07
23	Liver	28-Days	100	4.43E-07	4.92E-07	2.96E-07	2.96E-07	2.88E-07	3.15E-07	2.46E-07	2.46E-07	2.40E-07	2.40E-07	3.43E-07	3.92E-07	2.58E-07	3.04E-07	2.94E-07	2.94E-07	2.37E-07	2.37E-07	2.59E-07	2.83E-07	2.23E-07	2.23E-07	3.73E-07	3.98E-07	3.31E-07	3.31E-07	2.14E-07	2.14E-07	5.99E-07	5.99E-07	2.09E-07	2.09E-07	3.95E-07	3.95E-07	4.91E-07	4.91E-07	4.99E-07	4.99E-07	2.05E-07	2.05E-07
24	Liver	28-Days	100	3.32E-07	3.32E-07	4.20E-07	5.60E-07	9.00E-08	9.00E-08	2.76E-07	2.76E-07	5.93E-07	5.93E-07	4.07E-07	4.07E-07	4.75E-07	4.97E-07	2.62E-07	2.62E-07	3.40E-07	3.40E-07	3.09E-07	3.09E-07	3.38E-07	3.38E-07	3.88E-07	3.88E-07	4.05E-07	4.29E-07	1.79E-07	1.79E-07	2.93E-07	2.93E-07	3.29E-07	3.29E-07	3.51E-07	3.51E-07	2.97E-07	2.97E-07	3.57E-07	3.57E-07	1.48E-07	1.48E-07

Appendix D2 Table 11: Locus specific mutagenicity data of MutaMouse males post exposure to BbF for 90-days

Animal ID	Sample tissue	Exposure duration	BbF dose (mg/kg/day)	Chr. 1 MFmin	Chr. 1 MFmax	Chr. 1.2 MFmin	Chr. 1.2 MFmax	Chr. 2 MFmin	Chr. 2 MFmax	Chr. 3 MFmin	Chr. 3 MFmax	Chr. 4 MFmin	Chr. 4 MFmax
25	Bone Marrow	90-Days	0	0	0	3.59E-08	3.59E-08	0	0	0	0	0	0
26	Bone Marrow	90-Days	0	2.75E-08	2.75E-08	8.10E-08	8.10E-08	5.79E-08	5.79E-08	0	0	2.68E-08	2.68E-08
27	Bone Marrow	90-Days	0	2.91E-08	2.91E-08	2.89E-08	2.89E-08	0	0	0	0	2.96E-08	2.96E-08
28	Bone Marrow	90-Days	0	5.78E-08	5.78E-08	0	0	0	0	0	0	8.54E-08	8.54E-08
29	Bone Marrow	90-Days	3.125	4.70E-08	4.70E-08	9.39E-08	1.17E-07	5.00E-08	5.00E-08	7.04E-08	7.04E-08	4.66E-08	4.66E-08
30	Bone Marrow	90-Days	3.125	2.54E-08	2.54E-08	0	0	2.71E-08	2.71E-08	0	0	1.24E-07	1.48E-07
31	Bone Marrow	90-Days	3.125	4.11E-08	4.11E-08	0	0	4.28E-08	4.28E-08	8.08E-08	8.08E-08	0	0
32	Bone Marrow	90-Days	3.125	3.89E-08	5.83E-08	0	0	2.06E-08	2.06E-08	0	0	1.90E-08	1.90E-08
33	Bone Marrow	90-Days	6.25	4.12E-08	4.12E-08	1.95E-08	1.95E-08	1.09E-07	1.09E-07	8.32E-08	8.32E-08	8.01E-08	1.20E-07
34	Bone Marrow	90-Days	6.25	1.62E-08	1.62E-08	3.17E-08	6.34E-08	6.82E-08	6.82E-08	1.62E-08	1.62E-08	6.48E-08	1.94E-07
35	Bone Marrow	90-Days	6.25	7.44E-08	7.44E-08	4.85E-08	2.67E-07	1.32E-07	1.32E-07	1.24E-07	1.24E-07	2.42E-08	2.42E-08
36	Bone Marrow	90-Days	6.25	3.45E-08	3.45E-08	3.48E-08	3.48E-08	0	0	3.37E-08	3.37E-08	1.01E-07	1.01E-07
37	Bone Marrow	90-Days	12.5	2.68E-08	2.68E-08	5.23E-08	5.23E-08	8.40E-08	8.40E-08	0	0	5.25E-08	7.87E-08
38	Bone Marrow	90-Days	12.5	7.67E-08	7.67E-08	7.21E-08	7.21E-08	1.92E-07	4.94E-07	0	0	2.49E-08	2.49E-08
39	Bone Marrow	90-Days	12.5	5.12E-08	5.12E-08	1.04E-07	1.04E-07	1.10E-07	1.65E-07	2.53E-08	2.53E-08	2.51E-08	2.51E-08
40	Bone Marrow	90-Days	12.5	1.24E-07	1.24E-07	2.40E-08	2.40E-08	7.93E-08	1.06E-07	4.96E-08	4.96E-08	4.77E-08	4.77E-08
41	Bone Marrow	90-Days	25	7.87E-08	7.87E-08	2.55E-08	2.55E-08	8.11E-08	2.16E-07	1.03E-07	1.03E-07	2.54E-08	2.54E-08
42	Bone Marrow	90-Days	25	4.19E-08	4.19E-08	6.03E-08	8.04E-08	1.34E-07	1.34E-07	1.48E-07	1.48E-07	6.17E-08	1.23E-07
43	Bone Marrow	90-Days	25	2.05E-08	2.05E-08	3.95E-08	1.58E-07	6.48E-08	6.48E-08	1.23E-07	1.23E-07	6.03E-08	1.21E-07
44	Bone Marrow	90-Days	25	3.24E-08	3.24E-08	1.00E-07	1.67E-07	1.38E-07	7.26E-07	1.31E-07	2.94E-07	1.62E-07	1.62E-07
45	Bone Marrow	90-Days	50	9.98E-08	1.25E-07	1.76E-07	4.03E-07	1.30E-07	1.30E-07	4.96E-08	4.96E-08	1.97E-07	3.20E-07
46	Bone Marrow	90-Days	50	2.13E-07	2.13E-07	1.79E-07	1.79E-07	2.29E-07	2.29E-07	9.10E-08	9.10E-08	1.49E-07	1.49E-07
47	Bone Marrow	90-Days	50	8.37E-08	1.26E-07	1.00E-07	1.00E-07	1.10E-07	1.32E-07	1.46E-07	1.88E-07	1.84E-07	1.84E-07
48	Bone Marrow	90-Days	50	1.94E-07	2.91E-07	2.42E-08	2.42E-08	1.54E-07	2.31E-07	1.22E-07	1.47E-07	1.67E-07	2.15E-07
25	Liver	90-Days	0	2.74E-08	2.74E-08	4.60E-08	4.60E-08	2.88E-08	2.88E-08	5.63E-08	5.63E-08	0	0
26	Liver	90-Days	0	8.29E-08	8.29E-08	5.88E-08	5.88E-08	5.68E-08	5.68E-08	2.73E-08	2.73E-08	5.32E-08	5.32E-08
27	Liver	90-Days	0	3.19E-08	3.19E-08	1.65E-07	1.65E-07	3.37E-08	3.37E-08	6.35E-08	6.35E-08	3.05E-08	3.05E-08
28	Liver	90-Days	0	2.97E-08	2.97E-08	6.32E-08	2.53E-07	6.09E-08	6.09E-08	2.93E-08	2.93E-08	0	0
29	Liver	90-Days	3.125	2.69E-08	2.69E-08	2.80E-08	2.80E-08	0	0	5.33E-08	5.33E-08	1.03E-07	1.03E-07
30	Liver	90-Days	3.125	8.53E-08	8.53E-08	2.96E-08	2.96E-08	5.81E-08	5.81E-08	0	0	8.30E-08	8.30E-08
31	Liver	90-Days	3.125	9.70E-08	9.70E-08	4.72E-08	7.08E-08	0	0	9.38E-08	1.17E-07	1.83E-07	2.06E-07
32	Liver	90-Days	3.125	3.19E-08	3.19E-08	6.57E-08	6.57E-08	0	0	3.12E-08	3.12E-08	0	0
33	Liver	90-Days	6.25	0	0	1.49E-07	1.79E-07	8.79E-08	8.79E-08	2.85E-08	2.85E-08	5.62E-08	5.62E-08
34	Liver	90-Days	6.25	1.52E-07	1.52E-07	9.15E-08	9.15E-08	1.26E-07	1.26E-07	1.19E-07	1.19E-07	8.76E-08	8.76E-08
35	Liver	90-Days	6.25	1.92E-07	1.92E-07	3.75E-07	2.25E-06	3.91E-08	3.91E-08	1.52E-07	1.52E-07	1.90E-07	1.90E-07
36	Liver	90-Days	6.25	2.88E-08	2.88E-08	0	0	5.69E-08	5.69E-08	2.80E-08	2.80E-08	5.44E-08	5.44E-08
37	Liver	90-Days	12.5	1.21E-07	1.21E-07	8.90E-08	1.48E-07	6.14E-08	6.14E-08	1.75E-07	1.75E-07	1.75E-07	1.75E-07
38	Liver	90-Days	12.5	1.27E-07	1.27E-07	3.04E-07	4.74E-07	6.63E-08	6.63E-08	2.21E-07	2.21E-07	1.84E-07	1.84E-07
39	Liver	90-Days	12.5	2.03E-07	2.03E-07	1.31E-07	1.50E-07	1.00E-07	1.00E-07	2.55E-07	2.55E-07	1.15E-07	1.15E-07
40	Liver	90-Days	12.5	2.23E-07	2.23E-07	1.40E-07	1.40E-07	1.69E-07	1.69E-07	5.43E-08	5.43E-08	2.40E-07	2.67E-07
41	Liver	90-Days	25	3.69E-07	3.69E-07	5.63E-07	7.04E-07	3.23E-07	3.23E-07	3.03E-07	3.03E-07	2.36E-07	2.36E-07
42	Liver	90-Days	25	4.15E-07	4.15E-07	3.51E-07	5.74E-07	5.18E-07	5.18E-07	1.56E-07	1.56E-07	2.14E-07	2.14E-07
43	Liver	90-Days	25	5.65E-07	5.96E-07	3.18E-07	3.47E-07	3.48E-07	3.48E-07	3.33E-07	3.33E-07	7.57E-07	7.57E-07
44	Liver	90-Days	25	4.65E-07	4.65E-07	7.30E-07	1.05E-06	2.76E-07	2.76E-07	3.57E-07	3.57E-07	5.06E-07	5.38E-07
45	Liver	90-Days	50	1.55E-06	1.58E-06	1.47E-06	1.49E-06	7.28E-07	7.28E-07	1.29E-06	1.38E-06	1.27E-06	1.27E-06
46	Liver	90-Days	50	9.54E-07	9.88E-07	1.25E-06	1.52E-06	5.97E-07	5.97E-07	7.02E-07	7.02E-07	6.82E-07	6.82E-07
47	Liver	90-Days	50	7.57E-07	7.88E-07	1.15E-06	1.23E-06	5.49E-07	5.80E-07	1.04E-06	1.04E-06	8.45E-07	8.45E-07
48	Liver	90-Days	50	6.95E-07	6.95E-07	7.32E-07	7.64E-07	2.75E-07	2.75E-07	6.36E-07	6.36E-07	8.41E-07	8.41E-07

Animal ID	Sample tissue	Chr. 5 MFmin	Chr. 5 MFmax	Chr. 6 MFmin	Chr. 6 MFmax	Chr. 7 MFmin	Chr. 7 MFmax	Chr. 8 MFmin	Chr. 8 MFmax	Chr. 9 MFmin	Chr. 9 MFmax	Chr. 10 MFmin	Chr. 10 MFmax	Chr. 11 MFmin	Chr. 11 MFmax	Chr. 12 MFmin	Chr. 12 MFmax
25	Bone Marrow	6.71E-08	6.71E-08	9.99E-08	9.99E-08	0	0	1.56E-07	1.56E-07	0	0	0	0	3.54E-08	3.54E-08	0	0
26	Bone Marrow	5.30E-08	5.30E-08	7.65E-08	7.65E-08	7.23E-08	7.23E-08	1.15E-07	1.15E-07	5.41E-08	5.41E-08	0	0	5.71E-08	5.71E-08	2.71E-08	2.71E-08
27	Bone Marrow	1.37E-07	2.73E-07	5.45E-08	5.45E-08	2.44E-08	2.44E-08	0	0	0	0	0	0	2.88E-08	2.88E-08	5.85E-08	5.85E-08
28	Bone Marrow	0	0	2.70E-08	2.70E-08	1.01E-07	1.01E-07	9.08E-08	9.08E-08	2.85E-08	2.85E-08	2.75E-08	2.75E-08	3.02E-08	3.02E-08	2.91E-08	2.91E-08
29	Bone Marrow	4.53E-08	6.79E-08	1.97E-07	1.97E-07	2.08E-08	2.08E-08	1.72E-07	1.72E-07	2.29E-08	2.29E-08	0	0	1.21E-07	1.21E-07	2.36E-08	2.12E-07
30	Bone Marrow	4.95E-08	1.24E-07	7.10E-08	7.10E-08	4.48E-08	4.48E-08	1.06E-07	1.06E-07	2.52E-08	2.52E-08	2.38E-08	2.38E-08	5.34E-08	5.34E-08	1.01E-07	1.01E-07
31	Bone Marrow	7.69E-08	7.69E-08	7.79E-08	7.79E-08	7.25E-08	7.25E-08	1.29E-07	1.29E-07	0	0	0	0	0	0	0	0
32	Bone Marrow	5.64E-08	5.64E-08	5.44E-08	5.44E-08	1.70E-08	1.70E-08	6.23E-08	6.23E-08	9.47E-08	1.33E-07	1.81E-08	1.81E-08	1.01E-07	1.01E-07	3.83E-08	3.83E-08
33	Bone Marrow	1.21E-07	3.03E-07	5.71E-08	7.61E-08	3.60E-08	3.60E-08	1.11E-07	1.11E-07	0	0	7.58E-08	7.58E-08	1.29E-07	1.29E-07	7.97E-08	7.97E-08
34	Bone Marrow	6.27E-08	6.27E-08	3.07E-08	3.07E-08	2.83E-08	2.83E-08	1.25E-07	1.43E-07	3.12E-08	3.12E-08	1.53E-08	1.53E-08	1.16E-07	1.33E-07	1.61E-08	1.61E-08
35	Bone Marrow	2.40E-07	2.40E-07	2.28E-08	2.28E-08	2.15E-08	4.30E-08	1.56E-07	2.08E-07	2.44E-08	2.44E-08	0	0	3.64E-07	7.79E-07	2.46E-08	2.46E-08
36	Bone Marrow	3.24E-08	3.24E-08	0	0	6.08E-08	6.08E-08	1.44E-07	1.44E-07	1.34E-07	1.67E-07	3.27E-08	3.27E-08	1.75E-07	1.75E-07	6.95E-08	6.95E-08
37	Bone Marrow	1.82E-07	1.82E-07	0	0	9.35E-08	9.35E-08	3.67E-07	5.92E-07	1.04E-07	1.04E-07	0	0	2.76E-07	3.31E-07	2.66E-08	2.66E-08
38	Bone Marrow	7.48E-08	7.48E-08	9.29E-08	9.29E-08	1.34E-07	1.34E-07	5.12E-07	9.70E-07	1.74E-07	2.23E-07	4.59E-08	1.15E-07	4.01E-07	2.08E-06	1.24E-07	1.74E-07
39	Bone Marrow	7.30E-08	7.30E-08	4.78E-08	4.78E-08	4.51E-08	4.51E-08	1.07E-07	1.07E-07	2.02E-07	2.53E-07	2.43E-08	4.86E-08	2.91E-07	2.91E-07	5.25E-08	5.25E-08
40	Bone Marrow	1.43E-07	1.43E-07	2.28E-08	2.28E-08	6.55E-08	1.97E-07	3.33E-07	4.87E-07	9.81E-08	9.81E-08	6.85E-08	6.85E-08	4.92E-07	7.50E-07	7.30E-08	7.30E-08
41	Bone Marrow	2.49E-07	3.48E-07	9.77E-08	9.77E-08	4.53E-08	4.53E-08	2.83E-07	3.11E-07	1.25E-07	2.51E-07	2.45E-08	2.45E-08	4.48E-07	5.01E-07	1.79E-07	1.79E-07
42	Bone Marrow	1.43E-07	1.64E-07	1.96E-08	1.96E-08	5.58E-08	5.58E-08	5.00E-07	8.18E-07	8.14E-08	2.24E-07	1.17E-07	2.34E-07	3.02E-07	5.18E-07	6.17E-08	6.17E-08
43	Bone Marrow	1.59E-07	1.59E-07	2.11E-07	2.11E-07	1.08E-07	1.08E-07	2.46E-07	6.25E-07	1.59E-07	1.79E-07	5.73E-08	9.54E-08	4.21E-07	6.32E-07	4.01E-08	4.01E-08
44	Bone Marrow	3.44E-07	3.75E-07	3.11E-08	3.11E-08	5.86E-08	5.86E-08	2.40E-07	3.43E-07	1.29E-07	1.62E-07	0	0	1.68E-07	2.69E-07	0	0
45	Bone Marrow	1.91E-07	1.91E-07	7.08E-08	7.08E-08	1.77E-07	2.87E-07	6.24E-07	9.88E-07	3.43E-07	5.63E-07	7.24E-08	7.24E-08	4.83E-07	5.85E-07	7.60E-08	7.60E-08
46	Bone Marrow	1.18E-07	2.07E-07	8.36E-08	8.36E-08	2.68E-08	2.68E-08	3.45E-07	7.84E-07	1.50E-07	2.70E-07	8.46E-08	1.41E-07	5.11E-07	6.07E-07	9.13E-08	9.13E-08
47	Bone Marrow	1.21E-07	4.45E-07	9.71E-08	2.52E-07	9.12E-08	9.12E-08	5.55E-07	8.87E-07	2.84E-07	2.84E-07	9.71E-08	1.94E-07	4.49E-07	7.49E-07	1.63E-07	1.63E-07
48	Bone Marrow	1.65E-07	1.88E-07	2.29E-08	6.87E-08	1.29E-07	1.50E-07	3.07E-07	3.59E-07	1.20E-07	1.20E-07	2.32E-08	2.32E-08	6.02E-07	1.86E-06	4.86E-08	4.86E-08
25	Liver	5.71E-08	5.71E-08	0	0	0	0	3.31E-08	3.31E-08	0	0	2.39E-08	2.39E-08	2.70E-08	2.70E-08	2.48E-08	4.96E-08
26	Liver	0	0	2.74E-08	2.74E-08	2.53E-08	2.53E-08	1.11E-07	1.11E-07	7.79E-08	7.79E-08	2.87E-08	2.87E-08	1.06E-07	1.06E-07	8.89E-08	1.78E-07
27	Liver	3.10E-08	6.21E-08	0	0	2.89E-08	2.89E-08	3.02E-08	3.02E-08	6.13E-08	6.13E-08	0	0	0	0	2.02E-07	4.04E-07
28	Liver	0	0	8.75E-08	8.75E-08	0	0	1.15E-07	1.15E-07	8.33E-08	8.33E-08	6.09E-08	6.09E-08	8.55E-08	8.55E-08	3.17E-08	3.17E-08
29	Liver	2.62E-08	2.62E-08	5.22E-08	5.22E-08	4.98E-08	4.98E-08	1.29E-07	1.29E-07	2.57E-08	2.57E-08	1.63E-07	1.63E-07	7.91E-08	7.91E-08	2.83E-08	2.83E-08
30	Liver	0	0	0	0	2.57E-08	2.57E-08	1.44E-07	1.44E-07	5.35E-08	5.35E-08	0	0	8.18E-08	8.18E-08	3.01E-08	3.01E-08
31	Liver	4.64E-08	4.64E-08	0	0	2.15E-08	2.15E-08	7.44E-08	7.44E-08	4.49E-08	4.49E-08	0	0	9.30E-08	9.30E-08	2.41E-08	2.41E-08
32	Liver	1.53E-07	1.53E-07	3.15E-08	3.15E-08	2.85E-08	5.70E-08	1.61E-07	1.61E-07	8.85E-08	8.85E-08	3.28E-08	3.28E-08	1.21E-07	1.21E-07	0	0
33	Liver	5.58E-08	5.58E-08	2.86E-08	2.86E-08	5.24E-08	5.24E-08	1.46E-07	1.46E-07	1.09E-07	1.09E-07	0	0	1.40E-07	1.40E-07	0	0
34	Liver	5.90E-08	5.90E-08	5.91E-08	5.91E-08	8.03E-08	8.03E-08	1.80E-07	1.80E-07	5.74E-08	5.74E-08	3.10E-08	3.10E-08	3.22E-07	4.98E-07	3.09E-08	3.09E-08
35	Liver	3.64E-08	3.64E-08	1.17E-07	1.17E-07	3.49E-08	3.49E-08	2.71E-07	2.71E-07	7.16E-08	7.16E-08	8.14E-08	8.14E-08	7.32E-08	7.32E-08	8.28E-08	8.28E-08
36	Liver	5.70E-08	5.70E-08	1.06E-07	1.06E-07	4.75E-08	4.75E-08	1.85E-07	1.85E-07	2.65E-08	2.65E-08	7.73E-08	7.73E-08	1.12E-07	1.12E-07	1.03E-07	1.03E-07
37	Liver	1.16E-07	1.16E-07	1.46E-07	1.46E-07	1.05E-07	1.05E-07	2.42E-07	2.42E-07	1.13E-07	1.13E-07	1.84E-07	1.84E-07	1.16E-07	1.16E-07	9.12E-08	9.12E-08
38	Liver	1.21E-07	1.21E-07	1.25E-07	1.25E-07	5.85E-08	5.85E-08	2.47E-07	2.47E-07	1.20E-07	1.20E-07	3.36E-08	3.36E-08	2.47E-07	2.47E-07	2.71E-07	2.71E-07
39	Liver	1.18E-07	1.18E-07	1.53E-07	1.53E-07	3.50E-08	3.50E-08	1.86E-07	1.86E-07	1.33E-07	1.33E-07	9.68E-08	9.68E-08	1.37E-07	1.37E-07	1.53E-07	1.53E-07
40	Liver	1.85E-07	1.85E-07	2.17E-07	2.17E-07	9.84E-08	9.84E-08	3.42E-07	3.42E-07	1.80E-07	1.80E-07	1.40E-07	1.40E-07	1.32E-07	1.32E-07	2.27E-07	2.27E-07
41	Liver	4.00E-07	4.00E-07	2.90E-07	2.90E-07	3.56E-07	3.56E-07	3.91E-07	3.91E-07	4.43E-07	4.73E-07	2.37E-07	2.37E-07	2.43E-07	2.43E-07	2.94E-07	2.94E-07
42	Liver	2.13E-07	2.13E-07	2.15E-07	2.15E-07	2.52E-07	2.52E-07	4.40E-07	4.40E-07	4.78E-07	4.78E-07	1.62E-07	1.62E-07	5.49E-07	1.16E-06	5.86E-07	5.86E-07
43	Liver	4.82E-07	4.82E-07	2.70E-07	2.70E-07	2.03E-07	2.03E-07	6.59E-07	6.59E-07	4.36E-07	4.36E-07	1.22E-07	1.22E-07	8.33E-07	8.33E-07	3.04E-07	3.04E-07
44	Liver	4.92E-07	4.92E-07	5.37E-07	5.37E-07	3.98E-07	3.98E-07	8.92E-07	8.92E-07	5.99E-07	5.99E-07	1.24E-07	1.24E-07	7.51E-07	7.84E-07	5.90E-07	5.90E-07
45	Liver	1.40E-06	1.43E-06	7.15E-07	7.15E-07	9.82E-07	9.82E-07	1.83E-06	1.83E-06	1.44E-06	1.52E-06	3.17E-07	3.17E-07	2.00E-06	2.06E-06	1.38E-06	1.41E-06
46	Liver	6.89E-07	6.89E-07	5.25E-07	5.25E-07	4.79E-07	4.79E-07	9.26E-07	9.26E-07	8.96E-07	8.96E-07	3.11E-07	3.11E-07	1.10E-06	1.13E-06	5.64E-07	5.64E-07
47	Liver	8.22E-07	8.51E-07	6.41E-07	6.41E-07	4.89E-07	4.89E-07	1.08E-06	1.08E-06	9.95E-07	1.02E-06	6.36E-07	6.36E-07	1.26E-06	1.26E-06	8.12E-07	8.12E-07
48	Liver	7.90E-07	8.23E-07	4.65E-07	4.65E-07	6.76E-07	6.76E-07	8.51E-07	8.51E-07	7.99E-07	7.99E-07	2.34E-07	2.34E-07	9.79E-07	9.79E-07	6.33E-07	6.33E-07

