uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Caroline Healy
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (Biology)
GRADE / DEGREE

Department of Biology
FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Flow Cytometric Detection of Chemically Induced Tandem Repeat Mutations in Two Murine Cell Lines

TITRE DE LA THÈSE / TITLE OF THESIS

D. Johnson
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

S. Bennett

I. Lambert

M. Paulin-Levasseur

Gary W. Slater
LE DOYEN DE LA FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES /
DEAN OF THE FACULTY OF GRADUATE AND POSTDOCORAL STUDIES

Flow Cytometric Detection of Chemically Induced Tandem Repeat Mutations in Two Murine Cell Lines

by
Caroline Marie Thérèse Healy

A thesis submitted to the School of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of
Masters of Science in Biology

Department of Biology
Faculty of Science
University of Ottawa

October, 2005

Your file Votre référence
ISBN: 0-494-11289-1
Our file Notre référence
ISBN: 0-494-11289-1

CONTENTS

Acknowlegments

I thank Dr. R Slebos from the laboratory of Molecular Carcinogenesis at NIH Research Triangle Park for helpful advice on transfections, Dr. M. Liskay and Megan Nguyen from the laboratory of Molecular & Medical Genetics at Oregon Health & Science University for kindly providing the MC2a cell line, Dr. Mike Wade and Avril McMahon from the Systemic Toxicology and Pharmacokinetics Section at Health Canada for their contribution with flow cytometry analysis, Raymond Pilon from the Mutagenesis section at Health Canada for methodological advice on RNA work and Andrew Williams from Biostatistics and Epidemiology Division at Health Canada for technical support in statistical analyses. Also, a special thank you to my thesis supervisor Dr. Douglas Johnson (department of Biology, University of Ottawa) and research supervisor Dr. Craig Parfett (Mutagenesis section at Health Canada) for sharing their knowledge and insights and providing great guidance, to my Committee members for helpful suggestions and feedback as well as to all reviewers that gave useful advice.

Abstract

To facilitate detection of genotoxicity from environmental mutagen exposure we generated an *in vitro* green fluorescence protein (GFP) activation assay that quickly and effectively detects frameshift mutations in tandem repeat sequences (TRS). Two murine cell lines were stably transfected with GFP reporter plasmids in which the GFP constructs contain TRS that shift the GFP coding sequence out of frame. These included several 2-6bp repeat sequences, a control non-repetitive sequence and a human gene sequence with TRS. Transfected cultures were exposed to five model chemicals: hydrogen peroxide (H_2O_2), 12-*O*-tetradecanoyl-phorbol-13-acetate (TPA), benzo-*a*-pyrene-diol-epoxide (BPDE), ethyl nitrosourea (ENU), 9-aminoacridine (9AA). Frameshift mutations resulted in green fluorescent revertants, as determined by flow cytometry and confirmed by sequencing. All five treatments induced a statistically significant sequence- and dose-dependent response in both cell lines. Results from these experiments reveal that the assay responds robustly to various classes of mutagenic substances, as well as carcinogens that are inactive in conventional mutation assays.

Résumé

Pour faciliter la détection de la génotoxicité provenant de l'exposition à ces mutagènes environnementaux, nous avons produit un test i*n vitro* de reactivation d'une protéine fluorescente verte (GFP). Ce test, tel que réalisé, permet de détecter plus rapidement et plus efficacement les mutations causant un décalage du cadre de lecture dans les séquences répétées en tandem (SRT). Deux lignées cellulaires de souris ont été transfectées, avec des plasmides possédant le gène rapporteur GFP. Les gènes hybrides contenant la séquence codante de la GFP ont des SRT qui entraînent un décalage du cadre de lecture. Les séquences étudiées comprenaient plusieurs séquences répétées de 2 à 6 pb, une séquence non répétitive témoin et une séquence de gène humain. Les cultures transfectées ont été exposées à cinq substances chimiques modèles: un agent stressant oxydatif, le peroxyde d'hydrogène (H_2O_2); un promoteur de tumeur, le 12-*O*-tétradécanoyl-phorbol-13-acétate (TPA); un hydrocarbure polyaromatique, le diol époxide du benzo-*a*-pyrène (BPDE); un agent alkylant, l'éthylnitrosourée (ENU) et un agent intercalant, la 9-amino-acridine (9AA). Les mutations de décalage du cadre de lecture ont engendré l'expression de GFP révertants verts, mutations qui ont été détectées par la cytométrie de flux et qui ont aussi été confirmées par le séquençage. Les cinq traitements ont donné une réponse statistiquement significative, dépendamment de la séquence et de la dose dans les deux lignées cellulaires. Les résultats de ces expériences indiquent que l'essai est sensible à diverses catégories de mutagènes et de certains cancérogènes qui sont inactifs dans les essais conventionnels pour la détection de mutations.

List of Abbreviations

9AA	9 amino acridine
AFLP	amplified fragment length polymorphism
B-*a*-P	benzo-*a*-pyrene
BPDE	benzo-*a*-pyrene-diol epoxide
DMEM	Dulbecco's modified Eagle's medium
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
EDTA	ethylene diamine tetra-acetic acid
EGFP	enhanced green fluorescent protein
ENU	N-ethyl-N-nitrosourea
ESTR	extended simple tandem repeat
FBS	fetal bovine serum
FISH	fluorescence *in situ* hybridization
FSC	forward-angle light scatter
GFP	green fluorescent protein
HNPCC	hereditary non-polyposis colorectal cancer
HPRT	hypoxanthine-guanine phosphoribosyl transferase
H_2O_2	hydrogen peroxide
HPRT	hypoxanthine phosphoribosyltransferase
MMR	mismatch repair
MEM	minimal essential medium
MNNG	N-methyl-N-nitro-N-nitrosoguanidine
PAH	polycyclic aromatic hydrocarbons
PBS	phosphate-buffered saline
PCR	polymerace chain reaction
PKC	protein kinase C
RAPD	randomly amplified polymorphic DNA
RNA	ribonucleic acid
RNAse	ribonuclease
ROS	reactive oxygen species
RSI	repeat sequence instability
SDS	sodium dodecyl sulfate
SSC	side-angle light scatter
TBS	tris buffered saline
THF	tetrahydrofuran
TPA	twelve-*O*-tetradecanoyl-phorbol-13-acetate
TRS	tandem repeat sequences

List of Figures

List of Tables

Chapter 1

1.0 Introduction

Heritable DNA sequence alterations and structural changes commonly occur due to exposure to environmental stress as well as to endogenous cellular metabolism, e.g. oxidative stress (Barnes and Lindahl, 2004). DNA regions with tandem repeated sequences (TRS) are generally more prone to DNA sequence alterations than other parts of the genome (Debrauwere et al., 1997; Yauk, 1998). This is of particular concern since such genetic instability can lead to impairment of gene functions and to diseases if the damage is not properly repaired. Health regulations have been established to identify potential sources of environmental mutagens and to limit our exposure to them, e.g., the Canadian Environmental Protection Act, 1999. Many assays have been developed to determine the genotoxicity of environmental contaminants; however, as part of regulatory policy only conventional assays such as Ames bacterial reversion test (Ames, 1973; OECD, 1997a), *in vitro* and *in vivo* chromosome aberration (OECD, 1997a; OECD, 1997b) *in vivo* micronucleus tests (OECD, 1997d) and *in vitro* mammalian cell gene mutation assays (OECD, 1997e) have been internationally adopted to evaluate genotoxins. TRS represent relevant target sites to investigate mutagenesis and carcinogenesis: they are more mutable than heteropolymeric DNA, sensitive to stress exposure (Yauk, 1998), have functional roles in the genome (Li et al., 2004), sustain mutations that lead to diseases (Mori et al., 2001) and are minimally represented in toxicological screening assays. Assays such as DNA fingerprinting and PCR based test systems that measure size and sequence changes in TRS as biomarkers of exposure to contaminants are currently used as research tools but could prove effective alternative methods for chemical screening (Yauk, 2004). An additional approach currently under investigation is the development of reporter systems based upon chimeric green

fluorescent protein (GFP) constructs that are able to detect frameshift mutations in TRS joined upstream of the GFP coding DNA sequence (Gasche et al., 2001). To extend the range and utility of the frameshift detection assay, we generated novel GFP vectors containing TRS of 2-6bp, as well as a repetitive human gene sequence, to be used in toxicologically relevant cell lines that test diverse types of mutagens. Ultimately, the GFP frameshift detection system could be used as a quick, reliable and economic *in vitro* assay to detect mutations arising from alteration of cellular processes caused by exposure to environmental pollutants. Data generated from GFP activation assays would help to predict potential human health risk from exposure to contaminants and to assist in investigating the mechanism involved in environmentally-induced tandem repeat sequence instability.

1.1 Tandem repeats and genetic instability

Tandem repeats are widespread throughout most eukaryotic genomes (Toth et al., 2000). These repetitive regions contain sequence motifs that occur in both coding and non-coding DNA sequences. Long believed to be "junk DNA", recent molecular studies have indicated various functional roles for repetitive DNA including structural integrity of chromosomes, involvement in the regulation of transcription, splicing, and translation into repeated amino acids that form binding motifs or affect other biochemical properties of proteins (Nagashima et al., 2004). Some notable examples of this are: the essential short tandem repeats found in the telomeres of most eukaryotes (Bah et al., 2004), loss of striated muscle chloride channel CIC-1 from CTG repeat expansion induced alternative splicing (Mankodi et al., 2002) and frameshift mutations in the unstable poly A repeat in

the transforming-growth factor beta 1 type II receptor gene can lead to loss of function of the receptor (McCaffrey et al., 1997).

The term 'tandem repeat' includes microsatellites, minisatellites and expanded simple tandem repeats (ESTR). The specific definition for some of these categories can vary greatly from source to source (Sia et al., 1997), and for this reason we have adopted the general term TRS throughout this study. Typically microsatellites are short stretches of 1-5bp repeat units (Tautz and Schlotterer, 1994) and minisatellites contain 6-100bp repeat units that stretch from 500bp to several kb in length (Vergnaud and Denoeud, 2000). ESTR have repeat units of less than 10bp but the total length of the ESTR can range upwards to 18kb as found for some murine MMS10 (GGCAGA) arrays (Bois et al., 2001).

TRS are often found to be mutational hotspots. They are highly unstable or variable as frequent changes in repeat number and sequence can occur in these sequences during meiotic and mitotic processes. For somatic cells, it is currently believed that this increased sequence instability is partly due to unequal mitotic recombination (Bishop and Schiestl, 2002; Richard and Paques, 2000; Sia et al., 1997) and replication errors such as DNA slippage (Streisinger et al., 1966). Indeed, due to their repetitive nature, repeat units can be more easily misaligned or form hairpins, thus there is an increased probability of the introduction of errors during DNA replication by error prone polymerases (Tippin et al., 2004).

1.2 Mismatch repair

Repeat sequence instability is enhanced when the DNA mismatch repair (MMR) system is impaired. Indeed, in such circumstances, MMR proteins cannot efficiently correct replication errors. It has also been suggested recently that some MMR components are involved in regulating genetic recombination (Evans and Alani, 2000) and apoptosis (Young et al., 2004); thus MMR deficiency could promote repeat sequence instability (RSI). For instance, in the case of Hereditary Non-Polyposis Colorectal Cancer (HNPCC) syndrome heritable MMR protein defects are associated with mutator phenotypes giving rise to high RSI in cancer cells (Jiricny and Marra, 2003). Depending on which proteins are defective, different levels and types of RSI are observed (Baker et al., 1996; Prolla et al., 1998). Since the MMR system can have such an influential role in modulating genomic stability, it is crucial to understand MMR mechanisms and characterize MMR proteins. Briefly, in mammals MMR repair protein genes include two conserved gene families: MutL homologs (including MLH1-MLH3, PMS1, and PMS2) and MutS homologs (including MSH2-MSH6) (Aquilina and Bignami, 2001; Hsieh, 2001) that form complexes to function in DNA repair. Initially, either the heterodimer MutSα (MSH2/MSH6) or the heterodimer hMutSβ (MSH2/MSH3) bind to mismatched DNA, targeting base pair mismatches and 1-2 nucleotide insertion/deletion sites (MutSα) or longer insertion/deletion sites (MutSβ) (Hsieh, 2001). The next step for both complexes is to recruit a MutL heterodimer, either MutLα (MHL1/PMS2) or MutLβ (MHL1/PMS1) which enables activation of ATP hydrolysis (Aquilina and Bignami, 2001) leading to unwinding of the DNA by Helicase II, digestion of single-strand DNA at the damaged site by an exonuclease (usually Exo I), repair DNA synthesis by DNA

polymerase (usually DNA polymerase δ) and ligation by DNA ligase (Aquilina and Bignami, 2001; Hsieh, 2001) The proficiency of the MMR system is essential to ensure adequate protection of DNA integrity against environmental stress. As mentioned above, MMR proteins play an essential role in the detection and removal of DNA synthesis errors, consequently a loss or decrease in the function of these proteins can have catastrophic consequences, particularly in vulnerable repeat sequences that are hotspots for mutations. In addition, these proteins are not only involved in damage repair but also in signaling replication errors as well as apoptosis. Thus, preventing the removal of DNA synthesis errors allows mutant cells to remain in the population and accumulate mutations (Jiricny and Marra, 2003).

The MMR system can be indirectly disrupted. As demonstrated by Marra et al. (1998), modulating protein levels by overexpressing *MSH3* can elevate the level of MutSβ (MSH2/MSH3) being formed. In turn the relative level of the normally predominant MutSα (MSH2/MSH6) complex is decreased since MSH3 binds to the available MSH2, limiting binding by MSH6, thus leading to degradation of the monomer MSH6 protein. This decrease in MutSα levels functionally translates into a decrease in the ability of the MMR system to correct base pair mismatches and 1-2 bp insertions/deletions. Another example is silencing of *MHL*1 by promoter methylation indirectly leading to inactivation of repeat sequence rich genes such as *TGFβIIR* and *ECF*4 (Herman et al., 1998). Mismatch repair can also be inhibited through suppression of *MSH*2 expression. For instance, in humans, overexpression of apoptosis regulator Bcl-2 can decrease the transcriptional activity of the transcription factor E2F and in turn downregulate *MSH*2 expression (Youn et al., 2005). In addition, several MMR genes

contain tandem repeats making them also potentially vulnerable to spontaneous and stress-induced frameshift mutations (Jiricny and Marra, 2003).

1.3 Environmental stress exposure, tandem repeat sequences and health concerns

There is continuing concern regarding the risks and effects of anthropogenic pollutants on both human health and wildlife. Daily, millions of people are exposed to various anthropogenic contaminants that are potential mutagens and thus pose a genetic threat to both somatic and germline cells. Indeed, potentially hazardous substances are commonly found in polluted environments, consumer products, as well as in medical treatments. Exposure to these sources is dependent on work, diet, lifestyle and other factors and can lead to increased cancer risk as well as heritable genetic disorders, Specifically, somatic instability can lead to increased cancer risk (Boland and Goel, 2005; Jeong et al., 2003), and possibly neurological diseases (Hunter et al., 2005; Pamphlett, 2004) and may be associated with atherosclerosis (Inafuku et al., 2004).

Advances in understanding the effects of contaminants on human health as well as better access to information has increased public and scientific awareness of potential health risks associated with chemical exposure. In addition, the fact that new chemicals for industrial and consumer use are constantly being developed has made risk assessment a great priority. Indeed, in Canada, under CEPA, 1999, Health Canada must assess the public safety risk for over 1000 novel chemicals and polymers every year (Health-Canada., 2004a) in addition to evaluating health risk from approximately 23,000 existing substances and byproducts identified under CEPA,1988. Currently, humans are constantly exposed to potential mutagenic substances. For instance, polycyclic aromatic hydrocarbons (PAH's) produced by industry or by vehicles burning carbon-based fuels

are components of urban air (Kyrtopoulos et al., 2001). In addition, modern agriculture often depends on an arsenal of chemical pesticides and fertilizers, which can pose a threat if they enter water supplies or persist in the soil, if traces are left in foods, or if people come in direct contact with these substances (via air-born particles, handling, etc.). Potential threats can also come from the workplace (e.g. certain inks, metal particles, radiation and many others), from medical sources (e.g. chemotherapy), from pharmaceuticals (e.g. psolarens) and even in our diet (e.g. heterocyclic amines) and lifestyle, e.g., uv exposure, smoking, hair colouring, etc. There are other potential mutagenic risks from naturally occurring biological agents such as fungal toxins (aflatoxin from some *Aspergillis* sp), viruses with viral DNA that acts as a direct mutagen (human papiloma virus and hepatitis virus), and bacterial activity and toxins.

To improve regulatory abilities and risk assessment of novel and existing substances, it is essential to promote the development of novel mutation detection assays as well as identification of biomarkers for contaminant exposure to deduce and decrease the risks to human health. Indeed, although current regulatory tests are effective to measure specific end-points, a full array of regulatory tests is essential to fully evaluate the array of genetic changes that can be induced by exposure to a given chemical or substances. In particular, developing and validating tools to investigate mutations at TRS is essential for two main reasons. First, they were experimentally shown to represent sensitive biomarkers of exposure to damaging agent (Yauk, 2004). Notably, early studies carried out by Dubrova et al. (Dubrova et al., 1993) demonstrated that ionizing radiation could produce heritable sequence alterations at repeat sequences and thus are valid markers of exposure. Additional studies on humans (Kyrtopoulos et al., 2001), sentinel

animals (Somers et al., 2002) and cell cultures (Don Porto Carero et al., 2001) have provided further support for these observations. Second, recent research has demonstrated an association between repeat sequences instability and neurological diseases and disorders (Sun and Han, 2004), cancer (Mao et al., 1994), cholesterolaemia (Hubacek et al., 1999) and atherosclerosis (McCaffrey et al., 1997; Miniati et al., 2001)

1.4 Current methods for the detection of genotoxic agents

Establishing the genotoxic hazard of any given substance using just a few tests can be problematic since there are various mechanisms that can produce different types of genetic alteration. Indeed, genotoxity includes any deleterious genetic alteration from large scale chromosomal aberrations down to small point mutations. Some genotoxic agents are bioactivated, for example, some polycyclic aromatic hydrocarbons (PAH) present a danger when metabolized into active species that react with DNA to form premutagenic lesions. Other types of mutagenic agents, to mention only a few categories, include base analogs such as 5-bromouracil that when incorporated into DNA cause mispairing and base substitutions. Agents such as ethidium bromide that intercalate between bases, cause single base pair insertion or deletion events that lead to frameshifts. Substitutions can result from the addition of alkyl groups to bases by alkylating agents such as nitrogen mustard.

Methodologies have been developed and used to identify and characterize the genotoxicity of chemicals and radiation. By using different systems and tests, various toxicological endpoints can be targeted for assessment. The most common methods are based upon the direct detection of mutagenic events, including DNA sequence alterations, in coding and non-coding regions of the genome, DNA breakage, visible chromosomal

aberrations as well as functional assays. These can be divided into *in vitro* systems, such as bacteria, yeast cells, mammalian cells in culture and *in vivo* systems, such as transgenic and non-transgenic rodent assays that are commonly used in mutagenesis research. *In vivo* systems, particularly mammalian systems, are very useful in directly reflecting the effects of a mutagen on selected tissues and the intact animal. But for the initial screening of substances *in vitro* tests have proven to be very useful, ethical and less costly.

According to Health Canada (Health-Canada, 2004b) guidelines for genotoxicity testing, there are four categories of acceptable screening assays: 1/ bacterial tests, 2/ and 3/ are *in vitro* and *in vivo* tests measuring chromosomal effects in mammalian systems and 4/ i*n vitro* gene mutation assays in mammalian cells. All testing must be done according to internationally approved standards as defined by the OECD. The first category of genotoxicity assays for chemical regulation include functional *in vitro* tests such as the bacterial reversion assay/*Ames* test (Ames, 1973; OECD, 1997a). These evaluate mutation frequencies based on the observed frequency of reverse or forward mutation (restoration or loss of a function in a particular target) in bacterial systems such *Salmonella typhimurium* or *Escherichia coli*. The second category encompasses cytogenetic assays that examine chromosomal damage (breakage and abnormalities, presence of nuclear fragments) *in vitro* (OECD, 1997e). These are generally done in rodent cell lines that are grown and exposed to chemicals. Then, at the first division post-treatment metaphase chromosomes can be counted for aneuploidy and any aberrations observed. An *in vitro* micronuclei assay that examines aberrations during interphase, is pending OECD approval (Andreoli et al., 2003). Fluorescence *in situ* hybridization

(FISH) is another cytogenetic method that detects chromosomal damage, this time via hybridization to a DNA probe (Attia et al., 2003). Similarly, tests in the third category, examine chromosomal effects but *in vivo*. These include erythrocyte micronuclei assay (OECD, 1997c), in which bone marrow cells from rats or mouse (Trosic et al., 2004) are examined after chemical exposure to measure chromosome/mitotic spindle damage. The other *in vivo* regulatory test to measure chromosomal effects is the bone marrow chromosome aberration test (OECD, 1997d; e.g., (Siddique and Afzal, 2005). For the fourth category, *in vitro* gene mutation assays (OECD, 1997e), the typical methodology involves measuring the number of mutant cells that are resistant to a substrate metabolized by a gene of interest, e.g., trifluothymidine by thymidine kinase (Chang et al., 2002) after exposure to a test chemical. Normal cells have a functional gene, thus, the substrate is toxic for these cells. When the gene is inactivated by mutations from exposure to a mutagen, the mutants are conferred resistance to the substrate that they cannot metabolized.

There are several useful research tools that are not yet included in international regulatory guidelines including the comet assay test for DNA damage in a single cell. Briefly, cells are mixed with agarose, and the DNA from lysed cells is electrophoresed through the agarose on a microscope slide. The DNA is then stained after the short run, giving comet-like impression where the material that migrates farther represents damaged DNA (Garry et al., 2003) Also, DNA fingerprinting, Amplified Fragment Length Polymorphism (AFLP) and Randomly Amplified Polymorphic DNA (RAPD) technologies have proven very useful in detecting mutations from exposures to physical and chemical agents. (Bagley et al., 2001; Somers et al., 2002). These methods are based

on visualizing band mutations from PCR-amplified or restricted DNA on agarose or acrylamide gels. Ideally, these *in vitro* assays would all be quick, economical, sensitive, ethical, accurate and representative of human responses.

To investigate induced TRS instability, Yauk et al. (2002) developed a novel single molecule PCR assay that can detect tandem repeat mutations in non-coding loci. Based on the amplification of Ms6-hm1 (Hm1), a mouse ESTR sequence, it was shown that non-coding loci are very sensitive to mutagenic stresses and could represent a useful genotoxicity monitoring system. This assay can provide statistically robust estimates of mutation frequencies and decrease the need for animal sacrifice in germline mutation studies (Yauk et al., 2002). However, this assay has the disadvantage of being technically challenging and difficult to reproduce with reaction conditions needing to be re-optimized for each laboratory or new batch of PCR reagents. In addition, the sensitive single molecule PCR system detects such small length differences that PCR artifacts might potentially affect results.

A novel assay system uses a reverse mutation assay that detects microsatellite instability, deletions and insertions that lead to frameshifts that restore the proper reading frame of a reporter gene (Eckert and Hile, 1998; Eckert et al., 2002; Slebos et al., 2002). Detection of mutants can take advantage of antibiotic resistance with cells harboring constructs with out-of frame genes coding for antibiotic resistance. In that case, mutants with restored reading frames can grow on media containing the corresponding antibiotic for selection, and mutation frequencies can be calculated based on comparisons to control plates. The efficiency of this type of assay has been improved by the use of more easily detectable reporter proteins.

1.5 GFP reporter systems

Since its isolation from the jellyfish *Aequorea victoria*, green fluorescent protein (GFP) has been extensively used for studies of the cytoskeleton, and examining protein expression, localization and interaction. Gasche et al. (2001) from the Department of Medicine and Cancer Center at the University of California at San Diego developed a GFP fluorescence reversion assay to detect DNA damage in minisatellite sequences. Oligonucleotide sequences containing repeat dimer units were inserted upstream of a GFP sequence in a reporter plasmid vector, placing the gene sequence out of frame for protein translation. Frameshift mutations within the repeats can restore the reading frame; expression of GFP can be detected through flow cytometry and the mutants counted. An initial reporter system (Gasche et al., 2001), using transiently transfected cells, examined mutations in CA repeat sequences following exposure to oxidative stress from hydrogen peroxide. This transient assay system is limited by the short exposure times and the need to constantly normalize for transfection efficiency. A subsequent study (Gasche et al., 2001) used stably transfected MMR proficient and deficient cells, and examined mutagenesis in CA repeats. This work confirmed that MSI was higher in MMR deficient cells than in MMR proficient cells, but the study had looked at only a few transfected clones, that might have biased the results since a few transfected clones were unlikely to represent the entire cell population. An earlier study that took advantage of this type of GFP reporter construct was undertaken by Gaspar et al. (2000) who discovered that transcriptional or translational frameshifting in CAG repeats played a role in Machado-Joseph disease. Slebos et al. (2002) used a similar reporter system, inserting dinucleotide and tetranucleotide TRS into GFP reporter constructs, to test the impact of cigarette

smoking on MSI. Their study used populations of stably transfected cells that made the reporter study statistically more powerful (the sample size was larger as it included a pool of transfected clones) and biologically more meaningful.

Clearly, GFP reporter systems represent a relatively quick and powerful method for detecting frameshifts within repeat sequences. Here we propose to improve the GFP reporter assay by demonstrating its use as an efficient toxicological screening assay for a wide range of potential frameshift mutagens. As mentioned earlier, tests based on monitoring tandem repeat sequence variability such as DNA fingerprinting and single molecule PCR have shown that repetitive sequences are especially sensitive to a wide range of stresses (Parfett, 2003; Yauk et al., 2002). Although sensitive, these methods are labour intensive and time consuming. Indeed, although a wide variety mutation tests are currently used for regulatory testing, many are too slow, provide equivocal predictions or lack sensitivity to detect potential carcinogens or mutagens that act on specific DNA sequences, therefore mis-identifying harmful agents as non-genotoxic (Jackson et al., 1993). Clearly, GFP reporter system could provide a powerful tool for risk assessment.

The work in this thesis describes the construction of GFP reporter plasmids for toxicological screening and investigates their sensitivity and reproducibility in testing a battery of standard mutagens and a separate category of "non-genotoxic" carcinogens (hormone, diet, etc). The sensitivity of these assays is compared to the previously tested dinucleotide CA and tetranucleotide AAAG constructs described by Slebos et al. (2002). To identify what repeat sequences were the most sensitive biomarkers in the context of a GFP reporter construct, targets of 2-6bp repeat sequences were tested against a negative control construct of equal size, but with no repeats. A relevant human gene sequence

containing a repeat motif was also tested to examine whether more complex sequences could be used within the GFP reporter system as well as measure its sensitivity as a biomarker. This rapid *in vitro* assay could be another option for chemical screening in a regulatory framework. Here, two murine cell lines were selected: the C3H10T1/2 cell line used for various toxicological assays and the MC2a cell line which is MMR deficient.

1.6 The C3H10T1/2 cell line

The C3H10T1/2 murine cell line is well established in the literature as a model system for mutation and carcinogenesis research. Developed by Reznikoff et al. (Reznikoff et al., 1973b), it is a line of stable, pluripotent, embryonic C3H10T1/2 mouse cells from which clones are easily isolated and has an inherent susceptibility to postconfluence inhibition-meaning that cells grow less when they come in contact with each other (Taylor and Jones, 1979). Interestingly, PAH treatment induced malignant transformation of C3H10T1/2 cells (Reznikoff et al., 1973a). Similar results were obtained from treatments with monofunctional alkylating agents such as the direct-acting mutagen N-methyl-N'-nitro-N-nitrosoguanidine or MNNG (Bertram and Heidelberger, 1974). Other tested transforming agents include physical stresses (UV and X-rays), anti-tumour antibiotics, halogenated pyrimidines and pyrimidine nucleosides, cigarettes, and methotrexate to name a few (Boreiko et al., 1980; Fernandez et al., 1980; Mondal and Heidelberger, 1980). This cell line carries cytochrome P450 (Cyp1b1) which can activate PAHs so no additional extract is required to test PAHs in this murine system. (Shen et al., 1994). The two-stage cell transformation assay that involves induction and promotion of foci in C3H10T1/2 cells can even be used to detect some of the non-mutagenic carcinogens such

as dioxin and the tumour promoter 12-*O*-tetradecanoyl-phorbol-13-acetate (TPA) (Parfett, 2003)

1.7 MC2a cell line

The MC2a murine cell line, developed by Dr. Liskay's lab is a murine embryonic fibroblast cell that is mismatch repair deficient due to the absence of *MHL*1 gene. The MC2a cell line was established from muscle tissue from C57bl6/129SvEv mouse hybrid embryos of 15-17 days. In C57bl6/129SvEv, the Mlh1 gene was inactivated by a knock out protocol that used a hypoxanthine phosphoribosyltransferase (HPRT) minigene targeting vector derived from a construct used by Allan Bradley (Baker et al., 1996). This cell line has proven to be extremely useful because it is easily transfected and sensitive to our selection of chemical treatments. As mentioned earlier, the *MHL*1 protein is part of a MutL heterodimer that is involved in the activation of ATP hydrolysis during MMR (Hsieh, 2001). In MC2a cells, this deficiency translated into an increase in MSI for endogenous CA tandem repeats as measured by single molecule PCR at several loci as well as overall increased genomic instability as measured by increased mutation frequency in ouabain resistance assays (Baker et al., 1996). In addition, MMR deficient cells are expected to have a higher transfection efficiency and more suitable for stable transfection since they more readily integrate linear DNA than MMR proficient cells (Lin et al., 2001)

1.8 Test agents

To test the range of detection of the GFP reporter system, 5 different classes of damaging agents were employed: an intercalating agent, 9-amino-acridine (9AA); an oxidative stressor, hydrogen peroxide (H_2O_2); the polyaromatic hydrocarbon benzo-*a*-pyrene Diol Epoxide (BPDE); an alkylating agent, ethyl nitrosourea (ENU); and the tumour promoter, 12-O-tetradecanoyl-phorbol-13-acetate.

Since the GFP reporter assay is based on the detection of frameshift mutations, we used as a positive control 9AA, a recognized model frameshift mutagen in bacterial mutagen screening assays (Ferguson and Denny, 1991). The planar structure and dimensions of the 9AA molecule allow it to form a non-covalent bond with DNA by intercalating between bases (Neidle and Abraham, 1984). Based on Streisinger's slippage model, intercalation within repeat sequences enables single-strand slippage to occur where DNA unwinding has occurred (Streisinger et al., 1966). If not repaired, these errors remain as a duplication or a deletion of repeat units. A second mechanism that may account for acridine-induced frameshifting is through inhibition of topoisomerase II and stabilization of the cleavable DNA complex allowing erroneous additions and deletions to remain in the DNA (Masurekar et al., 1991). The third model proposed for 9AA mutagenicity suggests that after a double-strand break occurs, non-homologous end-joining can occur during mammalian cellular DNA repair thus leading to loss or gain of DNA fragments (Han et al., 1993).

Oxidative stress has been linked to premature aging (Dufour and Larsson, 2004) as well as cancer and atherosclerosis (Ross et al., 2001). The damaging oxyradicals are commonly generated from cellular metabolism (Marnett, 2000) but exogenous forms can

also be generated from cigarettes (Pryor et al., 1998) as well as airborne particulate matter (Squadrito et al., 2001) and other substances that be metabolized into reactive oxygen species. Endogenous hydrogen peroxide, as do other oxidative stressors, typically generates single-strand breaks, base damage and loss (Marnett, 2000) leading in some cases to frameshift mutations (Gasche et al., 2001).

Polyaromatic hydrocarbons (PAH), are found in urban air and are most commonly formed during incomplete combustion (IARC, 1983). One of the most notable PAHs is benzo-*a*-pyrene. It can be metabolized into benzo-*a*-pyrene diol epoxide (BPDE), a chemical that can bind to DNA, preferentially to purines, mostly guanine, thus forming bulky adducts (Boysen and Hecht, 2003). Frameshift mutations can arise when the bulky adducts enable misalignment of the two DNA strands to occur leading to slippage events (Zhuang et al., 2001).

Alkylating agents such as the monofunctional ENU generally act to transfer alkyl groups to DNA, most commonly yielding O^6-alkylguanine or in the case of ENU, O^6-ethylguanine (Harfe and Jinks-Robertson, 2000). These alkyl lesions can result in base substitutions but also frameshifting (Eckert and Hile, 1998; Slebos et al., 2002). The mechanisms associated with alkylation-induced frameshifting are still unclear; however, based on results from *in vitro* DNA synthesis with α and β polymerase, Ecker and Hile (1998) have proposed that the adducts formed by ENU treatments promote mispairing thus misalignment that allow slippage to occur during synthesis. Their results also showed that in the test tube, heteropolymeric DNA and monomeric repeat stretches of DNA are both prone to frameshifting with the type of DNA polymerase, in this case α and β, being the determining factor of mutational specificity. However, in the context of

reporter vectors transfected into mammalian cells, alkylation-induced frameshifting is more frequent in dimers and tetranucleotide repeat stretches than in heteropolymeric DNA (Slebos et al., 2002). This could mean that other endogenous cellular components such MMR proteins could also be involved in modulating mutational specificity.

Phorbol ester TPA is a tumour promoter that acts to modulate levels of cellular protein kinase C (PKC) by stimulating the TPA-responsive PKC responsive isoforms, which in turn triggers PKC proteolytic degradation and thus depletion of PKC levels in cells after extended exposure to TPA (Lu et al., 1998). Results from recent mutational studies (Parfett and Healy, submitted) suggest that in addition to promoting tumour formation (Lu et al., 1998), TPA might also indirectly induce mutations. The mechanisms involved in putative TPA mutagenicity are unclear; however, work by Humbert et al. (2002; 2003) demonstrated that TPA exposure can modulate expression of MMR protein MSH2. They proposed that the TPA-induced PKC activation is responsible for overexpression of *MSH2* as PKC activation leads to stimulation of transcriptional factor AP-1 that can bind to motifs in the promoter region of *MSH2.* In addition, PKC plays a role in MSH2 and MSH6 phosphorylation which increases binding of MutSα complexes formed by MSH6 and MSH2 in response to damaged DNA (Christmann et al., 2002). Interestingly, as pointed out by Humbert et al (2003) this type of AP-1 induction and increased PKC activity (Christmann et al., 2002) as triggered by TPA are typical cellular responses to genotoxic stress. Also, as mentioned above, after extended exposure to TPA, PKC depletion occurs. This secondary event can lead to impairment of the MMR system since decreased PKC levels can reduce MSH6 and MSH2 phosphorylation and therefore decrease MMR binding by MutSα (Christmann et al., 2002) making the cell more

vulnerable to endogenous oxidative stress from cellular metabolism and other exogenous genotoxic stresses. Two potential mechanisms could be involved in increased mutation levels induced by TPA: the indirect effects of TPA on mammalian MMR system promote fixation of mutations, or TPA simply promotes the growth of "aggressive" cells that are inherently more genetically unstable (Hornia et al., 1999).

1.9 Objectives and hypotheses.

1.9.1 Objectives:

The main objective of this thesis was to develop a quick, reliable *in vitro* fluorescence assay in mammalian cells to detect DNA sequence alterations from environmental mutagen exposure.

1.9.2 Underlying hypotheses:

1. It is believed that in somatic cells, the loss or gain of complete repeat units through DNA slippage makes tandem repeat sequences often more mutable than non repeat sequences. However it is not the only mechanism associated with repeat sequence instability. To a lesser extent imperfect gain or loss of units can occur, as well as rare mitotic recombination events, non-homologous end joining and other erroneous repairs. Thus we hypothesize that frame shift mutations can be detected in trinucleotide and hexanucleotide TRS.
2. Chemicals previously defined as "non-mutagenic" or carcinogenic, can potentially be mutagenic though indirect receptor-based mechanisms that generate mutagens by altering cell metabolism or MMR protein function. We hypothesize that these mutations can be detected by a sensitive GFP-based assay.

Chapter 2

2.0 Materials and Methods

2.1 Chemicals and reagents

Twelve-*O*-tetradecanoyl-phorbol-13-acetate (TPA; P-8139; CAS number16561-29-8), hydrogen peroxide (H-904; CAS number 7722-84-1), 9 amino acridine (9AA; A38401; CAS number 52417-22-8), dimethylsulfoxide (DMSO; D-2650; CAS number D-2650), N-Ethyl-N-nitrosourea (ENU; N8509; CAS number 759-73-9), tetrahydrofuran (40,175-7; CAS number 109-99-9) and acetone (179124; CAS number 67-64-1) were purchased from Sigma Aldrich (Oakville, ON) and benzo-*a*-pyrene-diol epoxide (477; CAS number 58917-67-2) from Midwest Research Institute (Kansas City, MO). Materials for cell culturing and harvesting, including Dulbecco's modified Eagle Medium (DMEM), fetal bovine serum (FBS), L-glutamine, penicillin/streptomycin sulphate, minimal essential medium (MEM) non-essential amino acids, gentamicin sulfate, Geneticin, Phosphate-Buffered Saline (PBS; pH 7.2 1.54 mM KH_2PO_4, 155.17 mM NaCl, 2.71 mM Na_2HPO_4-$7H_2O$), 0.25% trypsin and proteinase K were obtained from Gibco-Invitrogen (Burlington, ON). Seakem electrophoresis grade agarose was purchased from Mandel Scientific Company (Guelph, ON). All restriction enzymes and linkers were purchased from New England Biolabs (Missisauga, ON). The vectors pUC19 and pEGFP-N1 were obtained from BD Biosciences (San Jose, CA). Fugene transfection reagent was purchased from Roche Diagnostics (Laval, QC). All sequencing and PCR primers were ordered from Cortec (Kingston, ON) and all oligonucleotide inserts were ordered from Sigma Genosys (Oakville, ON). All other chemicals were purchased from Fisher Scientific (Nepean, ON) and were of analytical grade or higher.