Animal ID	Sample tissue	Chr. 13 MFmin	Chr. 13 MFmax	Chr. 14 MFmin	Chr. 14 MFmax	Chr. 15 MFmin	Chr. 15 MFmax	Chr. 16 MFmin	Chr. 16 MFmax	Chr. 17 MFmin	Chr. 17 MFmax	Chr. 18 MFmin	Chr. 18 MFmax	Chr. 19 MFmin	Chr. 19 MFmax
25	Bone Marrow	0	0	7.52E-08	7.52E-08	3.20E-08	3.20E-08	3.92E-08	3.92E-08	0	0	0	0	3.66E-08	3.66E-08
26	Bone Marrow	5.53E-08	5.53E-08	1.65E-07	1.65E-07	2.51E-08	2.00E-07	6.11E-08	6.11E-08	2.84E-08	2.84E-08	1.08E-07	1.08E-07	5.70E-08	5.70E-08
27	Bone Marrow	2.86E-08	2.86E-08	0	0	2.55E-08	2.55E-08	3.16E-08	3.16E-08	0	0	2.85E-08	2.85E-08	0	0
28	Bone Marrow	8.84E-08	8.84E-08	1.46E-07	1.46E-07	2.66E-08	2.66E-08	9.67E-08	9.67E-08	2.97E-08	2.97E-08	0	0	0	0
29	Bone Marrow	4.79E-08	4.79E-08	7.19E-08	7.19E-08	0	0	1.31E-07	1.57E-07	1.20E-07	2.64E-07	2.32E-08	2.32E-08	0	0
30	Bone Marrow	2.58E-08	2.58E-08	1.53E-07	1.53E-07	0	0	8.65E-08	1.15E-07	1.84E-07	2.10E-07	9.98E-08	1.25E-07	5.31E-08	2.12E-07
31	Bone Marrow	4.24E-08	4.24E-08	4.16E-08	4.16E-08	3.77E-08	3.77E-08	8.87E-08	8.87E-08	8.24E-08	8.24E-08	0	0	4.30E-08	4.30E-08
32	Bone Marrow	7.78E-08	9.72E-08	3.86E-08	1.35E-07	3.55E-08	3.55E-08	6.48E-08	6.48E-08	0	0	0	0	2.00E-08	4.00E-08
33	Bone Marrow	2.02E-08	2.02E-08	4.07E-08	4.07E-08	3.77E-08	5.66E-08	9.19E-08	1.15E-07	1.11E-07	1.11E-07	1.20E-07	1.20E-07	4.19E-08	4.19E-08
34	Bone Marrow	0	0	1.47E-07	1.47E-07	4.44E-08	5.92E-08	1.06E-07	1.77E-07	5.21E-08	1.22E-07	9.49E-08	9.49E-08	4.98E-08	4.98E-08
35	Bone Marrow	2.48E-08	4.96E-08	2.25E-07	4.25E-07	4.55E-08	4.55E-08	1.41E-07	1.41E-07	2.36E-07	2.36E-07	7.31E-08	7.31E-08	7.69E-08	7.69E-08
36	Bone Marrow	1.05E-07	1.41E-07	0	0	9.32E-08	9.32E-08	3.00E-07	3.00E-07	6.96E-08	6.96E-08	3.40E-08	3.40E-08	3.63E-08	7.26E-08
37	Bone Marrow	2.68E-08	2.68E-08	1.34E-07	1.34E-07	2.45E-08	2.45E-08	1.48E-07	1.48E-07	1.39E-07	4.17E-07	1.31E-07	1.31E-07	8.27E-08	8.27E-08
38	Bone Marrow	1.50E-07	1.50E-07	1.27E-07	1.27E-07	7.05E-08	7.05E-08	1.44E-07	1.44E-07	1.93E-07	2.76E-07	1.99E-07	1.99E-07	1.05E-07	1.05E-07
39	Bone Marrow	7.87E-08	7.87E-08	7.80E-08	7.80E-08	9.46E-08	1.18E-07	2.01E-07	3.15E-07	1.57E-07	1.57E-07	1.78E-07	1.78E-07	8.22E-08	8.22E-08
40	Bone Marrow	4.99E-08	4.99E-08	2.96E-07	3.45E-07	1.36E-07	1.36E-07	2.49E-07	2.77E-07	2.82E-07	3.08E-07	1.20E-07	1.20E-07	1.53E-07	1.53E-07
41	Bone Marrow	2.59E-08	2.59E-08	2.58E-07	2.84E-07	4.71E-08	9.41E-08	2.56E-07	1.05E-06	2.18E-07	6.00E-07	2.52E-08	2.52E-08	1.07E-07	1.07E-07
42	Bone Marrow	4.17E-08	4.17E-08	2.72E-07	2.93E-07	3.84E-08	3.84E-08	3.01E-07	3.93E-07	2.03E-07	2.03E-07	2.26E-07	2.26E-07	1.29E-07	4.31E-07
43	Bone Marrow	6.15E-08	3.89E-07	4.47E-07	5.69E-07	7.46E-08	7.46E-08	2.03E-07	2.93E-07	2.20E-07	2.64E-07	9.97E-08	1.20E-07	8.48E-08	8.48E-08
44	Bone Marrow	1.36E-07	2.38E-07	3.35E-07	3.68E-07	0	0	4.00E-07	4.73E-07	3.64E-07	9.28E-07	1.31E-07	1.31E-07	3.48E-08	3.48E-08
45	Bone Marrow	5.15E-08	5.15E-08	4.80E-07	6.82E-07	4.52E-08	4.52E-08	4.11E-07	4.38E-07	4.78E-07	5.78E-07	9.97E-08	9.97E-08	7.85E-08	7.85E-08
46	Bone Marrow	3.06E-08	3.06E-08	2.74E-07	2.74E-07	2.82E-08	2.82E-08	1.38E-07	2.75E-07	2.84E-07	3.15E-07	2.40E-07	2.70E-07	9.58E-08	1.60E-07
47	Bone Marrow	1.66E-07	3.52E-07	4.57E-07	1.56E-06	3.79E-08	5.69E-08	3.67E-07	8.03E-07	1.54E-07	2.42E-07	2.03E-07	2.43E-07	6.39E-08	1.07E-07
48	Bone Marrow	7.41E-08	7.41E-08	2.69E-07	3.17E-07	4.45E-08	1.11E-07	3.25E-07	5.41E-07	4.01E-07	5.76E-07	7.22E-08	7.22E-08	1.01E-07	1.01E-07
25	Liver	4.92E-08	7.38E-08	0	0	2.55E-08	2.55E-08	2.97E-08	2.97E-08	1.00E-07	1.34E-07	2.57E-08	2.57E-08	5.32E-08	5.32E-08
26	Liver	5.94E-08	5.94E-08	8.20E-08	8.20E-08	5.33E-08	5.33E-08	2.82E-08	2.82E-08	5.41E-08	5.41E-08	5.64E-08	8.47E-08	6.23E-08	6.23E-08
27	Liver	6.83E-08	5.81E-07	1.26E-07	1.26E-07	9.33E-08	9.33E-08	3.41E-08	3.41E-08	3.09E-08	1.24E-07	6.50E-08	9.74E-08	1.73E-07	2.76E-07
28	Liver	1.59E-07	3.81E-07	5.80E-08	5.80E-08	0	0	5.99E-08	5.99E-08	5.69E-08	5.69E-08	3.00E-08	3.00E-08	3.29E-08	3.29E-08
29	Liver	1.71E-07	5.70E-07	5.30E-08	5.30E-08	1.05E-07	1.05E-07	4.27E-07	5.41E-07	1.58E-07	3.42E-07	2.72E-08	2.72E-08	0	0
30	Liver	6.10E-08	6.10E-08	5.60E-08	5.60E-08	2.71E-08	2.71E-08	2.91E-08	2.91E-08	8.35E-08	8.35E-08	1.14E-07	1.14E-07	3.13E-08	3.13E-08
31	Liver	0	0	6.87E-08	6.87E-08	4.52E-08	4.52E-08	1.24E-07	1.24E-07	4.90E-08	4.90E-08	2.33E-08	2.33E-08	5.00E-08	5.00E-08
32	Liver	3.33E-08	3.33E-08	6.23E-08	6.23E-08	9.03E-08	9.03E-08	6.53E-08	6.53E-08	6.35E-08	6.35E-08	3.20E-08	3.20E-08	3.49E-08	3.49E-08
33	Liver	6.08E-08	2.13E-07	2.26E-07	2.26E-07	5.48E-08	5.48E-08	1.20E-07	1.20E-07	2.58E-07	2.58E-07	5.83E-08	5.83E-08	3.13E-08	3.13E-08
34	Liver	3.15E-08	3.15E-08	1.79E-07	1.79E-07	2.86E-08	2.86E-08	3.20E-08	3.20E-08	1.79E-07	1.79E-07	1.51E-07	1.51E-07	6.39E-08	6.39E-08
35	Liver	1.69E-07	2.03E-06	7.76E-08	7.76E-08	1.10E-07	1.10E-07	3.48E-07	3.48E-07	1.13E-07	1.13E-07	2.79E-07	2.79E-07	0	0
36	Liver	2.59E-08	2.59E-08	8.06E-08	8.06E-08	1.56E-07	1.56E-07	1.21E-07	1.21E-07	1.25E-07	1.56E-07	7.77E-08	1.04E-07	7.70E-08	7.70E-08
37	Liver	6.13E-08	6.13E-08	1.47E-07	1.47E-07	5.65E-08	5.65E-08	2.82E-07	2.82E-07	2.40E-07	3.15E-06	1.48E-07	1.48E-07	1.85E-07	1.85E-07
38	Liver	2.41E-07	7.59E-07	3.15E-07	3.15E-07	1.55E-07	1.55E-07	2.97E-07	2.97E-07	1.23E-07	1.23E-07	1.95E-07	1.95E-07	7.17E-08	7.17E-08
39	Liver	1.93E-08	1.93E-08	1.54E-07	1.54E-07	1.86E-08	1.86E-08	1.87E-07	1.87E-07	1.87E-07	2.07E-07	9.50E-08	9.50E-08	7.75E-08	7.75E-08
40	Liver	2.85E-08	2.85E-08	1.62E-07	1.62E-07	1.31E-07	1.31E-07	1.98E-07	1.98E-07	2.22E-07	2.22E-07	2.19E-07	2.19E-07	2.36E-07	2.36E-07
41	Liver	9.06E-08	3.02E-07	6.39E-07	6.39E-07	2.30E-07	2.30E-07	5.41E-07	6.08E-07	6.80E-07	6.80E-07	5.57E-07	5.86E-07	1.14E-07	1.14E-07
42	Liver	1.31E-07	1.64E-07	7.80E-07	8.11E-07	4.48E-07	4.48E-07	3.95E-07	3.95E-07	3.75E-07	3.75E-07	6.57E-07	6.57E-07	9.90E-08	9.90E-08
43	Liver	9.21E-08	9.21E-08	4.31E-07	4.31E-07	2.57E-07	2.57E-07	5.52E-07	5.52E-07	5.34E-07	5.34E-07	7.75E-07	7.75E-07	1.76E-07	1.76E-07
44	Liver	3.18E-07	7.32E-07	9.83E-07	9.83E-07	3.05E-07	3.05E-07	7.61E-07	7.61E-07	8.04E-07	8.04E-07	6.50E-07	6.81E-07	5.91E-08	5.91E-08
45	Liver	2.31E-07	2.31E-07	2.20E-06	2.35E-06	1.08E-06	1.08E-06	1.67E-06	1.70E-06	1.80E-06	1.80E-06	9.36E-07	9.36E-07	3.16E-07	3.16E-07
46	Liver	4.61E-07	7.08E-07	1.34E-06	1.38E-06	5.45E-07	5.77E-07	9.25E-07	9.60E-07	1.39E-06	1.42E-06	7.49E-07	7.83E-07	4.33E-07	5.41E-07
47	Liver	3.02E-08	3.02E-08	1.58E-06	1.67E-06	6.97E-07	6.97E-07	1.10E-06	1.13E-06	1.08E-06	1.11E-06	1.03E-06	1.06E-06	1.81E-07	1.81E-07
48	Liver	2.33E-07	2.66E-07	1.09E-06	1.13E-06	4.07E-07	4.07E-07	8.83E-07	9.18E-07	7.24E-07	7.24E-07	9.08E-07	9.40E-07	4.11E-07	4.11E-07

Appendix D2 Table 12: Locus specific mutagenicity data of MutaMouse males post exposure to BbF for 180-days