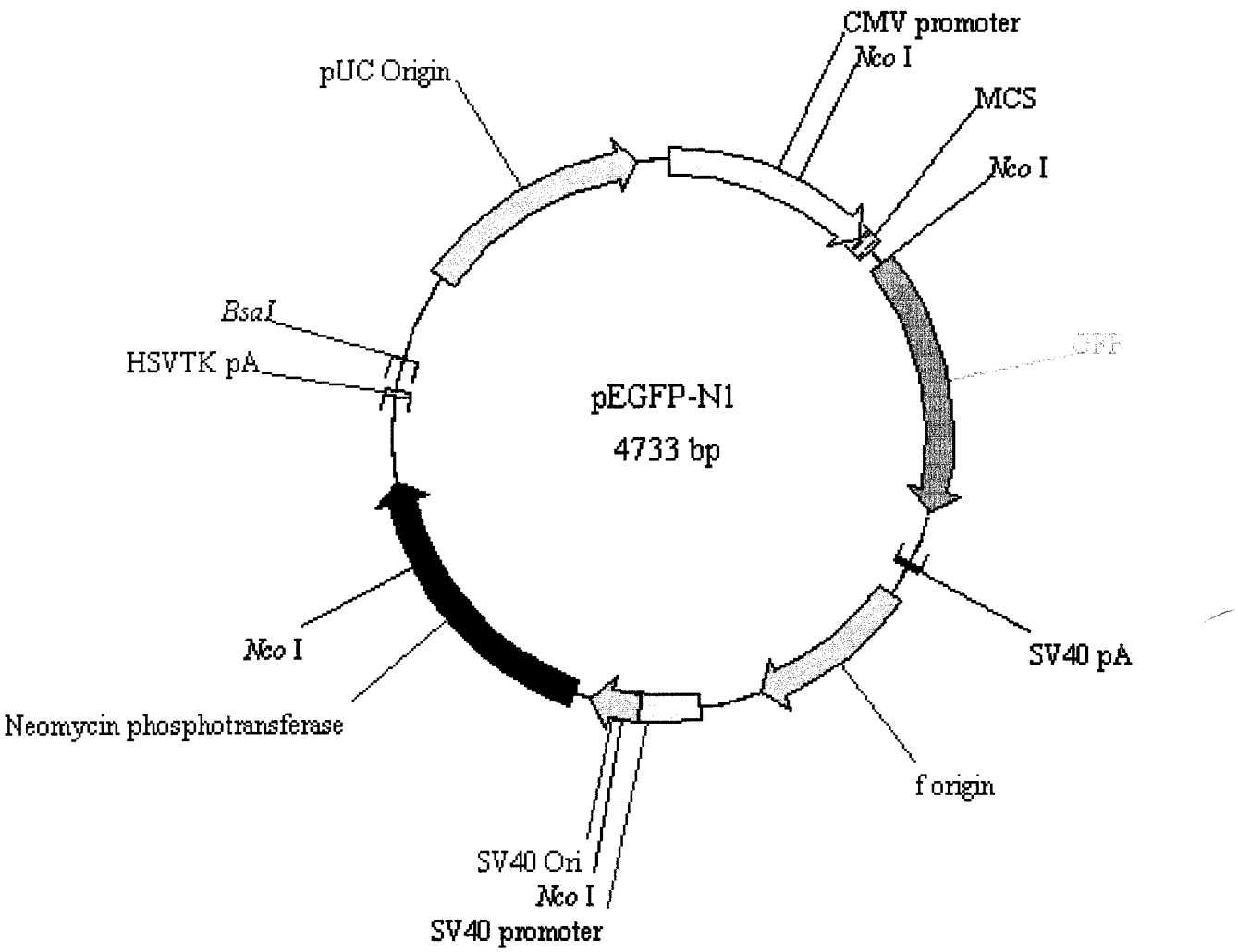
2.2 Plasmid constructs

2.2.1. Generating the initial plasmid

The p-EGFP-N1 plasmid (Figure 1) was chosen as the starting material because it encodes a red-shifted green fluorescent protein (GFP) that can be detected by flow cytometry and has amino acid codon sequence usage optimized for transfection of mammalian cells (BD. Biosciences Clontech, 2004). A flow diagram summarizes the construct (Appendix A). Briefly, to insert repeat sequences upstream of the enhanced GFP (EGFP) coding sequence, a novel restriction site was introduced by the insertion of an *Mlu* 8-mer linker d(GACGCGTC) into the *Nco*I site 5' to the EGFP gene. Since p-EGFP-N1 has 4 *Nco* I sites, this cannot be done directly. Thus the first step in the cloning procedure was to insert the *Pst* I-*Xba* I fragment of pEGFP-N1 containing the targeted *Nco* I site into pUC19. This enabled manipulation of sequences at the single *Nco* I site within the *Pst* I-*Xba* I fragment.

The plasmid pEGFP-N1 was first isolated from transformed *E. coli* JM110 *(dam)* to ensure cutting at the methylation-sensitive *Xba* I site. All plasmid DNA was isolated using Quiagen's mini or midi kits to obtain optimal purity. The isolated DNA was then cut with *Pst* I plus *Xba* I resulting in two fragments that were gel purified using GFX gel band purification kit (GE Healthcare; Piscataway, NJ). The larger band (NL) was saved and the smaller band cloned into pUC19. The resulting plasmid (pUC19GFP) was grown in RRI, a strain with high transformation efficiency (Sambrook et al. 1989).

Figure 1. Highlight of the main features of Clontech's EGFP-N1 plasmid used in the plasmid contructs. Features include a gene encoding a mammalian optimized sequence of red shifted green fluorescent protein (GFP) using the immediate early promoter of human cytomegalovirus (CMV), SV40 polyadenylation signals (SV40 pA) for GFP mRNA processing, key enzyme sites used for plasmid construct (*Nco I*, *Xba I* and *Pst I*) and plasmid linearization for transfection (*Bsa I*), for selection neomycin/kanamycin resistance gene (neomycin phosphotransferase) using the SV40 early promoter (SV40 promoter) and for mRNA processing Herpes simplex thymidine kinase gene polyadenylation signals (HSVTK pA) for G418 resistance in the mouse cells, and using bacterial promoter upstream of the neomycin phosphotransferase gene (not shown) as well as pUC19 origin of replication (pUC Origin) to propagate in transformed *E. coli.*



Adapted from Clontech's pEGFP-N1 vector (BD Biosciences Clontech 2004)

The plasmid was restricted at the unique *Nco* I site, blunt ended by a Klenow fill in reaction and the *Mlu* I linker was ligated into the site. Following transformation, plasmids containing the *Mlu* I linker were identified by restriction with *Mlu* I since the initial pUC19 vector does not have an *Mlu* I site. To prevent "translational leakage", i.e., possible translation initiating from the two inframe ATGs, the second methionine codon ATG, a potential initiation codon generated from the Klenow reaction, was changed into the leucine codon CTG (Kunz et al., 1992) by site directed mutagenesis using Stratagene's QuickChange XL site-directed mutagenesis kit. Briefly, the modified pUC19 vector template was used for mutagenesis along with two synthetic oligonucleotide primers from Cortec that contain the desired mutation (5'-GGTCGCCACCCTGCTGAGCAAGGGC-3'and 5'GCCCTTGCTCACCAGGGTGGCGACC-3').

The new plasmid was transformed into XL10-Gold ultracompetent cells (Statagene). This strain is recombination deficient (*rec*A) improving the stability of the plasmid sequence. The resulting plasmid was restricted with *Pst* I plus *Xba* I and the purified fragment containing the linker and EGFP sequences was ligated back into the NL fragment previously purified from pEGFP-N1 to give plasmid "NMC-E". The sequence of the construct was confirmed using a CMV primer 5'-TTGACGCAAATGGGCGGTAG-3' (position 516-536 of EGFP-N1) located upstream to the polylinker site (Cortec, Kingston, Ontario).

Figure 2. Cloning strategy for the chimeric GFP reporter gene constructs. Sequences were inserted downstream of the multiple cloning site (MCS) of Clontech's EGFP-N1 plasmid, immediately following a new ATG start codon of the GFP gene to render the coding sequence out of frame. To prevent translation leakage (false positives) the original ATG codon was changed into CTG. The human cytomegalovirus (CMV) promoter enhances transcription. When frameshift mutations occur, the reading frame is restored and the cell can produce a functional protein, giving a strong fluorescent green signal detectable by fluorescent microscopy and flow cytometry.

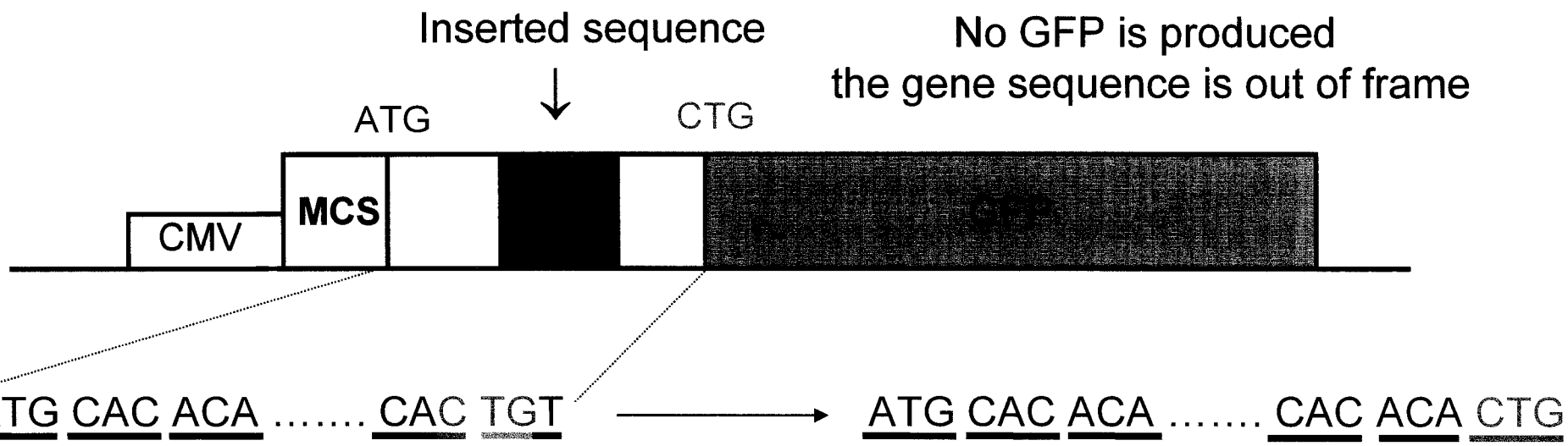

Inserted sequence
No GFP is produced
the gene sequence is out of frame
ATG
CTG
CMV
MCS
ATG CAC ACA ..... CAC TGT
ATG CAC ACA ..... CAC ACA CTG
Example of an out of frame sequence:
inserted sequence (CA)12
gives non functional GFP
Exposure to mutagens
can lead to
frameshift mutations
Example of a mutated in frame sequence:
inserted sequence (CA)12
a + 1 frameshift gives functional GFP

Table 1. List of plasmids used in this study.

Plasmid name	Plasmid description
CA	With GFP in –1 reading frame, contains dinucleotide $(CA)_{12}$
Trio	With GFP in –1 reading frame, contains trinucleotide $(CAG)_{12}$
AAAG	With GFP in –1 reading frame, contains tetranucleotide $(AAAG)_{12}$
Hm1	With GFP in –1 reading frame, contains pentanucleotide $(GGGCA)_{12}$
Per	With GFP in –1 reading frame, contains hexanucleotide $(ACAGGC)_{12}$
Hexa	With GFP in –1 reading frame, contains hexanucleotide $(ACCTCC)_{12}$
Hucy	With GFP in –1 reading frame, contains gene sequence for human aromatase with (TTTA) TRS
NLA	With GFP in –1 reading frame, contains a heteropolymeric insert
Posi	With GFP in correct reading frame, contains a heteropolymeric insert

2.2.2. Generating vectors with inserts

Vectors were designed to allow the detection of +1 frameshifts (Figure 2) in both negative control and repeat sequence constructs (see Table 1 and for a list of oligonucleotides see Appendix B). Oligonucleotides containing a repeat sequence were synthesized to include *Mlu* restriction-site overhangs flanking the repeat. To facilitate clone screening, all oligonucleotides also contained the sequence CGAT^CG, the restriction site for *Pvu* I, that is absent in the original vector. After transformation of XL10-Gold ultracompetent cells and selection on LB/kanamycin medium, plasmid DNA isolated from individual clones was restricted with *Pvu* I and electrophoresed in 0.8% agarose in 1 X TBE buffer (5 X TBE is 0.45M Tris, 0.45M boric acid, 10mM EDTA, pH 8.0) at 3 V/cm for 1 hr. Samples that had the restriction site were sent to Cortec to confirm the sequence of the constructs using an anti-sense EGFP primer (5'-GCAGATGAACTTCAGGGTC-3'; sequence 807-826 of plasmid EGFP-N1).

2.3 Cell culture

Pluripotent murine C3H10T1/2 cells (ATCC 226) were obtained from the American Type Culture Collection (Manassa, VA). Stock cultures were maintained in 150 mm tissue-culture dishes at 37 °C in high glucose DMEM, supplemented with 100 U/mL penicillin, 100 μg/mL streptomycin sulfate, and 10% (v/v) FBS (heat-inactivated at 56 °C for 30 min) in a humidified atmosphere containing 5% CO_2. The MMR deficient murine embryonic fibroblast cell line MC2a was generously given to us by M. Liskay (Molecular & Medical Genetics department at Oregon Health and Science University Portland, OR). Cells were maintained in low glucose DMEM enriched with 10 μM MEM non-essential amino acids, 50 μg/mL gentamicin sulfate and 10% FBS.

2.4 Generation of stably transfected populations

Sixteen hours prior to transfection, cells were seeded at $1.3x10^5$ cells per 35 mm dish in 2 mL of appropriate media for each cell line. Cells were transfected with 2 ug of plasmid DNA, previously linearised with *Bsa I* (see Figure 1) to increase the number of stable transfectants, using Fugene 6, a liposomal transfection reagent, according to the manufacturers protocol. Twenty four hours after transfection, cells were re-plated into 150mm dishes and grown in DMEM with 10% FBS and 1mg/mL or 0.5mg/mL of geneticin for selection of stable MC2a and C3H10T1/2 transfectants respectively. To decrease potential bias in GFP expression and TRS stability generated from clone to clone variation and small sample size, stably transfected populations were established for each construct by pooling over 100 individual cell colonies following the 1-2 weeks of antibiotic selection so each population is composed of over 100 founding clones.

2.5. RNA isolation and Northern blotting

To ensure that all cell populations had the same potential to generate GFP mutants, RNA levels expressed in each population arising from the different GFP constructs were compared. RNA from control, non-transfected cells and RNA from cells expressing GFP were used as negative and positive controls, respectively. RNA was isolated using the acidic phenol method (Edwards and Denhardt, 1985). For each transfected populations, total RNA was extracted from two 150 mm plates. Cells were scraped off the dish, resuspended in 3 mL of PBS containing 5.5 mM glucose and 0.1% EDTA. Following centrifugation at 400 x g and suspension in TSM (30mM Tris HCL pH 8.0, 150mM NaCl, 1.5 mM $MgCl_2$) with 1 mM vanadyl ribonucleosides, 4U RNAse inhibitor/mL TSN and 0.5% NP40, the cells were incubated on ice for 15min.

The solution was centrifuged for 5 min at 400xg, supernatant transferred to a fresh tube and 2.4 mL of urea lysis buffer (TUNES; 10mM Tris-HCl pH8, 7M Urea, 0.35 M NaCl, 1mM EDTA, 2% SDS) was added to complete the lysis. From this solution, RNA was extracted with phenol/chloroform, ethanol precipitated twice and suspended in TE buffer (10mM Tris pH 7.5, 1mM EDTA). RNA samples were heat denatured for 5 min at 65°C in 1 X RNA loading buffer (3.3X RNA loading buffer is 0.9% bromophenol blue, 10% Ficoll 400, 3.3 x MOPS) and chilled on ice. Samples (10 μg per lane as determined by UV absorbance) were electrophoresed on 1.0% agarose in 1 X MOPS denaturing buffer (1X MOPS, 5.6% formaldehyde) at 3 V/cm for 2 hrs (Parfett and Pilon, 1993). RNA was blotted onto a Zeta Probe® membrane at 45 mbar for 1 hour with a Vacugene unit (GE Healthcare) according to the manufacturer's recommendations. Membranes were washed in 5 X SSC (1 X SSC, 150 mM NaCl and 15 mM sodium citrate) and air-dried. RNA was cross-linked to the membrane using the BIO RAD UV chamber program C-L (125 mJ).

The heat-denatured GFP probe (PstI-Xba restricted fragment from pEGFP-N1 mentioned above, sequence 639-1412) and loading control poly(A) probe were labelled with [α-^{32}P] dCTP using Rediprime II DNA Labelling system and purified using ProbeQuant G-50 micro clean-up columns (GE Healthcare). Northern blots were hybridized 24 hrs at 55°C in 0.5M $NaHPO_4$, pH7.2, 1mM EDTA, 8%SDS, 5μg/mL polyA, 1mM ATP, 100 μg/mL salmon sperm DNA and 0.2% each BSA, Ficoll 400 and polyvinylpyrolidone. Blots were washed at 55°C for 15 min each in, sequentially, 1×SSC; 0.2×SSC; 0.1×SSC all with 0.1% SDS. Washed blots were wrapped in Saran Wrap and exposed to Kodak storage phosphor screens for 72 hours. The screens were

scanned using a Molecular Dynamic Storm Phosphoimager and bands quantified using Image Quant (Molecular Dynamics, Sunnyvale, CA).

2.6 Exposure of cells to test agents

To determine the range of sensitivity of the GFP reporter system, five standard hazardous test agents were examined, each at 3-4 concentrations selected on the basis of pre-evaluated killing curves (see below for description). Routinely, treatments comprised control (vehicle/solvent only), three concentrations below the 50% growth inhibition point and a fourth concentration that was slightly over the 50% growth inhibition point but still allowed reliable GFP counts after the 48 hr recovery period. For all treatments, cells were plated at 10^5 cells per 100mm dish in 10 mL of appropriate media plus Geneticin as mentioned above, 24 hours prior to exposures. For each concentration point, triplicate plates were seeded and two independent experiments were done. The tumour promoter twelve-*O*-tetradecanoyl-phorbol-13-acetate (TPA) was dissolved in DMSO and added to cultures once a day at 0, 100, 325, 650 and 1,325nM for 48 hours in DMEM (final concentration of DMSO in the medium was less than 0.2% for all treatments). For oxidative stressor hydrogen peroxide (H_2O_2), cells were exposed at 0. 0.001, 0.01, 0.1 and 1mM H_2O_2 in PBS for 1 hour followed by removal of treatment and addition of media. Benzo-*a*-pyrene diol epoxide (BPDE) was dissolved in the non-mutagenic solvent tetrahydrofuran (THF) and, all manipulations of BPDE-exposed cells were carried out in low light to minimize photolysis of BPDE. Cells were exposed to 0, 0.5, 1, 2 and 3μM of BPDE in PBS (the final concentration of THF was 0.2%) for 1 hour followed by aspiration of PBS and addition of fresh media. Cells were exposed to the alkylating agent ethyl nitrosourea (ENU) in Sorensen's buffer (pH 6.8 0.1 M Na_2HPO_4 and 0.1 M

KH_2PO_4; final concentration of acetone, used as initial solvent, was 0.075%) at 0, 1.07, 2.13, 4.27 and 6.40 mM for 6 hours followed by aspiration of treatment and the addition of fresh media. Cells were treated with nucleic acid intercalating agent 9-amino-acridine (9AA) at 0, 2, 8 and 16 μM in media for 4 hours followed by removal of treatment medium and addition of fresh media. All experiments were performed twice, with treatments in triplicate and cells allowed to recover 48hrs after treatment. All agents were tested on MC2a cells and as a point of comparison, 9AA and H_2O_2 treatments were tested on C3H10T1/2 cells.

Killing curves were generated from treating the cells as described above, except for that a larger range of doses was tested and 6 plates were used for each exposure point. After the 48 hour recovery, cell counts were obtained using a Coulter Counter®, fold increase corresponds to the ratio of cell count after recovery and cell plated. For example, a growth increase of 2 fold is a doubling of the cell population, and a fold increase below one such as 0.5 is a decrease by half of the population (loss of cells eg. from toxic treatment). For growth inhibition curves, exposure concentrations were 0, 100, 325, 650, 1300 and 2600 nM for TPA; 0, 0.001, 0.01, 0.1, 1 and 10 mM for H_2O_2; 0, 1, 2, 3, 4.5 and 6μM for BPDE; 0, 1.07, 2.13, 4.27, 6.40 mM for ENU and 0, 2, 8,16, 32 and 64 μM for 9AA..

2.7. Flow cytometry

Treated cells were washed with PBS, detached by treatment with PBS containing 5.5 mM glucose, 0.1% EDTA and 0.25% trypsin, then collected by centrifugation x 400g. The pellets were either suspended in 0.5 mL PBS for immediate analysis or suspended in media with 10% DMSO and stored at -20°C for later analysis. Prior to analysis frozen cells were thawed, collected by centrifugation x 400g and suspended in PBS. Cells were analysed using a FACSCalibur flow cytometer (BD Biosciences, Mississauga, ON) fitted with an argon ion laser (488 nm line excitation) and fluorescence emission of EGFP was reflected by a 550 dichroic long-pass filter and quantified after passage through a 530/30 nm bandpass filter. Data was analyzed using WinList 3D (Verity Software House, Topsham, ME). To only include viable cells and exclude dead/apoptotic cells, cellular debris, dust and particles, forward angle light scatter (FSC) was plotted versus side-angle light scatter (SSC). The FSC provided information on the relative size of particles and SSC in indicated granularity or porosity of the surface of the particles. Population gating is commonly performed for flow cytometric analyses. In this case, gating the viable cells is biologically relevant for this assay since we are interested in measuring mutations in live cells rather than in the dead ones. However, it should be noted that there was no significant difference between trends measured in gated and ungated populations.

To determine the frequency of green fluorescent cells in the cell samples, data was displayed on the green fluorescence (FL1-H) versus red fluorescence (FL2-H) axis to optimize clear detection of GFP mutants and minimize false positives (Pierce et al., 1999; Slebos and Taylor, 2001). Indeed, as seen on Figure 3 plotting FL1-H against FL2-H

permits discrimination between green and yellow fluorescence that cannot be distinguished when using solely the FL1-H parameter.

Using this methodology, calibration was done for each set of experiment and chemical treatment based on the distribution of a clonal population of cells highly expressing GFP to identify true positives (Figure 3A). The distribution of untransfected cells was used to measure autofluorescence, potential fluorescence from the treatment and minimize the number of false positives (Figure 3B). To ensure both positives and negative cells could be clearly distinguished, a mixed population was analysed for calibration (Figure 3C). In all figures, region R2 designates the cells counted as fluorescent green distributed along the FL1-H and FL2-H axes, while region R3 defines cells with general fluorescence (some green some yellow) along the FL1-H and FSC-H axes. As an example, as seen on Figure 3B, untransfected cells form a natural diagonal along the red and green axes and no unstransfected cell is found within the R2 region. Note however that some cells do emit yellow fluorescence as seen in region R3, particularly when exposed to high concentrations of H_2O_2, thus the need to use the FL2-H parameter to avoid false positives. As seen on Figure 3A (cells treated with a non toxic dose of 9AA) and 3B (untreated mixed population), green cells are recognized as deviating from this diagonal, forming a separate diagonal shifted towards FL1-H and are detected in both R2 and R3 regions. Based on this data, to determine true positives and limit the number of false positives, the results from counts in the R2 regions were used to obtain accurate green fluorescent cell counts. This approach was used to analyze all data (sample analysis Figure 4). In a typical analysis, counts of 10,000 cells were obtained for each sample.

Figure 3A. Characterization of GFP positive cells by flow cytometry. Plot showing the distribution of stably transfected MC2a population expressing GFP, after exposure to 2μM 9AA (low cytotoxicity), along the green fluorescence (FL1-H) vs red fluorescence (FL2-H) axes, FSC-H and SSC-H axes and FSC-H and FL1-H axes.

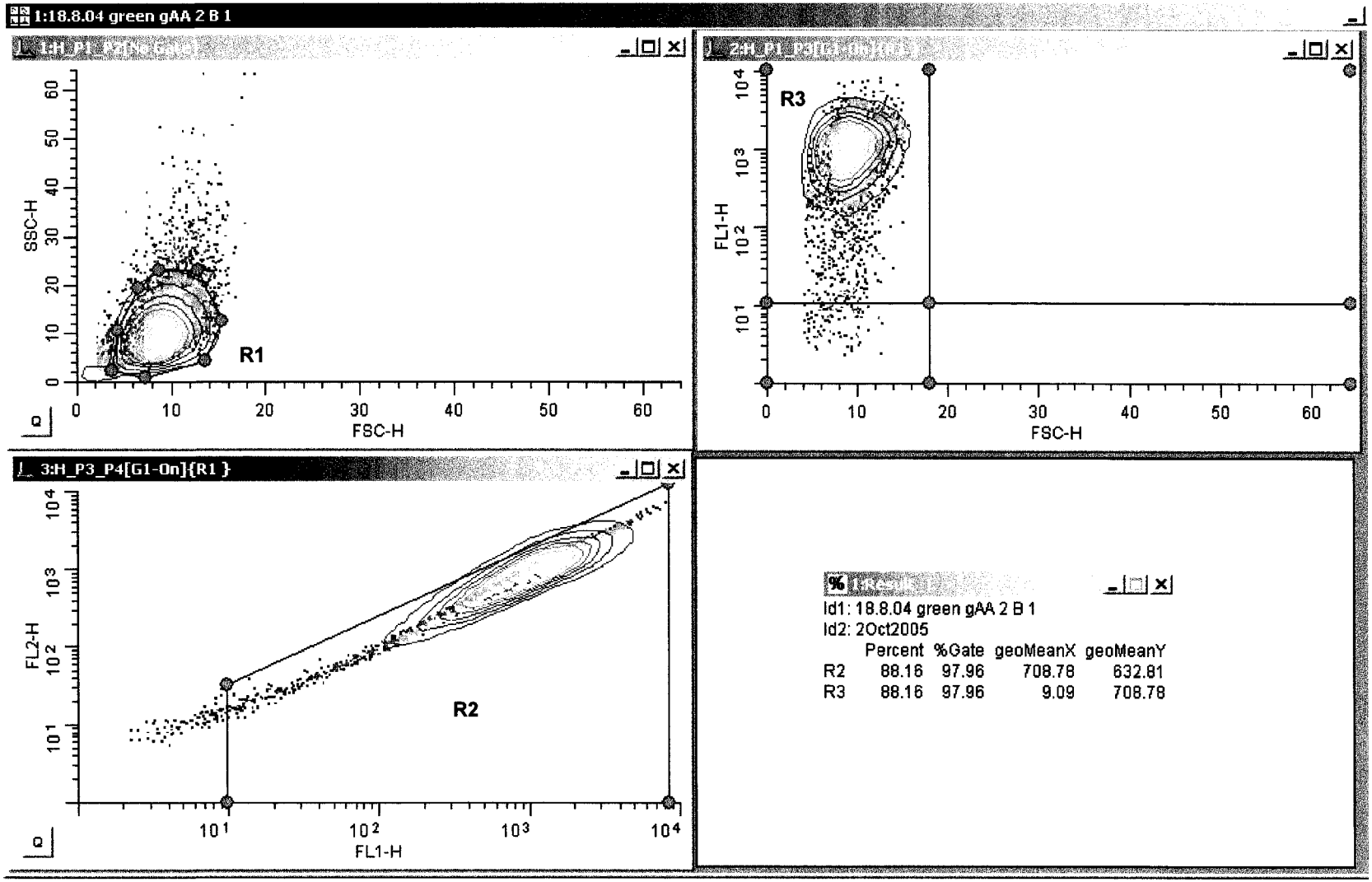

1:18.8.04 green gAA 2 B 1
SSC-H
FSC-H
R1
FL1-H
R3
FL2-H
R2
Id1: 18.8.04 green gAA 2 B 1
Id2: 2Oct2005
Percent %Gate geoMeanX geoMeanY
R2 88.16 97.96 708.78 632.81
R3 88.16 97.96 9.09 708.78

Figure 3B. Characterization of control (GFP negative) cells by flow cytometry. Plots showing the distribution of an untransfected MC2a population, after a 0.1mM exposure to H_2O_2 (cytotoxic concentration), along the green fluorescence (FL1-H) vs red fluorescence (FL2-H) axes, FSC-H and SSC-H axes and FSC-H and FL1-H axes.

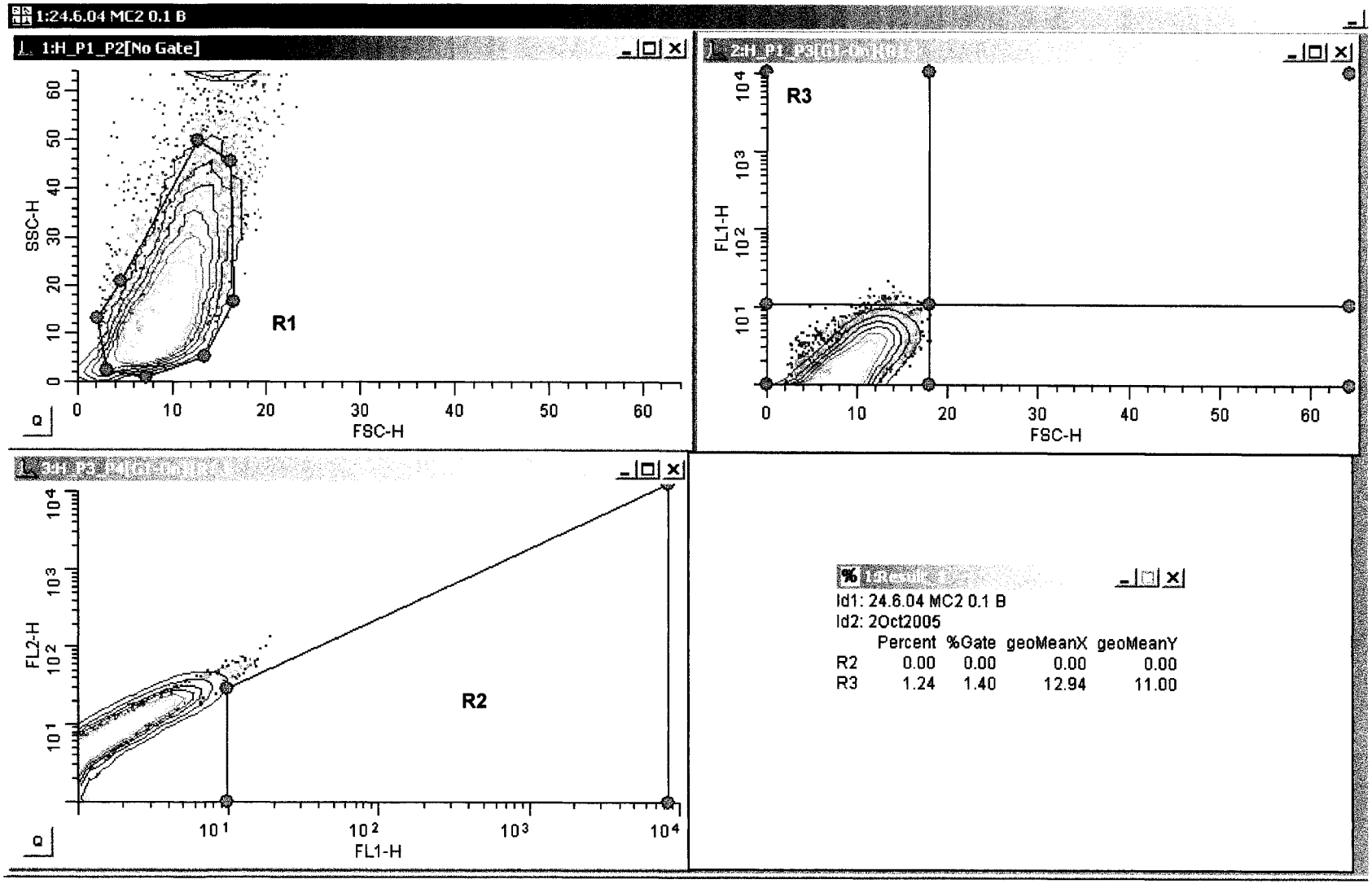

1:24.6.04 MC2 0.1 B
1:H_P1_P2[No Gate]
SSC-H
FSC-H
R1
2:H_P1_P3[No Gate]
FL1-H
FSC-H
R3
FL2-H
FL1-H
R2
Id1: 24.6.04 MC2 0.1 B
Id2: 2Oct2005
Percent %Gate geoMeanX geoMeanY
R2 0.00 0.00 0.00 0.00
R3 1.24 1.40 12.94 11.00

Figure 3C. Characterization of mixed sample with GFP positive cells and control (GFP negative) cells by flow cytometry. Plot showing the distribution of an untreated mixed sample with stably transfected MC2a population expressing GFP and untransfected MC2a cells, along the green fluorescence (FL1-H) vs red fluorescence (FL2-H) axes, FSC-H and SSC-H axes and FSC-H and FL1-H axes.

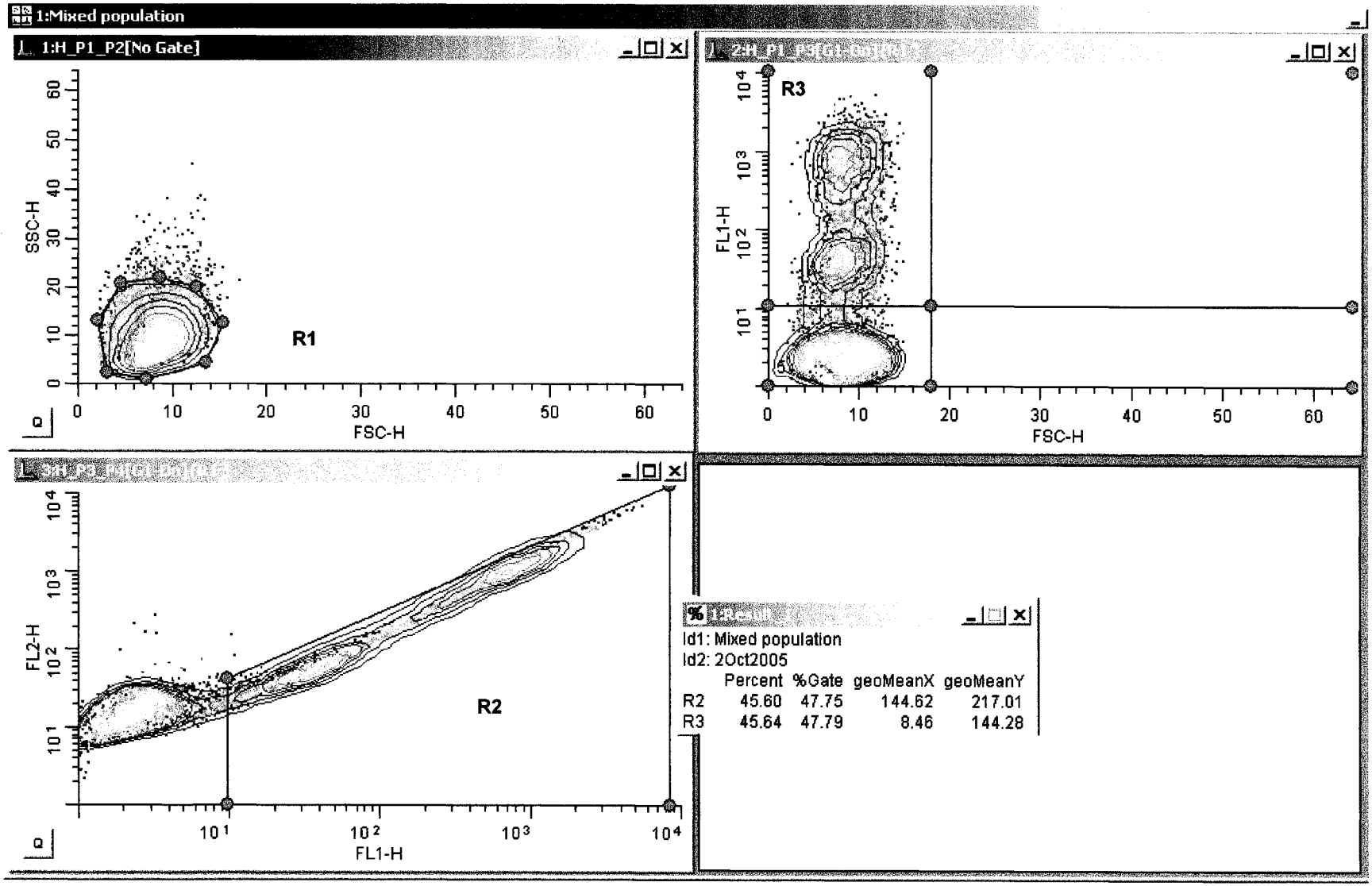
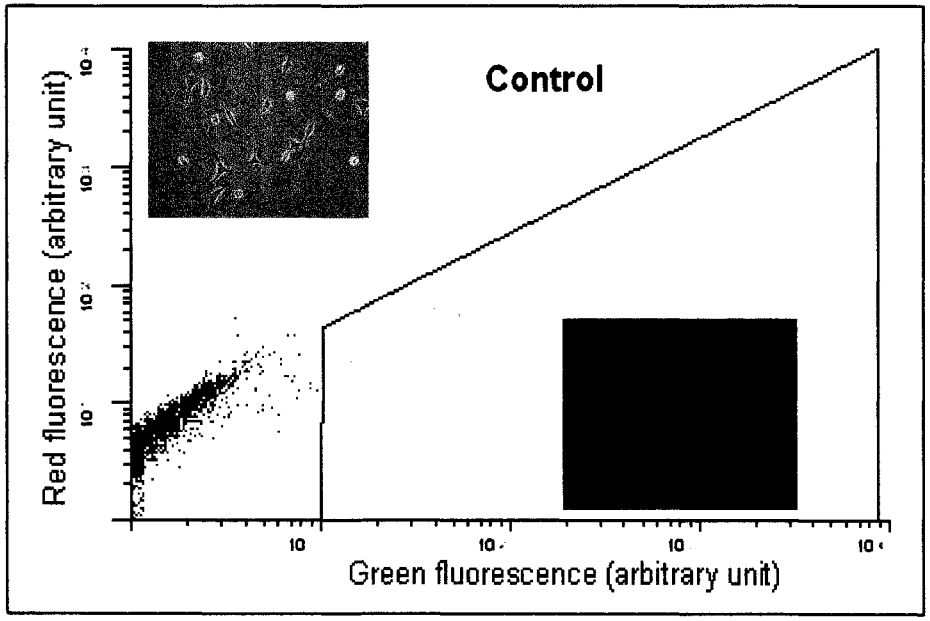
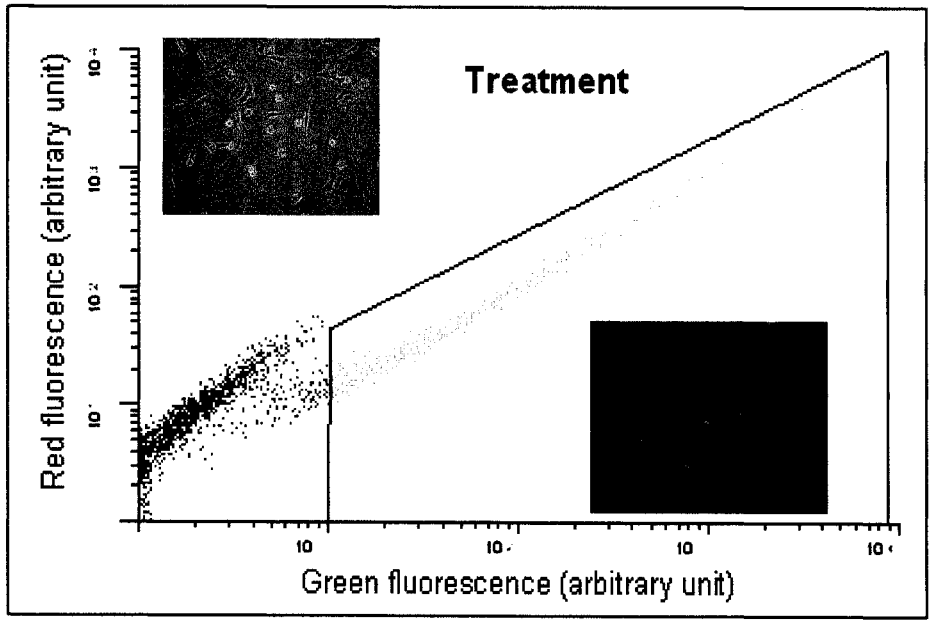

1:Mixed population
1:H_P1_P2[No Gate]
SSC-H
FSC-H
R1
R3
FL1-H
FL2-H
R2
Id1: Mixed population
Id2: 2Oct2005
Percent %Gate geoMeanX geoMeanY
R2 45.60 47.75 144.62 217.01
R3 45.64 47.79 8.46 144.28

Figure 4. Example of flow cytometric analyses of GFP levels in stably transfected MC2a populations with the Hm1 plasmid. The plot of red vs green fluorescence (A) shows background counts of GFP positive cells in an untreated population and control sample and (B) shows a sample treated with 16 μM 9 amino acridine. Normal cells are plotted in blue and GFP positive cells in green.

A


Control
Red fluorescence (arbitrary unit)
Green fluorescence (arbitrary unit)

B


Treatment
Red fluorescence (arbitrary unit)
Green fluorescence (arbitrary unit)

2.8. Statistical Analysis

Five different agents at 3 or 4 concentrations plus vehicle controls were tested in cells transfected with one of several repeat constructs. All experiments included triplicate incubations of 3 controls: 1) untransfected cells to measure potential treatment background, 2) EGFP-expressing cells to test for expression changes resulting from the treatment, and 3) cells transfected with the NLA plasmid with the EGFP sequence shifted out of phase by the insertion of a non-repeat sequence insert. Since all treatment protocols were different, statistical analysis for each of the exposures were treated independently. The mutation response corresponds to the average number of green fluorescent cell counts per 10,000 cells based on n=6 (triplicates for each exposure point for two independent experiments). To determine the statistical significance of exposure response for each plasmid, the analysis was performed using the GENMOD procedure of SAS/STAT software (Institute, 1989.) to fit a log-linear model to the Poisson distributed cell counts as described by (Slebos et al., 2002) using using Pearson Chi-square to estimate the dispersion parameter. For models which exhibited collinearity (parameter estimates with large standard errors) or did not converge the parameter for the lowest exposure was removed from the model. This process was repeated until the model converged and the parameter estimates displayed no significant collinearity. The parameter for the lowest exposure was removed because the counts for the control and the lower exposure groups were either the same or very similar for some plasmids.

2.9. Isolation and sequencing of GFP mutants

The 9AA exposures gave the most frequent mutants, thus this treatment was selected to confirm the occurrence of DNA mutations at TRS in green fluorescent cells with reporter genes. DNA was extracted from treated single green fluorescent clones and control non-fluorescent clones using lysis buffer supplemented with proteinase K (1mg/mL) followed by ethanol precipitation. A 786 bp fragment between the CMV and GFP region was amplified using primers G1 5'-CGTGGATAGCGGTTTGACTC- 3' (sequence 395-414) and G2 5'-TTCACCTTGATGCCGTCTT-3' (sequence 1153-1172). Briefly, the GFP PCR reaction consisted of 1 X PCR buffer (Jeffreys et al., 1990), 12mM Tris pH 7.5, 0.5 μM each of G1 and G2 primers, 50ng of genomic DNA template, 0.4 U Taq polymerase, 0.04 U of Pfu polymerase and deionized water to a final volume of 8 μL. The fragment was amplified by 30 cycles at 94 °C (30s), 59.5 °C (30 s), 72 °C (1 min) on a PTC-225 Peltier Thermal Cycler (MJ research). The PCR product was isolated on a 0.8 % agarose gel in 1XTBE, and purified using GFX gel band purification kit (GE Healthcare) and sequenced commercially (Cortec, Kingston, ON).

Chapter 3

3.0 Results

Instability in TRS is higher than in heteropolymeric DNA and has been associated with conditions such as cancer, neurological disorders, premature aging as well as atherosclerosis (e.g., Miniati et al., 2001; Dellarco et al., 1995). Consequently, TRS represent a potentially useful and relevant target for detecting exposure to stress agents. The current conventional assays, e.g., gene reversion test and chromosomal aberration assay, do not focus on measuring mutations in repetitive DNA. In addition they provide equivocal predictions or lack sensitivity to detect potential carcinogens or mutagens that act on specific DNA sequences, therefore, mis-identifying harmful agents as non-genotoxic (Jackson et al., 1993). Tests based on monitoring tandem repeat sequence variability such DNA fingerprinting and single molecule PCR have shown that repetitive sequences are especially sensitive to a wide range of stresses including classic mutagens as well as carcinogens such as tumour promoters. (Yauk et al., 2002; Parfett et al., 2003) Green fluorescent protein reporter systems have been developed to detect frameshift mutation in dinucleotide and tetranucleotide repeat sequences to compare their relative mutability and investigate their potential link diseases such as cancer (Slebos et al., 2002)

A primary goal of this study was to develop a quick, reliable and economic *in vitro* GFP frame-shift assay in mammalian cells as a biomarker of exposure to a wider range of environmental mutagens. The assay relies on the detection of frameshift mutations within inserted sequences, placed upstream of a GFP gene, that activate GFP expression. Specifically, a set of 9 test and control GFP plasmid constructs were built to ensure proper functioning of the reporter system as well as determine the most sensitive biomarkers among the tested sequences. Two control constructs were created. The first is a positive control (posi) with a heteropolymeric insert upstream of the GFP sequence but

positioning the coding sequence in frame. The second is an out-of-frame negative control NLA with a heteropolymeric insert upstream of the GFP sequence that places the gene out of frame. The 7 test constructs include representative repeat units of different sizes: CA with dinucleotides (CA), Trio with trinucleotides (CAG), AAAG with tetranucleotides (AAAG), Hm1 with pentanucleotides (GGGCA), Per with hexanucleotides (ACAGGC) and Hexa also with hexanucleotides (ACCTCC), as well as a human gene sequence containing tetranucleotide sequences (TTTA). Since mononucleotide repeat runs have already been extensively studied, they were not included in this study (eg. (Hienonen et al., 2005; Ruggiero et al., 2003)

To investigate the relevance of this assay for genotoxicity screening, a panel of 5 different classes of chemicals was selected: hydrogen peroxide (H_2O_2), 12-*O*-tetradecanoyl-phorbol-13-acetate (TPA), benzo-*a*-pyrene-diol-epoxide (BPDE), ethyl nitrosourea (ENU) 9-amino-acridine (9AA). Initial testing was primarily done with the MC2 cell line but as a point of comparison for reproducibility of response, the two most effective treatments (9AA and H_2O_2) were also tested in C3H10T1/2 harbouring the NLA, Trio and Hm1 plasmids.

A secondary objective was to test the underlying hypothesis that repeat sequence mutations are usually generated from Streisinger's slippage. The GFP assay detects frameshift mutations, consequently the loss or gain of 3 bases or 6 bases from the Trio, Per or Hexa repeat unit should not be detectable. However, if these TRS are hotspots for other mutational events then the mutation frequency observed in cells habouring Trio, Per or Hexa plasmids should be higher than that observed in cells carrying the negative control plasmid NLA with a heteropolymeric insert.

The first step to this study was characterizing the GFP reporter system prior to using it. The vectors were built and the accuracy of each construct was verified by sequencing. Transient transfections of the positive control and negative control were used to ensure functional expression of GFP from the positive control Posi and no expression from the out of frame negative control NLA. Preliminary testing with the positive control was used to determine the best conditions (see Materials and Methods) to generate stably transfected MC2a and C3H10T1/2 cell populations. Stable transfections with the other 8 plasmids were done accordingly. Since the out of frame constructs do not express functional GFP, stable functional integration of the plasmid was verified via the acquisition of antibiotic selection and Northern blots.

The second step involved was the exposure treatments and measurements of mutation frequency. Optimal testing concentrations were determined through growth inhibition curves and testing was done accordingly. To ensure accurate counts through flow cytometry, calibration with green fluorescent and unstransfected cells was performed.

The third step was to verify that the green cells counted were actually reflecting mutations at the DNA level within the inserted sequences. To achieve this, preliminary experiments were done to sequence the inserts from green mutant clones as well as non-expressing clones.

3.1 Characterization of the GFP reporter system

The GFP reporter system was developed to provide a quick and reliable way to test mutagenicity induced by a wide range of different mutagens and carcinogens in a toxicological assessment framework. This flow cytometric assay has the potential to

quickly detect the occurrence of frame shift mutation within repeat sequences. In addition to what was currently known from previous GFP reporter assays (Slebos et al., 2003; Gasche et al., 2002), this study aimed to apply this approach for toxicological assessment purposes. An additional goal was to test current hypothesis regarding how mutations might be generated in repeat sequences, e.g., according to models such as replicative slippage or non-homologous end joining models. Furthermore, the relative mutability of hexanucleotide and pentanucleotide repeats was examined, as well as human sequences that contain tetranucleotide repeated motifs. The latter had not been examined before in a GFP reporter system.

3.1.1 Plasmid constructs

All planned plasmid constructs were successfully generated as described in Materials and Methods, with sequences verified by sequencing. All out of frame constructs were designed as to place the GFP in the –1 reading frame. To restore the correct reading frame, a + 1 mutation must occur. Repeat sequence mutations that can lead to such an event include the perfect gain or loss of repeat units, or imperfect loss or gain of repeat unit or random insertion deletion (in/del) events of 3n+1 bases in the case of insertions or and 3n+ 2 bases in the case of deletions (see Table 2). The trinucleotide Trio and hexanucleotides Hexa and Per6 constructs have repeat units that are multiples of 3 thus the gain or loss of complete repeat units would not change the reading frame and not yield green mutants. These three constructs can only be used to detect the latter two types of mutations (in/del of imperfect repeats or random in/del).

Table 2. List of predicted frame-restoring mutations for each plasmid used in this study.

Plasmid name	Possible frame restoring mutations
CA	Gain of 2+3n repeats (eg. 2 repeats is 4 bp, 5 repeat is 10bp) or loss of 3n+1 repeats (eg. 1 repeat is 2bp, 4 repeats is 8 bp)
Trio	Gain of 3n+1 bp (eg. 1,4,7 bp) or loss of 2+3n bp (eg. 2,5,8 bp)
AAAG	Gain of 1+3n repeats (eg. 1 repeats is 4 bp, 4 repeat is 16bp) or loss of 3n+2 repeats (eg. 2 repeat is 8bp, 5 repeats is 20 bp)
Hm1	Gain of 2+3n repeats (eg. 2 repeats is 10 bp, 5 repeat is 25bp) or loss of 3n+1 repeats (eg. 1 repeat is 5bp, 4 repeats is 20 bp)
Per	Gain of 3n+1 bp (eg. 1,4,7 bp) or loss of 2+3n bp (eg. 2,5,8 bp)
Hexa	Gain of 3n+1 bp (eg. 1,4,7 bp) or loss of 2+3n bp (eg. 2,5,8 bp)
Hucy	Gain of 2+3n repeats (eg. 2 repeats is 4 bp, 5 repeat is 10bp) or loss of 3n+1 repeats (eg. 1 repeat is 2bp, 4 repeats is 8 bp)
NLA	Gain of 3n+1 bases (eg. 1,4,7 bp) or loss of 2+3n bases (eg. 2,5,8 bp)

The interest in having trinucleotide and hexanucleotide repeat constructs is to determine how frequently imperfect slippage or random in/del, e.g., possibly due to recombination, occur in repeat motif regions compared to non-repetitive DNA found in the negative repeat construct. The dinucleotide construct CA, tetranucleotide construct AAAG, pentanucleotide Hm1 and human sequence construct Hucy containing the tetranucleotide repeat TTTA can be used to detect the gain or loss of perfect repeats in the GFP reporter system. For the dinucleotide repeat CA mutations that restore the reading frame include the gain of 3n+2 repeat units or the loss of 3n+ 1 repeats. For the tetranucleotide, the AAAG and the TTTA repeat in Hucy, green mutants can be generated from the gain of 3n + 1 repeat units or the loss of 3n + 2 repeats. For the pentanucleotide repeat Hm1, we can detect the gain of 3n + 2 repeat units or the loss of 3n + 1 repeats.

To ensure functionality of the constructs (limited leakage and ability to potentially generate GFP) as well as feasibility to transfect both C3H10T1/2 and MC2a cells and obtain stably transfected populations, the in frame positive control and negative construct were tested.

3.1.2 Transient transfections

After 24 hours post transfection, both cell lines carrying the negative control (NLA) had little to no GFP expression (less than 0.02% of the population or 2 cells/10 000) thus negligible background from potential translational slippage, spontaneous mutation events or from mutations generated from integration of the sequence into the genome of the cells. Cultures carrying the positive control were able to transiently express GFP confirming that inserting sequences upstream of the GFP sequence, within the original pEGFP-N1 that left the correct reading frame, did not prevent GFP

expression. The cells that did not express GFP were either not carrying the plasmid (due to transfection efficiency not being a 100%) or were unable to express GFP due to possible vector damage during integration into the genome, possible gene silencing or metabolic reasons, e.g., suppressing translation or improper folding of the GFP). GFP could be generated by transient transfection with the in-frame construct (Posi) in both C3H10T1/2 and MC2a, thus frameshift mutations that restore the correct reading frame in the out of frame constructs should theoretically be able to activate GFP expression in transfected cells.

3.1.3 Stable transfections

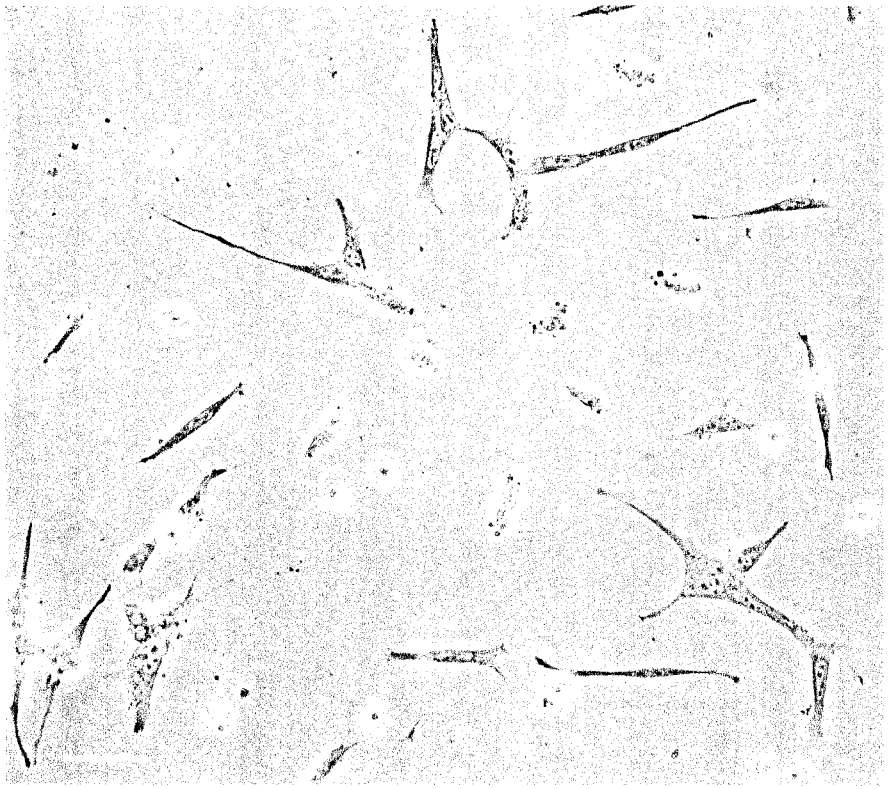
After 1 to 2 weeks of selections with G418, unstransfected cells did not survive since the natural populations of MC2a and C3H10T1/2 do not express the neomycin resistance gene, neomycin phosphotransferase. Regarding cultures transfected with the negative control plasmid, healthy G418 resistant colonies were observed in both MC2a and C3H10T1/2 suggesting that the plasmid was stably integrated into the genome of the cells and no green colony was observed visually in these negative control populations suggesting stability of the negative control construct. In the case of the positive control Posi, the MC2a and C3H10T1/2 cells behaved differently. Both cell lines yielded G418 resistant colonies but the growth rate and level of GFP expression differed between the 2 cell lines. The MC2a cultures had many stably transfected colonies expressing GFP (Figure 5). Over 75% of the cells show strong GFP expression as determined by flow cytometry, the other cells were GFP resistant and expressed lower levels of GFP. However in the case of C3H10T1/2, the surviving colonies were either green senescent cells or with limited to no GFP expression. These results might reflect cell line specific

cytotoxicity of GFP in the C3H10T1/2 cells. Cytotoxicity was also observed in a related cell line: NIH/3T3 by Liu et al. (1999).

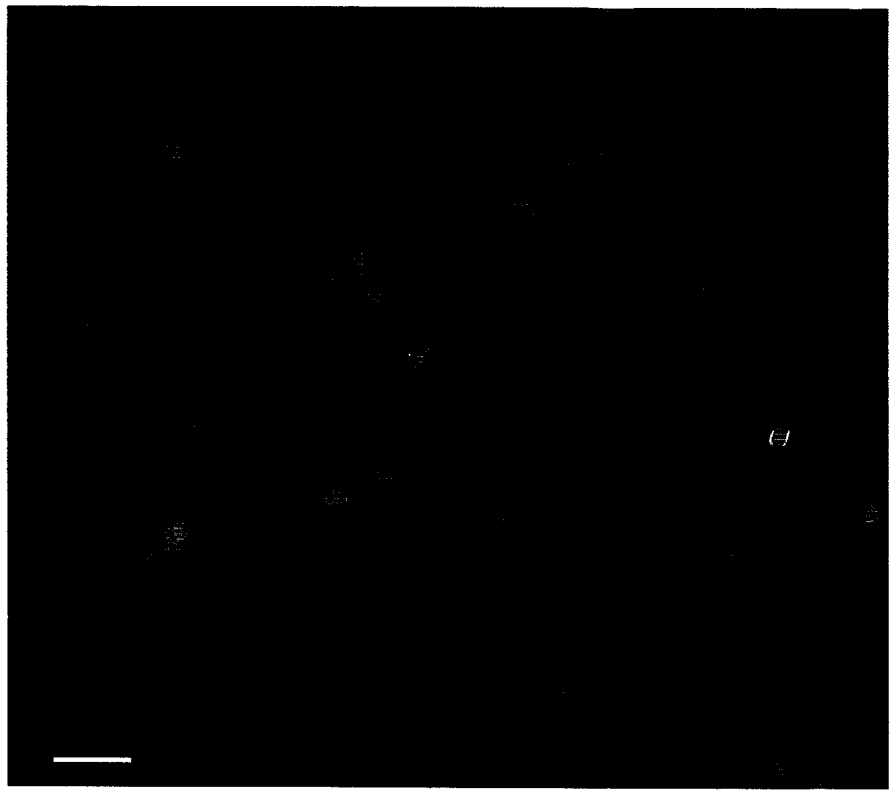
Further results showed that the GFP reporter system could be used in both cells lines as a stable assay. Since this is a GFP activation assay, both cells lines can be stably transfected with the out-of-frame GFP constructs. The C3H10T1/2 cells transfected with the negative control construct NLA formed healthy colonies under G418 selection. In addition, expression of any aberrant protein from the out-of-frame GFP sequences was limited to at most, a 30 amino acid fragment due to the –1 reading frame that introduces an opal stop codon TGA. Stable cultures with the out-of-frame constructs were easily generated in the C3H10T1/2 cell line without toxic effects. In addition, as mentioned in the methodology section, the counts of green mutants expressing GFP is measured 48 hrs after exposure to treatments. The results showed that C3H10T1/2 could transiently express GFP after 24 hrs and even after 48hrs without displaying toxic effects. Consequently, for these experiments, both cell lines were suitable for the use of the GFP reporter assay.

Figure 5. Photomicrograph of stably transfected MC2a murine cells carrying positive control plasmids. In this micrograph, after 2 weeks of selection with G418, cells are stably expressing GFP. Left panel is observed under phase contrast, right panel is in fluorescent light. Scale bar, 40μm.

Phase Contrast



Fluorescence



3.1.4 Characterization of the stably transfected populations

All out of frame repeat constructs and negative control plasmids were used to successfully generate stably transfected populations of pooled clones for both MC2a and C3H10T1/2 cells. Stable transfection was determined based on two criteria: G418 resistance and presence of a functional GFP sequence as determined by Northern blots. Based on the results with untransfected cells and cells transfected with the positive control plasmid, selection with G418 is effective in eliminating cells that do not carry neomycin resistance given by the plasmid sequence and based on phenotype, a high percentage of the remaining cell (over 75%) can strongly express the gene product of the positive control. It was logical to expect that a similar percentage of cells should have properly integrated the out-of-frame constructs and the potential to display frame shift mutations.

The second criterion, Northern blots testing the presence of GFP RNA was deemed necessary, since the out of frame construct do not generate functional GFP, thus, there was no visible marker to check for the presence of the plasmid. Also, inserts placed upstream of the out-of-frame construct places the GFP sequence in the – 1 reading frame generating an early stop codon limiting translation to 30 amino acid sequence. Each sequence would be different, therefore performing Western blots for each putative sequence would not have been efficient. Southern or Northern blots are alternative approaches. Northern blots were considered the better approach since cell populations with random integration of plasmids throughout cell genomes could generate smears in a Southern (see Gasche et al., 2002). In addition, Southerns do not provide information on

whether the plasmid sequence can be transcribed after random integration into the cellular genome.

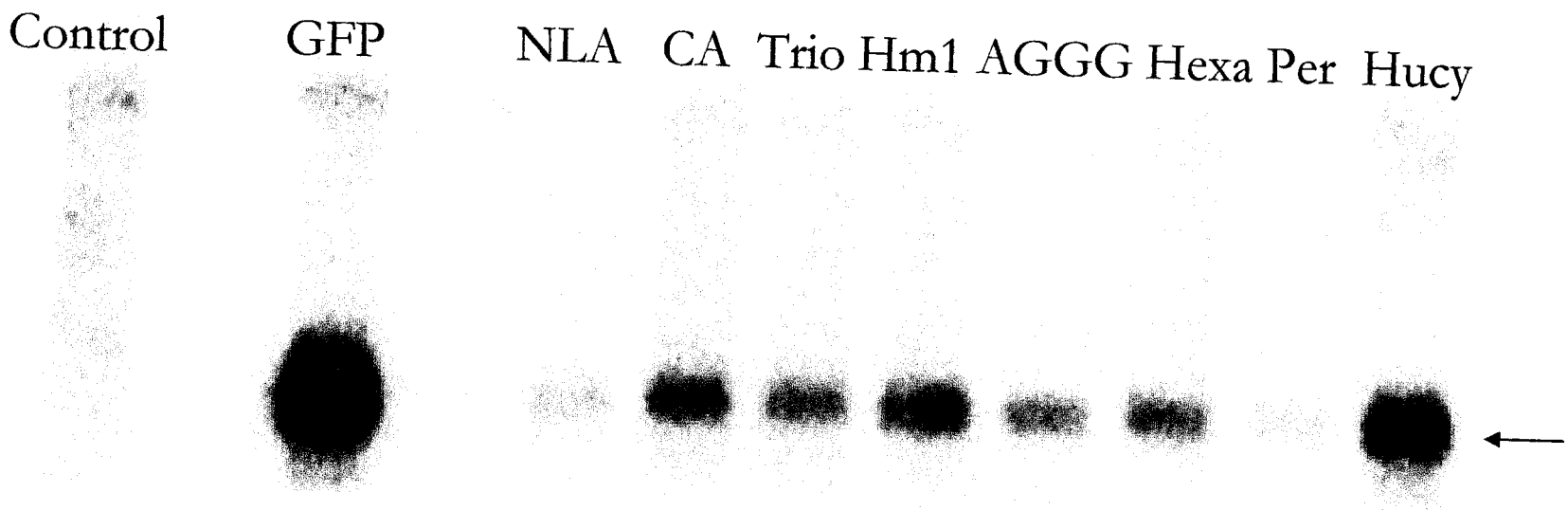
RNA was isolated from each individual G418 selected cell population, and analysed by Northern blot probed with the GFP sequence which is absent in the genome of control, untransfected MC2a cells but present in stably transfected cells (see Figure6). All constructs contained the same transcriptional consensus sequences which were unaffected by the presence of the inserted repeat sequences over 100 bp downstream. To effectively determine the relative levels GFP RNA expression for all constructs, all blots were hybridized with a poly(A) probe as a loading control. The positive control, RNA from a brightly fluorescent clone, was selected to characterize the GFP RNA band. For the out-of- frame constructs, populations were generated from pooled clones, thus RNA levels observed on the Northerns represent an average across all the pooled clones. Figures 6 shows poly(A) signals for all cell groups, the highest GFP RNA signal is observed in the green-fluorescent cells, no GFP RNA signal in the untransfected cells, and all transfected populations (other than the positive control) show similar levels of GFP-RNA expression relative to the polyA loading control. Thus all transfected populations do transcribe the insert/GFP sequence from the plasmid constructs. Presence of the TGA opal stop codon in these out-of-frame construct can potentially lead to nonsense-mediated mRNA decay thus a lower GFP signal in cells harbouring this constructs than in the GFP clone (Mo et al., 2005).

In the case of C3H10T1/2 Northerns (Figure 7), RNA from stably highly expressing green fluorescent MC2a clone was used as a positive control since no stable GFP transfected C3H10T1/2 cells were obtained due to putative cytotoxic effects or gene

silencing. As seen on Figure 7, the GFP probe hybes to the lower GFP RNA band in the positive control giving a bright signal, no specific hybridization is observed in the negative control RNA from untransfected C3H10T1/2 but slight non-specific hybridization was present (also seen in the other lanes). Specific hybridization was observed in the lanes loaded with RNA from C3H10T1/2 stably transfected with plasmids NLA, Trio and Hm1. Relative to the poly(A) signal, all 3 transfected C3H10T1/2 populations (NLA, Trio and Hm1) displayed similar levels of GFP RNA. The signal in these lanes was lower than in the positive MC2a GFP control lane. This, variations in GFP RNA signal strength could be attributable to differences between the metabolism of C3H10T1/2 and MC2a (eg. different transcription efficiency, levels of RNA degradation, gene silencing).

Figure 6. Molecular characterization of chimeric GFP expression by Northern blot analysis in MC2a populations. Ten µg of RNA as determined by UV absorbance, were loaded on a denaturing 1.0% agarose gel, transferred and hybridized with a 787-bp EGFP probe (see Materials and Methods). First loaded lane: unstransfected control MC2a cells, second loaded lane: stably transfected green fluorescent MC2a clone; following lanes are loaded with samples from MC2a populations transfected with test plasmid constructs (see Materials and Methods). A/ hybridized with a GFP probe, B/ hybridized with a loading control poly(A) probe. In A/ the slight smearing indicates non-specific hybridization, the arrow points to the specific GFP RNA band that is present in all test samples and seen as a bright band in the highly GFP expressing clone but absent in the untransfected control.

A


Control
GFP
NLA
CA
Trio
Hm1
AGGG
Hexa
Per
Hucy

B

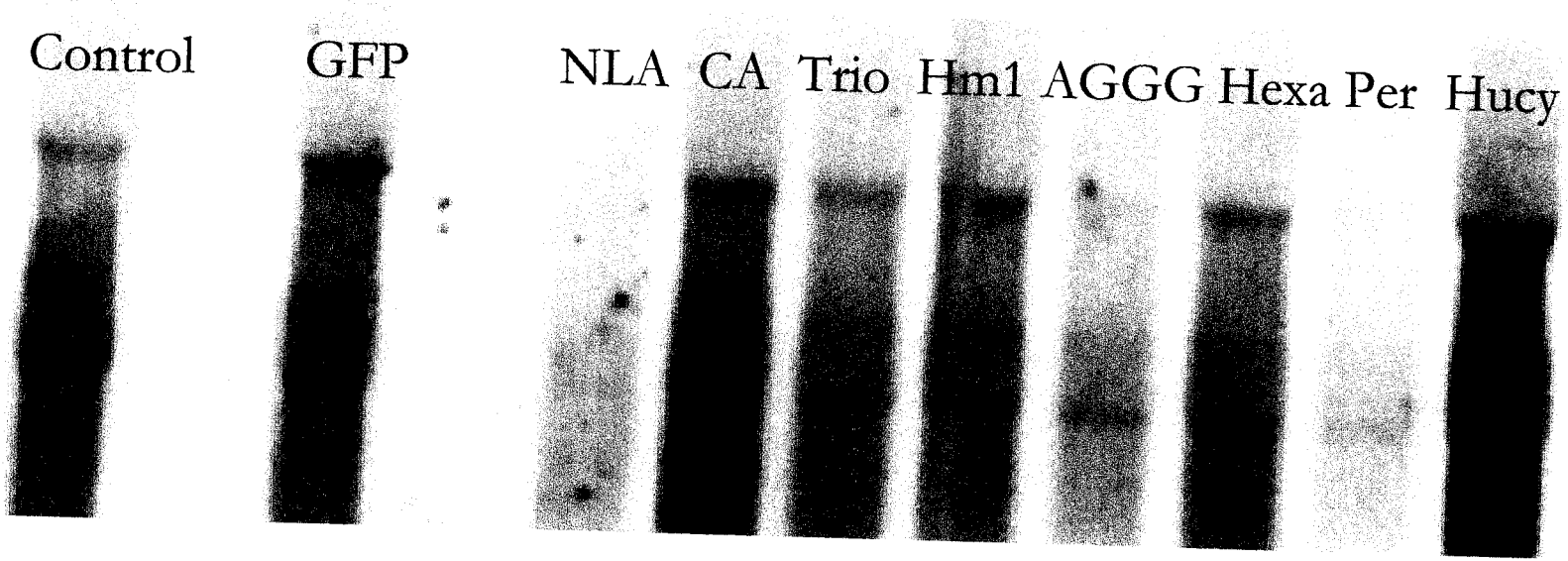
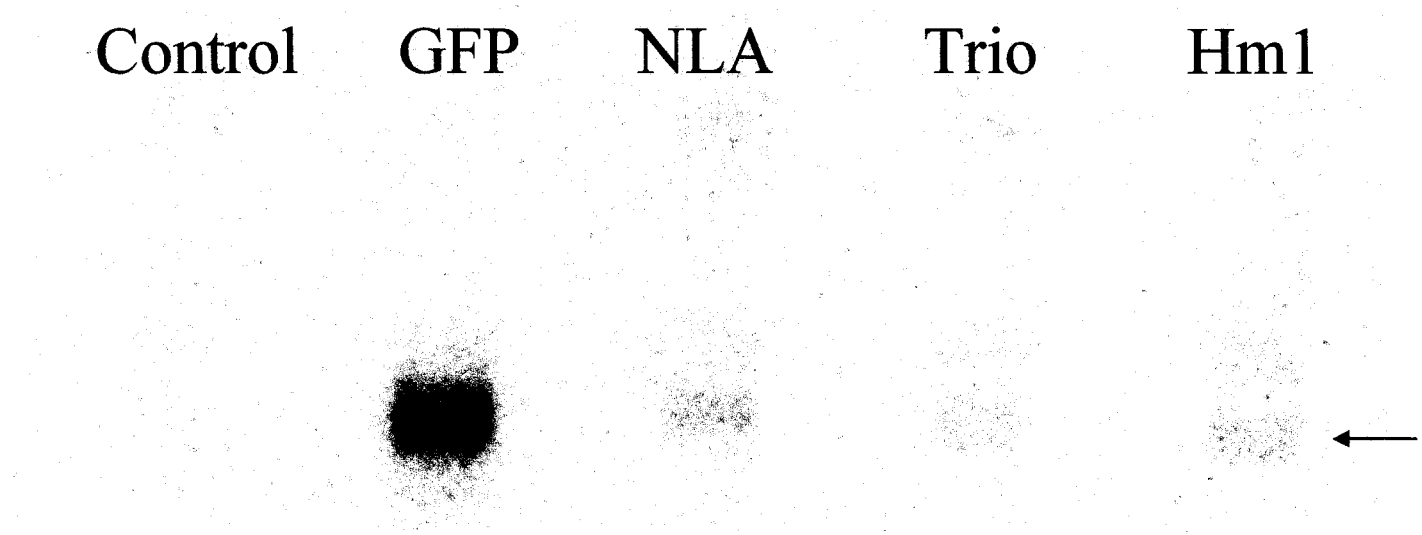
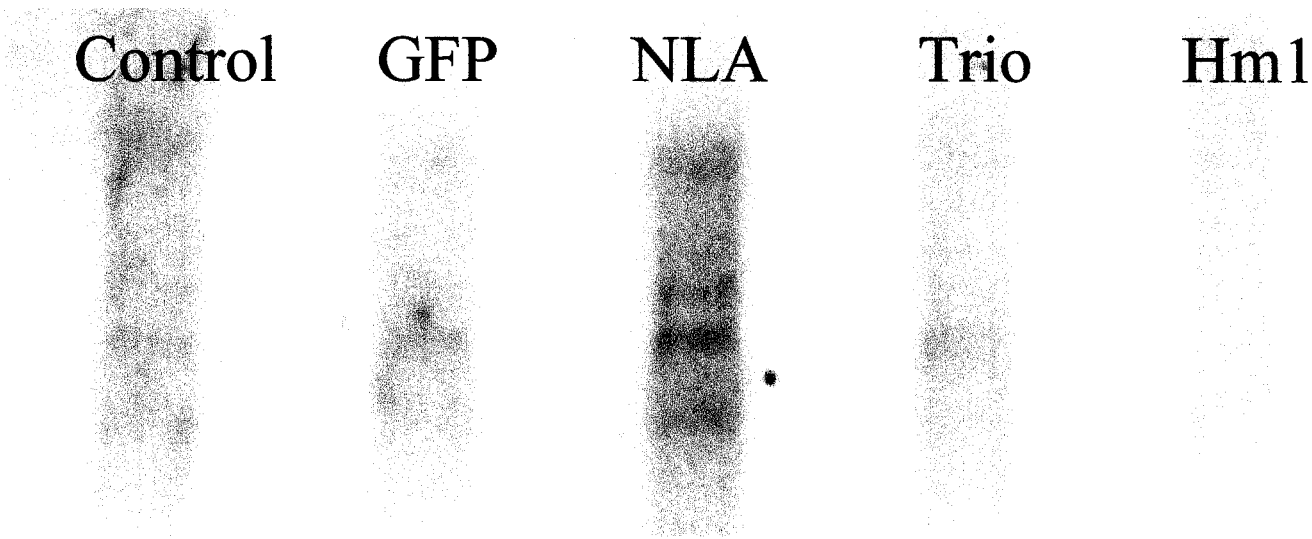

Control
GFP
NLA
CA
Trio
Hm1
AGGG
Hexa
Per
Hucy

Figure 7: Molecular characterization of chimeric GFP expression by Northern blot analysis in C3H10T1/2 populations. Ten μg of RNA as determined by UV absorbance, were loaded on a denaturing 1.0% agarose gel, transferred and hybridized with a 787-bp EGFP probe (see Materials and Methods).
First lane: stably transfected green fluorescent MC2a clone, second loaded lane: used as a positive control since C3H10T1/2 expressed GFP only transiently in these experiments, and C3H10T1/2 populations transfected with test plasmid constructs (see Materials and Methods). A/ hybridized with a GFP probe, B/ hybridized with a loading control poly(A) probe. In A/ the slight smearing indicates non-specific hybridization, the arrow points to the specific GFP RNA band that is present in all test samples and seen as a bright band in the highly GFP expressing clone but absent in the untransfected control.

A



B



3.2 Chemical treatments, flow cytometry and statistical analysis

3.2.1 Determination of concentration range for each chemical treatment

Appropriate concentrations were determined based on growth curves for MC2a cells. The curve is based on cell counts 48 hours after exposure to treatments. Therefore the survival measurements reflect a combination of growth rate and cell death as opposed to the colony formation assays that reflect the ability of surviving cells to replicate and form colonies. In the literature (Tomei et al., 1981), cell growth rate in C3H10T1/2 was found to be around 3.5-4.0 generations in 72 hrs (from initial plating to counting), which is a 12 to 16 fold increase if all cells undergo cell division. In this study, control cells showed 8 to 17 fold increases after 72 hrs in both C3H10T1/2 and MC2a suggesting generation times of 18-24 hrs in the same range as what was found by Tomei et al. (1981).

The lowest concentration of hydrogen peroxide (0.001mM) did not inhibit cell growth in MC2a (Figure 8) or C3H10T1/2 (Figure 9) with a fold increase around 8 which represents 4 generations in 72 hours. However, at 0.01mM, cell growth is decreased by about half (~ 4 fold increase or 2 generations) in both cell lines and down to a 2 fold at 1mM which is the highest concentration used in these experiments. At concentration below 1mM, the decrease in cell number reflects a combination of lag in cell growth as cells can undergo cell cycle arrest after exposure to oxidative stress until DNA damage is repaired as well as cell necrosis and apoptosis from toxic concentrations of H_2O_2 (Chang et al., 2003).

Figure 8. Killing curve of MC2a cells following H_2O_2 treatments. Cytotoxicity was determined as total adherent cell counts in MC2a cultures 48 hours after replacement of media containing test agent. Values are expressed relative to the number of cells seeded in each dish prior to treatments.