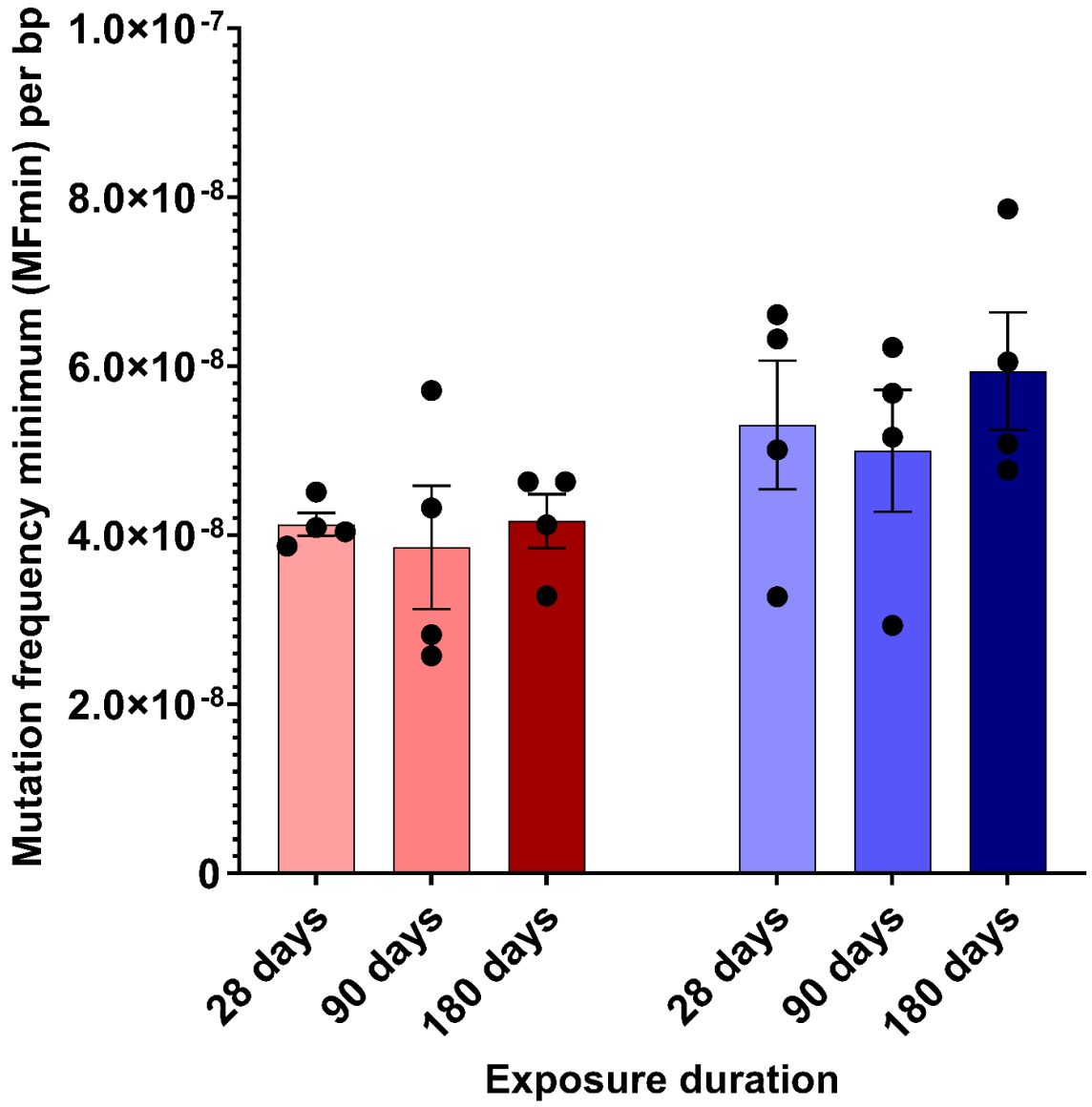
Animal ID	Sample tissue	Exposure duration	BbF dose (mg/kg/day)	Chr. 1 MFmin	Chr. 1 MFmax	Chr. 1.2 MFmin	Chr. 1.2 MFmax	Chr. 2 MFmin	Chr. 2 MFmax	Chr. 3 MFmin	Chr. 3 MFmax	Chr. 4 MFmin	Chr. 4 MFmax	Chr. 5 MFmin	Chr. 5 MFmax	Chr. 6 MFmin	Chr. 6 MFmax	Chr. 7 MFmin	Chr. 7 MFmax	Chr. 8 MFmin	Chr. 8 MFmax	Chr. 9 MFmin	Chr. 9 MFmax	Chr. 10 MFmin	Chr. 10 MFmax	Chr. 11 MFmin	Chr. 11 MFmax	Chr. 12 MFmin	Chr. 12 MFmax	Chr. 13 MFmin	Chr. 13 MFmax	Chr. 14 MFmin	Chr. 14 MFmax	Chr. 15 MFmin	Chr. 15 MFmax	Chr. 16 MFmin	Chr. 16 MFmax	Chr. 17 MFmin	Chr. 17 MFmax	Chr. 18 MFmin	Chr. 18 MFmax	Chr. 19 MFmin	Chr. 19 MFmax
49	Bone Marrow	180-Days	0	3.95E-08	3.95E-08	1.68E-08	1.68E-08	2.06E-08	2.06E-08	3.97E-08	3.97E-08	3.94E-08	3.94E-08	2.02E-08	2.02E-08	0	0	3.45E-08	3.45E-08	2.40E-08	2.40E-08	7.39E-08	1.29E-07	0	0	1.22E-07	1.22E-07	0	0	0	0	5.70E-08	7.60E-08	3.62E-08	3.62E-08	1.97E-07	1.97E-07	2.36E-08	2.36E-08	7.27E-08	7.27E-08	3.82E-08	3.82E-08
50	Bone Marrow	180-Days	0	0	0	3.06E-08	3.06E-08	5.20E-08	5.20E-08	5.05E-08	5.05E-08	8.47E-08	8.47E-08	5.06E-08	5.06E-08	1.58E-08	1.58E-08	4.44E-08	4.44E-08	7.74E-08	9.67E-08	3.29E-08	3.29E-08	1.52E-08	3.04E-08	1.23E-07	1.23E-07	6.27E-08	2.04E-07	4.67E-08	7.78E-08	5.03E-08	6.70E-08	1.54E-08	1.54E-08	3.74E-08	5.60E-08	9.41E-08	9.41E-08	6.40E-08	6.40E-08	0	0
51	Bone Marrow	180-Days	0	2.20E-08	2.20E-08	0	0	0	0	2.17E-08	2.17E-08	2.21E-08	2.21E-08	4.37E-08	4.37E-08	2.06E-08	2.06E-08	1.88E-08	1.88E-08	1.79E-07	1.79E-07	2.07E-08	2.07E-08	1.91E-08	1.91E-08	4.43E-08	4.43E-08	3.96E-08	3.96E-08	3.90E-08	9.76E-08	4.30E-08	1.51E-07	7.87E-08	7.87E-08	2.40E-08	2.40E-08	2.49E-08	2.49E-08	2.04E-08	2.04E-08	0	0
52	Bone Marrow	180-Days	0	8.76E-08	1.05E-07	1.60E-08	1.60E-08	3.55E-08	3.55E-08	1.73E-08	1.73E-08	5.36E-08	5.36E-08	3.49E-08	1.40E-07	3.37E-08	3.37E-08	1.06E-07	1.06E-07	0	0	6.65E-08	9.97E-08	3.18E-08	3.18E-08	8.78E-08	1.05E-07	3.28E-08	3.28E-08	1.63E-08	6.51E-08	5.25E-08	5.25E-08	1.58E-08	1.58E-08	9.45E-08	9.45E-08	5.90E-08	5.90E-08	5.01E-08	6.68E-08	3.40E-08	3.40E-08
53	Bone Marrow	180-Days	1.56	6.98E-08	6.98E-08	7.59E-08	9.11E-08	3.57E-08	3.57E-08	5.25E-08	5.25E-08	6.85E-08	8.56E-08	5.19E-08	1.73E-07	4.76E-08	4.76E-08	4.52E-08	4.52E-08	7.81E-08	7.81E-08	0	0	1.51E-08	3.02E-08	1.27E-07	1.27E-07	3.14E-08	2.04E-07	3.13E-08	3.13E-08	0	0	7.87E-08	1.10E-07	5.81E-08	5.81E-08	5.85E-08	9.75E-08	4.86E-08	4.86E-08	3.32E-08	4.99E-08
54	Bone Marrow	180-Days	1.56	1.56E-08	1.56E-08	2.83E-08	2.83E-08	1.59E-08	1.59E-08	6.19E-08	6.19E-08	3.13E-08	3.13E-08	7.77E-08	7.77E-08	1.46E-08	2.92E-08	1.35E-08	1.35E-08	1.80E-08	1.80E-08	4.51E-08	4.51E-08	4.21E-08	4.21E-08	1.28E-07	1.60E-07	5.78E-08	7.23E-08	2.87E-08	2.87E-08	4.62E-08	4.62E-08	2.82E-08	2.82E-08	5.15E-08	6.87E-08	3.47E-08	3.47E-08	2.94E-08	2.94E-08	4.52E-08	4.52E-08
55	Bone Marrow	180-Days	1.56	0	0	4.78E-08	4.78E-08	1.78E-08	1.78E-08	0	0	8.59E-08	1.20E-07	8.55E-08	8.55E-08	1.61E-08	1.61E-08	1.21E-07	1.36E-07	1.75E-07	1.94E-07	3.33E-08	3.33E-08	1.56E-08	3.13E-08	7.10E-08	1.07E-07	0	0	6.49E-08	6.49E-08	6.83E-08	6.83E-08	6.28E-08	6.28E-08	5.72E-08	5.72E-08	1.90E-08	1.90E-08	4.95E-08	4.95E-08	8.52E-08	8.52E-08
56	Bone Marrow	180-Days	1.56	0	0	3.01E-08	3.01E-08	3.48E-08	3.48E-08	6.78E-08	6.78E-08	0	0	6.77E-08	6.77E-08	6.24E-08	6.24E-08	2.94E-08	2.94E-08	7.70E-08	7.70E-08	1.15E-07	1.31E-07	0	0	1.24E-07	1.77E-07	7.69E-08	7.69E-08	7.68E-08	7.68E-08	9.98E-08	9.98E-08	0	0	7.54E-08	7.54E-08	3.82E-08	3.82E-08	4.74E-08	4.74E-08	4.85E-08	4.85E-08
57	Bone Marrow	180-Days	3.125	4.64E-08	4.64E-08	5.44E-08	5.44E-08	1.59E-08	1.59E-08	6.14E-08	6.14E-08	4.61E-08	4.61E-08	3.10E-08	3.10E-08	1.42E-08	1.42E-08	5.39E-08	1.08E-07	1.60E-07	1.60E-07	8.87E-08	1.04E-07	1.35E-08	1.35E-08	1.12E-07	3.05E-07	4.21E-08	4.21E-08	5.56E-08	3.61E-07	1.36E-07	2.41E-07	4.19E-08	6.99E-08	1.53E-07	1.87E-07	8.82E-08	8.82E-08	5.74E-08	1.86E-07	2.95E-08	2.95E-08
58	Bone Marrow	180-Days	3.125	3.43E-08	3.43E-08	5.72E-08	5.72E-08	8.64E-08	8.64E-08	3.43E-08	3.43E-08	5.15E-08	5.15E-08	3.52E-08	3.52E-08	6.49E-08	1.14E-07	2.94E-08	2.94E-08	8.60E-08	8.60E-08	7.85E-08	1.26E-07	2.88E-08	2.88E-08	1.39E-07	1.73E-07	1.53E-08	3.07E-08	1.47E-08	2.95E-08	9.80E-08	9.80E-08	4.62E-08	4.62E-08	1.32E-07	1.32E-07	1.47E-07	1.47E-07	1.56E-08	1.56E-08	3.22E-08	3.22E-08
59	Bone Marrow	180-Days	3.125	1.08E-07	1.24E-07	4.03E-08	5.37E-08	4.76E-08	4.76E-08	6.18E-08	6.18E-08	9.28E-08	9.28E-08	6.25E-08	1.56E-07	4.29E-08	5.72E-08	1.34E-08	1.34E-08	1.27E-07	1.27E-07	7.38E-08	7.38E-08	2.69E-08	2.69E-08	1.60E-07	1.60E-07	4.18E-08	4.18E-08	4.15E-08	4.15E-08	1.36E-07	1.81E-07	5.59E-08	5.59E-08	1.20E-07	1.20E-07	1.06E-07	1.06E-07	1.15E-07	1.44E-07	5.87E-08	8.81E-08
60	Bone Marrow	180-Days	3.125	1.59E-08	6.37E-08	1.46E-08	1.46E-08	3.18E-08	3.18E-08	9.37E-08	9.37E-08	3.22E-08	3.22E-08	1.55E-08	1.55E-08	1.52E-08	1.52E-08	6.92E-08	9.68E-08	1.10E-07	1.10E-07	7.57E-08	1.51E-07	1.16E-07	1.16E-07	1.43E-07	1.75E-07	3.00E-08	3.00E-08	5.95E-08	5.95E-08	6.36E-08	6.36E-08	7.12E-08	8.54E-08	1.36E-07	1.36E-07	3.47E-08	3.47E-08	7.60E-08	9.12E-08	1.54E-08	1.54E-08
61	Bone Marrow	180-Days	6.25	4.94E-08	4.94E-08	2.87E-08	2.87E-08	6.68E-08	6.68E-08	3.26E-08	3.26E-08	8.27E-08	7.28E-07	9.89E-08	9.89E-08	0	0	1.42E-08	1.42E-08	1.97E-07	1.97E-07	1.23E-07	1.39E-07	7.18E-08	7.18E-08	2.33E-07	3.67E-07	4.51E-08	1.20E-07	7.32E-08	7.32E-08	8.01E-08	9.61E-08	5.87E-08	5.87E-08	1.97E-07	1.97E-07	1.52E-07	2.28E-07	7.69E-08	7.69E-08	4.76E-08	4.76E-08
62	Bone Marrow	180-Days	6.25	5.57E-08	5.57E-08	5.02E-08	6.69E-08	1.12E-07	1.49E-07	3.63E-08	9.07E-08	5.53E-08	5.53E-08	1.09E-07	1.28E-07	1.21E-07	1.55E-07	4.79E-08	4.79E-08	1.05E-07	1.68E-07	1.78E-07	1.78E-07	4.96E-08	1.98E-07	3.43E-07	3.62E-07	1.37E-07	1.37E-07	5.09E-08	5.09E-08	1.28E-07	2.19E-07	3.34E-08	3.34E-08	1.63E-07	1.83E-07	2.04E-07	3.05E-07	6.93E-08	8.67E-08	5.36E-08	5.36E-08
63	Bone Marrow	180-Days	6.25	1.09E-07	1.09E-07	3.29E-08	3.29E-08	1.48E-07	3.33E-07	1.80E-08	1.80E-08	1.27E-07	1.27E-07	1.43E-07	3.41E-07	1.70E-08	1.70E-08	4.76E-08	4.76E-08	1.87E-07	2.08E-07	5.22E-08	5.22E-08	1.13E-07	1.29E-07	1.29E-07	1.29E-07	1.69E-08	1.69E-08	5.02E-08	5.02E-08	1.60E-07	3.21E-07	4.94E-08	9.87E-08	1.78E-07	1.98E-07	1.81E-07	4.23E-07	6.84E-08	6.84E-08	5.29E-08	8.82E-08
64	Bone Marrow	180-Days	6.25	1.29E-07	2.29E-07	2.54E-08	2.54E-08	1.47E-07	3.08E-07	1.14E-07	1.14E-07	1.28E-07	1.42E-07	8.54E-08	8.54E-08	2.65E-08	2.65E-08	0	0	1.64E-07	4.09E-07	1.36E-07	1.50E-07	6.31E-08	6.31E-08	2.49E-07	7.19E-07	5.22E-08	5.22E-08	5.21E-08	4.69E-07	1.97E-07	1.97E-07	3.87E-08	3.87E-08	2.85E-07	3.01E-07	2.24E-07	3.04E-07	1.08E-07	1.48E-07	4.10E-08	6.84E-08
65	Bone Marrow	180-Days	12.5	1.48E-07	2.40E-07	8.16E-08	1.79E-07	1.33E-07	1.33E-07	5.54E-08	5.54E-08	1.30E-07	1.49E-07	1.11E-07	1.11E-07	1.20E-07	1.20E-07	6.42E-08	6.42E-08	3.67E-07	5.40E-07	1.75E-07	2.10E-07	3.23E-08	3.23E-08	4.54E-07	9.27E-07	1.69E-08	1.69E-08	4.99E-08	4.99E-08	2.35E-07	8.12E-07	6.64E-08	6.64E-08	2.84E-07	3.04E-07	2.94E-07	3.78E-07	1.90E-07	7.43E-07	8.91E-08	8.91E-08
66	Bone Marrow	180-Days	12.5	8.60E-08	8.60E-08	1.37E-07	6.08E-07	1.76E-07	1.76E-07	8.51E-08	8.51E-08	2.38E-07	3.06E-07	1.03E-07	1.03E-07	3.17E-08	3.17E-08	1.05E-07	1.05E-07	2.54E-07	2.94E-07	1.32E-07	1.32E-07	3.01E-08	3.01E-08	3.74E-07	6.95E-07	7.84E-08	7.84E-08	1.55E-08	1.55E-08	2.85E-07	5.71E-07	1.55E-08	4.97E-07	1.90E-07	2.86E-07	2.89E-07	3.85E-07	7.99E-08	1.12E-07	4.92E-08	6.57E-08
67	Bone Marrow	180-Days	12.5	2.98E-07	2.98E-07	9.38E-08	1.09E-07	7.46E-08	7.46E-08	5.49E-08	5.49E-08	1.46E-07	1.46E-07	1.70E-07	1.89E-07	5.13E-08	5.13E-08	3.24E-08	3.24E-08	4.08E-07	6.44E-07	2.47E-07	3.00E-07	1.10E-07	1.10E-07	4.64E-07	8.31E-07	6.63E-08	8.28E-08	4.88E-08	4.88E-08	3.01E-07	3.90E-07	1.87E-07	1.87E-07	2.64E-07	3.05E-07	2.98E-07	3.61E-07	1.02E-07	1.02E-07	1.73E-08	1.73E-08
68	Bone Marrow	180-Days	12.5	6.75E-08	6.75E-08	1.10E-07	1.79E-07	2.11E-07	3.16E-07	1.20E-07	3.61E-07	1.01E-07	1.01E-07	1.57E-07	1.57E-07	9.26E-08	9.26E-08	1.48E-07	1.48E-07	3.81E-07	6.21E-07	1.44E-07	1.44E-07	1.40E-07	1.55E-07	5.11E-07	6.52E-07	8.85E-08	1.03E-07	8.79E-08	8.79E-08	2.92E-07	4.06E-07	6.13E-08	7.66E-08	3.58E-07	3.58E-07	3.48E-07	6.35E-07	1.24E-07	1.40E-07	4.74E-08	4.74E-08
69	Bone Marrow	180-Days	25	1.58E-07	1.75E-07	5.89E-08	5.89E-08	1.60E-07	3.56E-07	1.22E-07	1.22E-07	1.75E-07	4.02E-07	2.29E-07	2.64E-07	1.64E-07	1.64E-07	9.05E-08	1.06E-07	5.70E-07	1.26E-06	3.18E-07	1.15E-06	7.58E-08	9.09E-08	5.20E-07	7.17E-07	6.31E-08	6.31E-08	4.66E-08	4.66E-08	4.45E-07	2.19E-06	9.44E-08	2.36E-07	3.09E-07	1.23E-06	5.81E-07	7.22E-07	1.29E-07	1.45E-07	1.13E-07	2.59E-07
70	Bone Marrow	180-Days	25	1.79E-07	1.95E-07	2.16E-07	2.47E-07	2.41E-07	2.73E-07	1.42E-07	2.99E-07	1.96E-07	9.14E-07	1.89E-07	2.21E-07	1.08E-07	2.01E-07	9.85E-08	1.13E-07	3.67E-07	8.06E-07	2.95E-07	7.60E-07	1.21E-07	1.21E-07	6.79E-07	9.54E-07	2.02E-07	7.14E-07	1.54E-07	1.54E-07	2.78E-07	6.37E-07	7.21E-08	7.21E-08	3.09E-07	3.09E-07	3.08E-07	5.30E-07	1.71E-07	3.11E-07	7.94E-08	1.11E-07
71	Bone Marrow	180-Days	25	2.37E-07	2.54E-07	5.97E-08	1.64E-07	8.61E-08	8.61E-08	1.68E-08	1.68E-08	1.86E-07	2.19E-07	1.85E-07	3.03E-07	1.57E-08	3.14E-08	7.35E-08	7.35E-08	5.69E-07	2.25E-06	1.76E-07	2.40E-07	5.89E-08	5.89E-08	4.99E-07	2.22E-06	3.07E-08	3.07E-08	7.54E-08	9.05E-08	4.43E-07	1.92E-06	1.52E-08	1.52E-08	3.68E-07	5.89E-07	4.59E-07	6.69E-07	4.75E-08	4.75E-08	4.83E-08	1.13E-07
72	Bone Marrow	180-Days	25	1.51E-07	1.51E-07	8.76E-08	1.02E-07	1.67E-07	2.67E-07	9.94E-08	6.29E-07	1.68E-07	1.68E-07	1.36E-07	5.42E-07	1.27E-07	1.27E-07	1.01E-07	1.01E-07	4.12E-07	7.00E-07	2.81E-07	4.84E-07	1.03E-07	1.03E-07	4.23E-07	9.98E-07	9.12E-08	9.12E-08	7.49E-08	8.99E-08	3.90E-07	6.51E-07	1.04E-07	2.98E-07	1.99E-07	2.90E-07	3.53E-07	5.68E-07	2.50E-07	2.65E-07	4.72E-08	4.72E-08
49	Liver	180-Days	0	6.15E-08	6.15E-08	4.51E-08	6.01E-08	4.52E-08	4.52E-08	5.88E-08	7.35E-08	8.83E-08	8.83E-08	7.38E-08	7.38E-08	2.96E-08	2.96E-08	4.11E-08	4.11E-08	3.17E-08	3.17E-08	8.53E-08	8.53E-08	6.05E-08	6.05E-08	4.41E-08	4.41E-08	4.59E-08	4.59E-08	1.66E-07	1.96E-07	7.39E-08	8.87E-08	0	0	6.22E-08	6.22E-08	9.17E-08	1.07E-07	7.43E-08	7.43E-08	3.17E-08	3.17E-08
50	Liver	180-Days	0	7.55E-08	7.55E-08	7.33E-08	7.33E-08	7.52E-08	1.00E-07	2.44E-08	2.44E-08	0	0	0	0	0	0	4.55E-08	6.82E-08	9.92E-08	9.92E-08	4.77E-08	4.77E-08	0	0	7.32E-08	7.32E-08	0	0	9.79E-08	1.96E-07	2.42E-08	2.42E-08	7.15E-08	7.15E-08	1.04E-07	1.04E-07	2.48E-08	2.48E-08	9.78E-08	9.78E-08	2.60E-08	2.60E-08
51	Liver	180-Days	0	1.87E-07	2.08E-07	4.14E-08	4.14E-08	2.03E-08	2.03E-08	2.01E-08	2.01E-08	9.96E-08	9.96E-08	0	0	4.04E-08	6.06E-08	5.60E-08	1.68E-07	2.10E-08	2.10E-08	9.79E-08	1.18E-07	0	0	4.01E-08	4.01E-08	6.28E-08	6.28E-08	0	0	6.04E-08	6.04E-08	1.98E-08	1.98E-08	2.10E-08	2.10E-08	1.42E-07	1.42E-07	2.04E-08	2.04E-08	6.48E-08	8.64E-08
52	Liver	180-Days	0	8.05E-08	8.05E-08	3.96E-08	3.96E-08	4.07E-08	4.07E-08	7.86E-08	9.83E-08	1.16E-07	1.35E-07	9.88E-08	9.88E-08	3.85E-08	3.85E-08	7.44E-08	7.44E-08	8.06E-08	8.06E-08	5.75E-08	5.75E-08	8.42E-08	8.42E-08	9.94E-08	1.19E-07	1.61E-07	1.61E-07	5.93E-08	5.93E-08	1.17E-07	1.17E-07	7.76E-08	9.70E-08	1.26E-07	1.26E-07	0	0	7.93E-08	7.93E-08	6.47E-08	6.47E-08
53	Liver	180-Days	1.56	0	0	1.01E-07	1.41E-07	7.82E-08	9.77E-08	7.69E-08	7.69E-08	1.15E-07	1.34E-07	5.71E-08	5.71E-08	5.84E-08	1.17E-07	1.81E-08	1.81E-08	9.85E-08	9.85E-08	1.68E-07	1.68E-07	1.23E-07	1.23E-07	7.58E-08	7.58E-08	8.05E-08	8.05E-08	2.02E-08	2.02E-08	9.68E-08	1.16E-07	5.67E-08	5.67E-08	4.01E-08	4.01E-08	5.76E-08	5.76E-08	3.92E-08	3.92E-08	1.26E-07	1.26E-07
54	Liver	180-Days	1.56	8.29E-08	8.29E-08	8.21E-08	8.21E-08	0	0	5.99E-08	7.99E-08	5.90E-08	5.90E-08	6.01E-08	6.01E-08	8.00E-08	8.00E-08	7.53E-08	7.53E-08	2.06E-08	2.06E-08	7.87E-08	7.87E-08	4.12E-08	4.12E-08	8.01E-08	8.01E-08	6.21E-08	6.21E-08	4.12E-08	4.12E-08	2.02E-08	2.02E-08	5.93E-08	5.93E-08	2.13E-08	2.13E-08	6.10E-08	6.10E-08	8.09E-08	8.09E-08	8.66E-08	1.08E-07
55	Liver	180-Days	1.56	1.34E-07	1.34E-07	1.25E-07	1.87E-07	2.25E-08	2.25E-08	4.35E-08	4.35E-08	1.06E-07	1.06E-07	6.54E-08	6.54E-08	6.34E-08	6.34E-08	6.01E-08	6.01E-08	1.58E-07	1.58E-07	2.09E-08	2.09E-08	6.38E-08	6.38E-08	4.35E-08	4.35E-08	4.31E-08	4.31E-08	2.11E-08	4.23E-08	1.49E-07	1.49E-07	4.23E-08	4.23E-08	2.30E-08	2.30E-08	1.14E-07	1.37E-07	4.29E-08	4.29E-08	4.53E-08	4.53E-08
56	Liver	180-Days	1.56	3.72E-08	3.72E-08	1.84E-08	1.84E-08	7.40E-08	7.40E-08	1.81E-08	1.81E-08	1.78E-08	1.78E-08	1.81E-08	1.81E-08	3.56E-08	3.56E-08	6.75E-08	6.75E-08	1.30E-07	1.49E-07	1.06E-07	1.06E-07	9.11E-08	9.11E-08	1.81E-08	1.81E-08	5.57E-08	5.57E-08	5.50E-08	9.17E-08	1.63E-07	1.63E-07	3.57E-08	3.57E-08	3.84E-08	3.84E-08	1.84E-08	1.84E-08	9.09E-08	1.09E-07	3.84E-08	3.84E-08
57	Liver	180-Days	3.125	5.18E-08	5.18E-08	1.00E-07	1.17E-07	6.57E-08	6.57E-08	8.08E-08	8.08E-08	6.55E-08	6.55E-08	3.28E-08	3.28E-08	4.95E-08	4.95E-08	7.62E-08	7.62E-08	3.54E-08	3.54E-08	6.35E-08	7.94E-08	5.05E-08	5.05E-08	4.88E-08	4.88E-08	6.74E-08	8.43E-08	3.31E-08	3.31E-08	3.28E-08	3.28E-08	7.95E-08	7.95E-08	6.87E-08	6.87E-08	8.47E-08	8.47E-08	8.17E-08	8.17E-08	3.49E-08	3.49E-08
58	Liver	180-Days	3.125	6.31E-08	6.31E-08	2.04E-08	2.04E-08	4.21E-08	4.21E-08	6.09E-08	6.09E-08	4.00E-08	4.00E-08	4.06E-08	4.06E-08	2.00E-08	2.00E-08	1.90E-08	1.90E-08	1.25E-07	2.09E-07	9.98E-08	9.98E-08	1.22E-07	1.42E-07	4.11E-08	4.11E-08	1.03E-07	1.24E-07	4.07E-08	4.07E-08	2.02E-07	2.02E-07	1.20E-07	3.79E-07	4.32E-08	4.32E-08	6.26E-08	6.26E-08	1.01E-07	1.01E-07	2.16E-08	2.16E-08
59	Liver	180-Days	3.125	7.67E-08	7.67E-08	7.67E-08	7.67E-08	7.58E-08	7.58E-08	7.41E-08	7.41E-08	3.62E-08	3.62E-08	5.58E-08	5.58E-08	5.50E-08	5.50E-08	1.74E-08	1.74E-08	9.46E-08	9.46E-08	5.46E-08	5.46E-08	3.77E-08	3.77E-08	2.42E-07	8.19E-07	1.16E-07	1.16E-07	9.50E-08	9.50E-08	5.56E-08	5.56E-08	3.65E-08	5.47E-08	9.81E-08	9.81E-08	1.68E-07	1.68E-07	3.73E-08	3.73E-08	9.96E-08	1.39E-07
60	Liver	180-Days	3.125	1.04E-07	1.04E-07	0	0	0	0	6.03E-08	6.03E-08	1.79E-07	1.79E-07	2.01E-08	2.01E-08	5.98E-08	5.98E-08	1.88E-08	1.88E-08	4.17E-08	4.17E-08	1.37E-07	1.37E-07	4.15E-08	4.15E-08	4.03E-08	4.03E-08	1.25E-07	1.25E-07	6.16E-08	8.21E-08	1.21E-07	1.21E-07	3.97E-08	3.97E-08	1.27E-07	1.27E-07	4.13E-08	4.13E-08	6.09E-08	6.09E-08	0	0
61	Liver	180-Days	6.25	1.57E-07	1.57E-07	1.34E-07	1.34E-07	1.13E-07	1.13E-07	1.98E-07	1.98E-07	8.54E-08	8.54E-08	4.18E-07	4.18E-07	1.72E-07	1.72E-07	6.16E-08	6.16E-08	1.79E-07	1.79E-07	1.29E-07	1.29E-07	2.19E-08	2.19E-08	1.11E-07	1.11E-07	9.00E-08	9.00E-08	1.11E-07	1.11E-07	2.17E-07	2.17E-07	1.29E-07	1.29E-07	2.32E-07	2.32E-07	2.20E-07	2.20E-07	1.54E-07	1.54E-07	4.65E-08	4.65E-08
62	Liver	180-Days	6.25	1.38E-07	1.38E-07	9.34E-08	9.34E-08	2.49E-07	2.49E-07	6.66E-08	6.66E-08	1.33E-07	1.33E-07	1.33E-07	1.33E-07	2.25E-08	2.25E-08	1.06E-07	1.06E-07	1.15E-07	1.15E-07	1.09E-07	1.09E-07	4.64E-08	4.64E-08	2.00E-07	2.00E-07	4.72E-08	7.08E-08	9.33E-08	1.17E-07	2.70E-07	3.15E-07	8.84E-08	8.84E-08	1.40E-07	4.68E-07	2.01E-07	2.01E-07	1.35E-07	1.58E-07	9.65E-08	9.65E-08
63	Liver	180-Days	6.25	2.37E-07	2.37E-07	2.33E-07	2.33E-07	1.05E-07	1.05E-07	1.64E-07	1.64E-07	3.28E-07	3.28E-07	1.85E-07	5.76E-07	1.03E-07	1.03E-07	5.81E-08	5.81E-08	3.26E-07	3.26E-07	2.02E-07	2.02E-07	8.45E-08	8.45E-08	2.26E-07	2.26E-07	1.28E-07	1.28E-07	1.06E-07	1.27E-07	2.08E-07	2.08E-07	0	0	1.51E-07	1.51E-07	2.32E-07	2.32E-07	8.32E-08	8.32E-08	4.42E-08	4.42E-08
64	Liver	180-Days	6.25	1.90E-07	1.90E-07	1.39E-07	1.39E-07	6.32E-08	6.32E-08	2.24E-07	2.24E-07	2.39E-07	2.39E-07	8.26E-08	1.03E-07	7.95E-08	7.95E-08	7.56E-08	7.56E-08	2.35E-07	2.35E-07	1.19E-07	1.19E-07	5.99E-08	5.99E-08	2.49E-07	2.69E-07	1.44E-07	1.44E-07	4.03E-08	4.03E-08	3.03E-07	3.43E-07	3.99E-08	3.99E-08	1.53E-07	1.53E-07	1.70E-07	1.70E-07	1.82E-07	1.82E-07	2.14E-08	2.14E-08
65	Liver	180-Days	12.5	6.88E-07	7.32E-07	3.69E-07	3.69E-07	3.10E-07	3.10E-07	3.89E-07	3.89E-07	4.87E-07	4.87E-07	3.44E-07	3.44E-07	1.92E-07	1.92E-07	4.22E-07	4.22E-07	7.47E-07	7.47E-07	3.81E-07	4.02E-07	1.10E-07	1.10E-07	5.62E-07	5.62E-07	2.87E-07	2.87E-07	1.95E-07	1.95E-07	6.68E-07	7.32E-07	1.90E-07	1.90E-07	5.93E-07	5.93E-07	4.66E-07	4.66E-07	3.90E-07	4.33E-07	2.07E-07	2.07E-07
66	Liver	180-Days	12.5	4.64E-07	4.64E-07	3.86E-07	3.86E-07	2.67E-07	2.67E-07	3.28E-07	3.28E-07	4.77E-07	4.77E-07	3.44E-07	3.44E-07	2.71E-07	2.71E-07	4.00E-07	4.16E-07	5.84E-07	5.84E-07	3.03E-07	3.03E-07	1.23E-07	1.40E-07	4.53E-07	4.53E-07	3.01E-07	3.01E-07	7.03E-08	1.05E-07	5.99E-07	5.99E-07	2.02E-07	2.02E-07	3.87E-07	4.24E-07	6.84E-07	7.02E-07	3.45E-07	3.62E-07	1.12E-07	1.12E-07
67	Liver	180-Days	12.5	4.10E-07	4.10E-07	3.72E-07	3.94E-07	3.60E-07	4.24E-07	3.97E-07	4.39E-07	5.63E-07	5.84E-07	4.19E-07	4.40E-07	2.94E-07	3.14E-07	3.93E-07	3.93E-07	7.61E-07	7.61E-07	7.16E-07	7.37E-07	1.71E-07	1.71E-07	9.83E-07	1.00E-06	3.93E-07	3.93E-07	2.17E-07	3.04E-07	9.25E-07	9.46E-07	1.87E-07	1.87E-07	7.23E-07	7.66E-07	8.24E-07	8.45E-07	7.00E-07	7.22E-07	1.80E-07	2.02E-07
68	Liver	180-Days	12.5	4.72E-07	4.94E-07	2.55E-07	2.55E-07	2.85E-07	2.85E-07	4.78E-07	4.78E-07	6.75E-07	6.75E-07	3.87E-07	3.87E-07	1.33E-07	1.33E-07	2.66E-07	2.66E-07	4.42E-07	4.42E-07	5.06E-07	5.06E-07	2.50E-07	2.50E-07	6.20E-07	6.42E-07	3.49E-07	3.49E-07	4.57E-08	4.57E-08	5.35E-07	5.35E-07	8.61E-08	8.61E-08	5.83E-07	5.83E-07	5.03E-07	5.03E-07	3.81E-07	3.81E-07	7.15E-08	7.15E-08
69	Liver	180-Days	25	2.13E-06	2.18E-06	5.43E-07	5.43E-07	5.69E-07	5.69E-07	7.75E-07	7.75E-07	1.76E-06	1.81E-06	1.37E-06	1.37E-06	6.76E-07	7.80E-07	6.89E-07	6.89E-07	1.75E-06	1.75E-06	1.03E-06	1.03E-06	5.48E-07	5.48E-07	1.03E-06	1.03E-06	6.88E-07	6.88E-07	3.46E-07	3.46E-07	1.69E-06	1.95E-06	7.90E-07	7.90E-07	1.46E-06	1.46E-06	2.65E-06	2.65E-06	1.46E-06	1.46E-06	1.68E-07	1.68E-07
70	Liver	180-Days	25	1.23E-06	1.25E-06	8.56E-07	9.47E-07	1.05E-06	1.08E-06	7.95E-07	8.30E-07	1.14E-06	1.18E-06	8.85E-07	9.73E-07	8.43E-07	8.78E-07	6.46E-07	6.63E-07	1.85E-06	1.95E-06	1.39E-06	1.41E-06	4.70E-07	4.88E-07	1.54E-06	1.62E-06	1.08E-06	1.14E-06	2.16E-07	2.16E-07	1.71E-06	1.85E-06	9.93E-07	9.93E-07	1.29E-06	1.32E-06	2.05E-06	2.16E-06	1.14E-06	1.17E-06	3.41E-07	3.41E-07
71	Liver	180-Days	25	1.25E-06	1.36E-06	7.01E-07	7.01E-07	9.50E-07	9.50E-07	6.66E-07	6.66E-07	9.01E-07	9.26E-07	9.62E-07	9.87E-07	5.36E-07	5.36E-07	6.23E-07	6.46E-07	1.02E-06	1.05E-06	8.32E-07	8.56E-07	4.15E-07	4.15E-07	9.13E-07	9.61E-07	7.63E-07	8.16E-07	2.54E-07	2.54E-07	1.49E-06	1.51E-06	4.13E-07	4.13E-07	1.20E-06	1.23E-06	1.12E-06	1.12E-06	9.15E-07	9.15E-07	1.07E-07	1.07E-07
72	Liver	180-Days	25	1.08E-06	1.08E-06	7.34E-07	7.51E-07	6.35E-07	6.52E-07	8.18E-07	8.35E-07	1.19E-06	1.23E-06	8.06E-07	8.06E-07	4.67E-07	4.67E-07	5.75E-07	5.75E-07	1.43E-06	1.54E-06	1.02E-06	1.07E-06	5.03E-07	5.03E-07	1.16E-06	1.80E-06	8.42E-07	8.77E-07	2.19E-07	2.36E-07	1.37E-06	1.46E-06	6.59E-07	6.76E-07	1.43E-06	1.52E-06	1.15E-06	1.18E-06	1.26E-06	1.26E-06	2.68E-07	2.86E-07

Appendix D2 Table 13: SigProfiler analysis to determine similairties between induced mutation pattern and COSMIC signatures

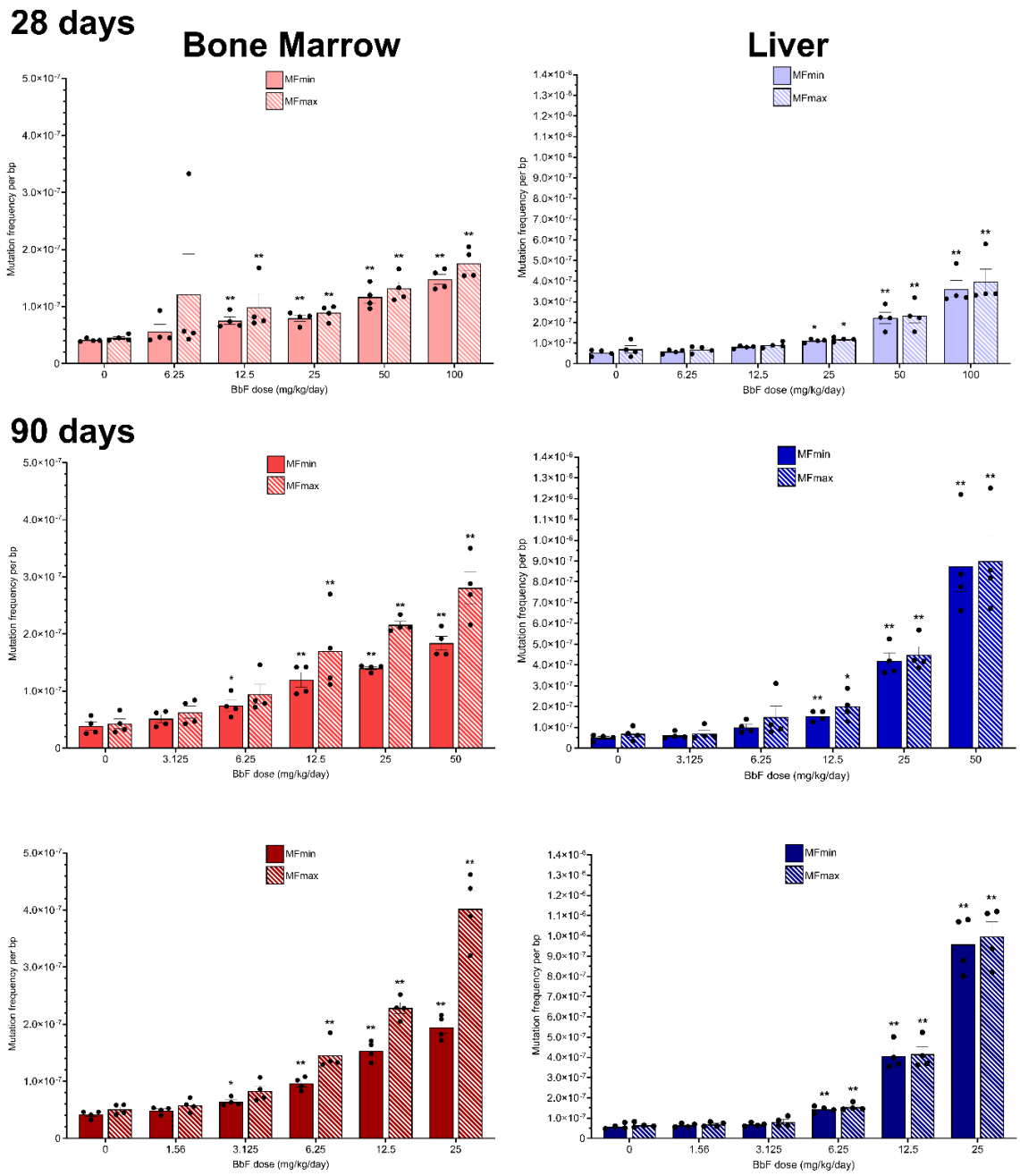
Sample tissue	Exposure duration	BbF dose (mg/kg/day)	SBS1 Mutations	SBS4 Mutations	SBS5 Mutations	SBS17a Mutations	SBS17b Mutations	SBS18 Mutations	SBS40a Mutations	Mutation count	SBS1 % Contribution	SBS4 % Contribution	SBS5 % Contribution	SBS17a % Contribution	SBS17b % Contribution	SBS18 % Contribution	SBS40a % Contribution	Cosine similarity	KL divergence
Bone Marrow	28 days	0	11	0	37	11	4	42	0	105	10.5	0.0	35.2	10.5	3.8	40.0	0.0	0.81	0.55
Bone Marrow	28 days	6.25	6	0	71	0	0	38	0	115	5.2	0.0	61.7	0.0	0.0	33.0	0.0	0.80	0.47
Bone Marrow	28 days	12.5	3	0	93	0	0	110	0	206	1.5	0.0	45.1	0.0	0.0	53.4	0.0	0.86	0.33
Bone Marrow	28 days	25	12	87	115	0	0	0	0	214	5.6	40.7	53.7	0.0	0.0	0.0	0.0	0.79	0.38
Bone Marrow	28 days	50	19	171	140	0	0	0	0	330	5.8	51.8	42.4	0.0	0.0	0.0	0.0	0.86	0.26
Bone Marrow	28 days	100	18	283	155	0	0	0	0	456	3.9	62.1	34.0	0.0	0.0	0.0	0.0	0.90	0.20
Bone Marrow	90 days	0	11	0	72	0	0	0	0	83	13.3	0.0	86.7	0.0	0.0	0.0	0.0	0.70	0.67
Bone Marrow	90 days	3.125	11	0	105	0	0	81	0	197	5.6	0.0	53.3	0.0	0.0	41.1	0.0	0.90	0.31
Bone Marrow	90 days	6.25	5	0	111	0	0	119	0	235	2.1	0.0	47.2	0.0	0.0	50.6	0.0	0.86	0.29
Bone Marrow	90 days	12.5	0	88	101	0	0	78	0	267	0.0	33.0	37.8	0.0	0.0	29.2	0.0	0.89	0.24
Bone Marrow	90 days	25	2	0	214	0	0	202	0	418	0.5	0.0	51.2	0.0	0.0	48.3	0.0	0.89	0.24
Bone Marrow	90 days	50	20	354	171	0	0	0	0	545	3.7	65.0	31.4	0.0	0.0	0.0	0.0	0.92	0.16
Bone Marrow	180 days	0	19	0	135	0	0	0	0	154	12.3	0.0	87.7	0.0	0.0	0.0	0.0	0.74	0.52
Bone Marrow	180 days	1.56	3	0	94	24	14	56	0	191	1.6	0.0	49.2	12.6	7.3	29.3	0.0	0.84	0.37
Bone Marrow	180 days	3.125	17	0	160	0	0	114	0	291	5.8	0.0	55.0	0.0	0.0	39.2	0.0	0.87	0.33
Bone Marrow	180 days	6.25	14	143	114	0	0	157	0	428	3.3	33.4	26.6	0.0	0.0	36.7	0.0	0.89	0.21
Bone Marrow	180 days	12.5	11	273	213	0	0	152	0	649	1.7	42.1	32.8	0.0	0.0	23.4	0.0	0.93	0.13
Bone Marrow	180 days	25	21	600	241	0	0	0	0	862	2.4	69.6	28.0	0.0	0.0	0.0	0.0	0.93	0.13
Liver	28 days	0	9	0	20	0	0	0	94	123	7.3	0.0	16.3	0	0	0	76.4	0.79	0.47
Liver	28 days	6.25	6	0	31	0	0	0	119	156	3.8	0.0	19.9	0	0	0	76.3	0.77	0.53
Liver	28 days	12.5	0	60	0	0	0	0	197	257	0.0	23.3	0.0	0	0	0	76.7	0.88	0.28
Liver	28 days	25	0	179	0	0	0	0	130	309	0.0	57.9	0.0	0	0	0	42.1	0.88	0.24
Liver	28 days	50	0	485	0	0	0	0	219	704	0.0	68.9	0.0	0	0	0	31.1	0.94	0.13
Liver	28 days	100	0	835	0	0	0	0	352	1187	0.0	70.3	0.0	0	0	0	29.7	0.95	0.10
Liver	90 days	0	12	0	88	0	0	0	0	100	12.0	0.0	88.0	0	0	0	0.0	0.72	0.67
Liver	90 days	3.125	10	0	51	0	0	0	89	150	6.7	0.0	34.0	0	0	0	59.3	0.79	0.41
Liver	90 days	6.25	8	122	90	0	0	0	0	220	3.6	55.5	40.9	0	0	0	0.0	0.88	0.27
Liver	90 days	12.5	0	236	0	0	0	0	185	421	0.0	56.1	0.0	0	0	0	43.9	0.90	0.20
Liver	90 days	25	11	825	177	0	0	0	0	1013	1.1	81.4	17.5	0	0	0	0.0	0.97	0.09
Liver	90 days	50	30	1822	254	0	0	0	0	2106	1.4	86.5	12.1	0	0	0	0.0	0.98	0.06
Liver	180 days	0	15	57	135	0	0	0	0	207	7.2	27.5	65.2	0	0	0	0.0	0.83	0.36
Liver	180 days	1.56	27	0	57	0	0	0	130	214	12.6	0.0	26.6	0	0	0	60.7	0.82	0.40
Liver	180 days	3.125	15	68	163	0	0	0	0	246	6.1	27.6	66.3	0	0	0	0.0	0.86	0.23
Liver	180 days	6.25	16	287	171	0	0	0	0	474	3.4	60.5	36.1	0	0	0	0.0	0.94	0.12
Liver	180 days	12.5	0	1059	0	0	0	0	390	1449	0.0	73.1	0.0	0	0	0	26.9	0.96	0.08
Liver	180 days	25	28	2549	444	0	0	0	0	3021	0.9	84.4	14.7	0	0	0	0.0	0.98	0.06

Appendix D2 Table 14: BMD modeling on MFmin, MFmax and MFmin of predominant mutation subtype (C:G>A:T) in BM and liver after prolonged exposure durations

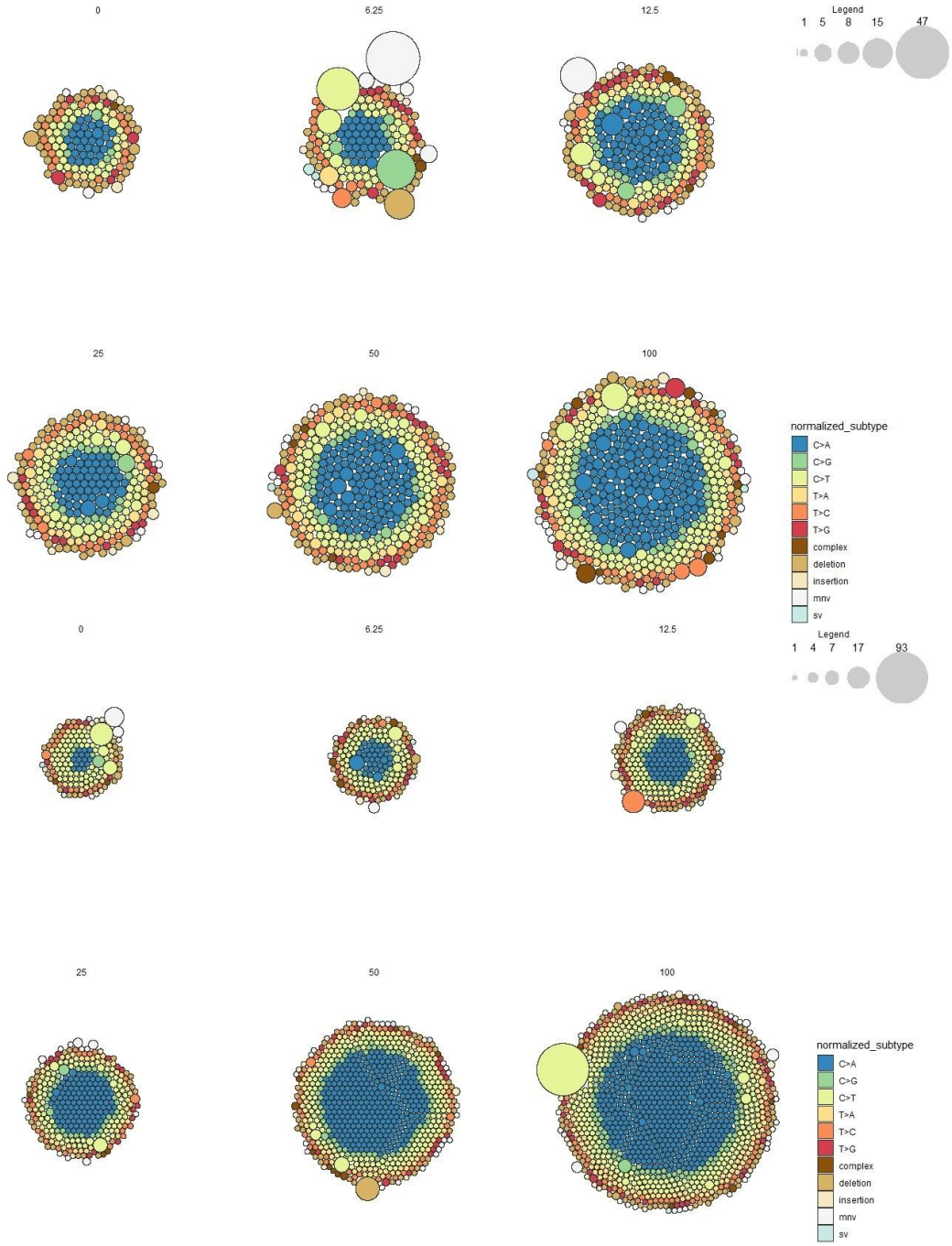
Exposure duration	Sample Tissue	Endpoint	n	CES	Bootstrap	BMDL	BMDU	Log10 BMDL	Log10 BMDU
28-Days	Bone Marrow	MFmin	24	0.5	200	3.60	12.80	0.56	1.11
	Bone Marrow	MFmax				0.07	26.50	-1.14	1.42
	Bone Marrow	MFmin C:G>A:T				0.68	15.90	-0.17	1.20
	Liver	MFmin	24	0.5	200	10.10	18.40	1.00	1.26
	Liver	MFmax				6.02	24.00	0.78	1.38
	Liver	MFmin C:G>A:T				1.36	5.96	0.13	0.78
90-Days	Bone Marrow	MFmin	24	0.5	200	1.56	5.82	0.19	0.76
	Bone Marrow	MFmax				1.00	5.31	0.00	0.73
	Bone Marrow	MFmin C:G>A:T				0.22	2.09	-0.67	0.32
	Liver	MFmin	24	0.5	200	2.34	6.31	0.37	0.80
	Liver	MFmax				1.21	7.64	0.08	0.88
	Liver	MFmin C:G>A:T				0.97	2.44	-0.01	0.39
180-Days	Bone Marrow	MFmin	24	0.5	200	2.29	3.75	0.36	0.57
	Bone Marrow	MFmax				1.63	3.42	0.21	0.53
	Bone Marrow	MFmin C:G>A:T				0.94	2.20	-0.03	0.34
	Liver	MFmin	24	0.5	200	3.14	4.61	0.50	0.66
	Liver	MFmax				2.91	4.57	0.46	0.66
	Liver	MFmin C:G>A:T				2.11	3.73	0.32	0.57



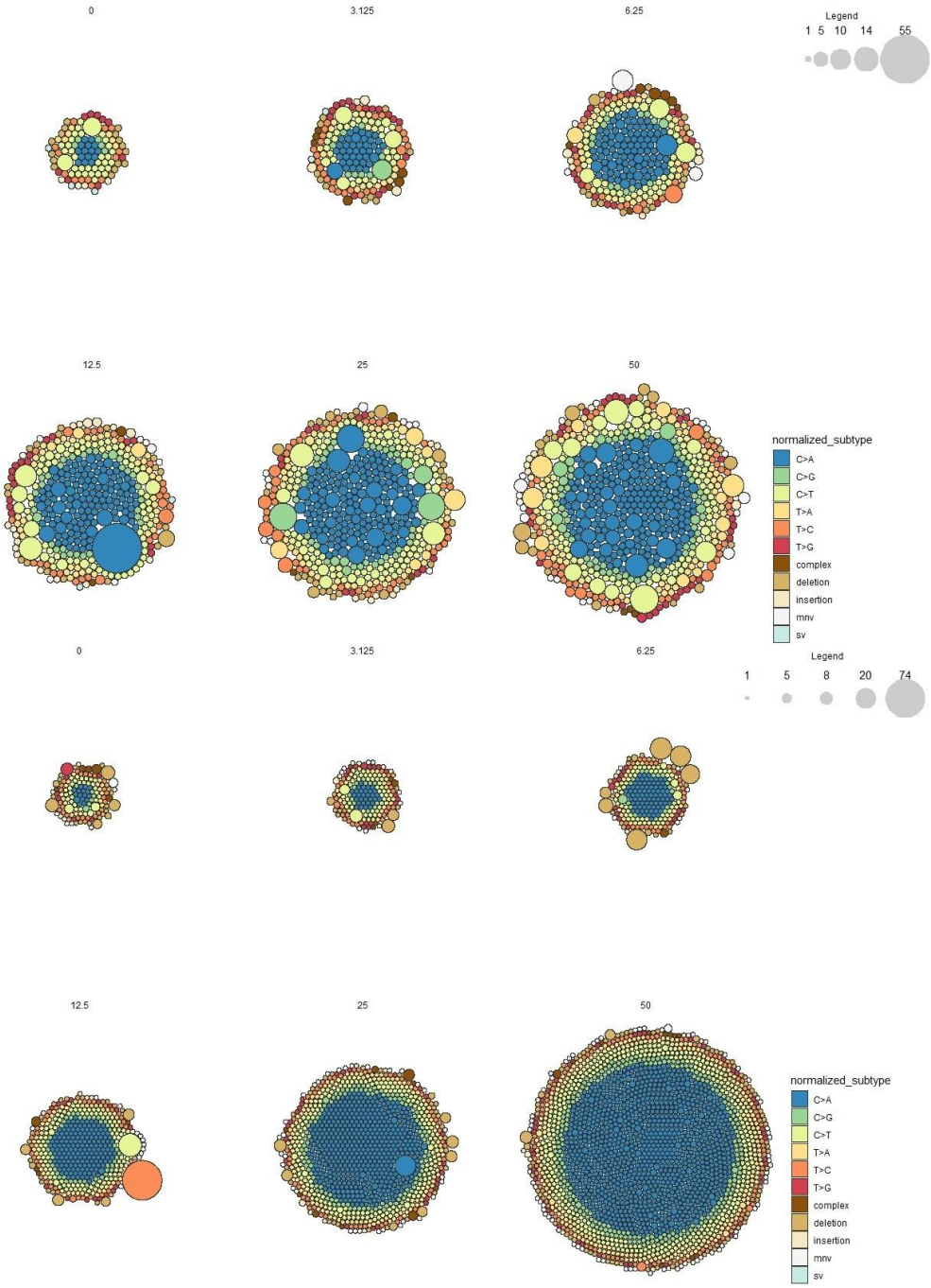
Appendix D2 Figure 1: Mutation frequency minimum estimates in DNA extracted from the control animals exposed for 28, 90 and 180 days to olive oil. The DNA was extracted from bone marrow (red) or liver (blue). No statistically significant difference between the three timepoints of each tissue were found