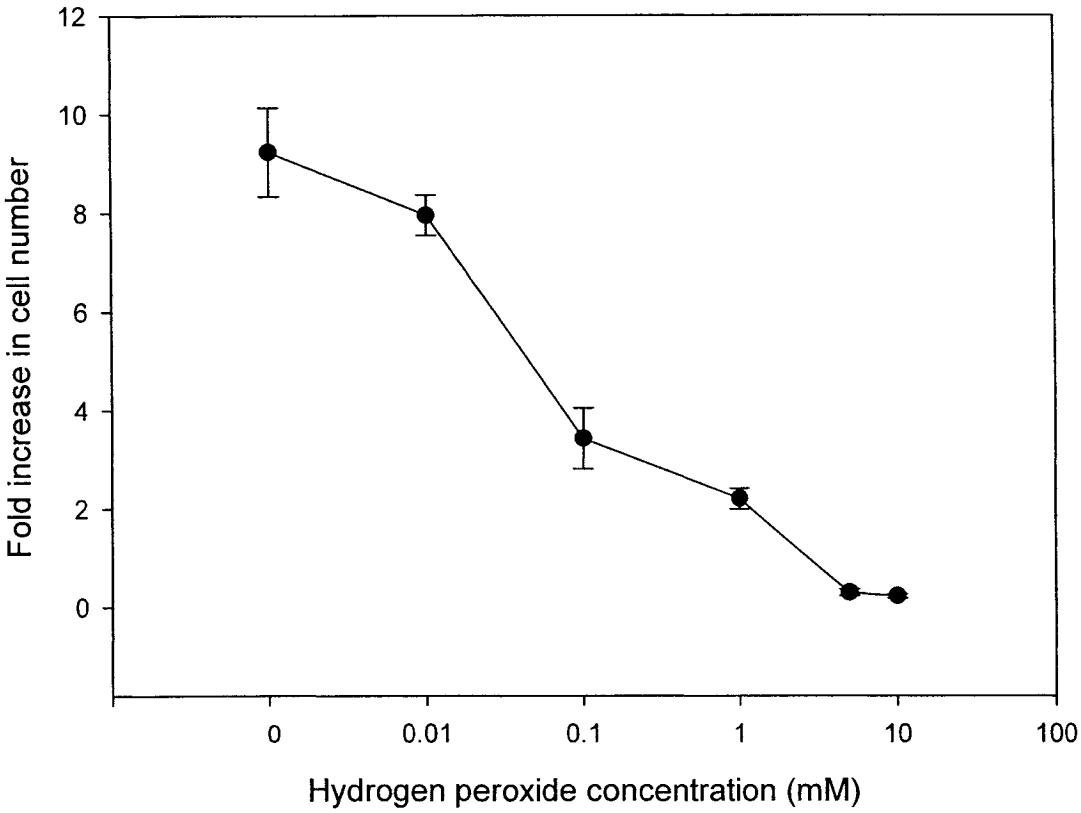
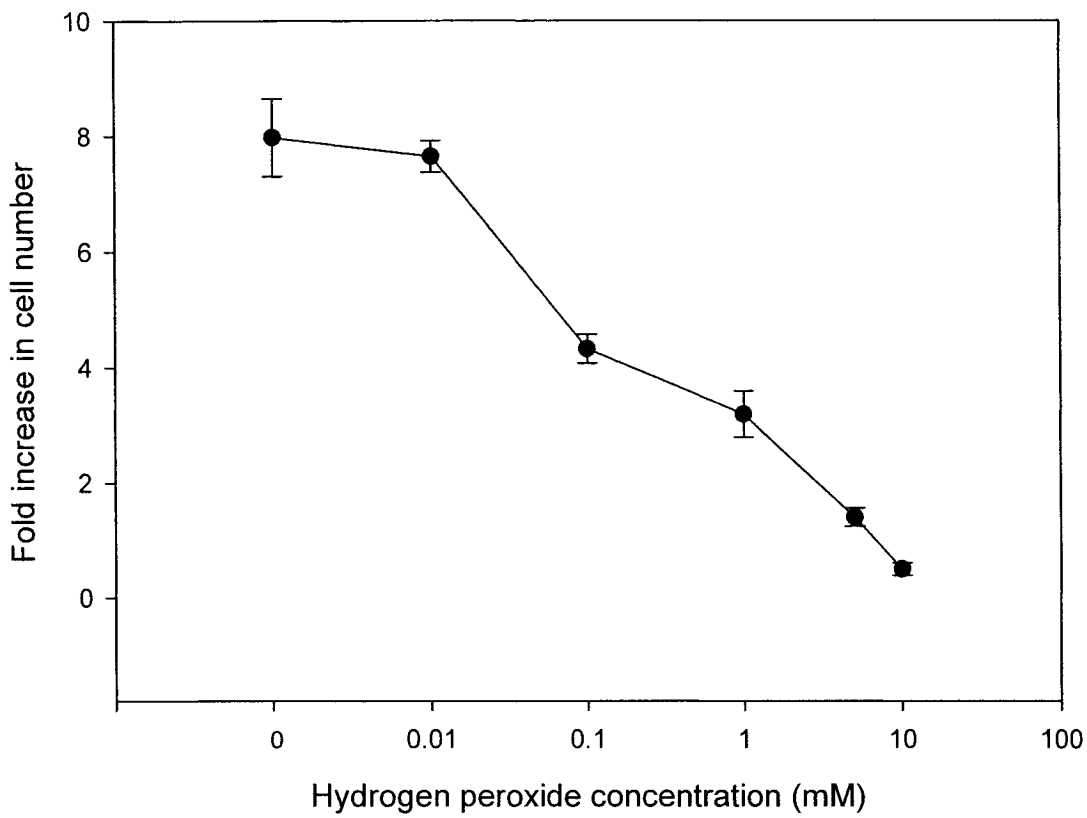

12
10
8
6
4
2
0
Fold increase in cell number
0
0.01
0.1
1
10
100
Hydrogen peroxide concentration (mM)

Figure 9: Killing curve of C3H10T1/2 cells following H_2O_2 treatments. Cytotoxicity was determined as total adherent cell counts in C3H10T1/2 cultures 48 hours after replacement of media containing test agent. Values are expressed relative to the number of cells seeded in each dish prior to treatments.


10
8
6
4
2
0
Fold increase in cell number
0
0.01
0.1
1
10
100
Hydrogen peroxide concentration (mM)

However at concentrations higher than 1mM, fold increase in cell number was below 1 for both cell lines (Figures 8 and 9) which reflects cell cycle arrest in addition to cell death. These results are consistent with Gasche et al.'s (2001) cell viability assays using propidium iodine staining to measure viability. Indeed, their findings showed that below 1mM H_2O_2, cell were relatively unaffected, at 1 mM, cytotoxicity was observed and at 10 mM H_2O_2, the viability was 0.3% reflecting high toxicity. The highest concentration used in this study for mutation experiments was 1mM, the highest concentration that still allowed accurate counts since a highly toxic treatment killed too many cells (Figures 8 and 9) to get accurate representative counts.

For 9AA, at a concentration of 2μM, the growth of MC2a cells (Figure 10) was not affected but for C3H10T1/2 (Figure 11) a slight effect was observed (2 folds or 1 generation less than the control). At 8 μM there was a 1/3 decrease of growth in both cell lines. At 16 μM, which was the highest concentration used for testing, the growth is decrease by about half that of the control in MC2a and C3H10T1/2 cells. At 32 μM, fold increase in cell numbers was down 3.5 in C3H10T1/2 and 1.5 in MC2a cells suggesting that MC2a cells were more sensitive to 9AA. The treatment became very toxic at 64 μM in both cell lines.

Alkylating agent ENU (Figure 12) inhibited growth by half at concentration between 1.07 and 4.27 mM in MC2a cells. At 6.40 mM, fold increase decreased to 3 fold (1.5 generation in 72hours). This was the highest concentration used in these experiments. At 8.53 mM ENU is highly toxic for MC2a cells with an increase below 1 fold.

Figure 10. Killing curve of MC2a cells following 9AA treatments.

Cytotoxicity was determined as total adherent cell counts in MC2a cultures 48 hours after replacement of media containing test agent. Values are expressed relative to the number of cells seeded in each dish prior to treatments.

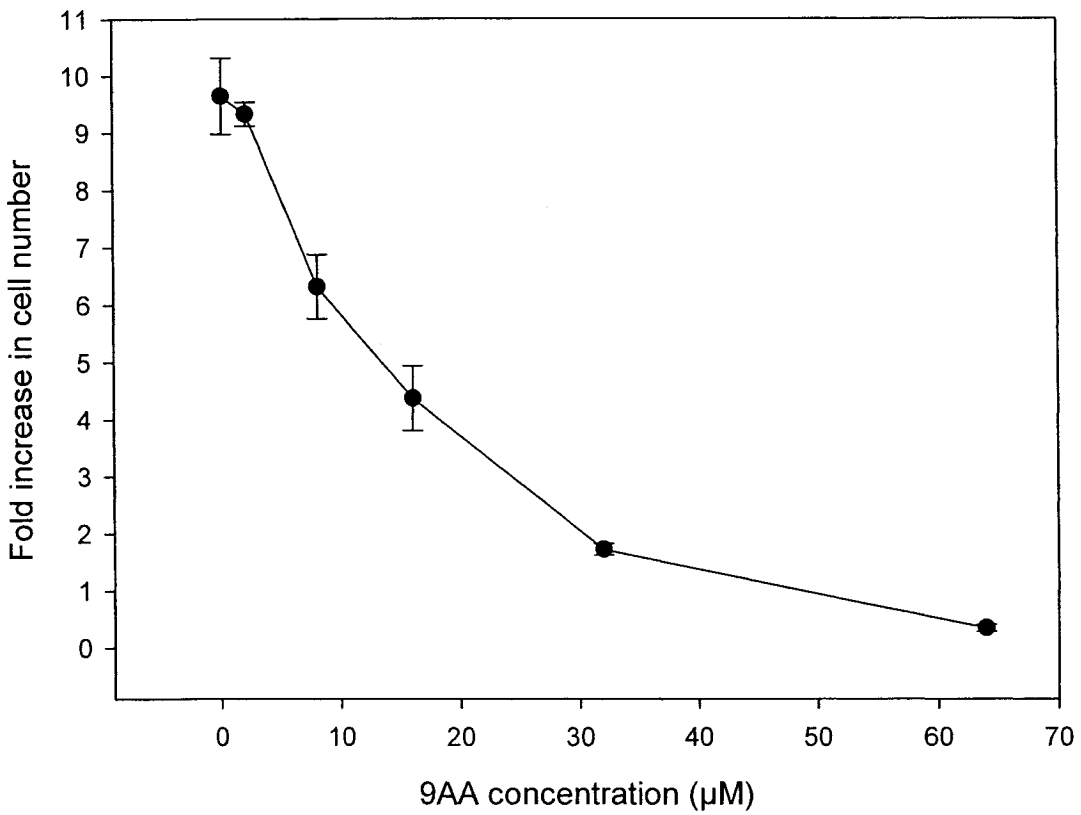

Fold increase in cell number
11
10
9
8
7
6
5
4
3
2
1
0
0
10
20
30
40
50
60
70
9AA concentration (μM)

Figure 11. Killing curve of C3H10T1/2 cells following 9AA treatments. Cytotoxicity was determined as total adherent cell counts in C3H10T1/2 cultures 48 hours after replacement of media containing test agent. Values are expressed relative to the number of cells seeded in each dish prior to treatments.

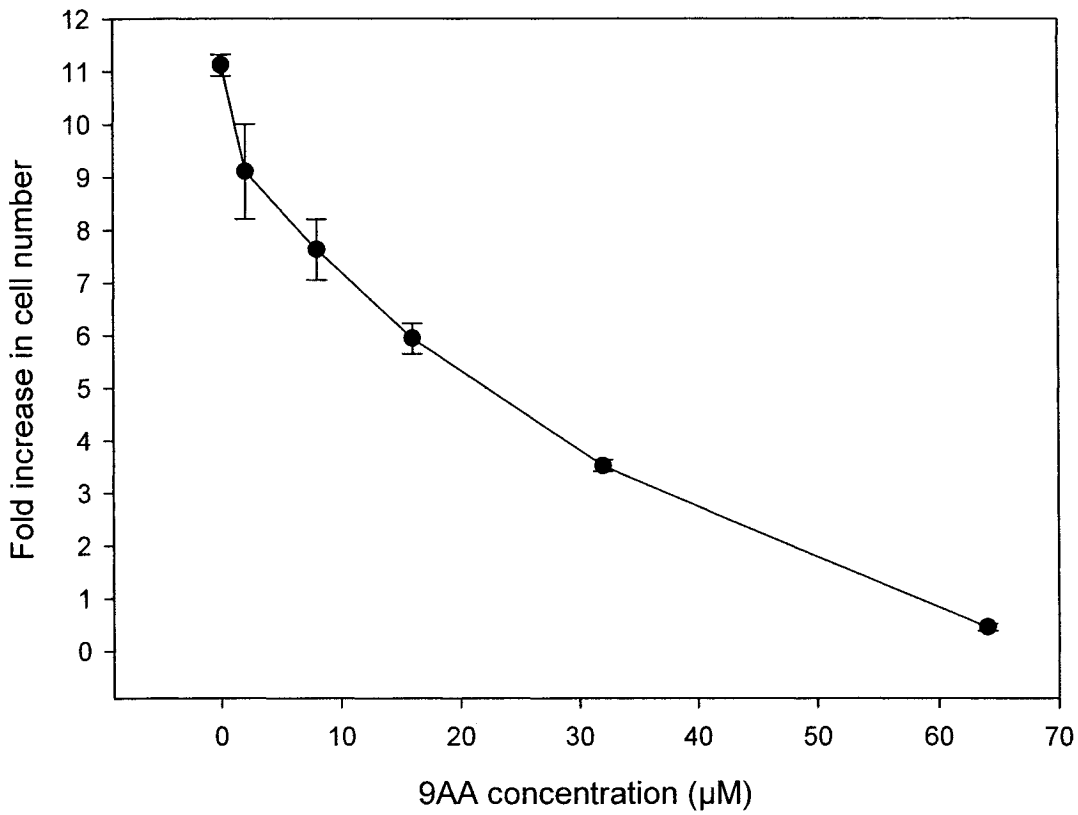
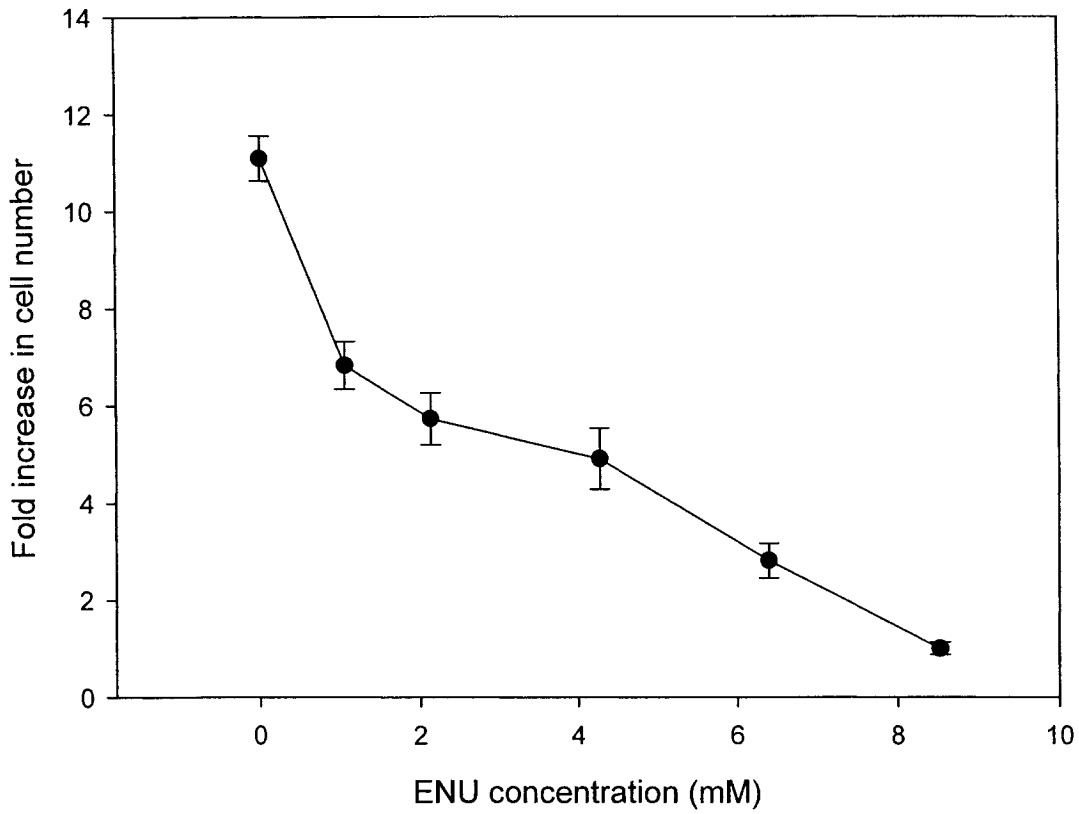

Fold increase in cell number
12
11
10
9
8
7
6
5
4
3
2
1
0
0
10
20
30
40
50
60
70
9AA concentration (μM)

Figure 12. Killing curve of MC2a cells following ENU treatments. Cytotoxicity was determined as total adherent cell counts in MC2a cultures 48 hours after replacement of media containing test agent. Values are expressed relative to the number of cells seeded in each dish prior to treatments.


14
12
10
8
6
4
2
0
Fold increase in cell number
0
2
4
6
8
10
ENU concentration (mM)

Tumour promoter TPA (Figure 13) had limited toxicity in MC2a mouse fibroblast cells. At the lowest concentration, 100 nM of TPA, growth was insignificantly enhanced, at concentrations between 325 up to 1300 nM, growth was decreases by less than two fold (less than one generation). TPA was used at non toxic concentrations with 1300 nM being the highest. At a double concentration of 2600 nM, growth decreased to around 6 fold increase.

In MC2a, BPDE (Figure 14) at 2 μM decreases growth by half, and at 3μM, the highest concentration used in these experiments, growth went down to a quarter of the fold increase observed in the controls. At 4.5 μM cells barely double and at 6 μM, BPDE became extremely toxic.

Figure 13. Killing curve of MC2a cells during TPA treatments. Cytotoxicity was determined as total adherent cell counts in MC2a cultures 48 hours after replacement of media containing test agent. Values are expressed relative to the number of cells seeded in each dish prior to treatments.

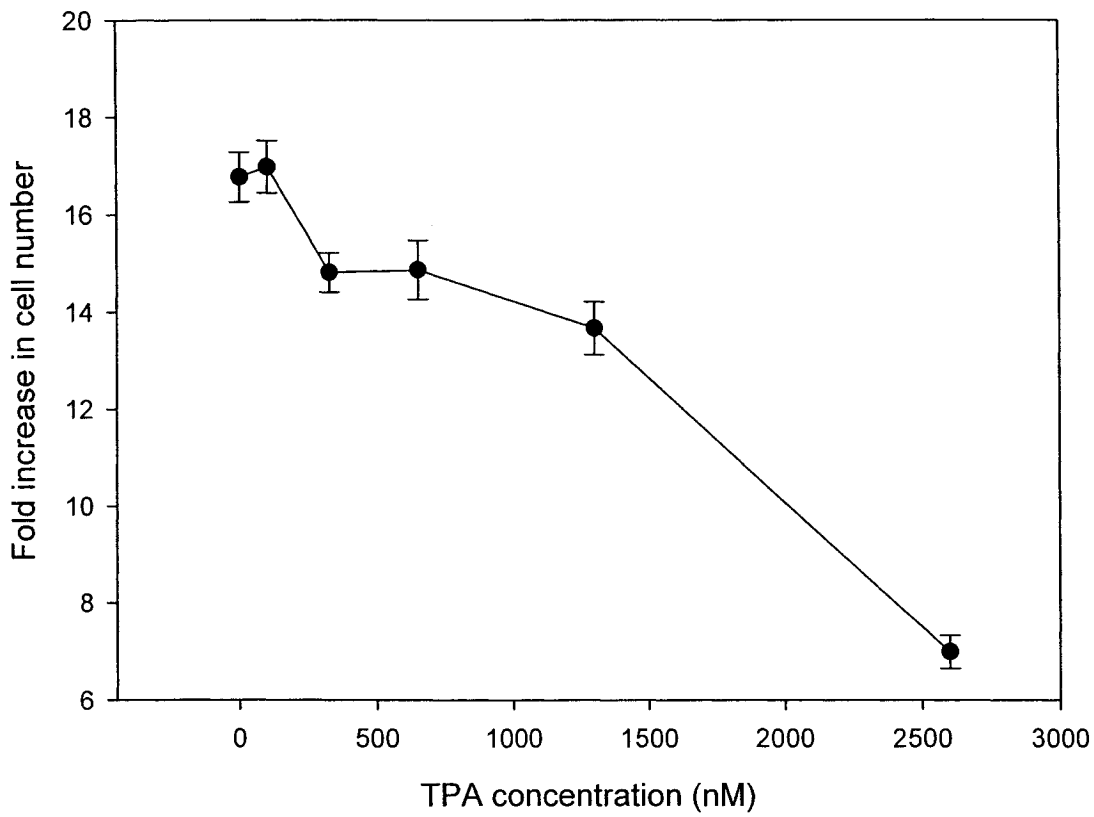
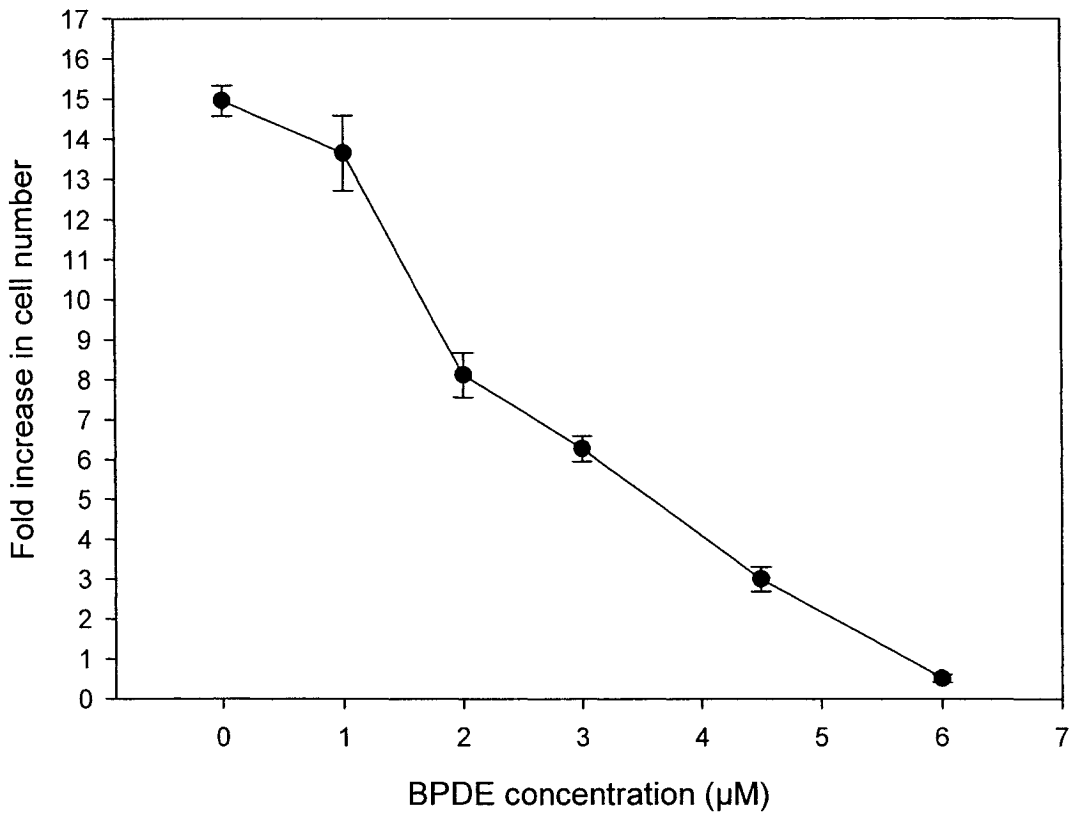

20
18
16
14
12
10
8
6
Fold increase in cell number
0
500
1000
1500
2000
2500
3000
TPA concentration (nM)

Figure 14. Killing curve of MC2a cells following BPDE treatments. Cytotoxicity was determined as total adherent cell counts in MC2a cultures 48 hours after replacement of media containing test agent. Values are expressed relative to the number of cells seeded in each dish prior to treatments.


Fold increase in cell number
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
0
0
1
2
3
4
5
6
7
BPDE concentration (μM)

3.2.2 Validation of the flow cytometry analysis.

To ensure quick, accurate and objective counts, frequency of green fluorescent mutants was determined by flow cytometry. Briefly, when the cell suspension travels through the flow cytometer, the individual cells are illuminated by a light source, the scanner within the instrument records measurements of light scattered and fluorescence emitted by these cells (Radcliff and Jaroszeski, 1998). Both untransfected cells and constitutively green fluorescent transfectants were used for the initial optimization of the parameters involved for flow cytometry. PBS without calcium, but supplemented with EDTA and glucose was determined to be an optimal buffer for suspending the cells for counts since the solution minimizes cell clumping and does not damage the cells prior to analysis.

For practical reasons, including availability of the flow cytometer and the amount of samples to be counted, it was necessary to develop an effective way to conserve samples without affecting results prior to cell counts. The viability of the cells and stability of the GFP signal after freezing were measured for a 2 week period (which represents the longest amount of time cells were kept prior to counting). A count for freshly harvested green cells was obtained (average of 6 readings: 95.17 % green), and aliquots of this population were frozen away and measured after 1 week (average of 6 readings: 95.07 % green) and after 2 weeks (average of 6 readings: 93.30 % green). There was no significant change in GFP distribution over this 2 week time period but there was a slight increase in the amount of cell debris in the sample as determined by

SSC/FSC. To ensure consistency in the results, for each set of treatments, all samples were measure on the same day.

3.2.3 Mutagenic effects due to chemical treatments

The numbers of spontaneous mutants in transfected MC2a cells varied between insert sequence with the highest frequency in cells with murine marker Ms6-Hm1 (Hm1), the dinucleotide repeat (CA), the tetranucleotide repeat (AAAG) and the human aromatase (Hucy). In contrast, the very low spontaneous mutation frequency for cells containing plasmids with trinucleotide or hexanucleotide repeats, roughly 1 in 10000 cells, were similar to the frequency in those containing the non-repeat insert demonstrating that, as expected, TRS of 3 or multiples of 3 nucleotides do not readily generate frame shift mutations.

When treated with test chemicals, the mutation responses were greatest for cells containing the murine marker Ms6-Hm1 (Hm1), the dinucleotide repeat (CA), the tetranucleotide repeat (AAAG) and the human aromatase (Hucy) sequences which displayed statistically significant exposure response for all 5 positive chemicals tested. Further, cells containing any of these 4 plasmids were more sensitive in detecting mutagenic activity than any of the trinucleotide (Trio) or hexanucleotide (Per6 and Hexa) repeat plasmid-containing cells, since all tended to show significantly increased levels of mutants at lower concentrations, considered across all test agents. The negative control plasmid NLA displayed limited significant changes in mutation frequency with the substances tested . There was no evidence of green fluorescence in untransfected parent cells nor any change in the proportion of cells expressing GFP in cell populations stably transfected with in-frame GFP regardless of any chemical exposure (data not shown).

Intercalating agent 9AA gave the highest mutation frequencies (5 to 27%), when considered across the di-, tetra- and pentanucleotide constructs (Figure 15 and 16) closely followed by oxidative stressor H_2O_2 (5 to 16%) (Figure 17 and 18) then the alkylating agent ENU (~2%) (Figure 19), and to the least extent, BPDE (~ 1%) (Figure 20). A significant response was also observed after exposure to tumour promoter TPA (~1.5%) (Figure 21). These agents generated the same order in the magnitudes of responses in the tri- and hexanucleotide contructs. The plasmids Hm1 and Trio displayed similar mutation responses to 9AA and H_2O_2 in both MC2a and C3H10T1/2 cells (Figures 15-18), although the response was slightly but consistently higher in the mismatch repair deficient cell line MC2a than in C3H10T1/2.

Figure 15: The frequency of green fluorescent mutant cells in MC2a cell populations after H_2O_2 treatments. A/ for cells harboring plasmids CA, AAAG, Hucy or Hm1, B/ for cells harboring NLA, Trio, Per, or Hexa. *Error bars*, SDs. Statistical significance is represented with * ($p<0.0001$), ‡ ($0.0001<p<0.01$) and † ($0.01<p< 0.05$)

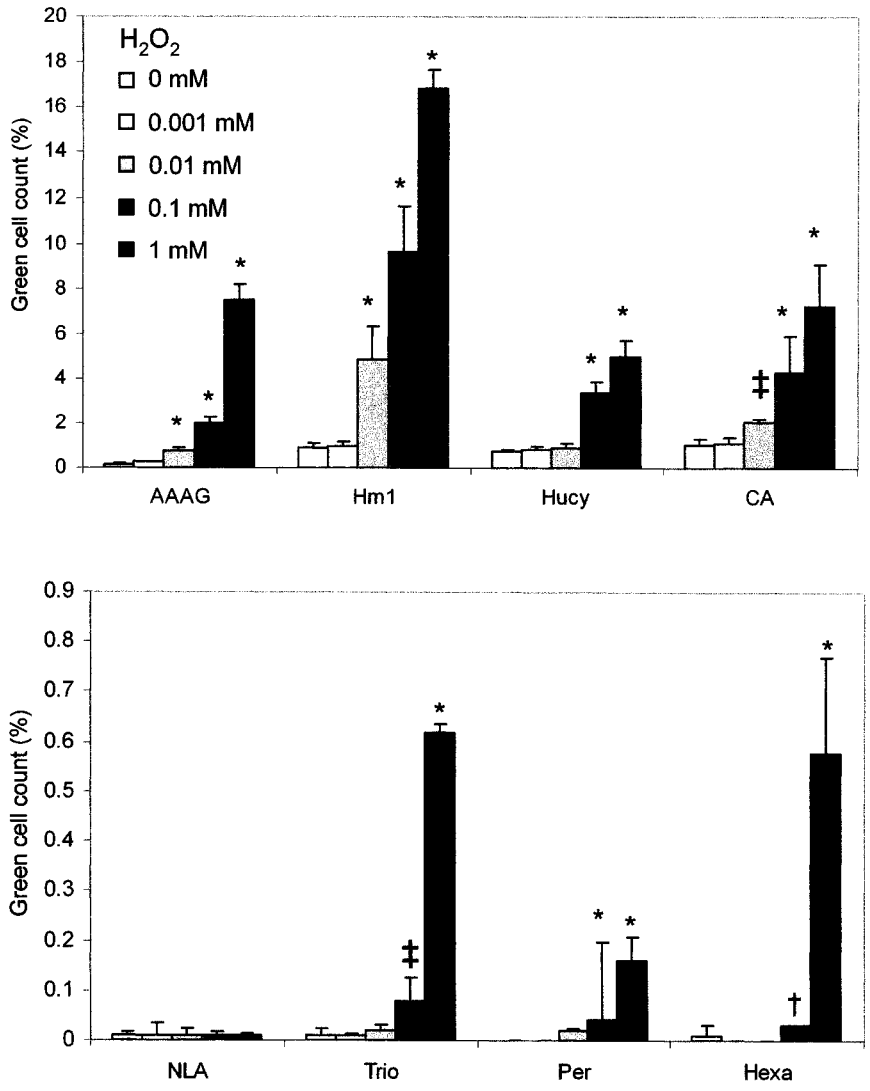

H_2O_2
0 mM
0.001 mM
0.01 mM
0.1 mM
1 mM
Green cell count (%)
AAAG
Hm1
Hucy
CA
NLA
Trio
Per
Hexa

Figure 16: The frequency of green fluorescent mutant cells in C3H10T1/2 cell populations after H_2O_2 treatments. A/ for cells harboring CA, AAAG, Hucy or Hm1, B/ for cells harboring NLA, Trio, Per, or Hexa. *Error bars*, SDs. Statistical significance is represented with * ($p<0.0001$), ‡ ($0.0001<p<0.01$) and † ($0.01<p< 0.05$)

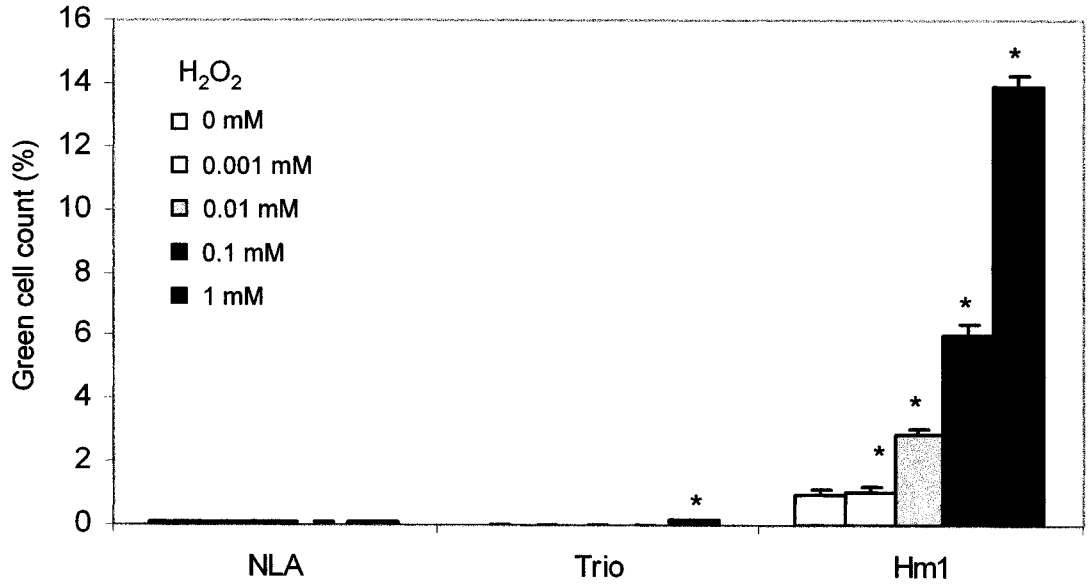
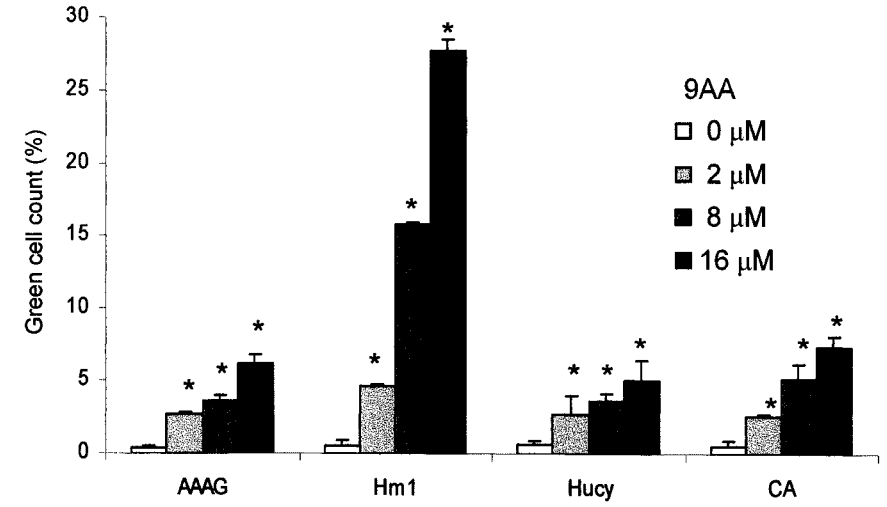

16
14
12
10
8
6
4
2
0
Green cell count (%)
H_2O_2
0 mM
0.001 mM
0.01 mM
0.1 mM
1 mM
*
NLA
Trio
Hm1

Figure 17: The frequency of green fluorescent mutant cells in MC2a cell populations after 9AA treatments A/ for cells harboring CA, AAAG, Hucy or Hm1, B/ for cells harboring NLA, Trio, Per, or Hexa. *Error bars*, SDs. Statistical significance is represented with * ($p<0.0001$), ‡ ($0.0001<p<0.01$) and † ($0.01<p< 0.05$)


30
25
20
15
10
5
0
Green cell count (%)
9AA
0 μM
2 μM
8 μM
16 μM
AAAG
Hm1
Hucy
CA

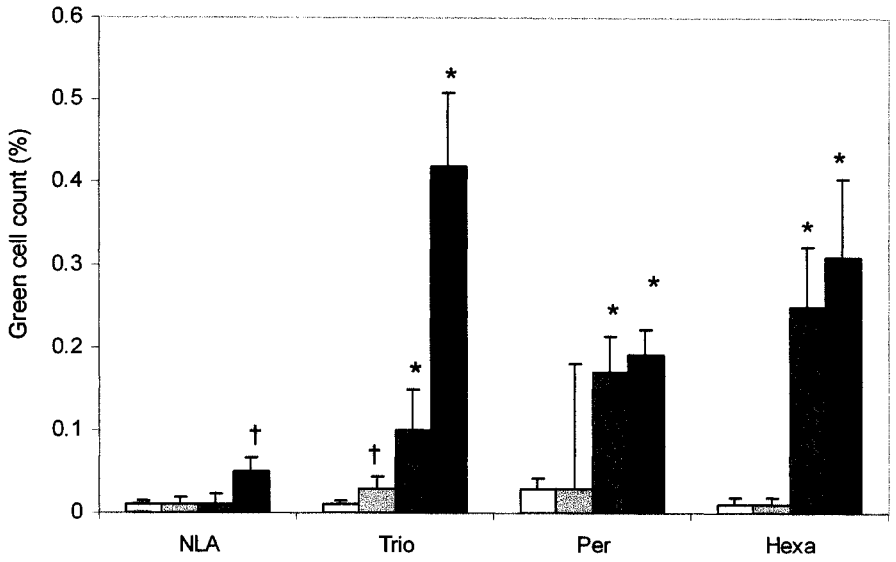

0.6
0.5
0.4
0.3
0.2
0.1
0
Green cell count (%)
NLA
Trio
Per
Hexa

Figure 18: The frequency of green fluorescent mutant cells in C3H10T1/2 cell populations after 9AA treatments A/ for cells harboring CA, AAAG, Hucy or Hm1, B/ for cells harboring NLA, Trio, Per, or Hexa. ***Error bars*****, SDs. Statistical significance is represented with * ($p<0.0001$), ‡ ($0.0001<p<0.01$) and † ($0.01<p< 0.05$)**