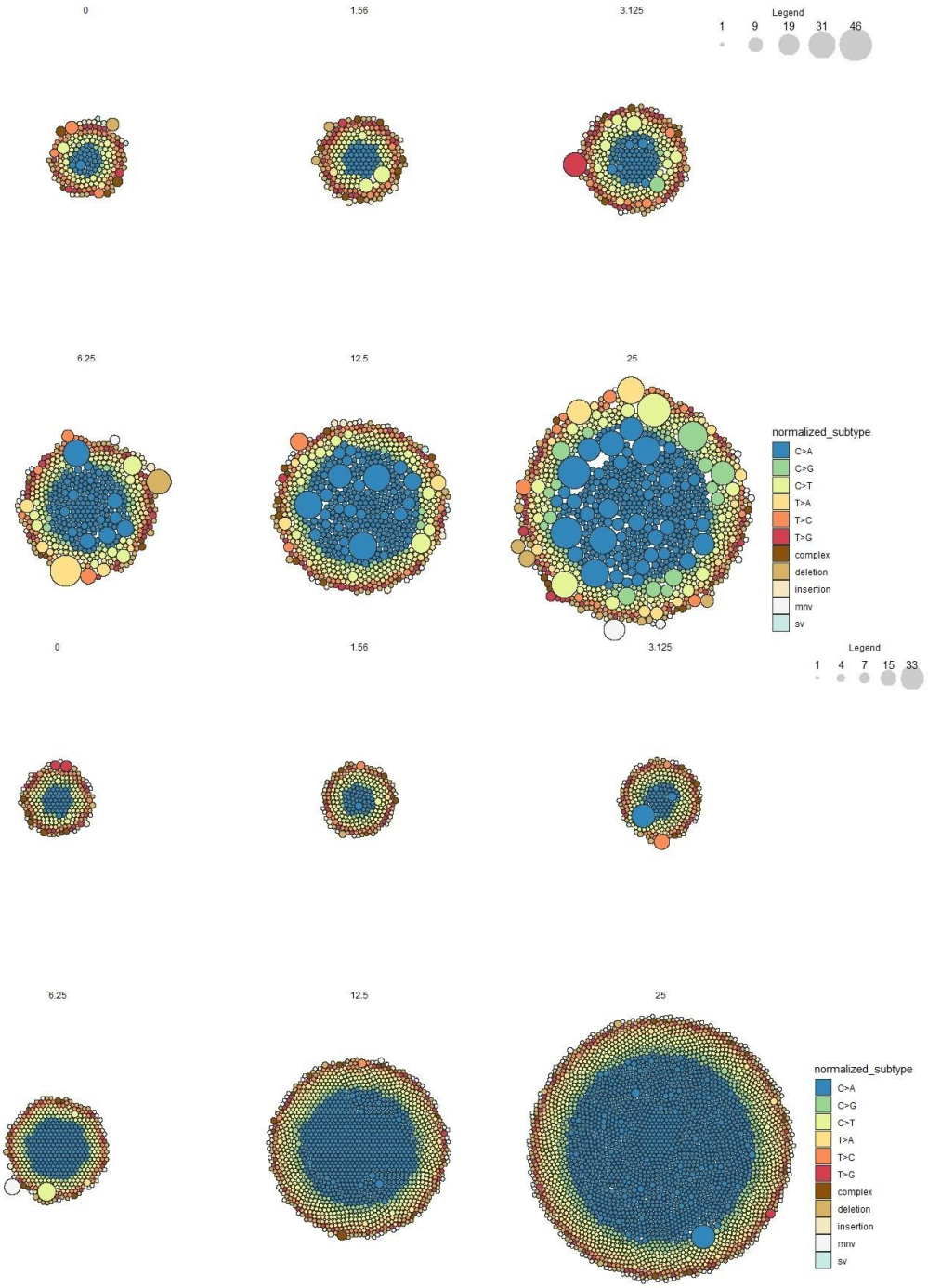
*Appendix D2 Figure 2: Estimated mutation frequency minimum (MFmin) and maximum (MFmax) in bone marrow (red) or liver (blue) after exposure to olive oil or increasing doses of benzo[b]fluoranthene. MFmin is displayed in solid bars whereas MFmax is represented by striped bars. Statistically significant changes of exposed cohorts towards the respective MFmin or MFmax of the vehicle controls are indicated through * = $p < 0.05$ and ** = $p < 0.01$*



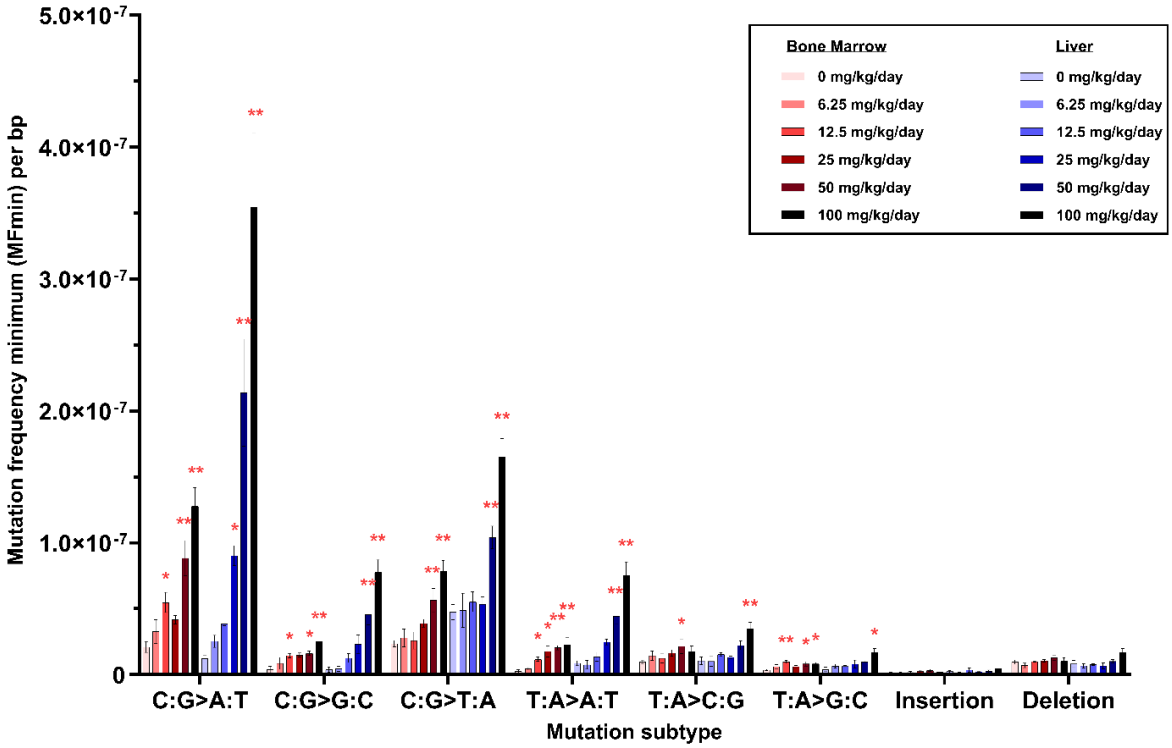
Appendix D2 Figure 3: Bubble plots indicating the alt_depth (size) of unique mutations (bubble) per dose after 28-days of exposure to BbF in bone marrow (top) and liver (bottom). Colours are indicative of the normalized subtype of mutations.



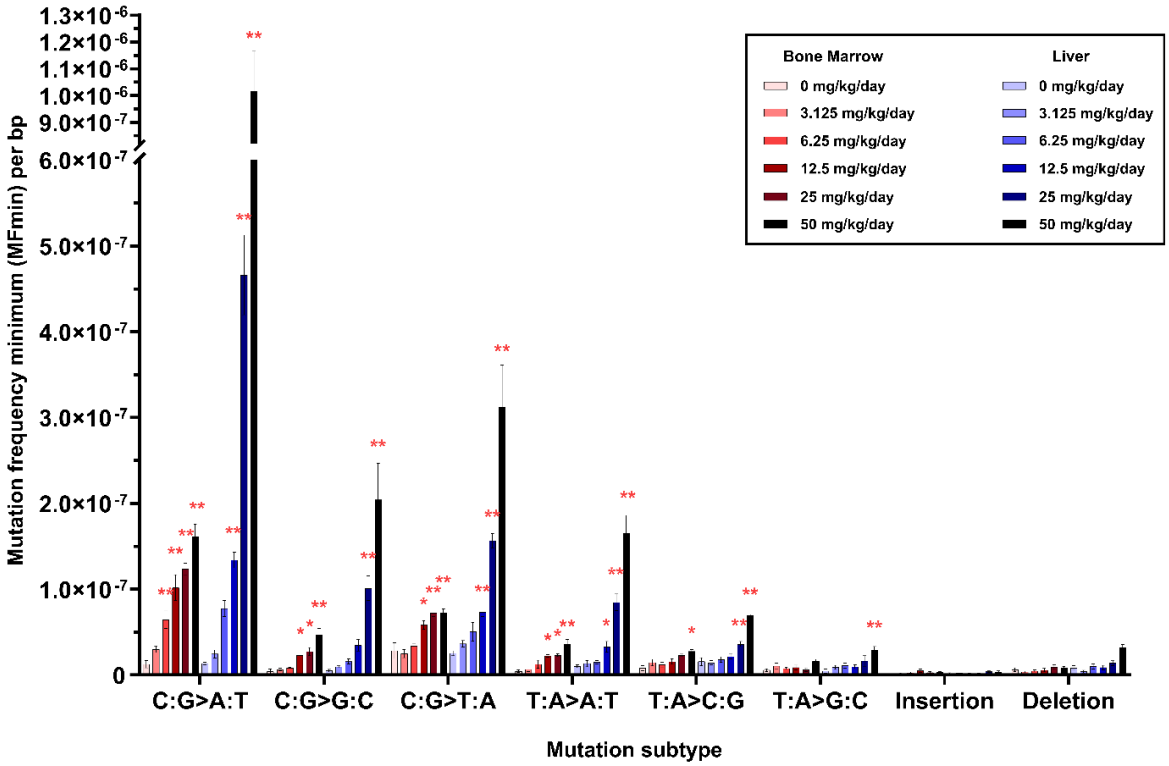
Appendix D2 Figure 4: Bubble plots indicating the alt_depth (size) of unique mutations (bubble) per dose after 90-days of exposure to BbF in bone marrow (top) and liver (bottom). Colours are indicative of the normalized subtype of mutations.



Appendix D2 Figure 5: Bubble plots indicating the alt_depth (size) of unique mutations (bubble) per dose after 180-days of exposure to BbF in bone marrow (top) and liver (bottom). Colours are indicative of the normalized subtype of mutations.



Appendix D2 Figure 6: Mean mutation frequency minimum (MFmin) ± SEM per subtype in bone marrow (red) and liver (blue) 28 days of exposure to olive oil or BbF. Statistical testing was done using generalized linear models for each dose and subtype. Significant differences between each dose and vehicle controls are indicated (p < 0;05; ** p < 0.01)*



Appendix D2 Figure 7: Mean mutation frequency minimum (MFmin) ± SEM per subtype in bone marrow (red) and liver (blue) 180 days of exposure to olive oil or BbF. Statistical testing was done using generalized linear models for each dose and subtype. Significant differences between each dose and vehicle controls are indicated (p < 0;05; ** p < 0.01)*

Appendix D3: Chapter 4

Appendix D3 Table 1: Assay metrics of the TGR assay on MutaMouse liver and BM samples after exposure to BbF.

Sample tissue	Exposure duration	BbF dose (mg/kg/day)	Total Mutants	Total Plaques	Mutant Frequency
Bone Marrow	28-Days	0	13	222159	5.85E-05
Bone Marrow	28-Days	0	10	294390	3.40E-05
Bone Marrow	28-Days	0	8	275007	2.91E-05
Bone Marrow	28-Days	0	11	244193	4.50E-05
Bone Marrow	28-Days	0	6	224478	2.67E-05
Bone Marrow	28-Days	0	5	404227	1.24E-05
Bone Marrow	28-Days	0	20	332659	6.01E-05
Bone Marrow	28-Days	0	7	281302	2.49E-05
Bone Marrow	28-Days	6.25	29	388654	7.46E-05
Bone Marrow	28-Days	6.25	18	334647	5.38E-05
Bone Marrow	28-Days	6.25	9	257777	3.49E-05
Bone Marrow	28-Days	6.25	25	266723	9.37E-05
Bone Marrow	28-Days	6.25	15	232596	6.45E-05
Bone Marrow	28-Days	6.25	37	413504	8.95E-05
Bone Marrow	28-Days	6.25	39	414167	9.42E-05
Bone Marrow	28-Days	6.25	23	247837	9.28E-05
Bone Marrow	28-Days	12.5	14	162850	8.60E-05
Bone Marrow	28-Days	12.5	44	412841	1.07E-04
Bone Marrow	28-Days	12.5	26	235744	1.10E-04
Bone Marrow	28-Days	12.5	24	212053	1.13E-04
Bone Marrow	28-Days	12.5	15	147112	1.02E-04
Bone Marrow	28-Days	12.5	25	225472	1.11E-04
Bone Marrow	28-Days	12.5	15	166495	9.01E-05
Bone Marrow	28-Days	12.5	16	198634	8.06E-05
Bone Marrow	28-Days	25	55	355852	1.55E-04
Bone Marrow	28-Days	25	41	221662	1.85E-04
Bone Marrow	28-Days	25	41	210397	1.95E-04
Bone Marrow	28-Days	25	51	232099	2.20E-04
Bone Marrow	28-Days	25	23	205427	1.12E-04
Bone Marrow	28-Days	25	30	245187	1.22E-04
Bone Marrow	28-Days	25	45	389311	1.15E-04
Bone Marrow	28-Days	25	29	200954	1.44E-04
Bone Marrow	28-Days	50	64	274344	2.33E-04
Bone Marrow	28-Days	50	22	150757	1.46E-04
Bone Marrow	28-Days	50	45	261753	1.72E-04
Bone Marrow	28-Days	50	58	221662	2.62E-04
Bone Marrow	28-Days	50	44	242370	1.82E-04
Bone Marrow	28-Days	50	74	276332	2.68E-04
Bone Marrow	28-Days	50	68	220502	3.08E-04
Bone Marrow	28-Days	50	93	343593	2.71E-04
Bone Marrow	28-Days	100	65	167158	3.89E-04
Bone Marrow	28-Days	100	91	247672	3.67E-04
Bone Marrow	28-Days	100	56	139823	4.01E-04
Bone Marrow	28-Days	100	60	211722	2.83E-04
Bone Marrow	28-Days	100	96	273019	3.52E-04
Bone Marrow	28-Days	100	93	191014	4.87E-04
Bone Marrow	28-Days	100	186	471156	3.95E-04
Bone Marrow	28-Days	100	100	289585	3.45E-04
Liver	28-Days	0	17	207415	6.27E-05
Liver	28-Days	0	8	168483	4.75E-05
Liver	28-Days	0	14	219011	6.39E-05
Liver	28-Days	0	24	255127	9.41E-05
Liver	28-Days	0	6	174613	3.44E-05
Liver	28-Days	0	9	126735	7.10E-05
Liver	28-Days	0	11	188032	5.85E-05
Liver	28-Days	0	15	234750	6.39E-05
Liver	28-Days	6.25	13	189357	6.87E-05
Liver	28-Days	6.25	9	137503	6.55E-05
Liver	28-Days	6.25	10	182565	5.48E-05
Liver	28-Days	6.25	18	209071	8.61E-05
Liver	28-Days	6.25	17	154733	1.10E-04
Liver	28-Days	6.25	22	186375	1.18E-04
Liver	28-Days	6.25	8	150591	5.31E-05
Liver	28-Days	6.25	9	141811	6.35E-05
Liver	28-Days	12.5	18	173619	1.04E-04
Liver	28-Days	12.5	18	121765	1.48E-04
Liver	28-Days	12.5	25	193830	1.29E-04
Liver	28-Days	12.5	26	152413	1.71E-04
Liver	28-Days	12.5	33	278651	1.18E-04
Liver	28-Days	12.5	16	202942	7.88E-05
Liver	28-Days	12.5	13	134356	9.68E-05
Liver	28-Days	12.5	14	125244	1.12E-04
Liver	28-Days	25	42	214041	1.96E-04
Liver	28-Days	25	38	164010	2.32E-04
Liver	28-Days	25	46	154898	2.97E-04
Liver	28-Days	25	39	135018	2.89E-04
Liver	28-Days	25	29	156721	1.85E-04
Liver	28-Days	25	39	238229	1.64E-04
Liver	28-Days	25	33	187866	1.76E-04
Liver	28-Days	25	51	174613	2.92E-04
Liver	28-Days	50	74	184387	4.01E-04
Liver	28-Days	50	42	136675	3.07E-04
Liver	28-Days	50	94	137835	6.82E-04
Liver	28-Days	50	76	175938	4.32E-04
Liver	28-Days	50	66	148272	4.45E-04
Liver	28-Days	50	124	249494	4.97E-04
Liver	28-Days	50	56	222987	2.51E-04
Liver	28-Days	50	233	447300	5.21E-04
Liver	28-Days	100	127	187535	9.56E-04
Liver	28-Days	100	152	159040	7.30E-04
Liver	28-Days	100	129	176766	5.06E-04
Liver	28-Days	100	75	148272	8.89E-04
Liver	28-Days	100	130	146284	7.43E-04
Liver	28-Days	100	395	147940	1.65E-03
Liver	28-Days	100	207	125078	6.73E-04
Liver	28-Days	100	103	153076	6.73E-04

Appendix D3 Table 2: Assay metrics of the Pig-a assay on MutaMouse blood samples after exposure to BbF.

Animal No.	Exposure duration	BbF Dose (mg/kg/day)	No. Counting Beads	RBC to Counting Bead Ratio	RET to Counting Bead Ratio	% RET	No. Counting Beads	Total RBC Equivalents	Total RET Equivalents	Mutant RBCs per 10^6 Total RBCs	Mutant RETs per 10^6 Total RETs
1	28 days	0	1593	121.495	3.660	3.01	4858	194,771,547	5,867,434	0.08	0.17
2	28 days	0	1338	251.225	11.671	4.65	3751	310,970,732	14,446,569	0.05	0.07
3	28 days	0	1320	220.520	7.369	3.34	7449	542,070,228	18,114,074	0.01	0.00
4	28 days	0	1267	207.403	8.760	4.22	4254	291,153,568	12,297,340	0.07	0.16
5	28 days	0	1278	183.026	5.480	2.99	3803	229,693,503	6,877,276	0.10	0.15
6	28 days	0	1270	202.718	8.666	4.27	3406	227,848,699	9,740,313	1.94	1.13
7	28 days	0	1276	181.603	6.839	3.77	3673	220,116,979	8,289,401	0.34	1.21
8	28 days	0	1224	228.391	11.687	5.12	4246	320,013,701	16,375,427	0.59	0.12
9	28 days	3.125	1442	178.800	5.626	3.15	4104	242,149,995	7,619,328	0.05	0.00
10	28 days	3.125	1163	281.932	12.171	4.32	3877	360,703,013	15,571,543	0.36	1.03
11	28 days	3.125	1296	176.589	6.678	3.78	4449	259,260,080	9,804,341	0.49	0.20
12	28 days	3.125	1186	248.321	11.561	4.66	4074	333,844,380	15,542,684	1.16	0.26
13	28 days	3.125	1292	164.861	6.197	3.76	3832	208,474,541	7,836,400	0.02	0.13
14	28 days	3.125	1234	196.238	9.181	4.68	3650	236,366,307	11,058,404	0.10	1.27
15	28 days	3.125	1218	154.974	7.227	4.66	4238	216,735,171	10,107,148	0.12	0.20
16	28 days	3.125	1216	207.921	9.476	4.56	4133	283,578,537	12,924,092	0.91	0.23
17	28 days	6.25	1379	171.479	7.127	4.16	4228	239,251,967	9,943,776	0.25	3.72
18	28 days	6.25	1288	239.382	11.686	4.88	3916	309,345,478	15,101,433	0.27	0.26
19	28 days	6.25	1384	172.172	5.995	3.48	4508	256,127,393	8,918,313	0.12	0.67
20	28 days	6.25	1174	249.932	11.868	4.75	4221	348,134,299	16,531,128	0.60	1.09
21	28 days	6.25	1277	175.036	7.168	4.09	3602	208,056,211	8,520,230	0.25	0.12
22	28 days	6.25	1120	259.818	11.296	4.35	4000	342,956,330	14,910,571	0.15	0.13
23	28 days	6.25	1426	102.699	4.750	4.63	4197	142,237,720	6,578,732	0.37	0.46
24	28 days	6.25	1379	128.560	6.365	4.95	4141	175,679,340	8,697,876	0.37	0.34
25	28 days	12.5	1336	169.354	7.452	4.40	4594	256,741,484	11,297,268	1.21	0.89
26	28 days	12.5	1175	280.835	14.760	5.26	3566	330,477,707	17,369,099	3.13	4.43
27	28 days	12.5	1273	154.392	6.339	4.11	4317	219,946,188	9,030,512	0.07	0.33
28	28 days	12.5	1255	235.224	9.462	4.02	3785	293,803,599	11,818,393	0.72	0.17
29	28 days	12.5	1385	120.344	6.619	5.50	3673	145,866,300	8,022,743	2.05	15.08
30	28 days	12.5	1176	219.140	12.178	5.56	3742	270,604,514	15,037,975	0.04	0.00
31	28 days	12.5	1306	138.453	5.871	4.24	4083	186,548,322	7,910,448	3.32	27.05
32	28 days	12.5	1443	189.965	10.924	5.75	3967	248,682,594	14,300,575	0.59	0.35
33	28 days	25	1329	173.910	7.543	4.34	4321	247,981,006	10,755,682	0.31	0.46
34	28 days	25	1184	239.831	14.209	5.92	3920	310,242,279	18,380,579	2.39	5.60
35	28 days	25	1379	139.297	7.076	5.08	4457	204,877,372	10,407,347	1.28	9.90
36	28 days	25	1183	223.619	10.990	4.91	3665	270,453,295	13,291,723	1.38	13.09
37	28 days	25	1326	145.719	6.892	4.73	4163	200,185,303	9,468,066	0.07	0.21
38	28 days	25	1331	144.119	9.361	6.50	4083	194,182,558	12,612,792	0.24	0.24
39	28 days	25	1363	133.681	6.707	5.02	4170	183,956,585	9,229,410	0.33	1.73
40	28 days	25	1221	217.269	11.888	5.47	4039	289,588,436	15,845,000	0.06	0.25
41	28 days	50	1316	174.025	8.553	4.91	4514	259,228,528	12,740,592	0.35	0.31
42	28 days	50	1192	235.363	12.588	5.35	4344	337,394,194	18,044,969	3.56	0.67
43	28 days	50	1432	147.866	6.131	4.15	4962	242,122,239	10,039,167	0.40	3.98
44	28 days	50	1414	162.484	9.906	6.10	4351	233,297,069	14,223,190	3.87	6.82
45	28 days	50	1303	145.704	7.480	5.13	3803	182,855,234	9,387,231	0.42	1.17
46	28 days	50	1318	175.334	11.614	6.62	3920	226,809,794	15,023,720	2.89	12.58
47	28 days	50	1252	183.782	9.196	5.00	4162	252,414,702	12,630,212	0.15	0.32
48	28 days	50	1260	177.903	12.856	7.23	4263	250,269,659	18,085,511	0.07	0.28
49	28 days	100	1293	186.343	9.423	5.06	4319	265,586,432	13,430,185	2.55	40.21
50	28 days	100	1333	173.723	9.478	5.46	5644	323,559,326	17,652,788	2.47	17.62
51	28 days	100	1411	123.702	7.739	6.26	4511	184,144,667	11,520,392	4.71	0.87
52	28 days	100	1270	199.337	8.610	4.32	3795	249,637,196	10,782,626	0.46	0.28
53	28 days	100	1212	173.141	7.753	4.48	4057	231,800,584	10,379,690	1.43	4.24
54	28 days	100	1319	116.453	7.992	6.86	4282	164,553,431	11,293,063	0.34	0.44
55	28 days	100	1269	138.878	9.151	6.59	4242	194,407,813	12,809,991	1.15	0.39
56	28 days	100	1242	168.068	11.422	6.80	4310	239,040,726	16,245,348	1.52	2.22

Appendix D3 Table 3: Mutation frequency metrics of DS on the MMP in liver and BM.