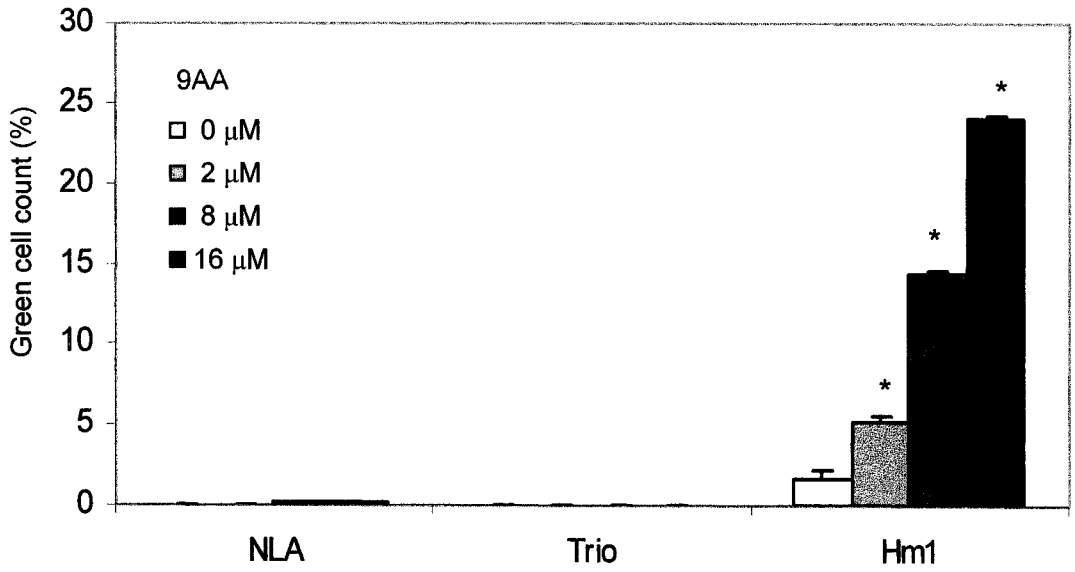

30
25
20
15
10
5
0
Green cell count (%)
9AA
0 μM
2 μM
8 μM
16 μM
*
*
*
NLA
Trio
Hm1

Figure 19: The frequency of green fluorescent mutant cells in MC2a cell populations after ENU treatments A/ for cells harboring CA, AAAG, Hucy or Hm1, B/ for cells harboring NLA, Trio, Per, or Hexa. *Error bars*, SDs. Statistical significance is represented with * ($p<0.0001$), ‡ ($0.0001<p<0.01$) and † ($0.01<p< 0.05$)

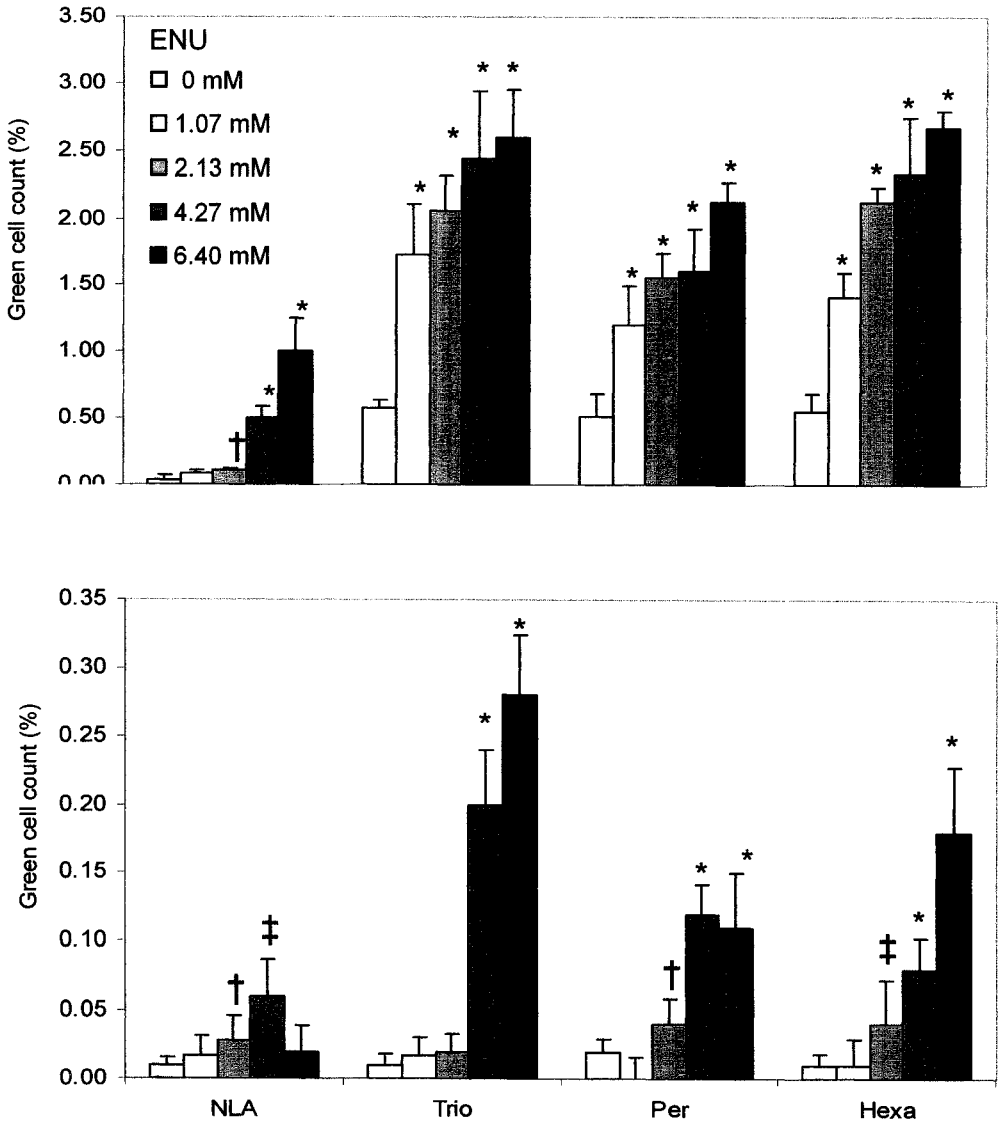

ENU
0 mM
1.07 mM
2.13 mM
4.27 mM
6.40 mM
Green cell count (%)
3.50
3.00
2.50
2.00
1.50
1.00
0.50
0.00
Green cell count (%)
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
NLA
Trio
Per
Hexa

Figure 20: The frequency of green fluorescent mutant cells in MC2a cell populations after TPA treatments A/ for cells harboring CA, AAAG, Hucy or Hm1, B/ for cells harboring NLA, Trio, Per, or Hexa. *Error bars*, SDs. Statistical significance is represented with * ($p<0.0001$), ‡ ($0.0001<p<0.01$) and † ($0.01<p< 0.05$)

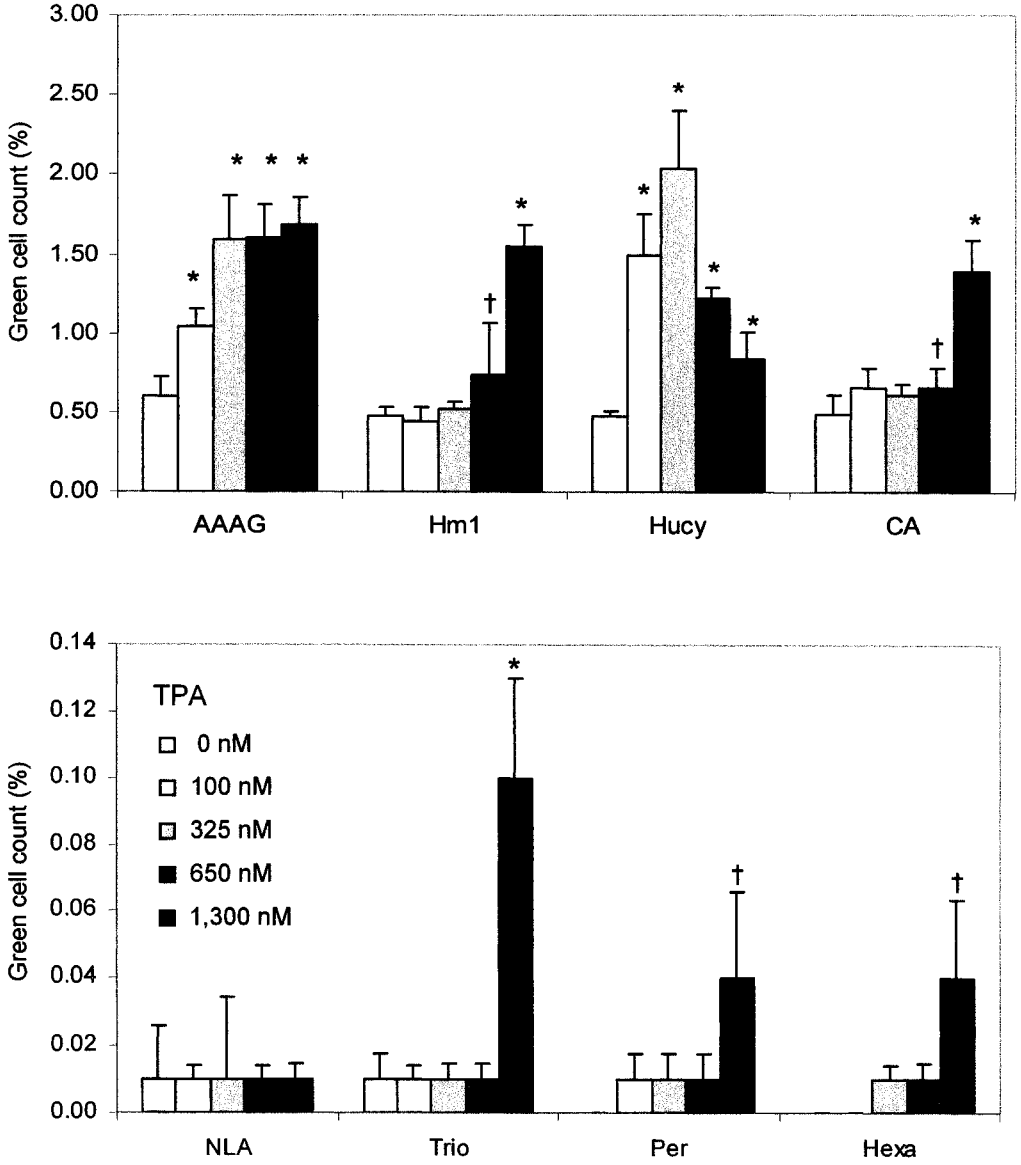
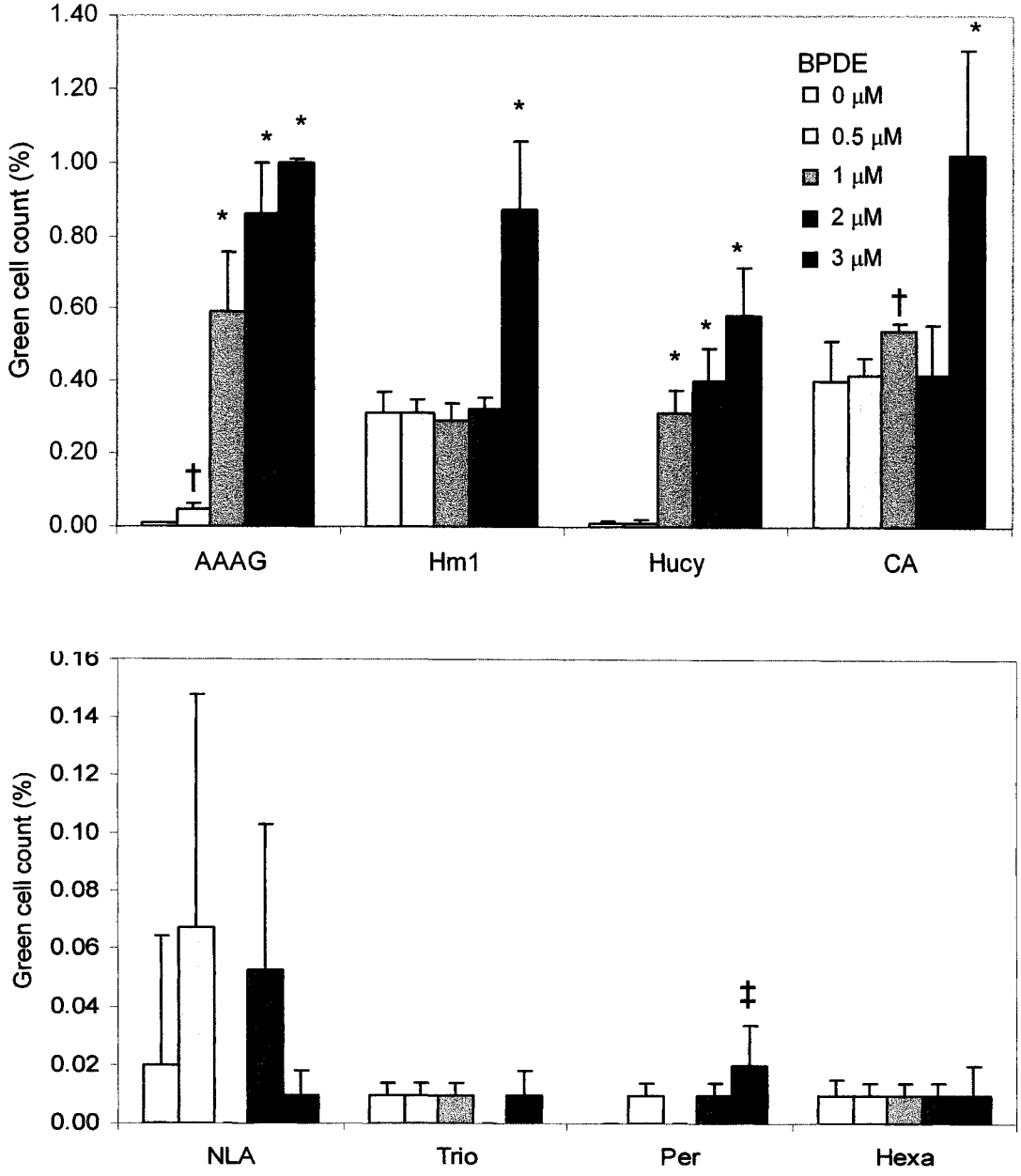

Green cell count (%)
3.00
2.50
2.00
1.50
1.00
0.50
0.00
AAAG
Hm1
Hucy
CA
TPA
0 nM
100 nM
325 nM
650 nM
1,300 nM
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
NLA
Trio
Per
Hexa

Figure 21: The frequency of green fluorescent mutant cells in MC2a cell populations after BPDE treatments. A/ for cells harboring CA, AAAG, Hucy or Hm1, B/ for cells harboring NLA, Trio, Per, or Hexa. *Error bars*, SDs. Statistical significance is represented with * ($p<0.0001$), ‡ ($0.0001<p<0.01$) and † ($0.01<p< 0.05$)


Green cell count (%)
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
BPDE
0 μM
0.5 μM
1 μM
2 μM
3 μM
AAAG
Hm1
Hucy
CA
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
NLA
Trio
Per
Hexa

A/ Effects of H_2O_2 exposure

Oxidative stress from hydrogen peroxide was generated from 1 hour treatments in PBS with varying concentrations of H_2O_2. MC2a cells harboring the plasmids CA, Hm1, AAAG and Hucy (Figure 15) showed a statistically significant exposure response over all growth inhibitory concentrations. Similar significant responses were also observed for plasmid Hm1 in C3H10T1/2 (Figure 16). In the case of the other plasmids in MC2a, Trio, Hexa and Per (Figure 15) all showed a trend of increased mutation frequency with increasing H_2O_2 exposure. In C3H10T1/2 cells, Trio gave a statistically significant responses restricted to the highly toxic concentration, 1mM H_2O_2 (Figure 16).

B/ Effects of 9AA exposure

Results for both positive (green fluorescent cells) and untransfected controls showed that the 9AA treatments did not significantly affect GFP levels in the GFP expressing population or generate any green fluorescence in non-transfected population for either C3H10T1/2 or MC2a (data not shown). The exposure-response pattern for 9AA was generally high in the cells transfected with repeat constructs except where the repeats were of 3 or 6 bases. Mutation frequency in CA, Hm1, AAAG and Hucy in MC2a was significantly higher for treatments ($p<0.0001$) than for vehicle controls at all concentrations (Figure 17), including at 2 μM, which did not significantly affect culture growth (Figure 10). Hm1 was similarly affected in C3H10T1/2 cultures (Figure 18). For Trio, Per and Hexa, there was also a clear trend towards increased mutation frequency with increasing 9AA concentration and to a much lesser extent a slight increase in the non-repeat control after treatments in MC2a. Similar trends were observed for NLA and Trio in the C3H10T1/2 (Figure 18). Other constructs were not tested in C3H10T1/2 cells.

C/ Effects of ENU exposure

Populations carrying the CA, Hm1 and Hucy repeats displayed mutation frequencies that were statistically significant for all ENU treatment levels (Figures 19). For AAAG, a similar trend was observed but statistical significance was not observed at the lowest concentration (1.07 mM). Trio, Per and Hexa plasmids all showed highly statistically significant increases in mutation frequency at the highest concentrations (4.24 and 6.36 mM) of ENU whereas NLA gave significant increases at the penultimate and one lower concentration (Figure 19).

D/ Effects of TPA exposure

There was an appearance of increasing mutation frequency with increasing concentration of TPA in cells transfected with any of the repeat-plasmids. Significant responses were observed in cells containing AAAG and Hucy plasmids across all concentrations tested while CA and Hm1 gave mutants at the highest concentrations tested (Figure 20). Cells containing the Trio, Per and hexa constructs also showed significant increases in mutation frequency at the highest concentration of TPA. AAAG and Hucy showed significant reversion at even the lowest concentration tested (100 nM) that had no effect on cell culture growth. At 1300 nM, the culture cell counts were reduced only 18% relative to the control.

E/ Effects of BPDE exposure

Benzo-*a*-pyrene (B*a*P) showed high fluorescence in non-transfected cells in the yellow range which can potentially mask GFP signal. Consequently, BPDE, a bulky adduct that is a known mutagenic metabolite of B*a*P was used instead since it did not display interfering fluorescence. BPDE did not significantly affect GFP fluorescence in

the GFP expressing population or lead to the detection of false positives post treatment. All cells harboring repeat constructs showed a trend (Figure 21). Statistically significant increases in mutant frequency were observed in response to the 3 highest concentrations (1, 2 and 3 μM) tested concentrations for AAAG and Hucy, but the magnitudes of the responses were 6-fold less than those measured for 9AA and H_2O_2. Hm1 and CA constructs were significantly positive only at the highest concentration. NLA, Trio and Hexa constructs were unresponsive, but at 3 μM BPDE, the Per construct gave significant response due to the low background (0 mutants) in these particular experiments. Although BPDE was highly cytotoxic, overall, the mutation frequencies from BPDE treatments were lower than H_2O_2. Although GFP expression was unaffected in the green fluorescent positive control, it is possible that these low mutation frequencies in the test constructs might be due to BPDE affecting protein folding genes. Since proper folding is essential for the detection of GFP fluorescence: only properly folded GFP protein will be detected. Alternatively, although BPDE generated a significant response, the bulky adduct might simply not induce as many frameshift mutations as the H_2O_2 oxidative stress from did.

3.3 Confirmation of frameshift mutations within GFP mutants

To verify that the expression of GFP was not the result of transcriptional or post-transcriptional modification but actually the result of DNA frameshift mutations, the repeat region of green fluorescent clones was sequenced. To fully validate this assay, large scale sequencing would be required but for the purposes of this study, 8 untreated non-green control clones and 8 9AA treated green fluorescent clones from populations of MC2a cells transfected with the Hm1 constructs were selected. As mentioned above

(section 3.1.3), unlike the C3H10T1/2 the MC2a cells can stably express GFP thus were more able to generate green fluorescent clonal cultures. Since the Hm1 plasmid has the highest mutation frequency and 9AA generated the highest mutation level, it was the most efficient combination to provide green fluorescent clones for sequencing. For each individual clone, DNA was extracted, and a 786 bp fragment between the CMV and GFP region was amplified and ran on an agarose gel to confirm the presence of the product. For all 8 control samples, sequencing results corresponded to the expected sequence suggesting that no PCR artifacts were generated in the sequenced region (see chromatograms in Appendix D). For all 8 green fluorescent mutants, results revealed loss of a single repeat unit and possibly occasional substitutions. The latter base changes would need to be confirmed by sequencing the opposite strand (the amount of sample limited the sequencing to one reaction). The loss of a 5 bp repeat unit is consistent with the green fluorescent phenotype. Indeed, the Hm1 plasmid has a 68 bp insert that places the GFP in the - 1 reading frame, losing 5 bp would leave the insert with 63 bp thus restoring the reading frame. The loss of a single pentanucleotide repeat unit is the simplest mutation that can occur within the Hm1 population which might explain its high occurrence. These results agree with the slippage model. Based on 8 samples it is difficult to say if that is the only or the most common mutation occurring in the 9AA treated populations but this data a least confirmed that green fluorescent mutants did reflect the occurrence of frameshift mutations at the DNA level within the insert.

Chapter 4

4.0 Discussion

The main objective of this work was to investigate the potential utility of TRS upstream of a GFP reporter system as a target in creating an assay of genotoxic stress in murine cells. To achieve this goal, we generated several different gene constructs in order to identify those that could function as the most sensitive markers when tested with representatives of standard classes of chemicals. The assay measures the frequency of + 1 frameshift mutations in a population of rodent cells stably transfected with reporter plasmids. Clearly, all possible combinations of DNA repeats could not be tested, so we selected relevant, representative repeat units of different sizes: dinucleotide CA, trinucleotide Trio (CAG), tetranucleotide AAAG, pentanucleotide Hm1 (GGGCA), hexanucleotides Per6 (ACAGGC) and Hexa (ACCTGC), as well as a human gene sequence Hucy containing repeated tetranucleotide motifs (TTTA). Mutants were detected by flow cytometry as fluorescent, green cells producing GFP. The presence of non-specific fluorescence generated from the chemical treatment, a potentially confounding factor, was controlled by examining the effects of each treatment on untransfected cells from the same parent cell population. Other controls included testing for effects on GFP expression or detection when green fluorescent cells harbouring the positive control plasmid were treated and measuring the mutation frequency of non-repetitive DNA, using plasmid NLA

A secondary objective of the current study was to determine whether repeat sequences were hot spots for types of mutations other than mutations predicted by Streisinger's slippage model. Indeed, the perfect gain or loss of trinucleotide (Trio) and hexanucleotide (Hexa and Per6) repeat units would not be detected in this assay and therefore lower frameshift mutation frequencys were expected in cells harbouring these

constructs. However, comparing the mutation levels in these repeats versus that of the NLA construct could help determine whether these tandem repeats are relative hotspots for other types of mutations, e.g., mutations resulting from non-homologous recombination, compared to non repeat sequences.

4.1. Transfections

To achieve stable transfection of cells, the plasmid DNA must integrate into the genome, the sequence of interest must be kept intact or with limited modifications and the reporter gene must remain functional (DNA methylation and sequence damage during integration could lead to a non-functional gene). To increase the frequency of integration of the foreign reporter gene DNA into the cell's genome, the plasmid was digested with *Bsa I* (see Figure 1, map of plasmid) at a strategic restriction site upstream of the coding gene that maximizes the chances of keeping the GFP reporter gene and neomycin resistance gene intact. After 2 weeks of antibiotic selection, only stably transfected cells remain. Plasmid DNA was expected to be integrated at random sites through the cells' genome. Since the site of integration and copy number could affect the level of GFP expression, instead of using a single transfected clone, populations of cells were used. When cells are pooled, an average response of the cells was measured thus limiting variation between cell populations. The exact copy number of the plasmid integrated in the cells genome was not determined. However, the results from Northern analyses using a GFP probe showed that the GFP sequence was equally transcribed into RNA in all transfected cell populations. It is unclear why the level of hybrid GFP transcription was lower in C3H10T1/2 than in MC2a cells. Lower levels of transcription in C3H10T1/2 for the reporter gene could be due to gene silencing or to cell metabolism. Further experiments

measuring the expression level of other reporter gene might yield some answers but this is outside the scope of this study

4.2. Hydrogen peroxide

The term "oxidative stress" refers to a state where the cells are exposed to reactive oxygen species (ROS) that exceed the protective capacity of the cellular antioxidant defence system (Halliwell and Whiteman, 2004). An imbalance can occur when there is a decrease in antioxidant enzyme activity arising from heritable defects in genes coding for these enzymes, e.g. glutathione peroxidase, as well as from induced effects such as nutritional deficiencies. In addition, an increase in ROS level due to increased cellular metabolism can increase oxidative stress as can exposure to toxins that are ROS or precursors of ROS. Oxidative stress can alter cellular metabolism, e.g., upregulation of antioxidant defences, cause oxidative damage to DNA, proteins and other biomolecules and even lead to cell death (Halliwell and Whiteman, 2004). Consequently, exposure to oxidative stress is a major health concern and molecular oxidative damage is associated with premature aging (Dufour and Larsson, 2004) atherosclerosis, diabetes, and obesity (Sies et al., 2005) as well as cancer and neurodegenerative diseases (McCord and Edeas, 2005).

In the current study, we used H_2O_2 to increase intracellular ROS levels as it is the most prevalent endogenous precursor of ROS in cells, is an intermediate in many oxidative pathways (Chang et al., 2003) and can readily cross cellular membranes (Klaunig and Kamendulis, 2004). An effective endpoint to measure oxidative stress is frameshift mutation in repeat sequences, as demonstrated by (Gasche et al., 2001) and (Slebos et al., 2002). In *E. coli*, oxidative stress from H_2O_2 has been reported to generate

DNA damage that can lead to increased mutation, rates particularly in repeat sequences (Jackson and Loeb, 2000). In the bacterial genome, expansions and deletions of dinuceotide repeats were most likely generated during repair of oxidative DNA damage, e.g., strand breaks. The results obtained from the GFP reporter assay in murine cells reported herein, were consistent with their proposed mechanism for the generation of mutations. Higher mutation frequencys were observed in constructs containing CA, AAAG, Hm1 or Hucy constructs, whereas lower mutation frequencys were observed in trinucleotide and hexanucleotide repeat constructs where perfect deletions or expansions of single or multiple repeat units do not result in frameshift mutations. The lowest mutation levels occurred in the non-repeat plasmid NLA since, lacking repeats, it is less prone to DNA slippage during repair. This result may suggest that DNA strand breaks are more frequent in TRS regions or that ROS-modified intermediate DNA structures are more stable in repeat regions than non-repetitive sequences (Gasche et al., 2001). The mutations observed could be the result of faulty repair of ROS induced base damage and loss (Marnett, 2000) or non-homologous end-joining during repair of the DNA strand breaks commonly observed in cells arrested in the cell cycle following oxidative stress (Heidenreich et al., 2003).

Hydrogen peroxide effectively induced frameshift mutations in both MMR proficient C3H10T1/2 and MMR deficient MC2a cells harbouring reporter plasmids. Significant concentration dependent responses were observed in the most sensitive constructs CA, AAAG, Hm1 and Hucy for all concentrations of H_2O_2, as well as in Trio at the highest toxic concentration in both cell lines. The mutation frequency observed in both murine cells for the dinucleotide repeat in the current study are similar to the values

reported at various concentrations of H_2O_2 in MMR deficient human colon cancer cells harbouring different GFP plasmid constructs containing CA and AAAG repeats (Gasche et al., 2001). Slebos et al.(Slebos et al., 2002) also observed an increase in frameshift mutations for tetranuceotide and dinucleotide repeats subjected to oxidative stress by the addition of t-butyl hydrogen peroxide. The endogenous Ms6-hm1 ESTR locus and other similar loci are also very responsive to oxidative stress. Co-treatment with xanthine and xanthine oxidase generated increased mutation frequencies measured in a DNA fingerprint study of C3H10T1/2 cells (Parfett and Healy, submitted).

4.3. 9-amino-acridine

In mammalian cells, the 9AA molecule intercalates between DNA bases where it acts as a catalytic inhibitor of DNA topoisomerase II by blocking DNA cleavage by means that do not stabilize 'strand-passing' DNA-TopoII reaction intermediates (Andoh and Ishida, 1998; Snyder, 2000). 9AA has been observed generate DNA strand breaks in treated cells (Henderson, 1998), but conflicting results from several in vitro mammalian cell mutagenicity tests indicate that it is, at best, a weak mutagen and clastogen, consistent with its ability to cause low levels of strand breaks (Ferguson and Denny, 1991). In bacterial systems, 9AA is a strong frameshift mutagen, acting preferentially to create (-1) frameshifts, at frequencies up to 0.1%, within short homopolymeric dG runs (Hoffmann et al., 2003). Stabilization of slipped-mispaired bases by 9AA during DNA replication is thought to shift the equilibrium between proper and improper base pairing to the latter. Using a xanthine-guanine phosphoribosyl transferase forward mutation assay in CHO cells, Ferguson et al., (Ferguson et al., 2000) demonstrated 9AA-induced mutations at low frequencies (4×10^{-6}) following exposure to a highly toxic

concentration. The majority of the mutations were found to be +1 and +3 frameshifts within or near to an area of high GC content.

The plasmid constructs most responsive to 9AA treatment in the current study were CA, AAAG, Hm1 and Hucy, with mutation frequencies in MC2a cells ranging from 5 to 27%. These results indicate that the di- tetra- and pentanucleotide sequences were hypermutable, both with respect to the heteropolymeric NLA construct (0.05% at the top, toxic concentration) and with respect to mutation frequencies in endogenous genes mentioned above. The other trinucleotide and hexanucleotide constructs also were reverted by 9AA, but gave lower frequencies in the range of 0.2 to 0.4%.

Sequencing data from the 9AA-induced mutants recovered from a treated cell population transfected with the Hm1 construct showed that all frameshifts were created by (-5) shifts due to loss of single GGGCA units from $(GGGCA)_{12}$ repeats. It seems likely that induced replicative slippage due to reduced fidelity of DNA replication across this sequence would more easily account for both the extremely high rate of occurrence of reversion and the consistent loss of single 5 bp repeats, than would targeted, topoisomerase II inhibition-mediated events. Whichever model is correct, it must also encompass the ability of 9AA to induce this type of frameshift at low, non-toxic concentrations (2 μM).

4.4 Ethyl nitrosourea

DNA damage from alkylating agents has been associated with carcinogenesis and mutagenesis. ENU is often used as a positive control in mutation studies since it is a powerful monofunctional ethylating agent that can induce mutations in a large spectrum of mutation assays (Shibuya and Morimoto, 1993). ENU generally ethylates DNA at O^6-

guanine and O^4-thymine that can result in base substitutions mainly transitions from GC to AT and to a lesser extent AT to GC (Eckert and Hile, 1998). Frameshift mutations from exposure to ENU (Watson et al., 1998) and other alkylating agents (Slebos et al., 2002) have also been observed particularly in repeat sequences but to a variable degree depending on the study. Results from experiments involving *in vitro* DNA synthesis with DNA polymerases α and β (Eckert and Hile, 1998) suggest that frameshift mutations occur during replication slippage from misalignment of mispaired DNA due to adducts formed by ENU treatments.

Our results showed a significant exposure response for all concentrations in CA, Hm1 and Hucy but only at the 2 highest concentrations for AAAG. The remaining constructs, NLA, Trio, Hexa and Per were less sensitive to ENU treatments contrary to what was predicted by results from Ecker and Hile (Eckert and Hile, 1998) which suggests that frameshifts from exposure to ENU were equally occurring in non-repetitive DNA and repetitive DNA. However those results were obtained from *in vitro* experiments in test tubes as opposed to the experiments conducted by Slebos et al. (Slebos et al., 2002) that were done in the context of GFP reporter constructs in mammalian cells. Our findings concur with the latter and suggest that the increase response in repeat sequence versus non-repetitive DNA might be due to modulation of mutation specificity by repair proteins that are present in the cells and absent in Ecker and Hile's studies (Eckert and Hile, 1998).

4.5 12-*O*-tetradecanoyl-phorbol-13-acetate

The phorbol ester TPA is a very effective tumor promoter and is a recognized carcinogen (Lu et al., 1998) as well as an inflammatory agent (Park et al., 2001). It is generally considered non-mutagenic since in most recognized regulatory assays it yields negative results, e.g., Ames assay, or mixed results, e.g. sister chromatid exchange (Kier et al., 1986; Tucker et al., 1993). But these assays may lack the sensitivity needed to detect mutations in specific highly-mutable mammalian DNA sequences upon which TPA and other putative non-genotoxic carcinogens may act (Jackson et al., 1993). Results from recent studies suggest that exposure to TPA can increase mutation frequencys. Specifically, DNA fingerprinting of C3H10T1/2 cells exposed to TPA shows significantly higher mutation frequencys than untreated controls in genomic restriction fragments that contain Hm1 sequences (Parfett and Healy, submitted).

Our results also revealed significant TPA induced increase of mutation frequencys in CA, Hm1, AAAG and Hucy in the absence of toxicity. Although significant exposure response was observed for all concentrations, the level of response was much lower than in the case of the other mutagens and response in the other constructs was very limited. However, results from the GFP assay do indicate that the assay is sensitive to putative non-genotoxic mutagens and gives results that concur with mutation frequencys of the endogenous Hm1 locus in DNA fingerprinting (Parfett and Healy, submitted).

The specific mechanisms involved in increased mutation frequencys following TPA exposure are unclear. TPA could indirectly promote increased mutation frequencys by increasing cellular growth rate thus increasing cell metabolism and endogenous oxidative stress. Alternatively, TPA could promote the selection of genetically unstable

cells with "aggressive" growth phenotypes (Hornia et al., 1999). Depletion of cellular levels of certain isoforms of protein kinase C (PKC) can also occur from extended exposure to TPA (Lu et al., 1998). Decreased levels of PKC can reduce expression of MMR protein MSH2 (Humbert et al., 2003; Humbert et al., 2002)to further impair the MMR system in MLH1-deficient MC2a cell. Effects on MMR and other DNA repair pathways may affect cells' ability to repair damage from endogenous oxidative stress arising from cellular metabolism and other exogenous genotoxic stresses is reduced and the mutation frequencys are increased.

4.6 Benzo(a)-pyrene diol epoxide

Polyaromatic hydrocarbons represent a common pollutant formed during incomplete combustion from a wide variety of point sources such as coal fired generators, coke ovens, cigarette smoke or cooking, and are found in urban air, industrial sites, etc. (IARC). This class of chemical is of particular concern since many of them, or their metabolites, are carcinogens or mutagens in humans. The most commonly studied is benzo-*a*-pyrene (B*a*P), a ubiquitous carcinogen that can be absorbed through the skin and metabolized to several oxygenated intermediates notably the highly mutagenic BPDE (Boysen and Hecht, 2003). Typically BPDE will bind to purines, especially guanine and generate GC to TA or GC to AT tranversions (de Boer et al., 1998; Meehan and Straub, 1979; Miller et al., 2000). Base substitution is the prevalent type of mutation arising from BPDE exposure (Denissenko et al., 1996; Miller et al., 2000) but frameshift events are frequent enough to be detected in non-repetitive DNA sequences as demonstrated by Zhu et al. (Zhu et al., 1994) in gene reversion mutation assays that exposed the adenine phosphoribosyltransferase (APRT) locus in human fibrosarcoma cells to BPDE.

The formation of BPDE adducts can give rise to frameshift mutations when the adduct causes misalignment of two DNA strands, thus facilitating slippage events during replication (Zhuang et al., 2001).

Our results showed that BPDE exposure led to a significant increase in mutants at all concentrations tested for the CA and Hm1 repeats. For AAAG and Hucy, concentrations of 1μM or higher were necessary to demonstrate significance. At 3μM the highest exposure response was obtained for AAAG but not Hm1. As the Hm1 repeat unit is composed of GGGCA and has a higher GC content than the tetranucleotide AAAG, this result suggests that frameshift mutations might be independent of GC content at least when these two sequences are compared. These results are consistent with the mutation frequencys detected for dinucleotide (CA) and tetranucleotide (AAAG, ATAG, and CAGT) repeats in a GFP reporter, by Slebos et al. (Slebos et al., 2002) where the rate of GFP activation following BPDE exposure was the same for all repeat constructs. The GC contents of the inserts in the present study ranged from 0% in the TTTA repeat of the Hucy construct to 25% in the AAAG construct, 54% in the negative control NLA, and up to 80% in the Hm1 repeat unit. Our results also showed that the trinucleotide and hexanucleotide constructs had no greater response to BPDE treatment, as was found for the non-repeat construct NLA, possibly due to undetectable perfect slippage events in trinucleotide and hexanucleotide repeats. Overall, the frame-restoring mutations in TRS from BPDE exposure seem to be independent of the nature of the covalent binding preference for dG. Mutations in repeat sequences caused by exposure to B[*a*]P were detected at Hm1 repeat sequences in C3H10T1/2 cells through DNA fingerprinting

(Parfett and Healy, unpublished data) that can detect mostly large insertion and deletion mutations.

4.7 Responses of heteropolymeric and of TRS constructs

4.7.1 The non-repeat construct NLA

The non-repeat construct NLA did not have any significant exposure response to any of the treatments; however, there were non-significant trends of increased mutation frequency from 9AA, ENU and BPDE treatments. These results are consistent with general views that heteropolymeric DNAs are generally less prone to frameshift mutations than repeat sequences (Slebos et al., 2002). Indeed, non-homologous end-joining or slippage from strand misalignment is more likely to occur in regions of high sequence similarity or repetitive runs. The NLA sequence was representative of a typical heteropolymeric DNA found throughout coding regions of the genome. Consequently, assays that solely depend on detecting mutations in non-repetitive DNA, and these include most gene reversion assays eg. TK and HPRT assays in mammalian cells (OECD, 1997e), can lack the sensitivity necessary to detect mutagenicity of novel and existing substances that nevertheless have carcinogenic effects (Jackson et al., 1993).

There was a slight, non-significant trend of increased mutation frequency from low concentrations of exposure to some of the substances included in this study. These trends might become significant if more repetitions of the experiments or higher cell counts are performed. These trends might suggest that the response of the more sensitive repeat constructs can be used to predict the fate of low frequency mutation events in heteropolymeric DNA after chronic exposure to low exposures of chemicals. Taking into account that chronic exposure to low levels of pollutants is common and that life

expectancy is increasing, large numbers of putative mutations in non-repetitive DNA can accumulate over time and increase odds of developing harmful mutations. However it should also be noted that repetitive sequences are widespread throughout the genome and many have functional roles (Subramanian et al., 2003) thus the importance to monitor how they are affected by chemicals.

4.7.2 The trinucleotide and hexanucleotide repeat constructs

Trinucleotide and hexanucleotide repeats are the most common repeat sequences found throughout the human genome (Subramanian et al., 2003). Somatic and germline mutations in these repeats are often associated with the development of diseases, consequently they represent an important genotoxin regulatory target. As mentioned earlier, the GFP reporter assay is geared towards the detection of frameshift mutations so perfect lost or gain of repeats would not be detected in these constructs. Nevertheless, we wanted to investigate whether frameshift mutations were more frequent in these repeats than in non-repetitive DNA and if they would be useful markers in a GFP reporter system.