Animal ID	Endpoint	Sample tissue	Exposure duration	BbF dose (mg/kg/day)	Mutation count unique	Mutation count all	Sample depth	Mutation frequency minimum	Mutation Frequency maximum
1	MMP	Bone Marrow	28-Days	0	37	43	820319269	4.51E-08	5.24E-08
2		Bone Marrow	28-Days	0	35	37	903268416	3.87E-08	4.10E-08
3		Bone Marrow	28-Days	0	34	35	841007628	4.04E-08	4.16E-08
4		Bone Marrow	28-days	0	39	43	953865140	4.09E-08	4.51E-08
5	MMP	Bone Marrow	28-Days	6.25	32	38	709601657	4.51E-08	5.36E-08
6		Bone Marrow	28-Days	6.25	53	190	570223160	9.29E-08	3.33E-07
7		Bone Marrow	28-Days	6.25	27	28	650976772	4.15E-08	4.30E-08
8		Bone Marrow	28-days	6.25	35	43	755934082	4.63E-08	5.69E-08
9	MMP	Bone Marrow	28-Days	12.5	67	75	996552403	6.72E-08	7.53E-08
10		Bone Marrow	28-Days	12.5	71	126	748281155	9.49E-08	1.68E-07
11		Bone Marrow	28-Days	12.5	72	80	975415700	7.38E-08	8.20E-08
12		Bone Marrow	28-days	12.5	47	51	714621726	6.58E-08	7.14E-08
13	MMP	Bone Marrow	28-Days	25	66	73	1034555738	6.38E-08	7.06E-08
14		Bone Marrow	28-Days	25	60	73	731411483	8.20E-08	9.98E-08
15		Bone Marrow	28-Days	25	49	52	587263813	8.34E-08	8.85E-08
16		Bone Marrow	28-days	25	90	99	1013524373	8.88E-08	9.77E-08
17	MMP	Bone Marrow	28-Days	50	72	84	747503560	9.63E-08	1.12E-07
18		Bone Marrow	28-Days	50	122	135	1015314151	1.20E-07	1.33E-07
19		Bone Marrow	28-Days	50	105	121	727894703	1.44E-07	1.66E-07
20		Bone Marrow	28-days	50	93	103	876901612	1.06E-07	1.17E-07
21	MMP	Bone Marrow	28-Days	100	143	176	859890285	1.66E-07	2.05E-07
22		Bone Marrow	28-Days	100	115	138	723027618	1.59E-07	1.91E-07
23		Bone Marrow	28-Days	100	149	176	1141474460	1.31E-07	1.54E-07
24		Bone Marrow	28-days	100	118	136	876503251	1.35E-07	1.55E-07
1	MMP	Liver	28-Days	0	44	47	749609387	6.32E-08	6.75E-08
2		Liver	28-Days	0	48	87	722340262	6.61E-08	1.20E-07
3		Liver	28-Days	0	41	47	627025630	5.01E-08	5.75E-08
4		Liver	28-days	0	28	28	677732398	3.39E-08	3.39E-08
5	MMP	Liver	28-Days	6.25	70	82	743455407	6.72E-08	7.87E-08
6		Liver	28-Days	6.25	40	58	704928576	5.70E-08	8.26E-08
7		Liver	28-Days	6.25	32	33	849818048	4.60E-08	4.75E-08
8		Liver	28-days	6.25	50	50	632701328	6.51E-08	6.51E-08
9	MMP	Liver	28-Days	12.5	92	95	693759043	8.38E-08	8.65E-08
10		Liver	28-Days	12.5	85	106	666531230	8.74E-08	1.09E-07
11		Liver	28-Days	12.5	68	72	517474384	7.11E-08	7.53E-08
12		Liver	28-days	12.5	71	79	738972435	8.02E-08	8.92E-08
13	MMP	Liver	28-Days	25	89	94	676238056	9.97E-08	1.05E-07
14		Liver	28-Days	25	102	119	626642186	1.11E-07	1.29E-07
15		Liver	28-Days	25	85	89	1029731981	1.14E-07	1.19E-07
16		Liver	28-days	25	76	76	731701362	1.20E-07	1.20E-07
17	MMP	Liver	28-Days	50	195	197	671817466	2.21E-07	2.23E-07
18		Liver	28-Days	50	275	304	639888577	2.90E-07	3.20E-07
19		Liver	28-Days	50	135	136	665939020	1.54E-07	1.55E-07
20		Liver	28-days	50	177	183	633802424	2.24E-07	2.31E-07
21	MMP	Liver	28-Days	100	250	269	690876985	3.15E-07	3.39E-07
22		Liver	28-Days	100	498	595	595451784	4.85E-07	5.80E-07
23		Liver	28-Days	100	263	272	678943393	3.21E-07	3.32E-07
24		Liver	28-days	100	297	305	605952416	3.30E-07	3.39E-07

Appendix D3 Table 4: Mutation frequency metrics of DS on lacZ in liver and BM.

Animal ID	Endpoint	Sample tissue	Exposure duration	BbF dose (mg/kg/day)	Mutation count unique	Mutation count all	Sample depth	Mutation frequency minimum	Mutation Frequency maximum
1	lacZ	Bone Marrow	28-Days	0	21	21	161810257	1.30E-07	1.30E-07
2		Bone Marrow	28-Days	0	61	74	290324467	2.10E-07	2.55E-07
3		Bone Marrow	28-Days	0	15	17	61849186	2.43E-07	2.75E-07
4		Bone Marrow	28-days	0	34	46	207008767	1.64E-07	2.22E-07
5	lacZ	Bone Marrow	28-Days	6.25	25	48	117390158	2.13E-07	4.09E-07
6		Bone Marrow	28-Days	6.25	41	47	224510942	1.83E-07	2.09E-07
7		Bone Marrow	28-Days	6.25	78	2401	403897853	1.93E-07	5.94E-06
8		Bone Marrow	28-days	6.25	8	36	44180898	1.81E-07	8.15E-07
9	lacZ	Bone Marrow	28-Days	12.5	35	42	129790694	2.70E-07	3.24E-07
10		Bone Marrow	28-Days	12.5	57	57	197647948	2.88E-07	2.88E-07
11		Bone Marrow	28-Days	12.5	30	32	129538110	2.32E-07	2.47E-07
12		Bone Marrow	28-days	12.5	27	35	82983382	3.25E-07	4.22E-07
13	lacZ	Bone Marrow	28-Days	25	15	33	45931880	3.27E-07	7.18E-07
14		Bone Marrow	28-Days	25	43	47	189999925	2.26E-07	2.47E-07
15		Bone Marrow	28-Days	25	93	109	368394389	2.52E-07	2.96E-07
16		Bone Marrow	28-days	25	50	441	149448380	3.35E-07	2.95E-06
17	lacZ	Bone Marrow	28-Days	50	42	43	120234448	3.49E-07	3.58E-07
18		Bone Marrow	28-Days	50	15	15	34018897	4.41E-07	4.41E-07
19		Bone Marrow	28-Days	50	64	382	134814238	4.75E-07	2.83E-06
20		Bone Marrow	28-days	50	104	114	165867250	6.27E-07	6.87E-07
21	lacZ	Bone Marrow	28-Days	100	96	106	148914836	6.45E-07	7.12E-07
22		Bone Marrow	28-Days	100	52	53	82091426	6.33E-07	6.46E-07
23		Bone Marrow	28-Days	100	120	132	214469159	5.60E-07	6.15E-07
24		Bone Marrow	28-days	100	278	311	193392372	1.44E-06	1.61E-06
1	lacZ	Liver	28-Days	0	28	29	133089805	2.10E-07	2.18E-07
2		Liver	28-Days	0	15	16	92685630	1.62E-07	1.73E-07
3		Liver	28-Days	0	88	111	498084122	1.77E-07	2.23E-07
4		Liver	28-days	0	20	22	149226891	1.34E-07	1.47E-07
5	lacZ	Liver	28-Days	6.25	52	58	217527560	2.39E-07	2.67E-07
6		Liver	28-Days	6.25	31	33	241967798	1.28E-07	1.36E-07
7		Liver	28-Days	6.25	34	1052	179140752	1.90E-07	5.87E-06
8		Liver	28-days	6.25	59	122	276937273	2.13E-07	4.41E-07
9	lacZ	Liver	28-Days	12.5	50	52	257218068	1.94E-07	2.02E-07
10		Liver	28-Days	12.5	41	41	144832441	2.83E-07	2.83E-07
11		Liver	28-Days	12.5	49	56	220072918	2.23E-07	2.54E-07
12		Liver	28-days	12.5	44	54	220755622	1.99E-07	2.45E-07
13	lacZ	Liver	28-Days	25	69	101	239689136	2.88E-07	4.21E-07
14		Liver	28-Days	25	40	43	151004525	2.65E-07	2.85E-07
15		Liver	28-Days	25	37	39	154366526	2.40E-07	2.53E-07
16		Liver	28-days	25	70	72	247935103	2.82E-07	2.90E-07
17	lacZ	Liver	28-Days	50	109	113	230776235	4.72E-07	4.90E-07
18		Liver	28-Days	50	48	56	99194787	4.84E-07	5.65E-07
19		Liver	28-Days	50	82	618	240904562	3.40E-07	2.57E-06
20		Liver	28-days	50	140	836	304059159	4.60E-07	2.75E-06
21	lacZ	Liver	28-Days	100	109	111	174006148	6.26E-07	6.38E-07
22		Liver	28-Days	100	324	345	310671524	1.04E-06	1.11E-06
23		Liver	28-Days	100	62	63	91393390	6.78E-07	6.89E-07
24		Liver	28-days	100	122	127	209870365	5.81E-07	6.05E-07

Appendix D3 Table 5: Mutation frequency metrics of DS on Pig-a in BM

Animal ID	Endpoint	Sample tissue	Exposure duration	BbF dose (mg/kg/day)	Mutation count unique	Mutation count all	Sample depth	Mutation frequency minimum	Mutation Frequency maximum
1	*Pig-a*	Bone Marrow	28-Days	0	12	200	131639652	9.12E-08	1.52E-06
2		Bone Marrow	28-Days	0	15	264	154609238	9.70E-08	1.71E-06
3		Bone Marrow	28-Days	0	12	86	135456006	8.86E-08	6.35E-07
4		Bone Marrow	28-days	0	8	205	155111250	5.16E-08	1.32E-06
5	*Pig-a*	Bone Marrow	28-Days	6.25	15	177	112760484	1.33E-07	1.57E-06
6		Bone Marrow	28-Days	6.25	11	113	85344758	1.29E-07	1.32E-06
7		Bone Marrow	28-Days	6.25	11	103	101616480	1.08E-07	1.01E-06
8		Bone Marrow	28-days	6.25	13	203	138257449	9.40E-08	1.47E-06
9	*Pig-a*	Bone Marrow	28-Days	12.5	13	137	175174353	7.42E-08	7.82E-07
10		Bone Marrow	28-Days	12.5	9	254	112201341	8.02E-08	2.26E-06
11		Bone Marrow	28-Days	12.5	14	186	161113118	8.69E-08	1.15E-06
12		Bone Marrow	28-days	12.5	12	123	129479446	9.27E-08	9.50E-07
13	*Pig-a*	Bone Marrow	28-Days	25	20	285	176808717	1.13E-07	1.61E-06
14		Bone Marrow	28-Days	25	8	95	135829900	5.89E-08	6.99E-07
15		Bone Marrow	28-Days	25	10	212	90160387	1.11E-07	2.35E-06
16		Bone Marrow	28-days	25	21	269	160525277	1.31E-07	1.68E-06
17	*Pig-a*	Bone Marrow	28-Days	50	18	121	127197850	1.42E-07	9.51E-07
18		Bone Marrow	28-Days	50	21	180	168699535	1.24E-07	1.07E-06
19		Bone Marrow	28-Days	50	10	161	112030452	8.93E-08	1.44E-06
20		Bone Marrow	28-days	50	22	232	146159434	1.51E-07	1.59E-06
21	*Pig-a*	Bone Marrow	28-Days	100	15	161	151879545	9.88E-08	1.06E-06
22		Bone Marrow	28-Days	100	16	223	125741338	1.27E-07	1.77E-06
23		Bone Marrow	28-Days	100	10	136	184742202	5.41E-08	7.36E-07
24		Bone Marrow	28-days	100	19	211	155379239	1.22E-07	1.36E-06

Appendix D3 Table 6: Subtype specific mutagenicity data on MMP in liver and bone marrow of MutaMouse males post exposure to BbF for 28-days

Animal ID	Endpoint	Sample Tissue	Exposure Duration	BbF dose (mg/kg/day)	C:G>A:T Count	C:G>A:T Count All	C:G>A:T MFmin	C:G>A:T MFmax	C:G>G:C Count	C:G>G:C Count All	C:G>G:C MFmin	C:G>G:C MFmax	C:G>T:A Count	C:G>T: A	C:G>T:A MFmin	C:G>T:A MFmax	T:A>A:T Count	T:A>A: T	T:A>A:T MFmin	T:A>A:T MFmax	T:A>C:G Count	T:A>C: G	T:A>C:G MFmin	T:A>C:G MFmax	T:A>G:C Count	T:A>G: C	T:A>G:C MFmin	T:A>G:C MFmax	Insertion Count Unique	Insertion Count All	Insertion MFmin	Insertion MFmax	Deletion Count	Deletio n	Deletion MFmin	Deletion MFmax
1	MMP	Bone Marrow	28-Days	0	9	10	2.59E-08	2.82E-08	3	4	8.45E-09	1.13E-08	7	7	1.97E-08	1.97E-08	1	1	2.15E-09	2.15E-09	4	4	8.60E-09	8.60E-09	2	5	4.30E-09	1.07E-08	0	0	0	0	11	12	1.34E-08	1.46E-08
2	MMP	Bone Marrow	28-Days	0	7	7	1.78E-08	1.78E-08	0	0	0	0	12	13	3.05E-08	3.31E-08	1	1	1.96E-09	1.96E-09	3	3	5.88E-09	5.88E-09	2	2	3.92E-09	3.92E-09	1	2	1.11E-09	2.21E-09	9	9	9.96E-09	9.96E-09
3	MMP	Bone Marrow	28-Days	0	4	4	1.10E-08	1.10E-08	0	0	0	0	8	8	2.20E-08	2.20E-08	3	3	6.29E-09	6.29E-09	6	6	1.26E-08	1.26E-08	1	1	2.10E-09	2.10E-09	3	4	3.57E-09	4.76E-09	8	8	9.51E-09	9.51E-09
4	MMP	Bone Marrow	28-Days	0	12	12	2.90E-08	2.90E-08	3	3	7.25E-09	7.25E-09	9	9	2.17E-08	2.17E-08	0	0	0	0	6	6	1.11E-08	1.11E-08	2	2	3.70E-09	3.70E-09	0	0	0	0	5	8	5.24E-09	8.39E-09
5	MMP	Bone Marrow	28-Days	6.25	9	9	2.94E-08	2.94E-08	1	3	3.27E-09	9.81E-09	11	11	3.60E-08	3.60E-08	2	2	4.95E-09	4.95E-09	4	4	9.91E-09	9.91E-09	1	2	2.48E-09	4.95E-09	0	0	0	0	3	3	4.23E-09	4.23E-09
6	MMP	Bone Marrow	28-Days	6.25	14	14	5.71E-08	5.71E-08	5	30	2.04E-08	1.22E-07	10	50	4.08E-08	2.04E-07	1	1	3.08E-09	3.08E-09	8	14	2.46E-08	4.31E-08	3	3	9.23E-09	9.23E-09	0	0	0	0	6	20	1.05E-08	3.51E-08
7	MMP	Bone Marrow	28-Days	6.25	4	5	1.43E-08	1.79E-08	3	3	1.07E-08	1.07E-08	3	3	1.07E-08	1.07E-08	1	1	2.70E-09	2.70E-09	5	5	1.35E-08	1.35E-08	3	3	8.09E-09	8.09E-09	1	1	1.54E-09	1.54E-09	6	6	9.22E-09	9.22E-09
8	MMP	Bone Marrow	28-Days	6.25	10	10	3.02E-08	3.02E-08	0	0	0	0	8	9	2.41E-08	2.71E-08	3	8	7.07E-09	1.89E-08	4	4	9.43E-09	9.43E-09	2	2	4.71E-09	4.71E-09	2	2	2.65E-09	2.65E-09	3	3	3.97E-09	3.97E-09
9	MMP	Bone Marrow	28-Days	12.5	15	17	3.44E-08	3.90E-08	7	7	1.61E-08	1.61E-08	14	16	3.21E-08	3.67E-08	6	6	1.07E-08	1.07E-08	8	8	1.43E-08	1.43E-08	4	6	7.13E-09	1.07E-08	1	1	1.00E-09	1.00E-09	10	11	1.00E-08	1.10E-08
10	MMP	Bone Marrow	28-Days	12.5	21	32	6.54E-08	9.97E-08	5	15	1.56E-08	4.67E-08	13	25	4.05E-08	7.79E-08	6	6	1.40E-08	1.40E-08	7	9	1.64E-08	2.11E-08	4	4	9.36E-09	9.36E-09	2	2	2.67E-09	2.67E-09	7	7	9.35E-09	9.35E-09
11	MMP	Bone Marrow	28-Days	12.5	22	25	5.19E-08	5.90E-08	7	7	1.65E-08	1.65E-08	5	7	1.18E-08	1.65E-08	8	9	1.45E-08	1.63E-08	9	10	1.63E-08	1.81E-08	7	7	1.27E-08	1.27E-08	3	4	3.08E-09	4.10E-09	9	9	9.23E-09	9.23E-09
12	MMP	Bone Marrow	28-Days	12.5	21	23	6.73E-08	7.37E-08	3	3	9.62E-09	9.62E-09	6	7	1.92E-08	2.24E-08	3	4	7.45E-09	9.93E-09	1	1	2.48E-09	2.48E-09	4	4	9.93E-09	9.93E-09	0	0	0	0	8	8	1.12E-08	1.12E-08
13	MMP	Bone Marrow	28-Days	25	17	17	3.77E-08	3.77E-08	6	6	1.33E-08	1.33E-08	17	22	3.77E-08	4.88E-08	4	5	6.86E-09	8.57E-09	6	6	1.03E-08	1.03E-08	3	3	5.14E-09	5.14E-09	1	1	9.67E-10	9.67E-10	9	10	8.70E-09	9.67E-09
14	MMP	Bone Marrow	28-Days	25	14	19	4.36E-08	5.92E-08	6	6	1.87E-08	1.87E-08	11	16	3.43E-08	4.98E-08	10	12	2.44E-08	2.92E-08	9	9	2.19E-08	2.19E-08	2	2	4.87E-09	4.87E-09	1	2	1.37E-09	2.73E-09	7	7	9.57E-09	9.57E-09
15	MMP	Bone Marrow	28-Days	25	9	10	3.57E-08	3.96E-08	4	4	1.59E-08	1.59E-08	12	12	4.76E-08	4.76E-08	5	5	1.49E-08	1.49E-08	6	7	1.79E-08	2.09E-08	2	2	5.97E-09	5.97E-09	3	4	5.11E-09	6.81E-09	7	7	1.19E-08	1.19E-08
16	MMP	Bone Marrow	28-Days	25	22	23	4.99E-08	5.22E-08	5	9	1.14E-08	2.04E-08	16	17	3.63E-08	3.86E-08	14	15	2.44E-08	2.62E-08	9	10	1.57E-08	1.75E-08	5	5	8.73E-09	8.73E-09	2	2	1.97E-09	1.97E-09	12	12	1.18E-08	1.18E-08
17	MMP	Bone Marrow	28-Days	50	32	42	9.84E-08	1.29E-07	4	5	1.23E-08	1.54E-08	10	11	3.07E-08	3.38E-08	7	7	1.66E-08	1.66E-08	3	3	7.11E-09	7.11E-09	4	4	9.47E-09	9.47E-09	2	2	2.68E-09	2.68E-09	8	8	1.07E-08	1.07E-08
18	MMP	Bone Marrow	28-Days	50	28	28	6.34E-08	6.34E-08	9	10	2.04E-08	2.26E-08	28	35	6.34E-08	7.92E-08	11	13	1.92E-08	2.27E-08	20	21	3.49E-08	3.66E-08	7	8	1.22E-08	1.39E-08	5	5	4.92E-09	4.92E-09	10	11	9.85E-09	1.08E-08
19	MMP	Bone Marrow	28-Days	50	38	46	1.21E-07	1.47E-07	5	5	1.60E-08	1.60E-08	23	26	7.35E-08	8.31E-08	9	10	2.17E-08	2.41E-08	9	9	2.17E-08	2.17E-08	2	2	4.82E-09	4.82E-09	3	4	4.12E-09	5.50E-09	14	17	1.92E-08	2.34E-08
20	MMP	Bone Marrow	28-Days	50	27	28	7.08E-08	7.35E-08	6	6	1.57E-08	1.57E-08	22	27	5.77E-08	7.08E-08	12	16	2.42E-08	3.23E-08	11	11	2.22E-08	2.22E-08	3	3	6.05E-09	6.05E-09	1	1	1.14E-09	1.14E-09	10	10	1.14E-08	1.14E-08
21	MMP	Bone Marrow	28-Days	100	43	57	1.14E-07	1.52E-07	11	14	2.93E-08	3.73E-08	32	35	8.52E-08	9.32E-08	18	20	3.72E-08	4.13E-08	9	9	1.86E-08	1.86E-08	4	9	8.26E-09	1.86E-08	3	4	3.49E-09	4.65E-09	14	15	1.63E-08	1.74E-08
22	MMP	Bone Marrow	28-Days	100	54	58	1.70E-07	1.83E-07	7	7	2.21E-08	2.21E-08	31	43	9.79E-08	1.36E-07	9	10	2.22E-08	2.46E-08	3	3	7.38E-09	7.38E-09	2	2	4.92E-09	4.92E-09	0	0	0	0	5	6	6.92E-09	8.30E-09
23	MMP	Bone Marrow	28-Days	100	57	71	1.15E-07	1.43E-07	15	17	3.02E-08	3.43E-08	31	33	6.25E-08	6.65E-08	14	16	2.17E-08	2.48E-08	10	14	1.55E-08	2.17E-08	6	7	9.29E-09	1.08E-08	3	3	2.63E-09	2.63E-09	6	7	5.26E-09	6.13E-09
24	MMP	Bone Marrow	28-Days	100	43	49	1.12E-07	1.28E-07	7	7	1.83E-08	1.83E-08	26	33	6.79E-08	8.61E-08	5	5	1.01E-08	1.01E-08	14	18	2.84E-08	3.65E-08	5	5	1.01E-08	1.01E-08	2	2	2.28E-09	2.28E-09	11	11	1.26E-08	1.26E-08
1	MMP	Liver	28-Days	0	2	2	6.64E-09	6.64E-09	0	0	0	0	19	19	6.31E-08	6.31E-08	5	5	1.26E-08	1.26E-08	4	6	1.01E-08	1.52E-08	0	0	0	0	3	3	4.31E-09	4.31E-09	10	11	1.44E-08	1.58E-08
2	MMP	Liver	28-Days	0	6	6	1.93E-08	1.93E-08	3	7	9.63E-09	2.25E-08	15	35	4.81E-08	1.12E-07	2	2	4.83E-09	4.83E-09	6	6	1.45E-08	1.45E-08	3	3	7.24E-09	7.24E-09	2	2	2.76E-09	2.76E-09	7	7	9.64E-09	9.64E-09
3	MMP	Liver	28-Days	0	3	3	8.45E-09	8.45E-09	1	1	2.82E-09	2.82E-09	15	21	4.23E-08	5.92E-08	5	5	1.08E-08	1.08E-08	7	7	1.51E-08	1.51E-08	3	3	6.48E-09	6.48E-09	1	1	1.22E-09	1.22E-09	6	6	7.34E-09	7.34E-09
4	MMP	Liver	28-Days	0	5	5	1.40E-08	1.40E-08	1	1	2.79E-09	2.79E-09	13	13	3.63E-08	3.63E-08	3	3	6.42E-09	6.42E-09	1	1	2.14E-09	2.14E-09	1	1	2.14E-09	2.14E-09	1	1	1.21E-09	1.21E-09	1	1	1.21E-09	1.21E-09
5	MMP	Liver	28-Days	6.25	15	16	3.35E-08	3.57E-08	2	2	4.47E-09	4.47E-09	26	30	5.81E-08	6.70E-08	0	0	0	0	11	12	1.85E-08	2.02E-08	4	4	6.73E-09	6.73E-09	1	1	9.60E-10	9.60E-10	9	12	8.64E-09	1.15E-08
6	MMP	Liver	28-Days	6.25	10	19	3.27E-08	6.21E-08	3	3	9.81E-09	9.81E-09	9	15	2.94E-08	4.90E-08	4	4	1.01E-08	1.01E-08	4	4	1.01E-08	1.01E-08	1	1	2.53E-09	2.53E-09	1	2	1.42E-09	2.85E-09	2	2	2.85E-09	2.85E-09
7	MMP	Liver	28-Days	6.25	6	6	2.00E-08	2.00E-08	1	1	3.34E-09	3.34E-09	8	8	2.67E-08	2.67E-08	6	6	1.52E-08	1.52E-08	0	0	0	0	3	4	7.58E-09	1.01E-08	0	0	0	0	7	7	1.01E-08	1.01E-08
8	MMP	Liver	28-Days	6.25	5	5	1.50E-08	1.50E-08	0	0	0	0	27	27	8.10E-08	8.10E-08	2	2	4.61E-09	4.61E-09	5	5	1.15E-08	1.15E-08	4	4	9.22E-09	9.22E-09	2	2	2.61E-09	2.61E-09	4	4	5.21E-09	5.21E-09
9	MMP	Liver	28-Days	12.5	19	19	4.02E-08	4.02E-08	6	7	1.27E-08	1.48E-08	35	36	7.40E-08	7.61E-08	5	5	8.00E-09	8.00E-09	8	8	1.28E-08	1.28E-08	4	4	6.40E-09	6.40E-09	2	2	1.82E-09	1.82E-09	9	9	8.19E-09	8.19E-09
10	MMP	Liver	28-Days	12.5	17	17	4.05E-08	4.05E-08	5	5	1.19E-08	1.19E-08	17	19	4.05E-08	4.53E-08	13	13	2.35E-08	2.35E-08	10	26	1.81E-08	4.71E-08	4	4	7.24E-09	7.24E-09	6	8	6.17E-09	8.23E-09	8	8	8.23E-09	8.23E-09
11	MMP	Liver	28-Days	12.5	14	14	3.41E-08	3.41E-08	2	2	4.88E-09	4.88E-09	19	19	4.63E-08	4.63E-08	6	6	1.10E-08	1.10E-08	8	8	1.47E-08	1.47E-08	4	4	7.33E-09	7.33E-09	6	6	6.28E-09	6.28E-09	8	8	8.37E-09	8.37E-09
12	MMP	Liver	28-Days	12.5	15	15	3.92E-08	3.92E-08	8	8	2.09E-08	2.09E-08	23	31	6.02E-08	8.11E-08	6	6	1.19E-08	1.19E-08	8	8	1.59E-08	1.59E-08	2	2	3.97E-09	3.97E-09	0	0	0	0	5	5	5.65E-09	5.65E-09
13	MMP	Liver	28-Days	25	32	32	8.30E-08	8.30E-08	5	6	1.30E-08	1.56E-08	15	17	3.89E-08	4.41E-08	11	11	2.17E-08	2.17E-08	6	6	1.18E-08	1.18E-08	7	7	1.38E-08	1.38E-08	1	1	1.12E-09	1.12E-09	8	8	8.96E-09	8.96E-09
14	MMP	Liver	28-Days	25	35	35	8.84E-08	8.84E-08	16	19	4.04E-08	4.80E-08	20	28	5.05E-08	7.07E-08	12	12	2.29E-08	2.29E-08	8	10	1.53E-08	1.91E-08	0	0	0	0	2	2	2.18E-09	2.18E-09	4	4	4.35E-09	4.35E-09
15	MMP	Liver	28-Days	25	25	27	7.77E-08	8.39E-08	9	9	2.80E-08	2.80E-08	19	21	5.91E-08	6.53E-08	10	10	2.34E-08	2.34E-08	5	5	1.17E-08	1.17E-08	3	3	7.02E-09	7.02E-09	2	2	2.67E-09	2.67E-09	9	9	1.20E-08	1.20E-08
16	MMP	Liver	28-Days	25	31	31	1.12E-07	1.12E-07	3	3	1.09E-08	1.09E-08	18	18	6.52E-08	6.52E-08	11	11	3.09E-08	3.09E-08	5	5	1.40E-08	1.40E-08	4	4	1.12E-08	1.12E-08	2	2	3.16E-09	3.16E-09	1	1	1.58E-09	1.58E-09
17	MMP	Liver	28-Days	50	75	76	1.96E-07	1.98E-07	17	17	4.43E-08	4.43E-08	42	42	1.10E-07	1.10E-07	21	21	4.22E-08	4.22E-08	10	10	2.01E-08	2.01E-08	5	5	1.00E-08	1.00E-08	1	1	1.13E-09	1.13E-09	12	12	1.36E-08	1.36E-08
18	MMP	Liver	28-Days	50	136	136	3.31E-07	3.31E-07	24	24	5.84E-08	5.84E-08	45	52	1.09E-07	1.27E-07	30	30	5.58E-08	5.58E-08	14	16	2.60E-08	2.98E-08	4	4	7.44E-09	7.44E-09	5	6	5.27E-09	6.33E-09	7	25	7.38E-09	2.64E-08
19	MMP	Liver	28-Days	50	54	54	1.43E-07	1.43E-07	9	9	2.38E-08	2.38E-08	30	30	7.94E-08	7.94E-08	15	15	3.00E-08	3.00E-08	7	7	1.40E-08	1.40E-08	6	7	1.20E-08	1.40E-08	1	1	1.14E-09	1.14E-09	9	9	1.03E-08	1.03E-08
20	MMP	Liver	28-Days	50	63	63	1.85E-07	1.85E-07	19	20	5.59E-08	5.89E-08	40	44	1.18E-07	1.30E-07	22	22	4.87E-08	4.87E-08	13	14	2.88E-08	3.10E-08	4	4	8.86E-09	8.86E-09	2	2	2.53E-09	2.53E-09	7	7	8.85E-09	8.85E-09
21	MMP	Liver	28-Days	100	107	109	3.13E-07	3.19E-07	28	33	8.20E-08	9.67E-08	56	63	1.64E-07	1.85E-07	28	28	6.19E-08	6.19E-08	9	10	1.99E-08	2.21E-08	5	5	1.11E-08	1.11E-08	3	3	3.78E-09	3.78E-09	6	6	7.56E-09	7.56E-09
22	MMP	Liver	28-Days	100	229	232	5.17E-07	5.24E-07	45	45	1.02E-07	1.02E-07	88	181	1.99E-07	4.09E-07	62	63	1.06E-07	1.08E-07	23	23	3.94E-08	3.94E-08	10	10	1.71E-08	1.71E-08	5	5	4.87E-09	4.87E-09	21	21	2.05E-08	2.05E-08
23	MMP	Liver	28-Days	100	117	120	3.30E-07	3.38E-07	21	21	5.92E-08	5.92E-08	47	48	1.32E-07	1.35E-07	30	30	6.47E-08	6.47E-08	17	17	3.66E-08	3.66E-08	6	7	1.29E-08	1.51E-08	4	6	4.89E-09	7.33E-09	13	14	1.59E-08	1.71E-08
24	MMP	Liver	28-Days	100	100	101	2.59E-07	2.62E-07	26	26	6.74E-08	6.74E-08	64	66	1.66E-07	1.71E-07	35	35	6.81E-08	6.81E-08	22	22	4.28E-08	4.28E-08	13	13	2.53E-08	2.53E-08	4	4	4.45E-09	4.45E-09	20	21	2.22E-08	2.33E-08