The frequently occurring trinucleotide repeat (CAG) used in the Trio construct can be found in sequences such as in the breast cancer gene 1 (*AIB1*) with an unstable polyglutamine tract (Dai and Wong, 2003) as well as PolgA, a nuclear gene required for mitochondrial DNA repair (Trifunovic et al., 2004). Non-functional or defective versions of this enzyme are associated with premature aging (Trifunovic et al., 2004) and male infertility (Jensen et al., 2004). Significant reversion activity was expressed only at the highest concentration of H_2O_2 and 9AA in cells harbouring Trio. Considering that trinucleotide tandem repeats are widespread throughout coding regions of the human

genome it raises a potential health risk from exposure to such chemicals when key TRS loci are affected by frameshift mutations.

A tandem repeat sequence common to man, mouse, rat and fruit fly is the Per6 locus (ACAGGC) used in DNA fingerprinting. This repeat is also found in mouse repetitive C-terminal binding protein r4A that interacts with the C-terminal of RNA polymerase II (Yuryev et al., 1996). Surprisingly, it showed limited response to any of the chemical treatments and no significant response was obtained at any of the exposures examined. Consequently, it represents a poor target in the context of a GFP reporter assay. The sequence ACCTGC placed in the Hexa construct is one of the most common hexamer repeat sequences in the human genome (Subramanian et al., 2003). It was more responsive than Per and reached a statistically significant exposure response for H_2O_2 at the highest exposure.

Results obtained from these constructs shows that frameshift mutations are more frequent in these tandem repeat sequences than in heteropolymeric DNA but less frequent than in the other TRS constructs. Consequently, this suggests that some fraction of the in/del events were not created by perfect DNA slippage. The higher mutation frequency from H_2O_2, 9AA and ENU treatments could suggest that these frameshifts were mostly likely the result of non-homologous and random in/del events possibly due to chemical specific effects since BPDE did not induce significant numbers of mutants in these constructs even at the most toxic concentrations. Conversely, TPA used in non-toxic or relatively non-toxic concentrations led to significant increase in the number of mutants in cells harbouring the Trio construct.

4.7.3 The di, tetra and pentanucleotide constructs

Stretches of tandem repeats containing the dinucleotide CA are found in many genes, for example in Spot-1, a nuclear localization signal binding protein that interacts with p53 through a domain encoded by (CA)n repeats (Elkind et al., 1995). Since one of the goals of this study was to identify biomarkers that are useful for estimating chemical exposure, the pentanucleotide sequence GGGCA is a logical choice. Indeed, the construct Hm1 has a GGGCA repeat unit found in the mouse *Ms6-hm* locus. This marker is extensively used in mutation assays particularly DNA fingerprinting and single molecule PCR (Parfett and Healy, submitted; Yauk et al., 2002).

Tetranucleotide tandem repeats are also common in both coding and non-coding DNA. In this study, AAAG was selected since it is found to be highly unstable in epithelial cells of smokers as measured by PCR of endogenous AAAG TRS from DNA of human lung tumours (Mao et al., 1994) thus potentially a good marker of stress exposure. Another tetranucleotide sequence of interest is TTTA which is found in human aromatase cytochrome P-450 gene (Dialyna et al., 2004). An 80 bp sequence from this gene was included in the Hucy construct to demonstrate that specific gene sequences that are not purely perfect repeats could be useful in the GFP reporter assay. In addition, Dialyna et al (Dialyna et al., 2004) have reported that some TTTA repeat allelic variants of human aromatase cytochrome P-450 gene are associated with increased risk for breast cancer.

The sensitivity of these constructs compared to that of the non-repeat construct NLA further demonstrates the potential usefulness of mutagenesis at tandem DNA repeats as a biomarker to predict mutation risk. However it should be noted that these

mutations occurred within a reporter gene stably transfected into cells and that endogenous genomic repeat sequences might respond differently to chemical exposure. Indeed, two important factors that can affect the mutation frequency of TRS reporter gene include the site of integration, state of the MMR system and cell line (Zhang et al., 1995).

4.8. Evaluation of the Assay

To ensure that GFP activation, as determined by flow cytometry truly resulted from frameshift mutations in the insert sequence, preliminary sequencing experiments were conducted. Since Hm1 was the most responsive plasmid and 9AA was the most mutagenic treatment and MC2a cells were able to stably express GFP, green clones from this group were selected. Untreated non-green MC2a cells harbouring the Hm1 construct were used for comparison. Sequencing of all 8 PCR products from the controls showed the expected sequence and all 8 green clones showed loss of a single repeat unit. Since all repeat units are identical, it is not possible to determine what specific unit was lost.

In addition, results obtained in these experiments reveal that the assay responds robustly to various classes of mutagenic substances yielding sensitive and reproducible results. This rapid, sensitive GFP reversion assay could potentially be used to screen for a wide range of genotoxic hazards using any desired DNA sequence target of interest and in various cell systems. To obtain the highest sensitivity, the constructs CA, Hm1, AAAG and Hucy are the most promising for use in chemical toxicological screening.

The drawback of this assay is that it only detects frameshift mutations therefore it would not detect events such as substitutions and loss of a multiple of a repeat unit that did not change the reading frame. When testing chemicals with aGFP activation assay the inherent fluorescence of the chemical, as well as potential staining effect it might have on

the cells must be investigated. To get accurate and sensitive measurements, background fluorescence from the cell or the exposure treatment or interaction of the two must be distinguishable from GFP fluorescence.

As for all *in vitro* assays, the results must be kept in perspective. Research insights can be gained from the use of these assays but the results must be compared to what occurs *in vivo* within animals and animal tissues. In the case of the GFP construct, comparisons to endogenous repeat sequences in their original genomic context must be made. As mentioned earlier, our data does match trends observed in single molecule PCR and DNA fingerprints, provides a larger sample size and can be economically done in less time. However, further testing would be required to validate this assay for the genotoxicity screening of many additional classes of chemical treatments under many conditions. Particularly interesting in this regard will be the analysis of the activities of the receptor-mediated carcinogens such as dioxin, phthalates, and growth hormones, which, like, phorbol ester, do not directly interact with DNA but may be able to induce instability via other pathways, e.g., by directly or indirectly affecting function and regulation of MMR proteins and other DNA repair proteins.

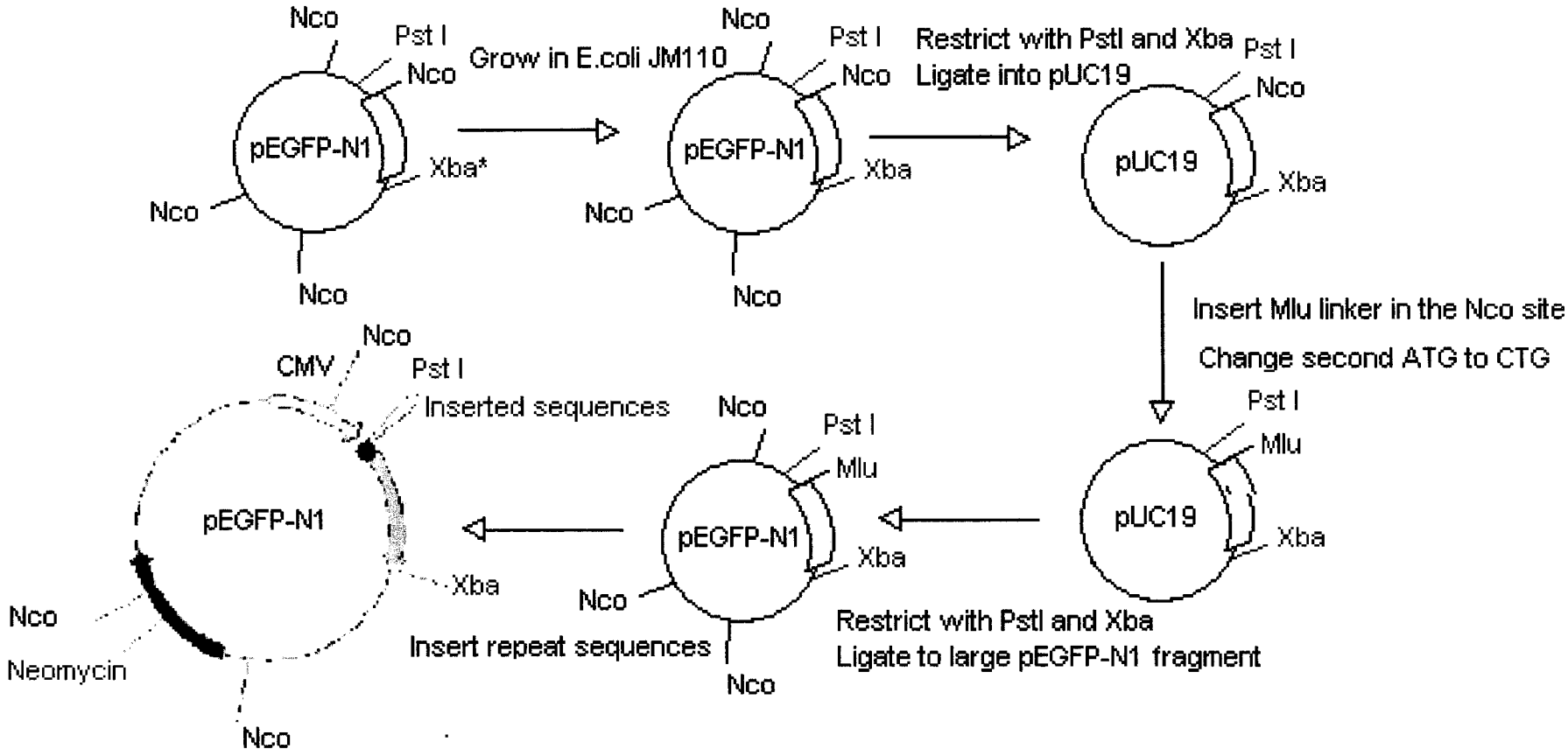
4.9 Conclusion

Results from this work show that GFP reporter assays are effective research tools in studying repeat sequence instability and could prove a promising alternative to genotoxicity screening assessment tools since they are quick, economic, robust and sensitive to a wide range of substances. The sensitivity of these assays has been further validated outside of the work of this thesis. When negative control chemicals acetone and ethanol were tested in MC2a cells carrying the negative control plasmid or Hm1,

the results revealed no significant concentration response from exposure to acetone or ethanol up to toxic concentrations (Healy et al., in preparation). Using frameshift mutations at TRS represents a very relevant endpoint considering that these sequences are widespread throughout the genome (Subramanian et al., 2003) and mutations at these loci can potentially lead to greater health risks. For instance, mutations in the polyglutamine tract of the mitochondrial DNA repair gene PolgA, could lead to non-functional or defective versions of this enzyme that are associated with premature aging (Trifunovic et al., 2004) and male infertility (Jensen et al., 2004).

This assay represents a quick method of detecting frameshift mutagens. Coupled with large-scale sequencing, this assay provides a good tool to characterize the mutational spectra induced from exposure to various environmental stress agents. As demonstrated by Vo et al. (2005), the use of gene silencing and knockout cell lines in combination with GFP reporter constructs can potentially be used to dissect and evaluate the pathways involved in generating TRS mutations. In conclusion, the detection of TRS mutations using the GFP reporter assays represents a powerful tool for the detection of mutations and a very pertinent approach to addressing mechanistic questions.

Appendix A. Flow chart of the cloning strategy


Nco
Pst I
Nco
pEGFP-N1
Xba*
Nco
Nco
Grow in E.coli JM110
Nco
Pst I
Nco
pEGFP-N1
Xba
Nco
Nco
Restrict with PstI and Xba
Ligate into pUC19
Pst I
Nco
pUC19
Xba
Insert Mlu linker in the Nco site
Change second ATG to CTG
Pst I
Mlu
pUC19
Xba
Restrict with PstI and Xba
Ligate to large pEGFP-N1 fragment
Nco
Pst I
Mlu
pEGFP-N1
Xba
Nco
Insert repeat sequences
Nco
CMV
Pst I
Inserted sequences
pEGFP-N1
Xba
Nco
Neomycin
Nco

Appendix B. List of oligonuclotides used for plasmid constructs in this study.

Plasmid name	Repeat unit	Sequence of inserted oligonucleotides
Positive	N/A	5'CGCGATCGACCATTGGAATTG 3'
NLA	N/A	5'CGCGATCGCGATTCAGATTGCAAGCACGAGACTACCGACCAAG TTCTACGAGCACCTTCTACGGCCTCAAGTAC 3'
CA	$(CA)_{x12}$	5'CGCGATCG CCATTCCAGATTGCAAGC $(CA)_{x12}$CATCTTACGAGA CTACCG 3'
Trio	$(CAG)_{x12}$	5'CGCGATCGCCATTGGAATTG (CAG) x12GTCAACCTATGT 3'
AAAG	$(AAAG)_{x12}$	5'CGCGATCG CCATTC $(AAAG)_{x12}$CATCTT 3'
Hm1	$(GGGCA)_{x12}$	5'CGCGATCG $(GGGCA)_{x12}$ 3'
Hexa	$(ACCTCC)_{x12}$	5'CGCG ATCG $(ACCTCC)_{x12}$ 3'
Per	$(ACAGGC)_{x12}$	5'CGCGATCG $(ACAGGC)_{x12}$ 3'
Hucy	$(TTTA)_{x8}$	5'CGCGATCGGAAATCATATTTTTAAAATATTTTATTTATTTATTTA TTTATTTATTTATTTATTGCATC 3'

Appendix C. List of SAS programs used.

SAS Codes for the GENMOD Procedure for the H2O2 treatment in MC2a.

```
proc import
      datafile = "C:\Documents and Settings\chealy\Desktop\Caro GFP
Flow2\H202green.xls"
      out = cell
      dbms = excel
      replace;
run;
proc contents;
run;
Data cell;
      set cell;
      Percentage_of_green_cells = _COL4;
      Count_of_green_cells = _COL5;
      offset_t = log(10000);
run;

proc sort data = cell;
by plasmid dose;
run;
proc means data = cell noprint;
by plasmid dose;
var Percentage_of_green_cells;
output out = factorplot mean = rate;
run;

proc export
      data = factorplot
      outfile = "C:\Documents and Settings\chealy\Desktop\Caro GFP
Flow2\ factorplot.xls"
      dbms = excel
      replace
;
run;

%macro myint(dat, plasmid, plasmidId, varlist, outdat);

      data &outdat;
            set &dat;

            array temp{*} &varlist;

            do k = 1 to 4;
                  temp[k] = 0;
            end;
            &plasmidId = 0;
            if plasmid = &plasmid then do;
                  &plasmidId = 1;
                  if dose = 0.001 then temp[1] = 1;
                  if dose = 0.01 then temp[2] = 1;
                  if dose = 0.1 then temp[3] = 1;
                  if dose = 1 then temp[4] = 1;
            end;
      run;
%mend;
```

```
%myint(cell , "CA  ", plasmid1, dose_ca1-dose_ca4, cell1);
%myint(cell1, "Green", plasmid2, dose_gr1-dose_gr4, cell);
%myint(cell , "HEXA", plasmid3, dose_hx1-dose_hx4, cell1);
%myint(cell1, "HM  ", plasmid4, dose_hm1-dose_hm4, cell);
%myint(cell , "HUCY", plasmid5, dose_hy1-dose_hy4, cell1);
%myint(cell1, "NLA ", plasmid6, dose_nl1-dose_nl4, cell);
%myint(cell , "PER ", plasmid7, dose_pr1-dose_pr4, cell1);
%myint(cell1, "TRIO", plasmid8, dose_tr1-dose_tr4, cell);
%myint(cell, "AAAG", plasmid9, dose_aa1-dose_aa4, cell1);
data cell1;
 set cell1;
offset_t = log(10000);
run;

proc genmod data = cell1;
            ods output modelFit = GenFit
            ParameterEstimates = GenEst;
            class dose plasmid;
            model Count_of_green_cells =
                  dose_ca1-dose_ca4
                  dose_gr1-dose_gr4
                  dose_hx1-dose_hx4
                  dose_hm1-dose_hm4
                  dose_hy1-dose_hy4
                  dose_nl1-dose_nl4
                  dose_pr1-dose_pr4
                  dose_tr1-dose_tr4
                  dose_aa1-dose_aa4
                  plasmid2
                  /offset = offset_t dist = poisson link = log pscale;
run;

proc genmod data = cell1;
            ods output modelFit = GenFit
            ParameterEstimates = GenEst;
            class dose plasmid;

            model Count_of_green_cells =
                  dose_ca1-dose_ca4
                  dose_gr1-dose_gr4
                  dose_hx3-dose_hx4
                  dose_hm1-dose_hm4
                  dose_hy1-dose_hy4
                  dose_nl1-dose_nl4
                  dose_pr2-dose_pr4
                  dose_tr1-dose_tr4
                  dose_aa1-dose_aa4              plasmid2
                  /offset = offset_t dist = poisson link = log pscale;
run;


data cell2;
 set cell1;
offset_t = log(10000);
run;
```

```
proc genmod data = cell2;
            ods output modelFit = GenFit
            ParameterEstimates = GenEst;
            class dose plasmid;

            model Count_of_green_cells =
                  dose_ca1-dose_ca4
                  dose_gr1-dose_gr4
                  dose_hx3-dose_hx4
                  dose_hm1-dose_hm4
                  dose_hy1-dose_hy4
                  dose_nl1-dose_nl4
                  dose_pr2-dose_pr4
                  dose_tr1-dose_tr4
                  dose_aa1-dose_aa4              plasmid2
                  /offset = offset_t dist = poisson link = log pscale;
run;
```

SAS Codes for the GENMOD Procedure for the H2O2 treatment in C3H10T1/2.

```
proc import
      datafile = "C:\Documents and Settings\chealy\Desktop\Caro GFP
Flow2\H20210Tgreen.xls"
      out = cell
      dbms = excel
      replace;
run;
proc contents;
run;
Data cell;
      set cell;
      Percentage_of_green_cells = _COL4;
      Count_of_green_cells = _COL5;
      offset_t = log(10000);
run;

proc sort data = cell;
by plasmid dose;
run;
proc means data = cell noprint;
by plasmid dose;
var Percentage_of_green_cells;
output out = factorplot mean = rate;
run;

proc export
      data = factorplot
      outfile = "C:\Documents and Settings\chealy\Desktop\Caro GFP
Flow2\ factorplot.xls"
      dbms = excel
      replace
;
run;

%macro myint(dat, plasmid, plasmidId, varlist, outdat);

      data &outdat;
            set &dat;

            array temp{*} &varlist;

            do k = 1 to 4;
                  temp[k] = 0;
            end;
            &plasmidId = 0;
            if plasmid = &plasmid then do;
                  &plasmidId = 1;
                  if dose = 0.001 then temp[1] = 1;
                  if dose = 0.01 then temp[2] = 1;
                  if dose = 0.1 then temp[3] = 1;
                  if dose = 1 then temp[4] = 1;
            end;
      run;
%mend;
```

```
%myint(cell1, "Green", plasmid2, dose_gr1-dose_gr4, cell);
%myint(cell1, "HM  ", plasmid4, dose_hm1-dose_hm4, cell);
%myint(cell1, "NLA ", plasmid6, dose_nl1-dose_nl4, cell);
%myint(cell1, "TRIO", plasmid8, dose_tr1-dose_tr4, cell);
data cell1;
 set cell1;
offset_t = log(10000);
run;

proc genmod data = cell1;
            ods output modelFit = GenFit
            ParameterEstimates = GenEst;
            class dose plasmid;
            model Count_of_green_cells =
                  dose_gr1-dose_gr4
                  dose_hm1-dose_hm4
                  dose_nl1-dose_nl4
                  dose_tr1-dose_tr4
                  plasmid2
                  /offset = offset_t dist = poisson link = log pscale;
run;

proc genmod data = cell1;
            ods output modelFit = GenFit
            ParameterEstimates = GenEst;
            class dose plasmid;

            model Count_of_green_cells =
                  dose_hm1-dose_hm4
                  dose_nl1-dose_nl4
                  dose_tr1-dose_tr4
                  plasmid2
                  /offset = offset_t dist = poisson link = log pscale;
run;


data cell2;
 set cell1;
offset_t = log(10000);
run;

proc genmod data = cell2;
            ods output modelFit = GenFit
            ParameterEstimates = GenEst;
            class dose plasmid;

            model Count_of_green_cells =
                        dose_hm1-dose_hm4
                  dose_nl1-dose_nl4
                  dose_tr1-dose_tr4
                  plasmid2

                  /offset = offset_t dist = poisson link = log pscale;
run;
```

SAS Codes for the GENMOD Procedure for the 9AA treatment in MC2a.

```
proc import
      datafile = "C:\Documents and Settings\chealy\Desktop\Caro GFP
Flow2\9AAgreen.xls"
      out = cell
      dbms = excel
      replace;
run;
proc contents;
run;
Data cell;
      set cell;
      Percentage_of_green_cells = _COL4;
      Count_of_green_cells = _COL5;
      offset_t = log(10000);
run;


proc sort data = cell;
by plasmid dose;
run;
proc means data = cell noprint;
by plasmid dose;
var Percentage_of_green_cells;
output out = factorplot mean = rate;
run;

proc export
      data = factorplot
      outfile = "C:\Documents and Settings\chealy\Desktop\Caro GFP
Flow2\ factorplot.xls"
      dbms = excel
      replace
;
run;

%macro myint(dat, plasmid, plasmidId, varlist, outdat);

      data &outdat;
            set &dat;

            array temp{*} &varlist;

            do k = 1 to 3;
                  temp[k] = 0;
            end;
            &plasmidId = 0;
            if plasmid = &plasmid then do;
                  &plasmidId = 1;
                  if dose = 2 then temp[1] = 1;
                  if dose = 8 then temp[2] = 1;
                  if dose = 16 then temp[3] = 1;

            end;
      run;
```

```
%mend;

%myint(cell , "CA  ", plasmid1, dose_ca1-dose_ca3, cell1);
%myint(cell1, "Green", plasmid2, dose_gr1-dose_gr3, cell);
%myint(cell , "HEXA", plasmid3, dose_hx1-dose_hx3, cell1);
%myint(cell1, "HM  ", plasmid4, dose_hm1-dose_hm3, cell);
%myint(cell , "HUCY", plasmid5, dose_hy1-dose_hy3, cell1);
%myint(cell1, "NLA ", plasmid6, dose_nl1-dose_nl3, cell);
%myint(cell , "PER ", plasmid7, dose_pr1-dose_pr3, cell1);
%myint(cell1, "TRIO", plasmid8, dose_tr1-dose_tr3, cell);
%myint(cell, "AAAG", plasmid9, dose_aa1-dose_aa3, cell1);
data cell1;
 set cell1;
offset_t = log(10000);

run;

proc genmod data = cell1;
            ods output modelFit = GenFit
            ParameterEstimates = GenEst;
            class dose plasmid;
                  model Count_of_green_cells =
                  dose_ca1-dose_ca3
                  dose_gr1-dose_gr3
                  dose_hx1-dose_hx3
                  dose_hm1-dose_hm3
                  dose_hy1-dose_hy3
                  dose_nl1-dose_nl3
                  dose_pr1-dose_pr3
                  dose_tr1-dose_tr3
                  dose_aa1-dose_aa3
                  plasmid2
                  /offset = offset_t dist = poisson link = log pscale;
run;

proc genmod data = cell1;
            ods output modelFit = GenFit
            ParameterEstimates = GenEst;
            class dose plasmid;

            model Count_of_green_cells =
                  dose_ca1-dose_ca3
                  dose_gr1-dose_gr3
                  dose_hx3-dose_hx3
                  dose_hm1-dose_hm3
                  dose_hy1-dose_hy3
                  dose_nl1-dose_nl3
                  dose_pr2-dose_pr3
                  dose_tr1-dose_tr3
                  dose_aa1-dose_aa3              plasmid2
                  /offset = offset_t dist = poisson link = log pscale;
run;


data cell2;
 set cell1;
offset_t = log(10000);
```

```
run;

proc genmod data = cell2;
            ods output modelFit = GenFit
            ParameterEstimates = GenEst;
            class dose plasmid;

            model Count_of_green_cells =
                  dose_ca1-dose_ca3
                  dose_gr1-dose_gr3
                  dose_hx3-dose_hx3
                  dose_hm1-dose_hm3
                  dose_hy1-dose_hy3
                  dose_nl1-dose_nl3
                  dose_pr2-dose_pr3
                  dose_tr1-dose_tr3
                  dose_aa1-dose_aa3              plasmid2
                  /offset = offset_t dist = poisson link = log pscale;
run;
```

SAS Codes for the GENMOD Procedure for the 9AA treatment in C3H10T1/2.

```
proc import
      datafile = "C:\Documents and Settings\chealy\Desktop\Caro GFP
Flow2\9AA10Tgreen.xls"
      out = cell
      dbms = excel
      replace;
run;
proc contents;
run;
Data cell;
      set cell;
      Percentage_of_green_cells = _COL4;
      Count_of_green_cells = _COL5;
      offset_t = log(10000);
run;

proc sort data = cell;
by plasmid dose;
run;
proc means data = cell noprint;
by plasmid dose;
var Percentage_of_green_cells;
output out = factorplot mean = rate;
run;

proc export
      data = factorplot
      outfile = "C:\Documents and Settings\chealy\Desktop\Caro GFP
Flow2\ factorplot.xls"
      dbms = excel
      replace
;
run;

%macro myint(dat, plasmid, plasmidId, varlist, outdat);

      data &outdat;
            set &dat;

            array temp{*} &varlist;


            do k = 1 to 3;
                  temp[k] = 0;
            end;
            &plasmidId = 0;
            if plasmid = &plasmid then do;
                  &plasmidId = 1;
                  if dose = 2 then temp[1] = 1;
                  if dose = 8 then temp[2] = 1;
                  if dose = 16 then temp[3] = 1;

            end;
      run;
```

```
%mend;

%myint(cell1, "Green", plasmid2, dose_gr1-dose_gr3, cell);
%myint(cell1, "HM  ", plasmid4, dose_hm1-dose_hm3, cell);
%myint(cell1, "NLA ", plasmid6, dose_nl1-dose_nl3, cell);
myint(cell1, "TRIO", plasmid8, dose_tr1-dose_tr3, cell);
data cell1;
 set cell1;
offset_t = log(10000);

run;

proc genmod data = cell1;
            ods output modelFit = GenFit
            ParameterEstimates = GenEst;
            class dose plasmid;
                  model Count_of_green_cells =
                  dose_gr1-dose_gr3
                  dose_hm1-dose_hm3
                  dose_nl1-dose_nl3
                  dose_tr1-dose_tr3
                  plasmid2
                  /offset = offset_t dist = poisson link = log pscale;
run;

proc genmod data = cell1;
            ods output modelFit = GenFit
            ParameterEstimates = GenEst;
            class dose plasmid;

            model Count_of_green_cells =
            dose_gr1-dose_gr3
                  dose_hm1-dose_hm3
                  dose_nl1-dose_nl3
                  dose_tr1-dose_tr3
                  plasmid2
                  /offset = offset_t dist = poisson link = log pscale;
run;

data cell2;
 set cell1;
offset_t = log(10000);
run;

proc genmod data = cell2;
            ods output modelFit = GenFit
            ParameterEstimates = GenEst;
            class dose plasmid;

            model Count_of_green_cells =
                  dose_gr1-dose_gr3
                  dose_hm1-dose_hm3
                  dose_nl1-dose_nl3
                  dose_tr1-dose_tr3
                  plasmid2
                  /offset = offset_t dist = poisson link = log pscale;
run;
```

SAS Codes for the GENMOD Procedure for the ENU treatment in MC2a.

```
proc import
      datafile = "C:\Documents and Settings\chealy\Desktop\Caro GFP
Flow2\ENUgreen.xls"
      out = cell
      dbms = excel
      replace;
run;
proc contents;
run;
Data cell;
      set cell;
      Percentage_of_green_cells = _COL4;
      Count_of_green_cells = _COL5;
      offset_t = log(10000);
run;

proc sort data = cell;
by plasmid dose;
run;
proc means data = cell noprint;
by plasmid dose;
var Percentage_of_green_cells;
output out = factorplot mean = rate;
run;

proc export
      data = factorplot
      outfile = "C:\Documents and Settings\chealy\Desktop\Caro GFP
Flow2\ factorplot.xls"
      dbms = excel
      replace
;
run;

%macro myint(dat, plasmid, plasmidId, varlist, outdat);

      data &outdat;
            set &dat;

            array temp{*} &varlist;

            do k = 1 to 4;
                  temp[k] = 0;
            end;
            &plasmidId = 0;
            if plasmid = &plasmid then do;
                  &plasmidId = 1;
                  if dose = 1.07 then temp[1] = 1;
                  if dose = 2.13 then temp[2] = 1;
                  if dose = 4.27 then temp[3] = 1;
                  if dose = 6.40 then temp[4] = 1;
            end;
      run;
%mend;
```

```
%myint(cell , "CA  ", plasmid1, dose_ca1-dose_ca4, cell1);
%myint(cell1, "Green", plasmid2, dose_gr1-dose_gr4, cell);
%myint(cell , "HEXA", plasmid3, dose_hx1-dose_hx4, cell1);
%myint(cell1, "HM  ", plasmid4, dose_hm1-dose_hm4, cell);
%myint(cell , "HUCY", plasmid5, dose_hy1-dose_hy4, cell1);
%myint(cell1, "NLA ", plasmid6, dose_nl1-dose_nl4, cell);
%myint(cell , "PER ", plasmid7, dose_pr1-dose_pr4, cell1);
%myint(cell1, "TRIO", plasmid8, dose_tr1-dose_tr4, cell);
%myint(cell, "AAAG", plasmid9, dose_aa1-dose_aa4, cell1);
data cell1;
 set cell1;
offset_t = log(10000);
run;

proc genmod data = cell1;
            ods output modelFit = GenFit
            ParameterEstimates = GenEst;
            class dose plasmid;
            model Count_of_green_cells =
                  dose_ca1-dose_ca4
                  dose_gr1-dose_gr4
                  dose_hx1-dose_hx4
                  dose_hm1-dose_hm4
                  dose_hy1-dose_hy4
                  dose_nl1-dose_nl4
                  dose_pr1-dose_pr4
                  dose_tr1-dose_tr4
                  dose_aa1-dose_aa4 /
                  plasmid2
                  /offset = offset_t dist = poisson link = log pscale;
run;

proc genmod data = cell1;
            ods output modelFit = GenFit
            ParameterEstimates = GenEst;
            class dose plasmid;

            model Count_of_green_cells =
                  dose_ca1-dose_ca4
                  dose_gr1-dose_gr4
                  dose_hx3-dose_hx4
                  dose_hm1-dose_hm4
                  dose_hy1-dose_hy4
                  dose_nl1-dose_nl4
                  dose_pr2-dose_pr4
                  dose_tr1-dose_tr4
                  dose_aa1-dose_aa4              plasmid2
                  /offset = offset_t dist = poisson link = log pscale;
run;


data cell2;
 set cell1;
offset_t = log(10000);
run;

proc genmod data = cell2;
```

```
proc genmod data = cell2;
            ods output modelFit = GenFit
            ParameterEstimates = GenEst;
            class dose plasmid;

            model Count_of_green_cells =
                  dose_ca1-dose_ca4
                  dose_gr1-dose_gr4
                  dose_hx3-dose_hx4
                  dose_hm1-dose_hm4
                  dose_hy1-dose_hy4
                  dose_nl1-dose_nl4
                  dose_pr2-dose_pr4
                  dose_tr1-dose_tr4
                  dose_aa1-dose_aa4              plasmid2
                  /offset = offset_t dist = poisson link = log pscale;
run;
```

SAS Codes for the GENMOD Procedure for the TPA treatment in MC2a.

```
proc import
      datafile = "C:\Documents and Settings\chealy\Desktop\Caro GFP
Flow2\TPAgreen.xls"
      out = cell
      dbms = excel
      replace;
run;
proc contents;
run;
Data cell;
      set cell;
      Percentage_of_green_cells = _COL4;
      Count_of_green_cells = _COL5;
      offset_t = log(10000);
run;

proc sort data = cell;
by plasmid dose;
run;
proc means data = cell noprint;
by plasmid dose;
var Percentage_of_green_cells;
output out = factorplot mean = rate;
run;

proc export
      data = factorplot
      outfile = "C:\Documents and Settings\chealy\Desktop\Caro GFP
Flow2\ factorplot.xls"
      dbms = excel
      replace
;
run;

%macro myint(dat, plasmid, plasmidId, varlist, outdat);

      data &outdat;
            set &dat;

            array temp{*} &varlist;

            do k = 1 to 4;
                  temp[k] = 0;
            end;
            &plasmidId = 0;
            if plasmid = &plasmid then do;
                  &plasmidId = 1;
                  if dose = 100 then temp[1] = 1;
                  if dose = 325 then temp[2] = 1;
                  if dose = 650 then temp[3] = 1;
                  if dose = 1,325 then temp[4] = 1;
            end;
      run;
%mend;
```

```
%myint(cell , "CA  ", plasmid1, dose_ca1-dose_ca4, cell11);
%myint(cell11, "Green", plasmid2, dose_gr1-dose_gr4, cell);
%myint(cell , "HEXA", plasmid3, dose_hx1-dose_hx4, cell11);
%myint(cell11, "HM  ", plasmid4, dose_hm1-dose_hm4, cell);
%myint(cell , "HUCY", plasmid5, dose_hy1-dose_hy4, cell11);
%myint(cell11, "NLA ", plasmid6, dose_nl1-dose_nl4, cell);
%myint(cell , "PER ", plasmid7, dose_pr1-dose_pr4, cell11);
%myint(cell11, "TRIO", plasmid8, dose_tr1-dose_tr4, cell);
%myint(cell, "AAAG", plasmid9, dose_aa1-dose_aa4, cell11);
data cell11;
 set cell11;
offset_t = log(10000);
run;

proc genmod data = cell11;
            ods output modelFit = GenFit
            ParameterEstimates = GenEst;
            class dose plasmid;
            model Count_of_green_cells =
                  dose_ca1-dose_ca4
                  dose_gr1-dose_gr4
                  dose_hx1-dose_hx4
                  dose_hm1-dose_hm4
                  dose_hy1-dose_hy4
                  dose_nl1-dose_nl4
                  dose_pr1-dose_pr4
                  dose_tr1-dose_tr4
                  dose_aa1-dose_aa4 /
                  plasmid2
                  /offset = offset_t dist = poisson link = log pscale;
run;

proc genmod data = cell11;
            ods output modelFit = GenFit
            ParameterEstimates = GenEst;
            class dose plasmid;

            model Count_of_green_cells =
                  dose_ca1-dose_ca4
                  dose_gr1-dose_gr4
                  dose_hx3-dose_hx4
                  dose_hm1-dose_hm4
                  dose_hy1-dose_hy4
                  dose_nl1-dose_nl4
                  dose_pr2-dose_pr4
                  dose_tr1-dose_tr4
                  dose_aa1-dose_aa4              plasmid2
                  /offset = offset_t dist = poisson link = log pscale;
run;


data cell2;
 set cell11;
offset_t = log(10000);
run;

proc genmod data = cell2;
```

```
proc genmod data = cell2;
            ods output modelFit = GenFit
            ParameterEstimates = GenEst;
            class dose plasmid;

            model Count_of_green_cells =
                  dose_ca1-dose_ca4
                  dose_gr1-dose_gr4
                  dose_hx3-dose_hx4
                  dose_hm1-dose_hm4
                  dose_hy1-dose_hy4
                  dose_nl1-dose_nl4
                  dose_pr2-dose_pr4
                  dose_tr1-dose_tr4
                  dose_aa1-dose_aa4                plasmid2
                  /offset = offset_t dist = poisson link = log pscale;
run;
```

SAS Codes for the GENMOD Procedure for the BPDE treatment in MC2a.