Appendix D3 Table 7: Subtype specific mutagenicity data on lacZ in liver and bone marrow of MutaMouse males post exposure to BbF for 28-days

Animal ID	Endpoint	Sample Tissue	Exposure Duration	BbF dose (mg/kg/day)	C:G>A:T Count	C:G>A:T Count All	C:G>A:T MFmin	C:G>A:T MFmax	C:G>G:C Count	C:G>G:C Count All	C:G>G:C MFmin	C:G>G:C MFmax	C:G>T:A Count	C:G>T: A	C:G>T:A MFmin	C:G>T:A MFmax	T:A>A:T Count	T:A>A: T	T:A>A:T MFmin	T:A>A:T MFmax	T:A>C:G Count	T:A>C: G	T:A>C:G MFmin	T:A>C:G MFmax	T:A>G:C Count	T:A>G: C	T:A>G:C MFmin	T:A>G:C MFmax	Insertion Count Unique	Insertion Count All	Insertion MFmin	Insertion MFmax	Deletion Count	Deletio n	Deletion MFmin	Deletion MFmax
1	lacZ	Bone Marrow	28-Days	0	3	3	3.56E-08	3.56E-08	0	0	0	0	12	12	1.42E-07	1.42E-07	0	0	0	0	3	3	3.87E-08	3.87E-08	1	1	1.29E-08	1.29E-08	0	0	0	0	2	2	1.24E-08	1.24E-08
2	lacZ	Bone Marrow	28-Days	0	6	6	3.95E-08	3.95E-08	3	3	1.98E-08	1.98E-08	40	51	2.63E-07	3.36E-07	2	2	1.44E-08	1.44E-08	1	1	7.22E-09	7.22E-09	0	0	0	0	2	2	6.89E-09	6.89E-09	6	6	2.07E-08	2.07E-08
3	lacZ	Bone Marrow	28-Days	0	4	4	1.24E-07	1.24E-07	1	1	3.10E-08	3.10E-08	8	10	2.48E-07	3.10E-07	1	1	3.37E-08	3.37E-08	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1.62E-08	1.62E-08
4	lacZ	Bone Marrow	28-Days	0	4	4	3.71E-08	3.71E-08	0	0	0	0	24	36	2.23E-07	3.34E-07	0	0	0	0	3	3	3.03E-08	3.03E-08	1	1	1.01E-08	1.01E-08	1	1	4.83E-09	4.83E-09	1	1	4.83E-09	4.83E-09
5	lacZ	Bone Marrow	28-Days	6.25	2	20	3.27E-08	3.27E-07	3	3	4.91E-08	4.91E-08	16	17	2.62E-07	2.78E-07	2	3	3.55E-08	5.33E-08	0	0	0	0	1	1	1.78E-08	1.78E-08	0	0	0	0	1	4	8.52E-09	3.41E-08
6	lacZ	Bone Marrow	28-Days	6.25	10	13	8.54E-08	1.11E-07	5	5	4.27E-08	4.27E-08	17	20	1.45E-07	1.71E-07	2	2	1.86E-08	1.86E-08	1	1	9.31E-09	9.31E-09	1	1	9.31E-09	9.31E-09	0	0	0	0	4	4	1.78E-08	1.78E-08
7	lacZ	Bone Marrow	28-Days	6.25	17	19	8.04E-08	8.98E-08	3	3	1.42E-08	1.42E-08	36	1581	1.70E-07	7.48E-06	3	3	1.56E-08	1.56E-08	4	779	2.08E-08	4.05E-06	2	2	1.04E-08	1.04E-08	1	1	2.48E-09	2.48E-09	8	8	1.98E-08	1.98E-08
8	lacZ	Bone Marrow	28-Days	6.25	1	1	4.35E-08	4.35E-08	0	0	0	0	5	5	2.17E-07	2.17E-07	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	30	4.53E-08	6.79E-07
9	lacZ	Bone Marrow	28-Days	12.5	8	8	1.18E-07	1.18E-07	4	5	5.91E-08	7.39E-08	14	15	2.07E-07	2.22E-07	3	3	4.83E-08	4.83E-08	1	1	1.61E-08	1.61E-08	1	1	1.61E-08	1.61E-08	0	0	0	0	2	7	1.54E-08	5.39E-08
10	lacZ	Bone Marrow	28-Days	12.5	16	16	1.55E-07	1.55E-07	4	4	3.88E-08	3.88E-08	23	23	2.23E-07	2.23E-07	4	4	4.23E-08	4.23E-08	4	4	4.23E-08	4.23E-08	1	1	1.06E-08	1.06E-08	0	0	0	0	5	5	2.53E-08	2.53E-08
11	lacZ	Bone Marrow	28-Days	12.5	10	10	1.48E-07	1.48E-07	3	3	4.45E-08	4.45E-08	11	13	1.63E-07	1.93E-07	2	2	3.22E-08	3.22E-08	2	2	3.22E-08	3.22E-08	0	0	0	0	0	0	0	0	1	1	7.72E-09	7.72E-09
12	lacZ	Bone Marrow	28-Days	12.5	9	9	2.08E-07	2.08E-07	2	2	4.63E-08	4.63E-08	10	12	2.31E-07	2.78E-07	1	1	2.52E-08	2.52E-08	1	1	2.52E-08	2.52E-08	2	2	5.03E-08	5.03E-08	0	0	0	0	2	8	2.41E-08	9.64E-08
13	lacZ	Bone Marrow	28-Days	25	4	5	1.68E-07	2.09E-07	1	1	4.19E-08	4.19E-08	7	7	2.93E-07	2.93E-07	2	2	9.07E-08	9.07E-08	0	0	0	0	0	0	0	0	0	0	0	0	1	18	2.18E-08	3.92E-07
14	lacZ	Bone Marrow	28-Days	25	11	13	1.11E-07	1.31E-07	5	5	5.05E-08	5.05E-08	21	23	2.12E-07	2.32E-07	1	1	1.10E-08	1.10E-08	2	2	2.20E-08	2.20E-08	0	0	0	0	0	0	0	0	2	2	1.05E-08	1.05E-08
15	lacZ	Bone Marrow	28-Days	25	23	25	1.19E-07	1.29E-07	12	12	6.20E-08	6.20E-08	43	56	2.22E-07	2.89E-07	1	1	5.72E-09	5.72E-09	4	5	2.29E-08	2.86E-08	2	2	1.14E-08	1.14E-08	0	0	0	0	6	6	1.63E-08	1.63E-08
16	lacZ	Bone Marrow	28-Days	25	16	16	2.05E-07	2.05E-07	8	9	1.03E-07	1.15E-07	18	408	2.31E-07	5.23E-06	1	1	1.40E-08	1.40E-08	3	3	4.20E-08	4.20E-08	1	1	1.40E-08	1.40E-08	0	0	0	0	3	3	2.01E-08	2.01E-08
17	lacZ	Bone Marrow	28-Days	50	15	15	2.40E-07	2.40E-07	1	1	1.60E-08	1.60E-08	16	17	2.55E-07	2.71E-07	2	2	3.47E-08	3.47E-08	4	4	6.94E-08	6.94E-08	2	2	3.47E-08	3.47E-08	1	1	8.32E-09	8.32E-09	1	1	8.32E-09	8.32E-09
18	lacZ	Bone Marrow	28-Days	50	5	5	2.82E-07	2.82E-07	0	0	0	0	7	7	3.95E-07	3.95E-07	2	2	1.23E-07	1.23E-07	1	1	6.13E-08	6.13E-08	0	0	0	0	0	0	0	0	0	0	0	0
19	lacZ	Bone Marrow	28-Days	50	28	29	3.98E-07	4.13E-07	10	10	1.42E-07	1.42E-07	12	16	1.71E-07	2.28E-07	0	0	0	0	4	316	6.20E-08	4.90E-06	2	2	3.10E-08	3.10E-08	1	2	7.42E-09	1.48E-08	4	4	2.97E-08	2.97E-08
20	lacZ	Bone Marrow	28-Days	50	35	37	4.05E-07	4.28E-07	20	20	2.31E-07	2.31E-07	22	24	2.54E-07	2.78E-07	8	8	1.01E-07	1.01E-07	7	12	8.82E-08	1.51E-07	2	2	2.52E-08	2.52E-08	0	0	0	0	8	9	4.82E-08	5.43E-08
21	lacZ	Bone Marrow	28-Days	100	36	37	4.64E-07	4.77E-07	13	13	1.68E-07	1.68E-07	29	36	3.74E-07	4.64E-07	3	3	4.21E-08	4.21E-08	6	6	8.41E-08	8.41E-08	1	1	1.40E-08	1.40E-08	1	1	6.72E-09	6.72E-09	6	8	4.03E-08	5.37E-08
22	lacZ	Bone Marrow	28-Days	100	23	23	5.37E-07	5.37E-07	8	8	1.87E-07	1.87E-07	11	12	2.57E-07	2.80E-07	5	5	1.27E-07	1.27E-07	1	1	2.55E-08	2.55E-08	2	2	5.09E-08	5.09E-08	0	0	0	0	0	0	0	0
23	lacZ	Bone Marrow	28-Days	100	40	49	3.57E-07	4.38E-07	13	13	1.16E-07	1.16E-07	44	45	3.93E-07	4.02E-07	4	4	3.90E-08	3.90E-08	3	3	2.93E-08	2.93E-08	2	2	1.95E-08	1.95E-08	2	3	9.33E-09	1.40E-08	9	10	4.20E-08	4.66E-08
24	lacZ	Bone Marrow	28-Days	100	38	43	3.76E-07	4.25E-07	3	3	2.97E-08	2.97E-08	78	89	7.72E-07	8.81E-07	76	83	8.23E-07	8.99E-07	60	68	6.50E-07	7.37E-07	21	23	2.27E-07	2.49E-07	0	0	0	0	1	1	5.17E-09	5.17E-09
1	lacZ	Liver	28-Days	0	1	1	1.44E-08	1.44E-08	0	0	0	0	21	22	3.02E-07	3.17E-07	2	2	3.14E-08	3.14E-08	0	0	0	0	3	3	4.71E-08	4.71E-08	0	0	0	0	1	1	7.51E-09	7.51E-09
2	lacZ	Liver	28-Days	0	1	1	2.06E-08	2.06E-08	0	0	0	0	8	9	1.65E-07	1.86E-07	1	1	2.26E-08	2.26E-08	1	1	2.26E-08	2.26E-08	2	2	4.53E-08	4.53E-08	0	0	0	0	2	2	2.16E-08	2.16E-08
3	lacZ	Liver	28-Days	0	4	4	1.53E-08	1.53E-08	1	1	3.81E-09	3.81E-09	67	88	2.55E-07	3.36E-07	0	0	0	0	5	5	2.12E-08	2.12E-08	3	3	1.27E-08	1.27E-08	0	0	0	0	7	7	1.41E-08	1.41E-08
4	lacZ	Liver	28-Days	0	1	1	1.28E-08	1.28E-08	0	0	0	0	13	13	1.67E-07	1.67E-07	0	0	0	0	2	2	2.80E-08	2.80E-08	0	0	0	0	0	0	0	0	4	6	2.68E-08	4.02E-08
5	lacZ	Liver	28-Days	6.25	11	11	9.67E-08	9.67E-08	2	2	1.76E-08	1.76E-08	28	33	2.46E-07	2.90E-07	2	2	1.93E-08	1.93E-08	5	6	4.82E-08	5.78E-08	1	1	9.63E-09	9.63E-09	0	0	0	0	2	2	9.19E-09	9.19E-09
6	lacZ	Liver	28-Days	6.25	5	5	3.95E-08	3.95E-08	1	1	7.90E-09	7.90E-09	19	21	1.50E-07	1.66E-07	3	3	2.60E-08	2.60E-08	0	0	0	0	0	0	0	0	0	0	0	0	3	3	1.24E-08	1.24E-08
7	lacZ	Liver	28-Days	6.25	8	8	8.54E-08	8.54E-08	1	1	1.07E-08	1.07E-08	21	649	2.24E-07	6.93E-06	1	1	1.17E-08	1.17E-08	1	391	1.17E-08	4.57E-06	1	1	1.17E-08	1.17E-08	0	0	0	0	1	1	5.58E-09	5.58E-09
8	lacZ	Liver	28-Days	6.25	4	4	2.76E-08	2.76E-08	0	0	0	0	48	111	3.32E-07	7.67E-07	1	1	7.57E-09	7.57E-09	1	1	7.57E-09	7.57E-09	2	2	1.51E-08	1.51E-08	1	1	3.61E-09	3.61E-09	1	1	3.61E-09	3.61E-09
9	lacZ	Liver	28-Days	12.5	10	11	7.41E-08	8.15E-08	2	2	1.48E-08	1.48E-08	33	34	2.45E-07	2.52E-07	1	1	8.17E-09	8.17E-09	2	2	1.63E-08	1.63E-08	1	1	8.17E-09	8.17E-09	1	1	3.89E-09	3.89E-09	0	0	0	0
10	lacZ	Liver	28-Days	12.5	11	11	1.45E-07	1.45E-07	2	2	2.64E-08	2.64E-08	23	23	3.04E-07	3.04E-07	0	0	0	0	2	2	2.89E-08	2.89E-08	1	1	1.45E-08	1.45E-08	0	0	0	0	1	1	6.90E-09	6.90E-09
11	lacZ	Liver	28-Days	12.5	10	10	8.69E-08	8.69E-08	2	2	1.74E-08	1.74E-08	30	37	2.61E-07	3.22E-07	3	3	2.86E-08	2.86E-08	1	1	9.52E-09	9.52E-09	1	1	9.52E-09	9.52E-09	0	0	0	0	2	2	9.09E-09	9.09E-09
12	lacZ	Liver	28-Days	12.5	4	6	3.47E-08	5.20E-08	4	4	3.47E-08	3.47E-08	33	41	2.86E-07	3.55E-07	2	2	1.90E-08	1.90E-08	1	1	9.49E-09	9.49E-09	0	0	0	0	0	0	0	0	0	0	0	0
13	lacZ	Liver	28-Days	25	20	23	1.60E-07	1.84E-07	4	4	3.19E-08	3.19E-08	36	65	2.87E-07	5.19E-07	0	0	0	0	3	3	2.62E-08	2.62E-08	3	3	2.62E-08	2.62E-08	0	0	0	0	1	1	4.17E-09	4.17E-09
14	lacZ	Liver	28-Days	25	11	11	1.40E-07	1.40E-07	3	3	3.81E-08	3.81E-08	19	22	2.41E-07	2.79E-07	4	4	5.54E-08	5.54E-08	0	0	0	0	0	0	0	0	0	0	0	0	1	1	6.62E-09	6.62E-09
15	lacZ	Liver	28-Days	25	10	10	1.24E-07	1.24E-07	3	3	3.73E-08	3.73E-08	21	23	2.61E-07	2.86E-07	1	1	1.35E-08	1.35E-08	1	1	1.35E-08	1.35E-08	0	0	0	0	1	1	6.48E-09	6.48E-09	0	0	0	0
16	lacZ	Liver	28-Days	25	26	27	2.01E-07	2.09E-07	4	4	3.09E-08	3.09E-08	31	32	2.40E-07	2.48E-07	1	1	8.43E-09	8.43E-09	0	0	0	0	1	1	8.43E-09	8.43E-09	0	0	0	0	3	3	1.21E-08	1.21E-08
17	lacZ	Liver	28-Days	50	48	51	3.98E-07	4.23E-07	6	6	4.97E-08	4.97E-08	39	40	3.23E-07	3.32E-07	4	4	3.63E-08	3.63E-08	1	1	9.08E-09	9.08E-09	2	2	1.82E-08	1.82E-08	1	1	4.33E-09	4.33E-09	3	3	1.30E-08	1.30E-08
18	lacZ	Liver	28-Days	50	21	21	4.05E-07	4.05E-07	7	7	1.35E-07	1.35E-07	17	24	3.28E-07	4.63E-07	0	0	0	0	1	1	2.11E-08	2.11E-08	0	0	0	0	0	0	0	0	0	0	0	0
19	lacZ	Liver	28-Days	50	28	28	2.22E-07	2.22E-07	7	7	5.56E-08	5.56E-08	36	37	2.86E-07	2.94E-07	3	3	2.61E-08	2.61E-08	3	538	2.61E-08	4.68E-06	1	1	8.70E-09	8.70E-09	0	0	0	0	0	0	0	0
20	lacZ	Liver	28-Days	50	60	65	3.77E-07	4.09E-07	14	14	8.81E-08	8.81E-08	46	54	2.89E-07	3.40E-07	5	5	3.45E-08	3.45E-08	6	688	4.14E-08	4.74E-06	2	2	1.38E-08	1.38E-08	1	1	3.29E-09	3.29E-09	4	5	1.32E-08	1.64E-08
21	lacZ	Liver	28-Days	100	54	55	5.95E-07	6.06E-07	16	16	1.76E-07	1.76E-07	26	27	2.87E-07	2.98E-07	6	6	7.20E-08	7.20E-08	2	2	2.40E-08	2.40E-08	0	0	0	0	1	1	5.75E-09	5.75E-09	2	2	1.15E-08	1.15E-08
22	lacZ	Liver	28-Days	100	159	173	9.76E-07	1.06E-06	36	36	2.21E-07	2.21E-07	85	90	5.22E-07	5.53E-07	13	15	8.80E-08	1.01E-07	4	4	2.71E-08	2.71E-08	4	4	2.71E-08	2.71E-08	1	1	3.22E-09	3.22E-09	6	6	1.93E-08	1.93E-08
23	lacZ	Liver	28-Days	100	34	34	7.13E-07	7.13E-07	9	9	1.89E-07	1.89E-07	9	10	1.89E-07	2.10E-07	2	2	4.58E-08	4.58E-08	2	2	4.58E-08	4.58E-08	3	3	6.87E-08	6.87E-08	0	0	0	0	1	1	1.09E-08	1.09E-08
24	lacZ	Liver	28-Days	100	58	60	5.29E-07	5.48E-07	10	10	9.13E-08	9.13E-08	42	45	3.83E-07	4.11E-07	2	2	1.99E-08	1.99E-08	1	1	9.97E-09	9.97E-09	1	1	9.97E-09	9.97E-09	1	1	4.76E-09	4.76E-09	2	2	9.53E-09	9.53E-09

Appendix D3 Table 8: Subtype specific mutagenicity data on Pig-a in liver and bone marrow of MutaMouse males post exposure to BbF for 28-days

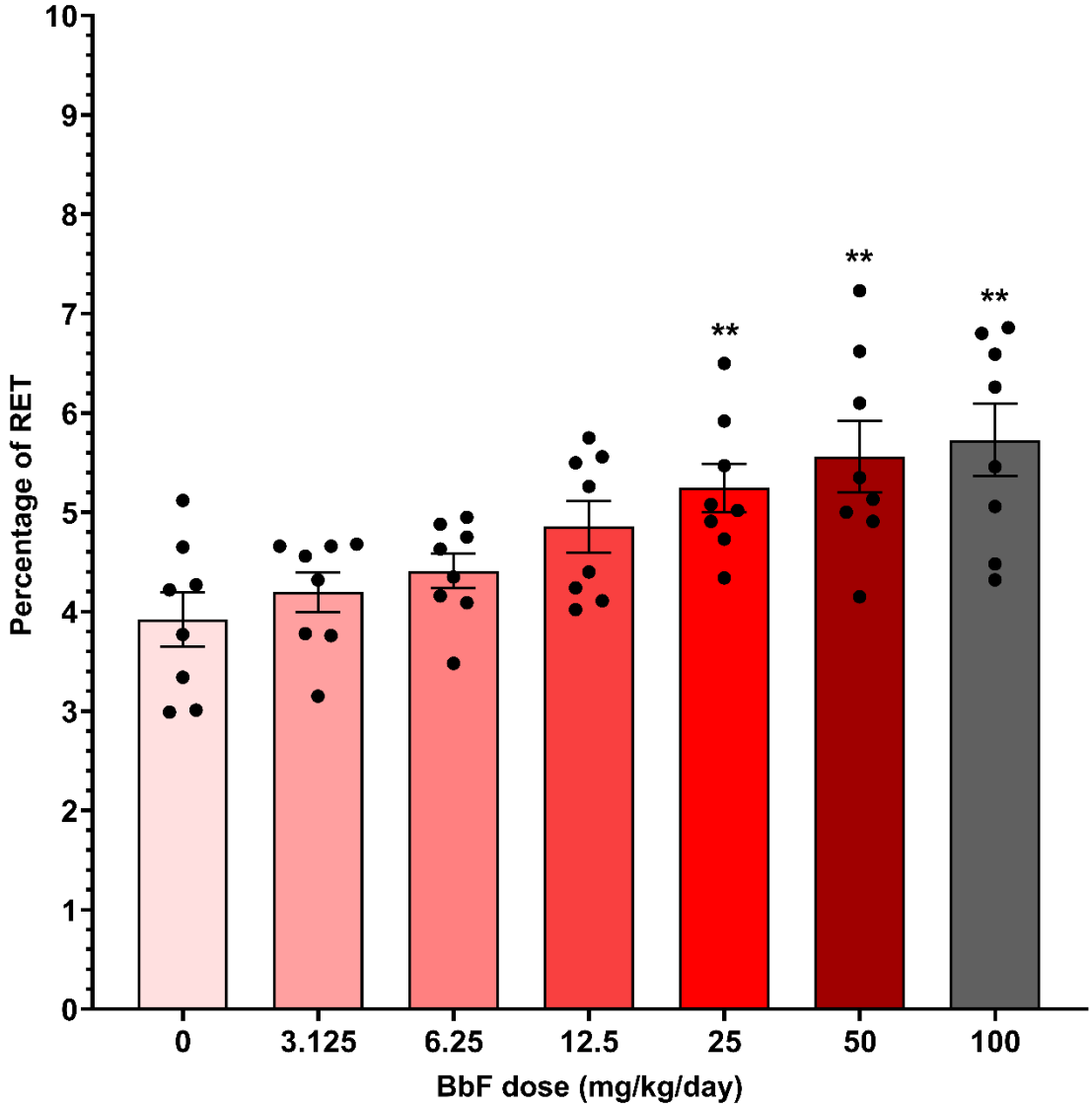
Animal ID	Sample Tissue	Exposure Duration	BbF dose (mg/kg/day)	C:G>A:T Count	C:G>A:T Count All	C:G>A:T MFmin	C:G>A:T MFmax	C:G>G:C Count	C:G>G:C Count All	C:G>G:C MFmin	C:G>G:C MFmax	C:G>T:A Count	C:G>T: A	C:G>T:A MFmin	C:G>T:A MFmax	T:A>A:T Count	T:A>A: T	T:A>A:T MFmin	T:A>A:T MFmax	T:A>C:G Count	T:A>C: G	T:A>C:G MFmin	T:A>C:G MFmax	T:A>G:C Count	T:A>G: C	T:A>G:C MFmin	T:A>G:C MFmax	Insertion Count Unique	Insertion Count All	Insertion MFmin	Insertion MFmax	Deletion Count	Deletio n	Deletion MFmin	Deletion MFmax
1	Bone Marrow	28-Days	0	2	2	3.67E-08	3.67E-08	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	51	3.80E-08	3.87E-07	5	147	3.80E-08	1.12E-06
2	Bone Marrow	28-Days	0	0	0	0	0	0	0	0	0	2	2	3.12E-08	3.12E-08	0	0	0	0	0	0	0	0	1	1	1.11E-08	1.11E-08	5	115	3.23E-08	7.44E-07	7	146	4.53E-08	9.44E-07
3	Bone Marrow	28-Days	0	1	1	1.78E-08	1.78E-08	1	1	1.78E-08	1.78E-08	0	0	0	0	1	1	1.26E-08	1.26E-08	1	1	1.26E-08	1.26E-08	0	0	0	0	5	49	3.69E-08	3.62E-07	3	33	2.21E-08	2.44E-07
4	Bone Marrow	28-Days	0	1	1	1.55E-08	1.55E-08	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	95	1.93E-08	6.12E-07	4	109	2.58E-08	7.03E-07
5	Bone Marrow	28-Days	6.25	1	1	2.14E-08	2.14E-08	0	0	0	0	3	3	6.42E-08	6.42E-08	0	0	0	0	0	0	0	0	0	0	0	0	4	99	3.55E-08	8.78E-07	7	74	6.21E-08	6.56E-07
6	Bone Marrow	28-Days	6.25	0	0	0	0	1	1	2.81E-08	2.81E-08	1	1	2.81E-08	2.81E-08	0	0	0	0	0	0	0	0	0	0	0	0	3	34	3.52E-08	3.98E-07	6	77	7.03E-08	9.02E-07
7	Bone Marrow	28-Days	6.25	2	2	4.77E-08	4.77E-08	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	39	1.97E-08	3.84E-07	7	62	6.89E-08	6.10E-07
8	Bone Marrow	28-Days	6.25	0	0	0	0	0	0	0	0	2	2	3.47E-08	3.47E-08	0	0	0	0	0	0	0	0	0	0	0	0	5	85	3.62E-08	6.15E-07	6	116	4.34E-08	8.39E-07
9	Bone Marrow	28-Days	12.5	0	0	0	0	0	0	0	0	1	1	1.37E-08	1.37E-08	0	0	0	0	3	3	2.93E-08	2.93E-08	0	0	0	0	3	84	1.71E-08	4.80E-07	6	49	3.43E-08	2.80E-07
10	Bone Marrow	28-Days	12.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	120	3.57E-08	1.07E-06	5	134	4.46E-08	1.19E-06
11	Bone Marrow	28-Days	12.5	0	0	0	0	2	2	2.99E-08	2.99E-08	0	0	0	0	0	0	0	0	1	5	1.06E-08	5.30E-08	1	1	1.06E-08	1.06E-08	5	104	3.10E-08	6.46E-07	5	74	3.10E-08	4.59E-07
12	Bone Marrow	28-Days	12.5	1	2	1.86E-08	3.72E-08	0	0	0	0	2	2	3.72E-08	3.72E-08	0	0	0	0	0	0	0	0	0	0	0	0	4	74	3.09E-08	5.72E-07	5	45	3.86E-08	3.48E-07
13	Bone Marrow	28-Days	25	2	2	2.72E-08	2.72E-08	1	1	1.36E-08	1.36E-08	5	5	6.81E-08	6.81E-08	0	0	0	0	0	0	0	0	0	0	0	0	4	105	2.26E-08	5.94E-07	8	172	4.52E-08	9.73E-07
14	Bone Marrow	28-Days	25	1	1	1.77E-08	1.77E-08	0	0	0	0	1	3	1.77E-08	5.31E-08	0	0	0	0	0	0	0	0	0	0	0	0	3	64	2.21E-08	4.71E-07	3	27	2.21E-08	1.99E-07
15	Bone Marrow	28-Days	25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	114	4.44E-08	1.26E-06	6	98	6.65E-08	1.09E-06
16	Bone Marrow	28-Days	25	2	3	3.00E-08	4.50E-08	1	1	1.50E-08	1.50E-08	2	4	3.00E-08	6.00E-08	2	2	2.13E-08	2.13E-08	1	1	1.07E-08	1.07E-08	1	1	1.07E-08	1.07E-08	9	126	5.61E-08	7.85E-07	3	131	1.87E-08	8.16E-07
17	Bone Marrow	28-Days	50	3	3	5.70E-08	5.70E-08	0	0	0	0	2	6	3.80E-08	1.14E-07	0	0	0	0	2	2	2.68E-08	2.68E-08	1	6	1.34E-08	8.05E-08	4	75	3.14E-08	5.90E-07	6	29	4.72E-08	2.28E-07
18	Bone Marrow	28-Days	50	2	2	2.86E-08	2.86E-08	1	1	1.43E-08	1.43E-08	5	5	7.14E-08	7.14E-08	2	2	2.03E-08	2.03E-08	0	0	0	0	0	0	0	0	5	83	2.96E-08	4.92E-07	5	86	2.96E-08	5.10E-07
19	Bone Marrow	28-Days	50	0	0	0	0	2	2	4.32E-08	4.32E-08	2	5	4.32E-08	1.08E-07	1	1	1.52E-08	1.52E-08	0	0	0	0	0	0	0	0	2	97	1.79E-08	8.66E-07	3	56	2.68E-08	5.00E-07
20	Bone Marrow	28-Days	50	3	3	4.95E-08	4.95E-08	1	1	1.65E-08	1.65E-08	3	4	4.95E-08	6.60E-08	2	3	2.34E-08	3.51E-08	2	2	2.34E-08	2.34E-08	0	0	0	0	6	99	4.11E-08	6.77E-07	5	120	3.42E-08	8.21E-07
21	Bone Marrow	28-Days	100	2	2	3.16E-08	3.16E-08	0	0	0	0	5	6	7.90E-08	9.48E-08	0	0	0	0	2	2	2.26E-08	2.26E-08	0	0	0	0	4	91	2.63E-08	5.99E-07	2	60	1.32E-08	3.95E-07
22	Bone Marrow	28-Days	100	3	4	5.75E-08	7.67E-08	0	0	0	0	0	0	0	0	1	3	1.36E-08	4.08E-08	0	0	0	0	0	0	0	0	6	92	4.77E-08	7.32E-07	6	124	4.77E-08	9.86E-07
23	Bone Marrow	28-Days	100	0	0	0	0	0	0	0	0	0	0	0	0	1	1	9.26E-09	9.26E-09	0	0	0	0	0	0	0	0	3	53	1.62E-08	2.87E-07	6	82	3.25E-08	4.44E-07
24	Bone Marrow	28-Days	100	2	2	3.10E-08	3.10E-08	2	2	3.10E-08	3.10E-08	1	2	1.55E-08	3.10E-08	0	0	0	0	1	1	1.10E-08	1.10E-08	0	0	0	0	4	87	2.57E-08	5.60E-07	8	116	5.15E-08	7.47E-07

Appendix D3 Table 9: Overview of Pearson correlation analysis between DS on MMP or lacZ vs. TGR assay

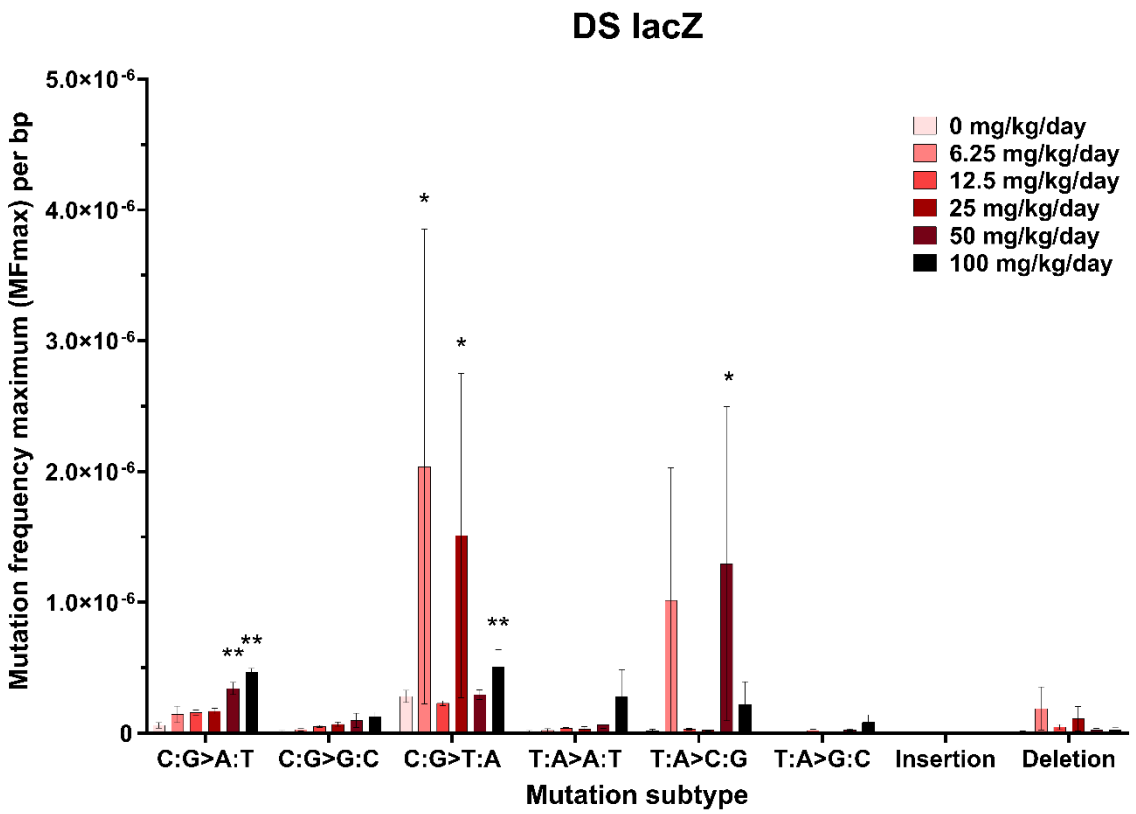
Tissue	Endpoint 1	Endpoint 2	Pearson r (linear)	p (linear)	Pearson r (log10)	p (log10)	Spearman rho	p (Spearman)	95% CI linear	95% CI log10
Bone marrow	DS lacZ MFMin	TGR Mutant frequency	0.696	0.000159	0.822	0.000000819	0.891	5.17E-09	0.43-0.85	0.63-0.92
Bone marrow	DS lacZ MFmax	TGR Mutant frequency	0.091	0.674	0.441	0.0311	0.58	0.00297	-0.8	0.05-0.72
Bone marrow	DS MMP MFMin	TGR Mutant frequency	0.893	4.29E-09	0.843	0.000000228	0.875	2.27E-08	0.77-0.95	0.67-0.93
Bone marrow	DS MMP MFmax	TGR Mutant frequency	0.455	0.0254	0.587	0.00257	0.702	0.000133	0.06-0.73	0.24-0.80
Liver	DS lacZ MFMin	TGR Mutant frequency	0.867	4.37E-08	0.929	5.43E-11	0.926	8.78E-11	0.70-0.94	0.84-0.97
Liver	DS lacZ MFmax	TGR Mutant frequency	0.071	0.743	0.488	0.0157	0.737	0.0000394	-0.8	0.10-0.74
Liver	DS MMP MFMin	TGR Mutant frequency	0.911	6.66E-10	0.928	6.47E-11	0.904	1.35E-09	0.80-0.96	0.84-0.97
Liver	DS MMP MFmax	TGR Mutant frequency	0.876	2.01E-08	0.868	4.06E-08	0.82	0.000000938	0.73-0.95	0.71-0.94

Appendix D3 Table 10: Overview of BMD modeling DS MMP, DS lacZ and TGR assay data in bone marrow and liver

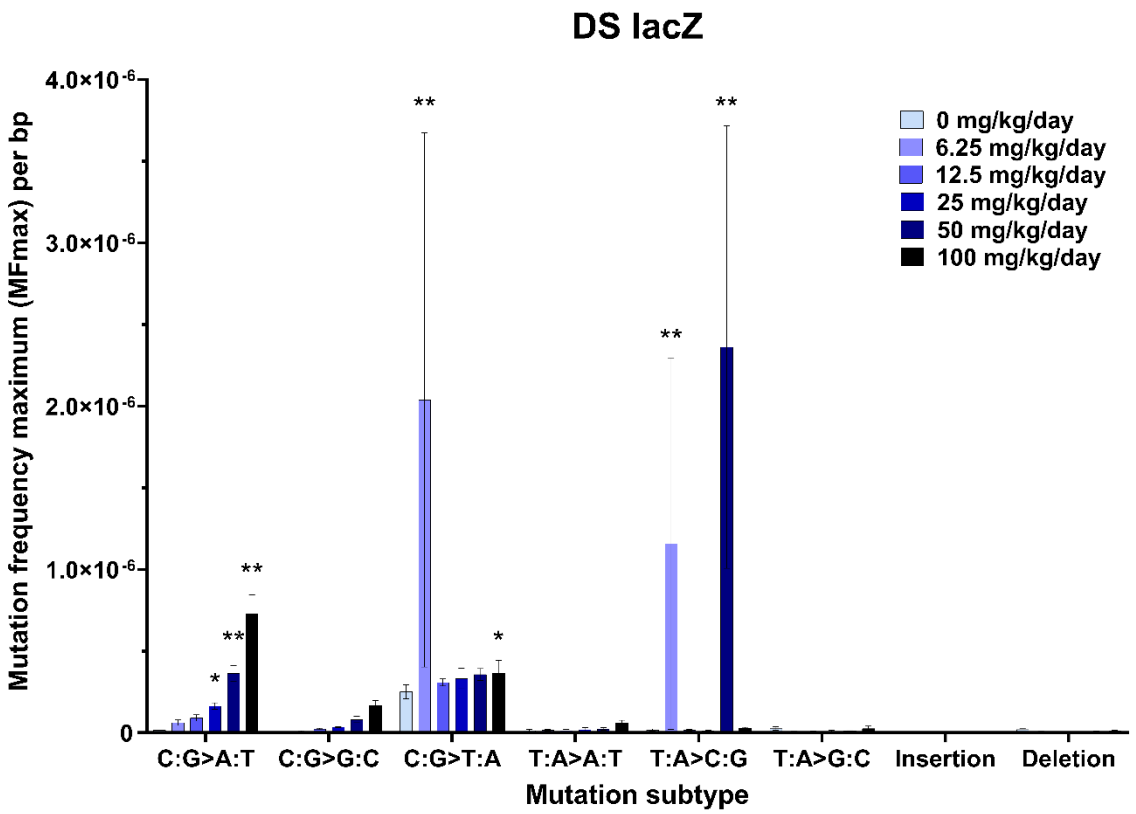
Exposure duration	Sample Tissue	Endpoint	n	CES	Bootstrap	BMDL	BMDU	Log10 BMDL	Log10 BMDU
28-Days	Bone Marrow	DS MMP	24	0.5	200	3.60	12.80	0.56	1.11
	Bone Marrow	DS lacZ				7.62	32.50	0.88	1.51
	Bone Marrow	TGR assay	48			1.11	7.79	0.05	0.89
	Liver	DS MMP	24	0.5	200	10.10	18.40	1.00	1.26
	Liver	DS lacZ				10.60	29.70	1.03	1.47
	Liver	TGR assay	48			5.08	15.40	0.71	1.19



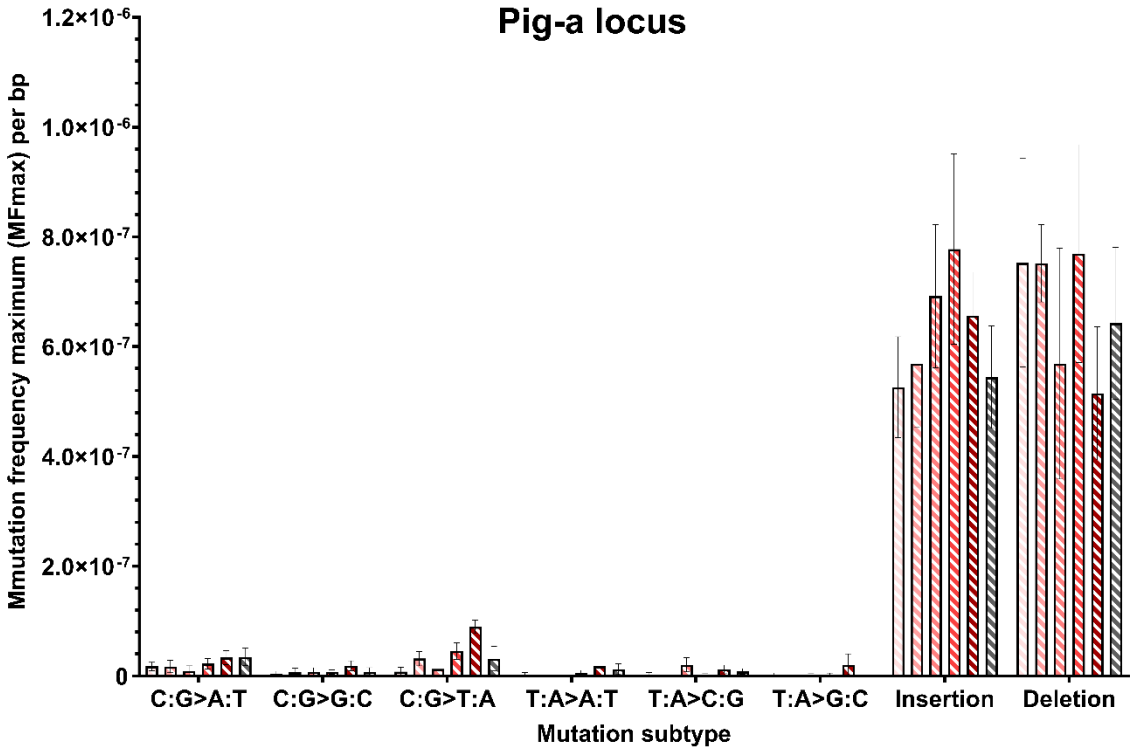
*Appendix D3 Figure 1:%RET ± SEM measured using the Pig-a assay on male MutaMouse blood extracted after 28-days of exposure to increasing doses of BbF or olive oil via oral gavage. Statistically significant increases compared to the respective vehicle controls were determined after Box-Cox transformation of the mutant frequencies followed by ANOVA with post-hoc Dunnett test and are indicated as * $p < 0.05$ and ** $p < 0.01$.*



*Appendix D3 Figure 2: Mutation frequency maximum (MFmax) ± SEM estimates per mutation subtype measured by duplex sequencing in MutaMouse males exposed to increasing doses of BbF or olive oil for 28+3 days via oral gavage. Results are shown for the lacZ locus, in bone marrow with n = 4 per group. Statistically significant increases compared to respective vehicle controls were determined using generalized linear mixed models and are indicated as * p < 0.05 and ** p < 0.01.*



*Appendix D3 Figure 3: : Mutation frequency maximum (MFmax) ± SEM estimates per mutation subtype measured by duplex sequencing in MutaMouse males exposed to increasing doses of BbF or olive oil for 28+3 days via oral gavage. Results are shown for the lacZ locus, in liver with n = 4 per group. Statistically significant increases compared to respective vehicle controls were determined using generalized linear mixed models and are indicated as * p < 0.05 and ** p < 0.01.*



*Appendix D3 Figure 4: Mutation frequency maximum (MFmax) ± SEM estimates per mutation subtype measured by duplex sequencing in MutaMouse males exposed to increasing doses of BbF or olive oil for 28+3 days via oral gavage. Results are shown for the Pig-a locus, in bone marrow with n = 4 per group. Statistically significant increases compared to respective vehicle controls were determined using generalized linear mixed models and are indicated as * $p < 0.05$ and ** $p < 0.01$.*