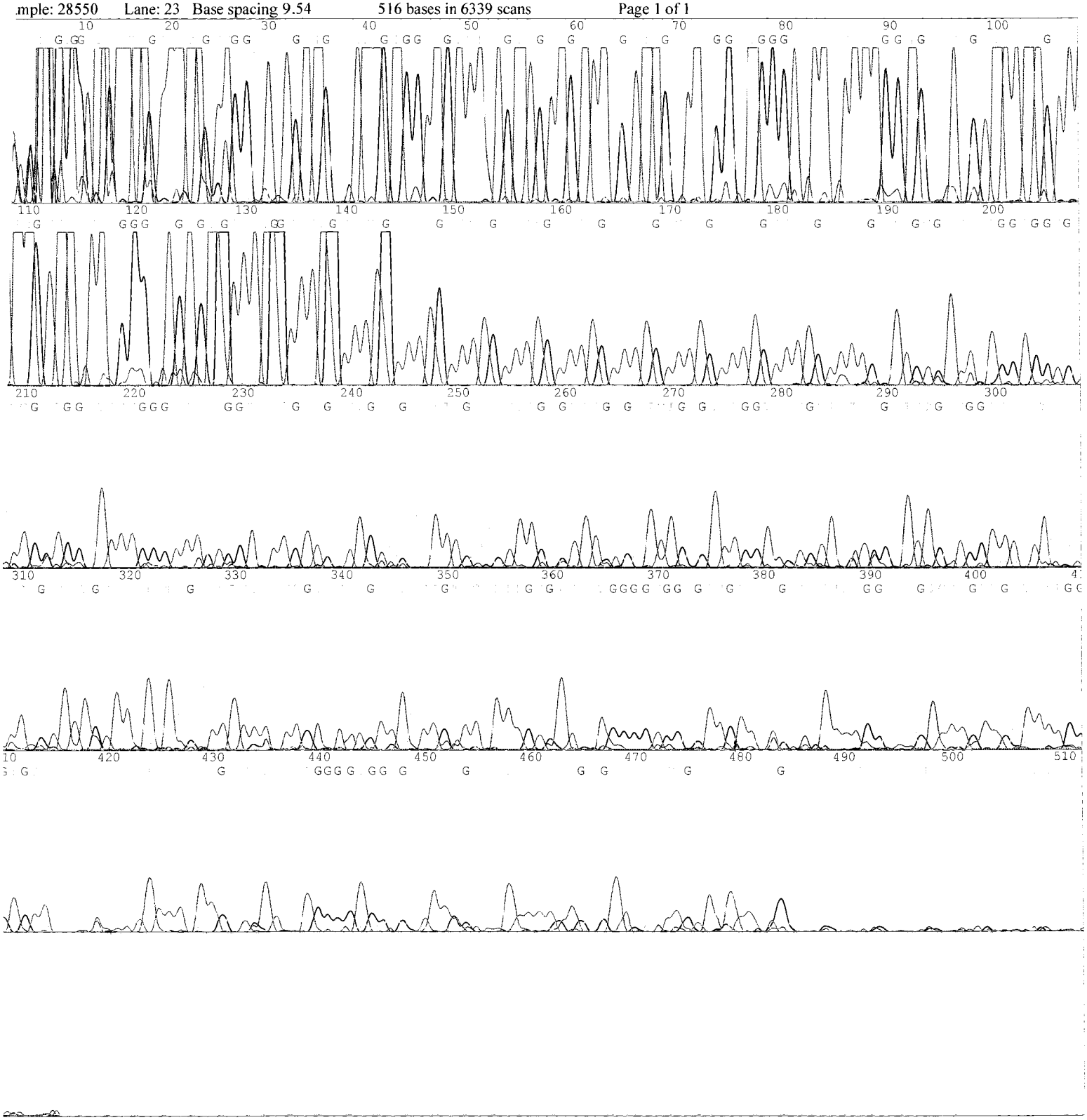
```
proc import
      datafile = "C:\Documents and Settings\chealy\Desktop\Caro GFP
Flow2\BPDEgreen.xls"
      out = cell
      dbms = excel
      replace;
run;
proc contents;
run;
Data cell;
      set cell;
      Percentage_of_green_cells = _COL4;
      Count_of_green_cells = _COL5;
      offset_t = log(10000);
run;

proc sort data = cell;
by plasmid dose;
run;
proc means data = cell noprint;
by plasmid dose;
var Percentage_of_green_cells;
output out = factorplot mean = rate;
run;

proc export
      data = factorplot
      outfile = "C:\Documents and Settings\chealy\Desktop\Caro GFP
Flow2\ factorplot.xls"
      dbms = excel
      replace
;
run;

%macro myint(dat, plasmid, plasmidId, varlist, outdat);

      data &outdat;
            set &dat;

            array temp{*} &varlist;

            do k = 1 to 4;
                  temp[k] = 0;
            end;
            &plasmidId = 0;
            if plasmid = &plasmid then do;
                  &plasmidId = 1;
                  if dose = 0.5 then temp[1] = 1;
                  if dose = 1 then temp[2] = 1;
                  if dose = 2 then temp[3] = 1;
                  if dose = 3 then temp[4] = 1;
            end;
      run;
%mend;
```

```
%myint(cell , "CA  ", plasmid1, dose_ca1-dose_ca4, cell11);
%myint(cell11, "Green", plasmid2, dose_gr1-dose_gr4, cell);
%myint(cell , "HEXA", plasmid3, dose_hx1-dose_hx4, cell11);
%myint(cell11, "HM  ", plasmid4, dose_hm1-dose_hm4, cell);
%myint(cell , "HUCY", plasmid5, dose_hy1-dose_hy4, cell11);
%myint(cell11, "NLA ", plasmid6, dose_nl1-dose_nl4, cell);
%myint(cell , "PER ", plasmid7, dose_pr1-dose_pr4, cell11);
%myint(cell11, "TRIO", plasmid8, dose_tr1-dose_tr4, cell);
%myint(cell, "AAAG", plasmid9, dose_aa1-dose_aa4, cell11);
data cell11;
 set cell11;
offset_t = log(10000);
run;

proc genmod data = cell11;
            ods output modelFit = GenFit
            ParameterEstimates = GenEst;
            class dose plasmid;
            model Count_of_green_cells =
                  dose_ca1-dose_ca4
                  dose_gr1-dose_gr4
                  dose_hx1-dose_hx4
                  dose_hm1-dose_hm4
                  dose_hy1-dose_hy4
                  dose_nl1-dose_nl4
                  dose_pr1-dose_pr4
                  dose_tr1-dose_tr4
                  dose_aa1-dose_aa4 /
                  plasmid2
                  /offset = offset_t dist = poisson link = log pscale;
run;

proc genmod data = cell11;
            ods output modelFit = GenFit
            ParameterEstimates = GenEst;
            class dose plasmid;

            model Count_of_green_cells =
                  dose_ca1-dose_ca4
                  dose_gr1-dose_gr4
                  dose_hx3-dose_hx4
                  dose_hm1-dose_hm4
                  dose_hy1-dose_hy4
                  dose_nl1-dose_nl4
                  dose_pr2-dose_pr4
                  dose_tr1-dose_tr4
                  dose_aa1-dose_aa4              plasmid2
                  /offset = offset_t dist = poisson link = log pscale;
run;


data cell12;
 set cell11;
offset_t = log(10000);
run;

proc genmod data = cell12;
```

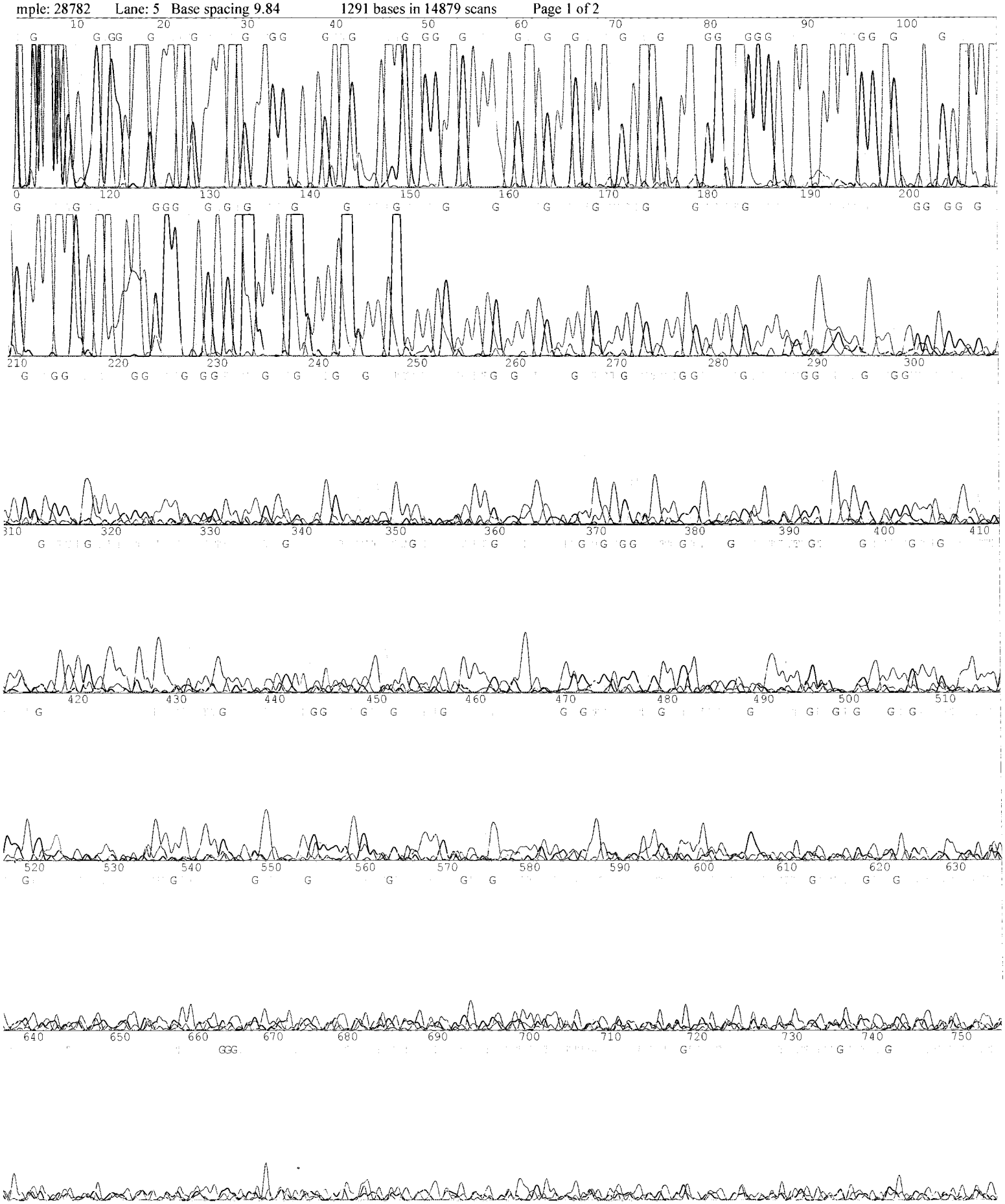
```
ods output modelFit = GenFit
ParameterEstimates = GenEst;
class dose plasmid;

model Count_of_green_cells =
      dose_ca1-dose_ca4
      dose_gr1-dose_gr4
      dose_hx3-dose_hx4
      dose_hm1-dose_hm4
      dose_hy1-dose_hy4
      dose_nl1-dose_nl4
      dose_pr2-dose_pr4
      dose_tr1-dose_tr4
      dose_aa1-dose_aa4              plasmid2
      /offset = offset_t dist = poisson link = log pscale;
run;
```

Appendix D. Chromatograms from sequencing reactions.

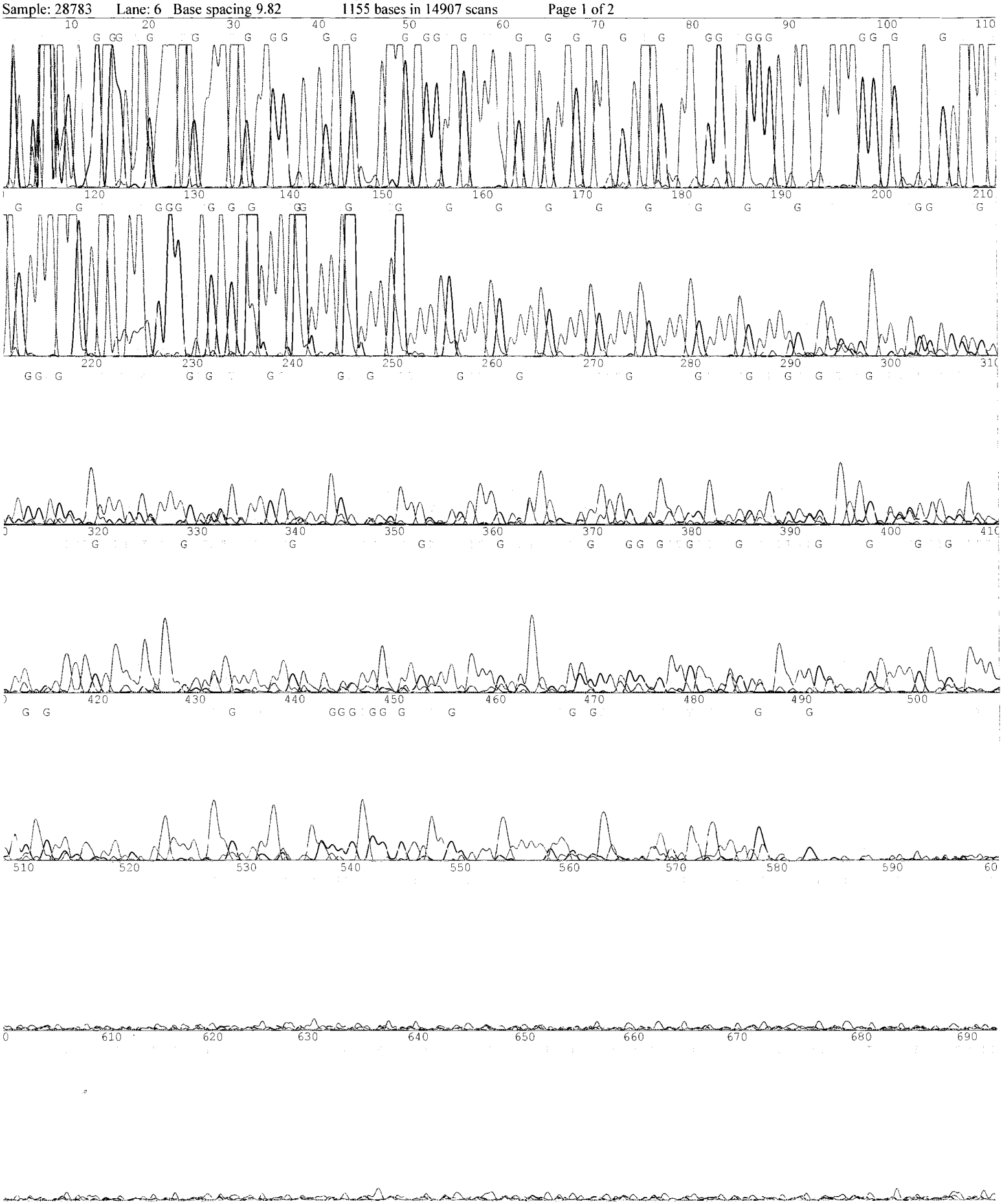
le: 28550 Run Ended: Nov 19, 2004, 4:11:24 Signal G:250 T:295 A:290 C:288 Comment:
mple: 28550 Lane: 23 Base spacing 9.54 516 bases in 6339 scans Page 1 of 1



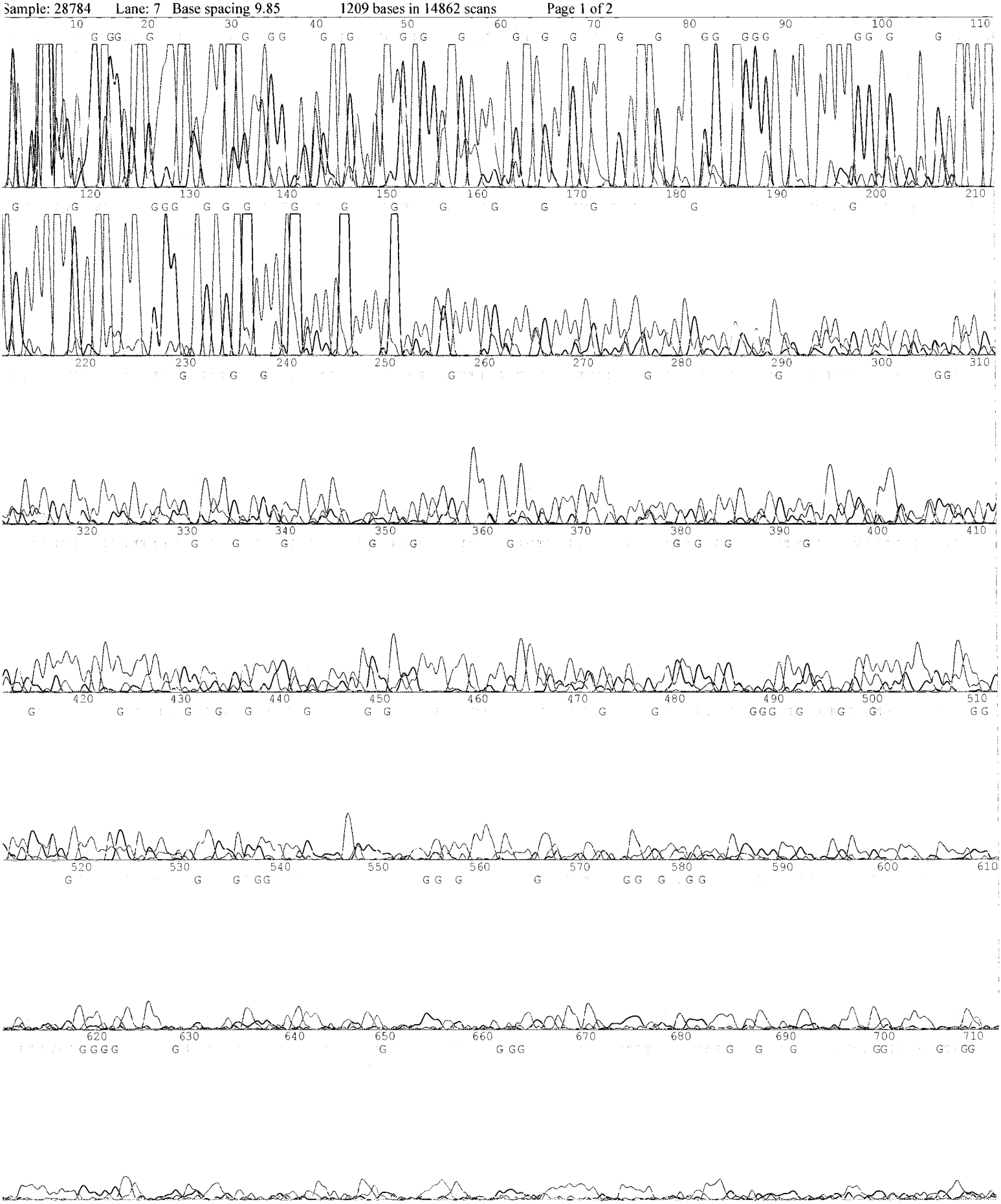
e: 28782 Run Ended: Dec 7, 2004, 2:29:57 Signal G:43 T:51 A:48 C:39 Comment:
mple: 28782 Lane: 5 Base spacing 9.84 1291 bases in 14879 scans Page 1 of 2



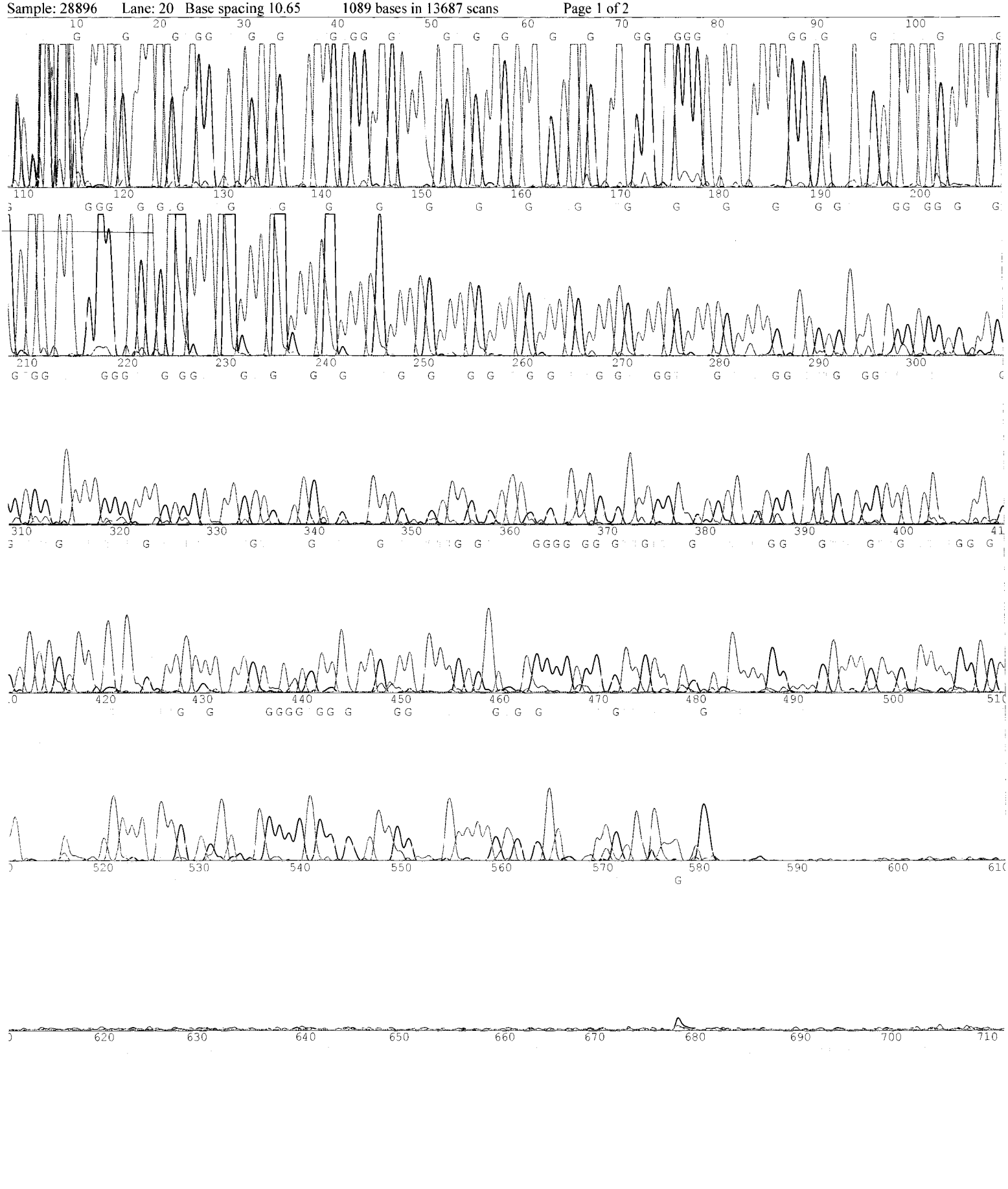
File: 28783 Run Ended: Dec 7, 2004, 2:29:57 Signal G:85 T:101 A:100 C:89 Comment:
Sample: 28783 Lane: 6 Base spacing 9.82 1155 bases in 14907 scans Page 1 of 2



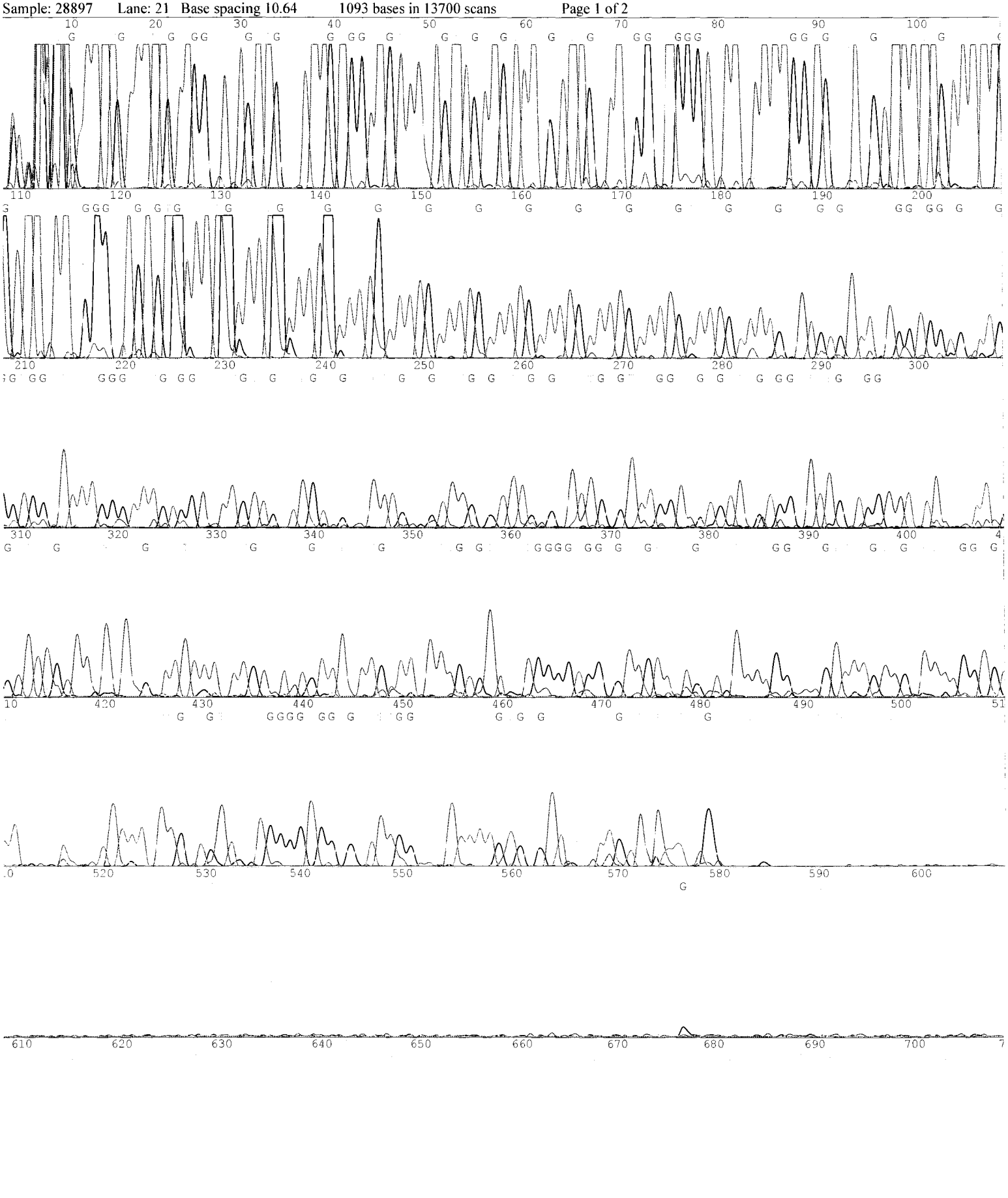
File: 28784 Run Ended: Dec 7, 2004, 2:29:57 Signal G:74 T:111 A:129 C:81 Comment:
Sample: 28784 Lane: 7 Base spacing 9.85 1209 bases in 14862 scans Page 1 of 2



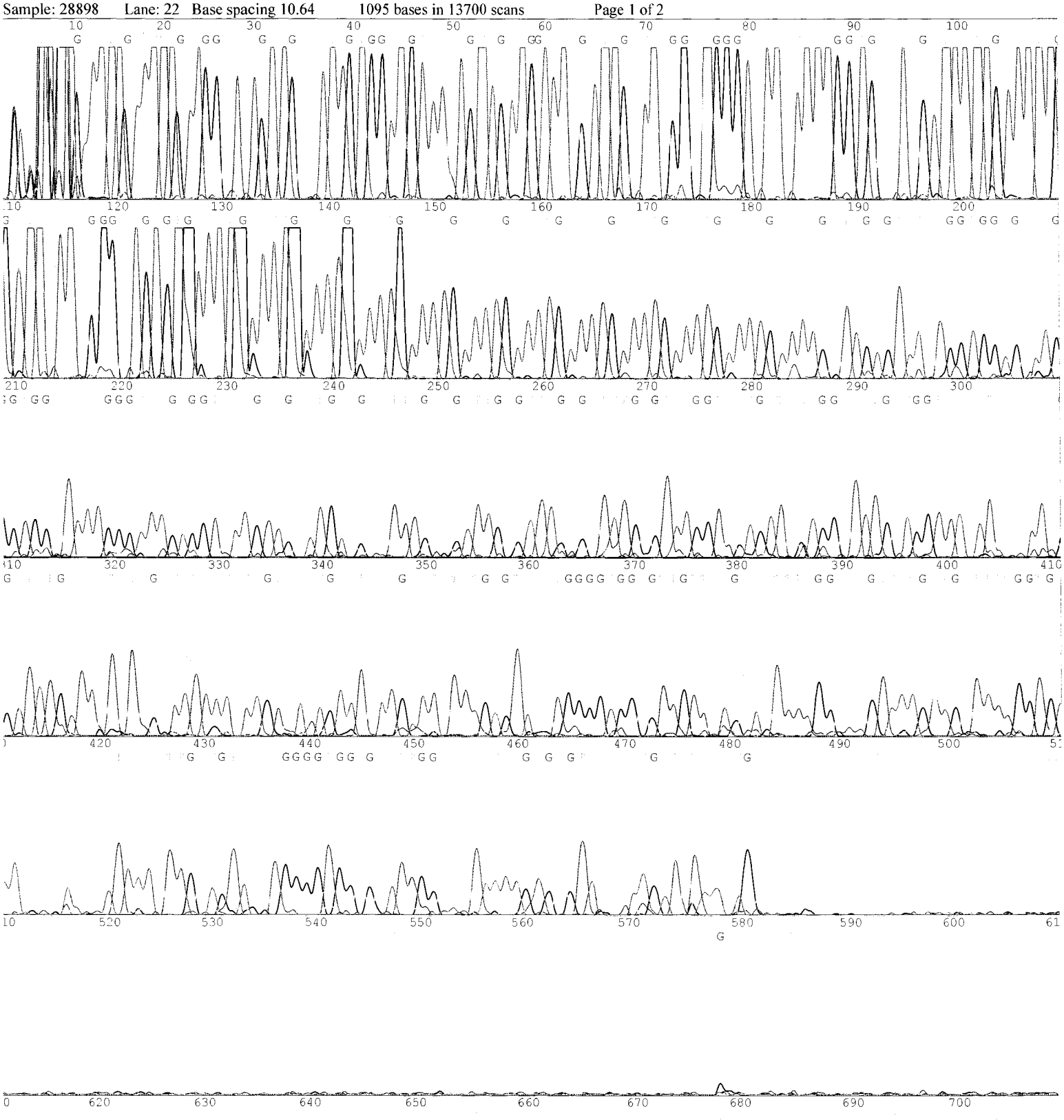
File: 28896 Run Ended: Jan 7, 2005, 4:56:16 Signal G:131 T:178 A:175 C:162 Comment:
Sample: 28896 Lane: 20 Base spacing 10.65 1089 bases in 13687 scans Page 1 of 2



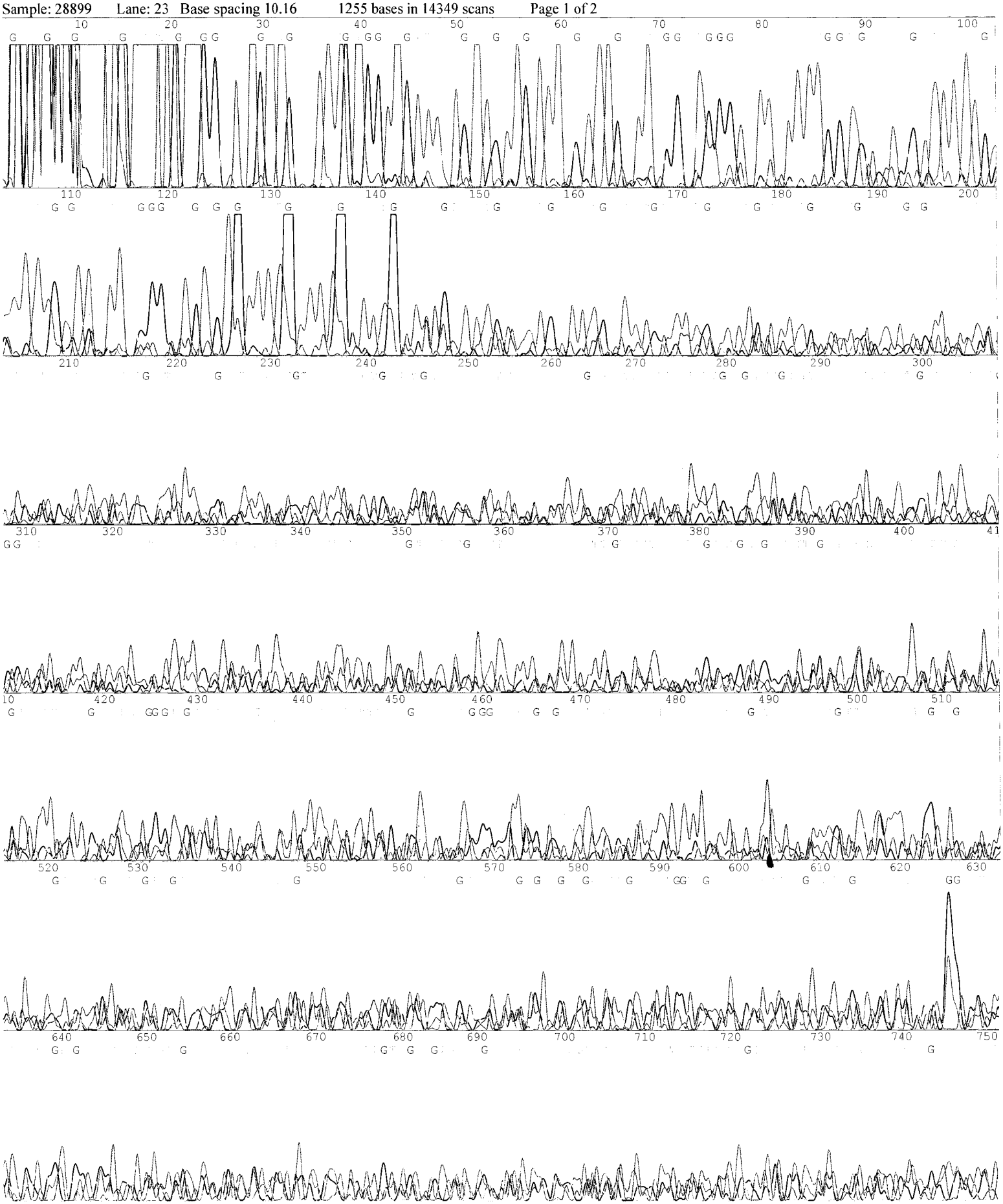
File: 28897 Run Ended: Jan 7, 2005, 4:56:16 Signal G:168 T:220 A:218 C:203 Comment:
Sample: 28897 Lane: 21 Base spacing 10.64 1093 bases in 13700 scans Page 1 of 2



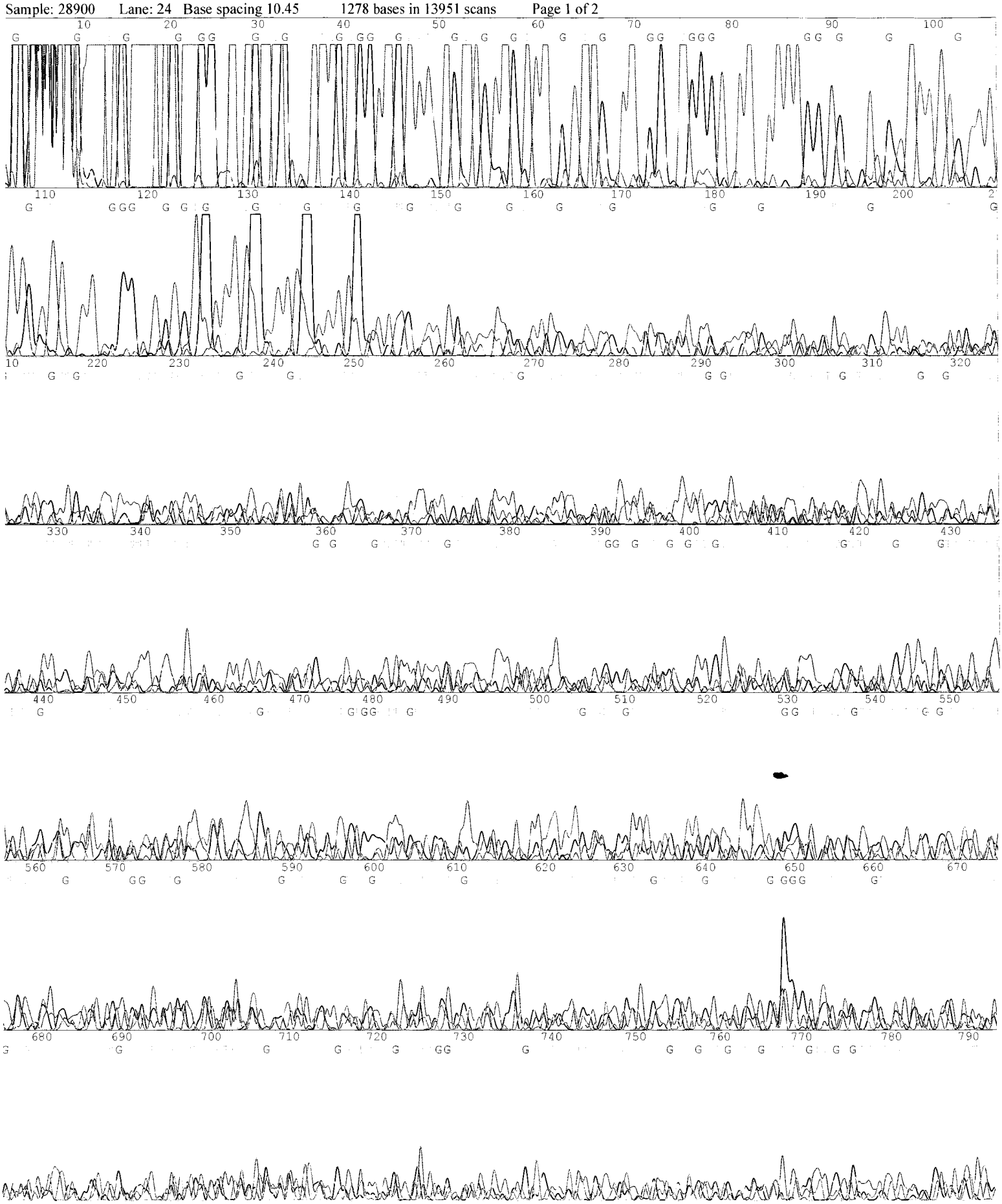
File: 28898 Run Ended: Jan 7, 2005, 4:56:16 Signal G:135 T:190 A:190 C:169 Comment:
Sample: 28898 Lane: 22 Base spacing 10.64 1095 bases in 13700 scans Page 1 of 2



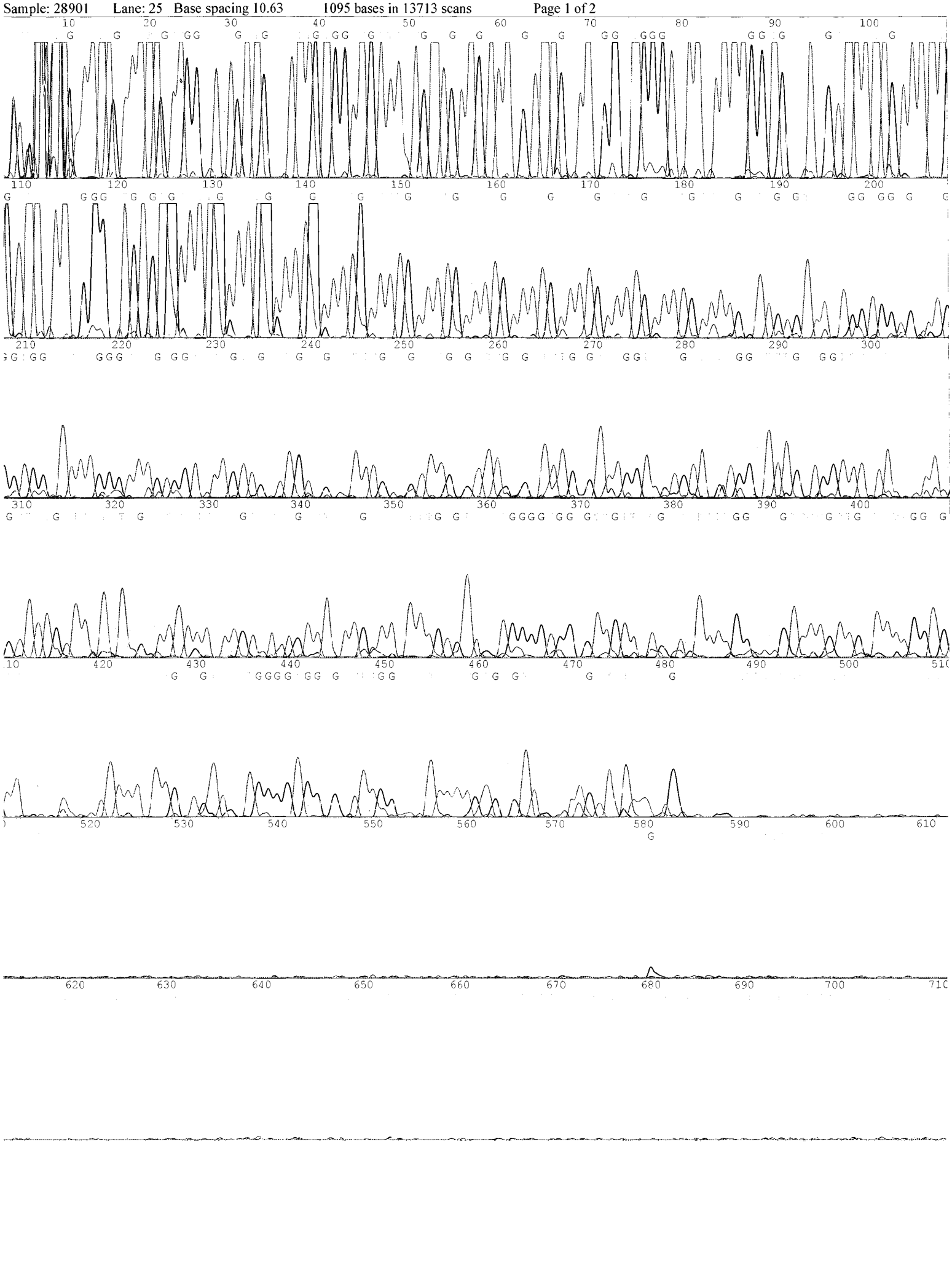
File: 28899 Run Ended: Jan 7, 2005, 4:56:16 Signal G:23 T:32 A:28 C:23 Comment:
Sample: 28899 Lane: 23 Base spacing 10.16 1255 bases in 14349 scans Page 1 of 2



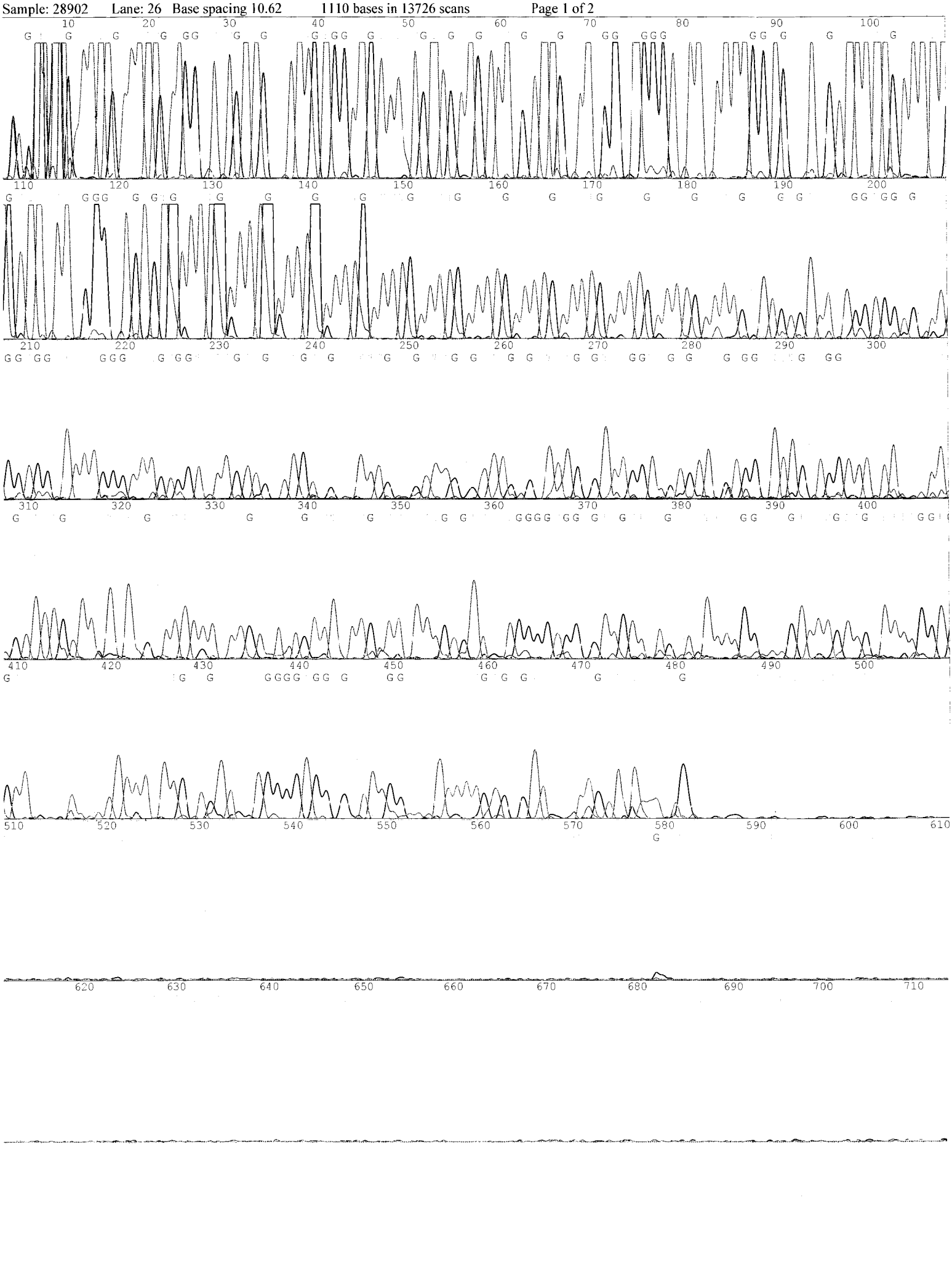
File: 28900 Run Ended: Jan 7, 2005, 4:56:16 Signal G:27 T:37 A:32 C:28 Comment:
Sample: 28900 Lane: 24 Base spacing 10.45 1278 bases in 13951 scans



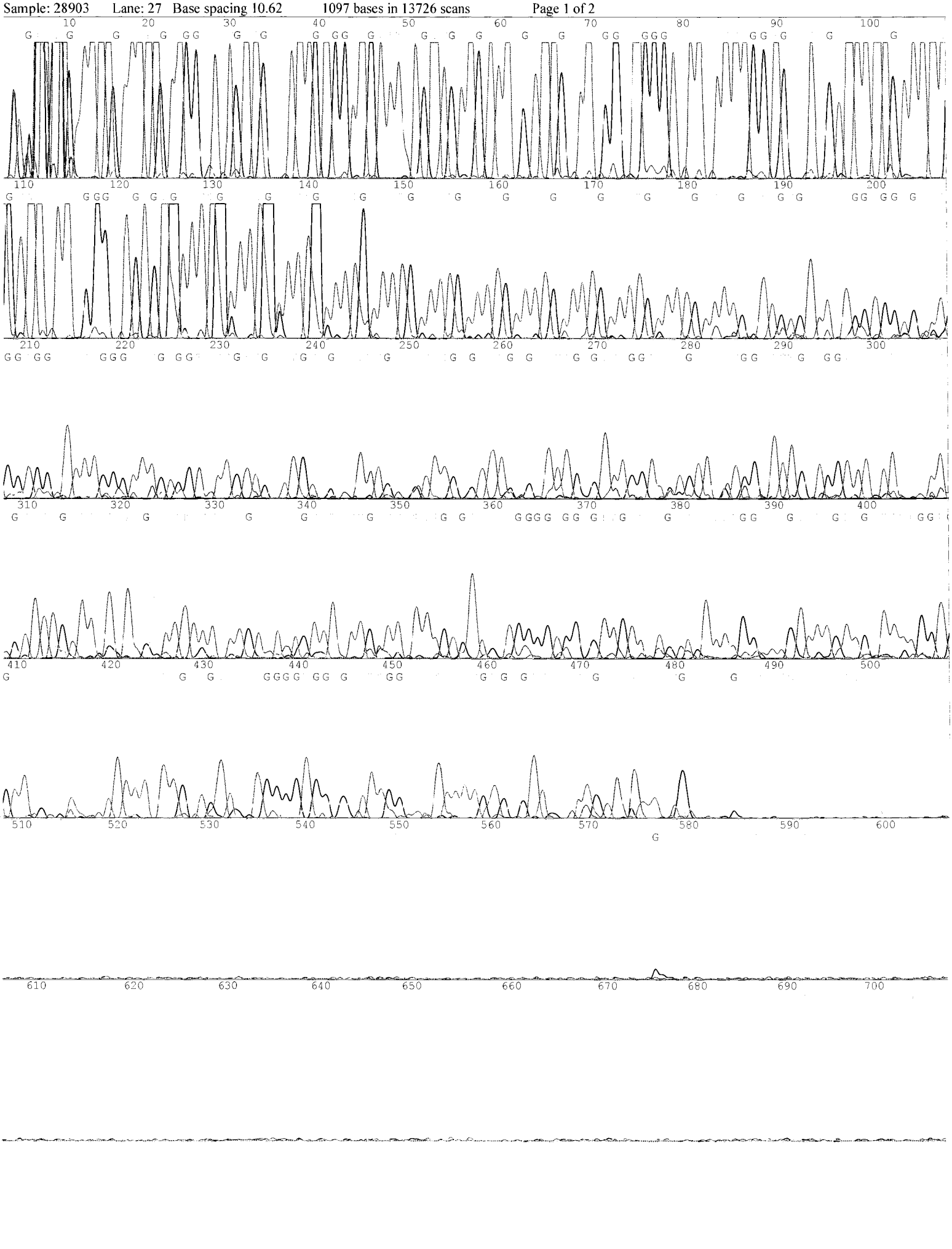
File: 28901 Run Ended: Jan 7, 2005, 4:56:16 Signal G:157 T:215 A:215 C:199 Comment:
Sample: 28901 Lane: 25 Base spacing 10.63 1095 bases in 13713 scans Page 1 of 2



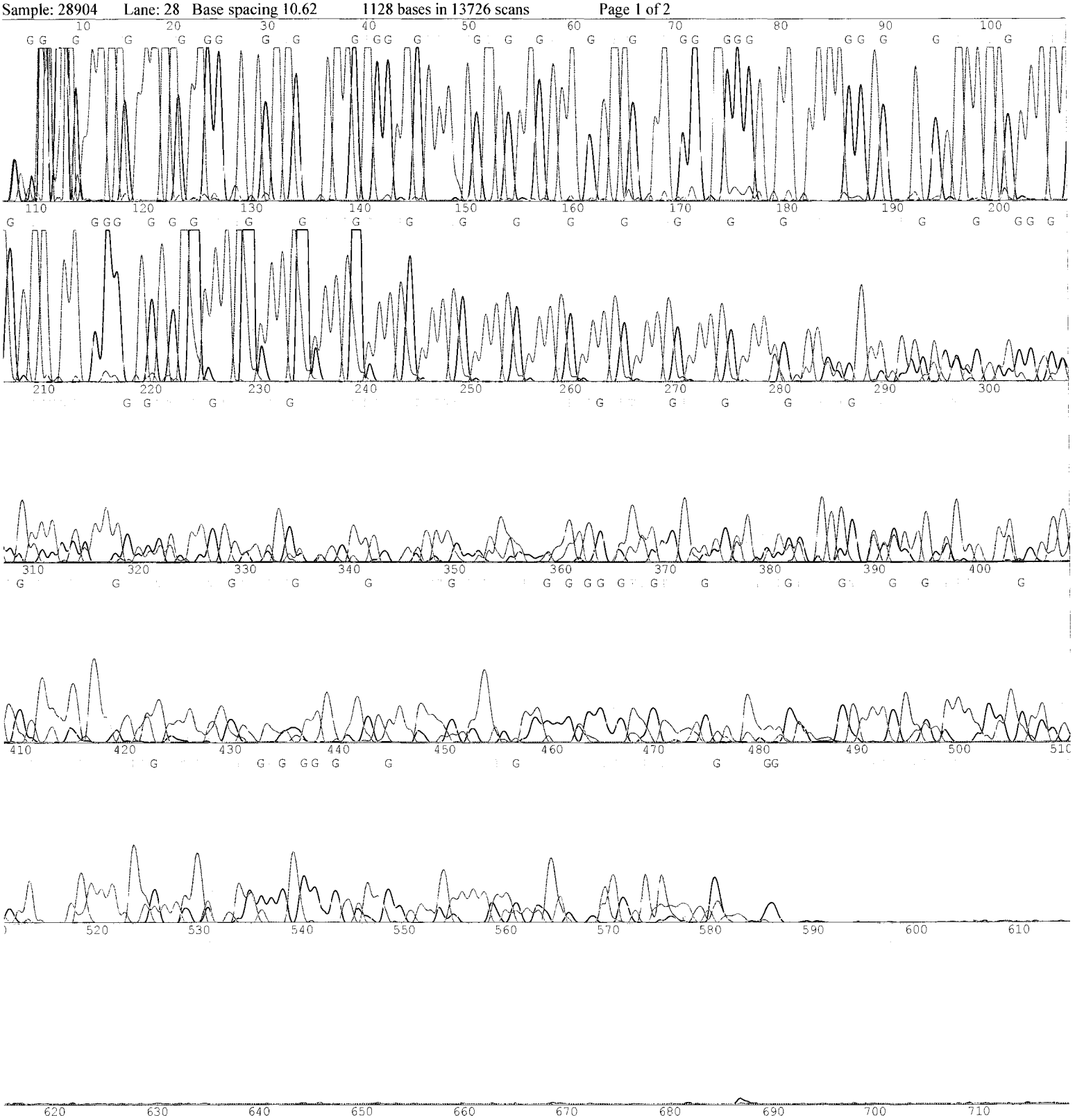
File: 28902 Run Ended: Jan 7, 2005, 4:56:16 Signal G:191 T:254 A:264 C:240 Comment:
Sample: 28902 Lane: 26 Base spacing 10.62 1110 bases in 13726 scans Page 1 of 2



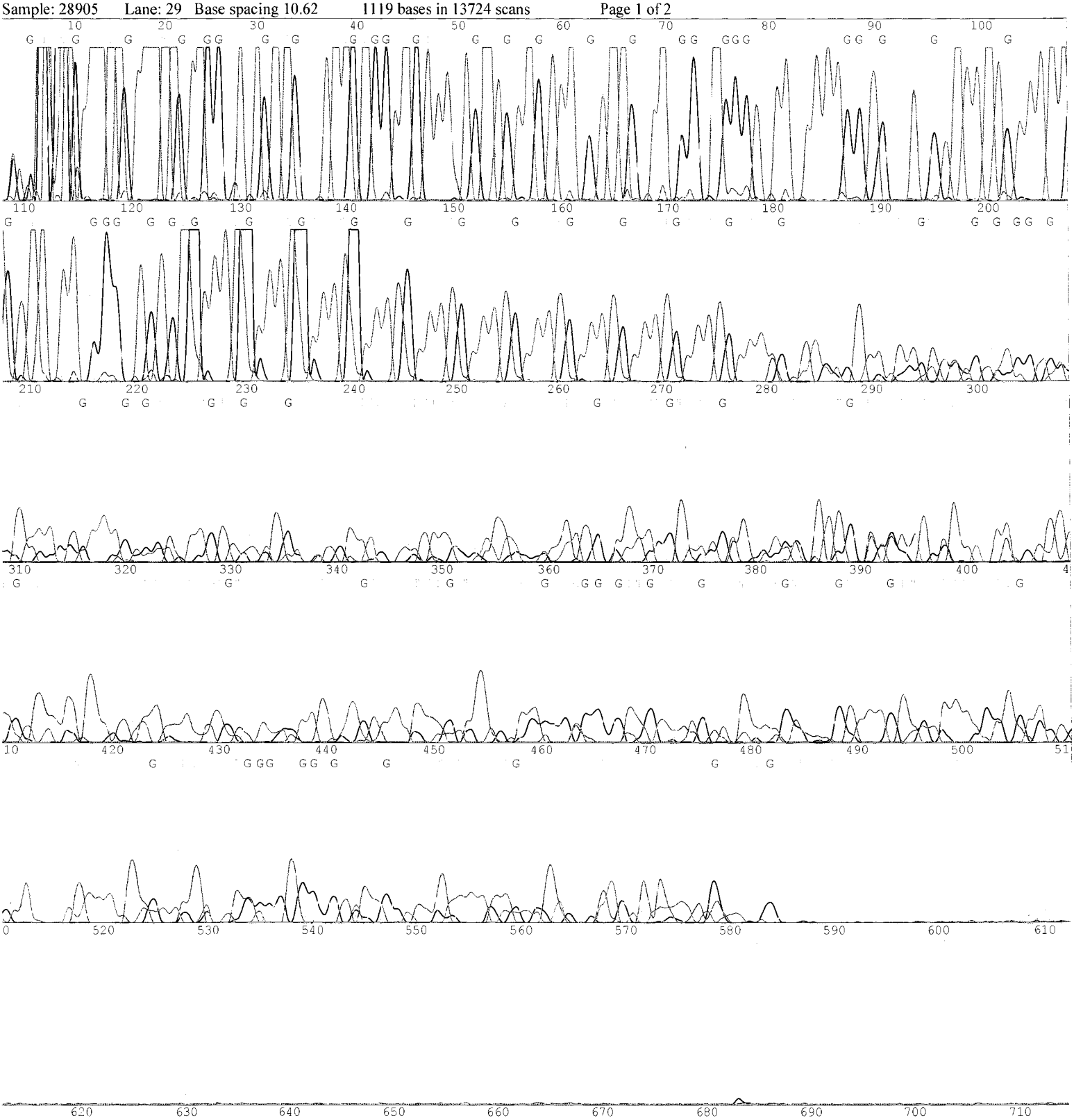
File: 28903 Run Ended: Jan 7, 2005, 4:56:16 Signal G:150 T:206 A:212 C:193 Comment:
Sample: 28903 Lane: 27 Base spacing 10.62 1097 bases in 13726 scans Page 1 of 2



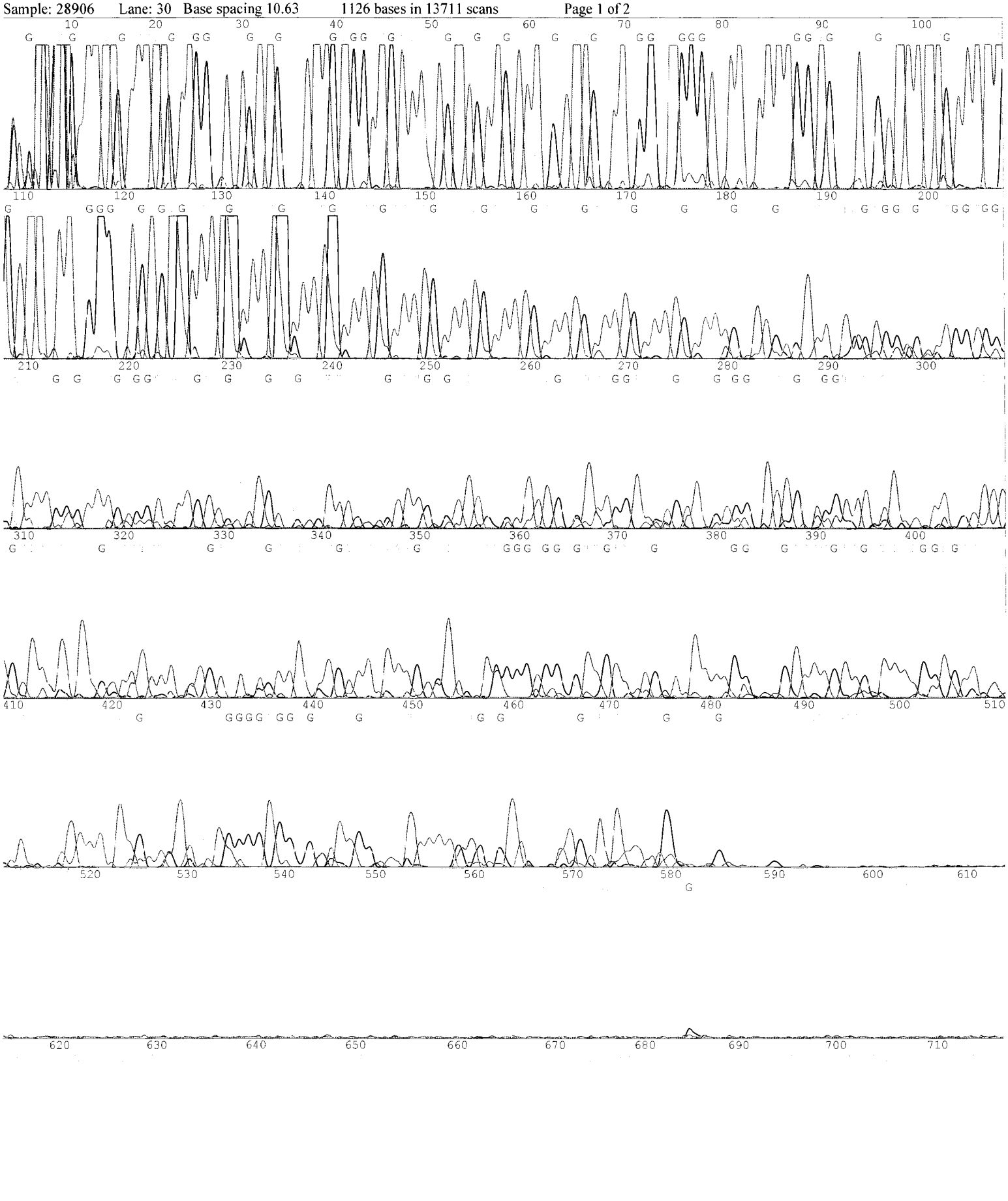
File: 28904 Run Ended: Jan 7, 2005, 4:56:16 Signal G:300 T:406 A:422 C:387 Comment:
Sample: 28904 Lane: 28 Base spacing 10.62 1128 bases in 13726 scans Page 1 of 2



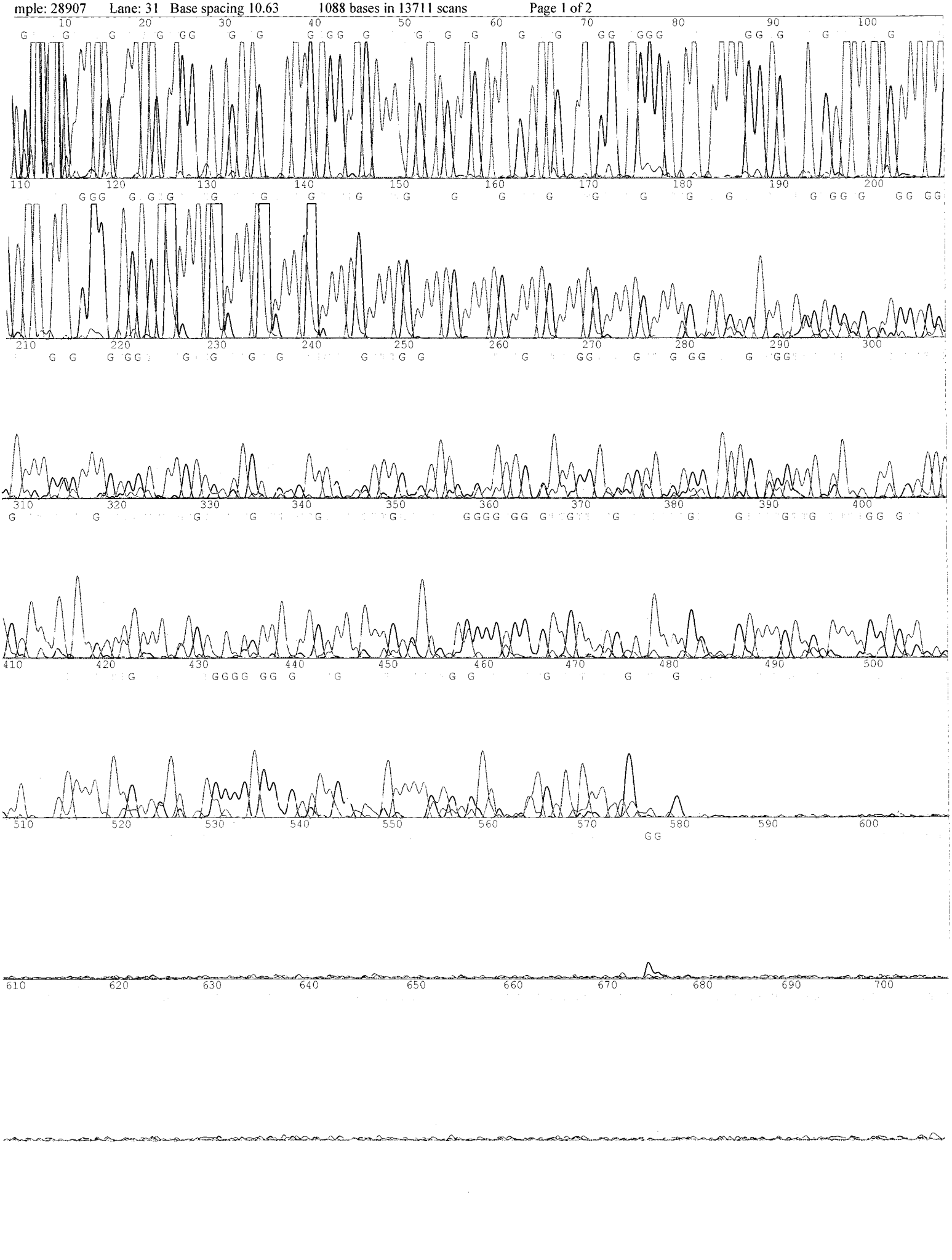
File: 28905 Run Ended: Jan 7, 2005, 4:56:16 Signal G:265 T:355 A:361 C:339 Comment:
Sample: 28905 Lane: 29 Base spacing 10.62 1119 bases in 13724 scans Page 1 of 2



File: 28906 Run Ended: Jan 7, 2005, 4:56:16 Signal G:166 T:215 A:217 C:201 Comment:
Sample: 28906 Lane: 30 Base spacing 10.63 1126 bases in 13711 scans Page 1 of 2



e: 28907 Run Ended: Jan 7, 2005, 4:56:16 Signal G:109 T:136 A:138 C:123 Comment:
mple: 28907 Lane: 31 Base spacing 10.63 1088 bases in 13711 scans Page 1 of 2



References

Ames, B.N. (1973) Carcinogens are mutagens: their detection and classification. *Environ Health Perspect*, **6**, 115-118.

Andoh, T. and Ishida, R. (1998) Catalytic inhibitors of DNA topoisomerase II. *Biochim Biophys Acta*, **1400**, 155-171.

Andreoli, C., Gigante, D. and Nunziata, A. (2003) A review of in vitro methods to assess the biological activity of tobacco smoke with the aim of reducing the toxicity of smoke. *Toxicol In Vitro*, **17**, 587-594.

Aquilina, G. and Bignami, M. (2001) Mismatch repair in correction of replication errors and processing of DNA damage. *J Cell Physiol*, **187**, 145-154.

Attia, S.M., Kliesch, U., Schriever-Schwemmer, G., Badary, O.A., Hamada, F.M. and Adler, I.D. (2003) Etoposide and merbarone are clastogenic and aneugenic in the mouse bone marrow micronucleus test complemented by fluorescence in situ hybridization with the mouse minor satellite DNA probe. *Environ Mol Mutagen*, **41**, 99-103.

Bagley, M.J., Anderson, S.L. and May, B. (2001) Choice of methodology for assessing genetic impacts of environmental stressors: polymorphism and reproducibility of RAPD and AFLP fingerprints. *Ecotoxicology*, **10**, 239-244.

Bah, A., Bachand, F., Clair, E., Autexier, C. and Wellinger, R.J. (2004) Humanized telomeres and an attempt to express a functional human telomerase in yeast. *Nucleic Acids Res*, **32**, 1917-1927.

Baker, S.M., Plug, A.W., Prolla, T.A., Bronner, C.E., Harris, A.C., Yao, X., Christie, D.M., Monell, C., Arnheim, N., Bradley, A., Ashley, T. and Liskay, R.M. (1996) Involvement of mouse Mlh1 in DNA mismatch repair and meiotic crossing over. *Nat Genet*, **13**, 336-342.

Barnes, D.E. and Lindahl, T. (2004) Repair and genetic consequences of endogenous DNA base damage in mammalian cells. *Annu Rev Genet*, **38**, 445-476.

Bertram, J.S. and Heidelberger, C. (1974) Cell cycle dependency of oncogenic transformation induced by N-methyl-N'-nitro-N-nitrosoquanidine in culture. *Cancer Res*, **34**, 526-537.

Bishop, A.J. and Schiestl, R.H. (2002) Homologous Recombination and Its Role in Carcinogenesis. *J Biomed Biotechnol*, **2**, 75-85.

Bois, P.R., Southgate, L. and Jeffreys, A.J. (2001) Length of uninterrupted repeats determines instability at the unstable mouse expanded simple tandem repeat family MMS10 derived from independent SINE B1 elements. *Mamm Genome*, **12**, 104-111.

Boland, C.R. and Goel, A. (2005) Somatic evolution of cancer cells. *Semin Cancer Biol.*

Boreiko, C., Mondal, S., Narayan, K.S. and Heidelberger, C. (1980) Effect of 12-O-tetradecanoylphorbol-13-acetate on the morphology and growth of C3H/10T1/2 mouse embryo cells. *Cancer Res*, **40**, 4709-4716.

Boysen, G. and Hecht, S.S. (2003) Analysis of DNA and protein adducts of benzo[a]pyrene in human tissues using structure-specific methods. *Mutat Res*, **543**, 17-30.

Chang, D.K., Goel, A., Ricciardiello, L., Lee, D.H., Chang, C.L., Carethers, J.M. and Boland, C.R. (2003) Effect of H(2)O(2) on cell cycle and survival in DNA mismatch repair-deficient and -proficient cell lines. *Cancer Lett*, **195**, 243-251.

Chang, P.L., Tucker, M.A., Hicks, P.H. and Prince, C.W. (2002) Novel protein kinase C isoforms and mitogen-activated kinase kinase mediate phorbol ester-induced osteopontin expression. *Int J Biochem Cell Biol*, **34**, 1142-1151.

Christmann, M., Tomicic, M.T. and Kaina, B. (2002) Phosphorylation of mismatch repair proteins MSH2 and MSH6 affecting MutSalpha mismatch-binding activity. *Nucleic Acids Res*, **30**, 1959-1966.

Dai, P. and Wong, L.J. (2003) Somatic instability of the DNA sequences encoding the polymorphic polyglutamine tract of the AIB1 gene. *J Med Genet*, **40**, 885-890.

de Boer, J.G., Provost, S., Gorelick, N., Tindall, K. and Glickman, B.W. (1998) Spontaneous mutation in lacI transgenic mice: a comparison of tissues. *Mutagenesis*, **13**, 109-114.

Debrauwere, H., Gendrel, C.G., Lechat, S. and Dutreix, M. (1997) Differences and similarities between various tandem repeat sequences: minisatellites and microsatellites. *Biochimie*, **79**, 577-586.

Denissenko, M.F., Pao, A., Tang, M. and Pfeifer, G.P. (1996) Preferential formation of benzo[a]pyrene adducts at lung cancer mutational hotspots in P53. *Science*, **274**, 430-432.

Dialyna, I., Tzanakakis, G., Dolapsakis, G. and Tsatsakis, A. (2004) A tetranucleotide repeat polymorphism in the CYP19 gene and breast cancer susceptibility in a Greek population exposed and not exposed to pesticides. *Toxicol Lett*, **151**, 267-271.

Don Porto Carero, A., Hoet, P.H., Verschaeve, L., Schoeters, G. and Nemery, B. (2001) Genotoxic effects of carbon black particles, diesel exhaust particles, and urban air particulates and their extracts on a human alveolar epithelial cell line (A549) and a human monocytic cell line (THP-1). *Environ Mol Mutagen*, **37**, 155-163.

Dubrova, Y.E., Jeffreys, A.J. and Malashenko, A.M. (1993) Mouse minisatellite mutations induced by ionizing radiation. *Nat Genet*, **5**, 92-94.

Dufour, E. and Larsson, N.G. (2004) Understanding aging: revealing order out of chaos. *Biochim Biophys Acta*, **1658**, 122-132.

Eckert, K.A. and Hile, S.E. (1998) Alkylation-induced frameshift mutagenesis during in vitro DNA synthesis by DNA polymerases alpha and beta. *Mutat Res*, **422**, 255-269.

Eckert, K.A., Mowery, A. and Hile, S.E. (2002) Misalignment-mediated DNA polymerase beta mutations: comparison of microsatellite and frame-shift error rates using a forward mutation assay. *Biochemistry*, **41**, 10490-10498.

Edwards, D.R. and Denhardt, D.T. (1985) A study of mitochondrial and nuclear transcription with cloned cDNA probes. Changes in the relative abundance of mitochondrial transcripts after stimulation of quiescent mouse fibroblasts. *Exp Cell Res*, **157**, 127-143.

Elkind, N.B., Goldfinger, N. and Rotter, V. (1995) Spot-1, a novel NLS-binding protein that interacts with p53 through a domain encoded by p(CA)n repeats. *Oncogene*, **11**, 841-851.

Evans, E. and Alani, E. (2000) Roles for mismatch repair factors in regulating genetic recombination. *Mol Cell Biol*, **20**, 7839-7844.

Ferguson, L.R. and Denny, W.A. (1991) The genetic toxicology of acridines. *Mutat Res*, **258**, 123-160.

Ferguson, L.R., Turner, P.M. and Denny, W.A. (2000) The mutagenic spectrum of acridine-linked aniline nitrogen mustards in AS52 cells: implications of DNA targeting with high selectivity for adenine or guanine bases. *Mutat Res*, **469**, 115-126.

Fernandez, A., Mondal, S. and Heidelberger, C. (1980) Probabilistic view of the transformation of cultured C3H/10T1/2 mouse embryo fibroblasts by 3-methylcholanthrene. *Proc Natl Acad Sci U S A*, **77**, 7272-7276.

Garry, S., Nesslany, F., Aliouat, E., Haguenoer, J.M. and Marzin, D. (2003) Assessment of genotoxic effect of benzo[a]pyrene in endotracheally treated rat using the comet assay. *Mutat Res*, **534**, 33-43.

Gasche, C., Chang, C.L., Rhees, J., Goel, A. and Boland, C.R. (2001) Oxidative stress increases frameshift mutations in human colorectal cancer cells. *Cancer Res*, **61**, 7444-7448.

Halliwell, B. and Whiteman, M. (2004) Measuring reactive species and oxidative damage in vivo and in cell culture: how should you do it and what do the results mean? *Br J Pharmacol*, **142**, 231-255.

Han, Y.H., Austin, M.J., Pommier, Y. and Povirk, L.F. (1993) Small deletion and insertion mutations induced by the topoisomerase II inhibitor teniposide in CHO cells and comparison with sites of drug-stimulated DNA cleavage in vitro. *J Mol Biol*, **229**, 52-66.

Harfe, B.D. and Jinks-Robertson, S. (2000) DNA mismatch repair and genetic instability. *Annu Rev Genet*, **34**, 359-399.

Health-Canada. (2004b) VICH GL 23 - Studies to Evaluate the Safety of Residues of Veterinary Drugs in Human Food: Genotoxicity Testing.

Health-Canada. (2004a) *Assessing and Managing the Health Risks of New Substances under The Canadian Environmental Protection Act (CEPA).* ***http://www.hc-sc.gc.ca/english/iyh/environment/cepa_new_substances.html.***

Healy, C., Wade, M., McMahon, A., Williams, A., Johnson, D. and Parfett, C. (in preparation) Flow Cytometric Detection of Tandem Repeat Mutations Induced by Various Chemical Classes. *Mut. Res.*

Heidenreich, E., Novotny, R., Kneidinger, B., Holzmann, V. and Wintersberger, U. (2003) Non-homologous end joining as an important mutagenic process in cell cycle-arrested cells. *Embo J*, **22**, 2274-2283.

Henderson, C.A. (1998) Therapeutic options for the treatment of colorectal cancer following 5-fluorouracil failure. *Semin Oncol*, **25**, 29-38.

Herman, J.G., Umar, A., Polyak, K., Graff, J.R., Ahuja, N., Issa, J.P., Markowitz, S., Willson, J.K., Hamilton, S.R., Kinzler, K.W., Kane, M.F., Kolodner, R.D., Vogelstein, B., Kunkel, T.A. and Baylin, S.B. (1998) Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma. *Proc Natl Acad Sci U S A*, **95**, 6870-6875.

Hienonen, T., Sammalkorpi, H., Enholm, S., Alhopuro, P., Barber, T.D., Lehtonen, R., Nupponen, N.N., Lehtonen, H., Salovaara, R., Mecklin, J.P., Jarvinen, H.,

Koistinen, R., Arango, D., Launonen, V., Vogelstein, B., Karhu, A. and Aaltonen, L.A. (2005) Mutations in two short noncoding mononucleotide repeats in most microsatellite-unstable colorectal cancers. *Cancer Res*, **65**, 4607-4613.

Hoffmann, G.R., Calciano, M.A., Lawless, B.M. and Mahoney, K.M. (2003) Frameshift mutations induced by three classes of acridines in the lacZ reversion assay in Escherichia coli: potency of responses and relationship to slipped mispairing models. *Environ Mol Mutagen*, **42**, 111-121.

Hornia, A., Lu, Z., Sukezane, T., Zhong, M., Joseph, T., Frankel, P. and Foster, D.A. (1999) Antagonistic effects of protein kinase C alpha and delta on both transformation and phospholipase D activity mediated by the epidermal growth factor receptor. *Mol Cell Biol*, **19**, 7672-7680.

Hsieh, P. (2001) Molecular mechanisms of DNA mismatch repair. *Mutat Res*, **486**, 71-87.

Hubacek, J.A., Pistulkova, H., Valenta, Z. and Poledne, R. (1999) (TTA)n repeat polymorphism in the HMG-CoA reductase gene and cholesterolaemia. *Vasa*, **28**, 169-171.

Humbert, O., Achour, I., Lautier, D., Laurent, G. and Salles, B. (2003) hMSH2 expression is driven by AP1-dependent regulation through phorbol-ester exposure. *Nucleic Acids Res*, **31**, 5627-5634.

Humbert, O., Hermine, T., Hernandez, H., Bouget, T., Selves, J., Laurent, G., Salles, B. and Lautier, D. (2002) Implication of protein kinase C in the regulation of DNA mismatch repair protein expression and function. *J Biol Chem*, **277**, 18061-18068.

Hunter, J.M., Crouse, A.B., Lesort, M., Johnson, G.V. and Detloff, P.J. (2005) Verification of somatic CAG repeat expansion by pre-PCR fractionation. *J Neurosci Methods*, **144**, 11-17.

IARC, P.a.c., Part 1, chemical, environmental and experimental data, IARC Monographs on the Evaluations of the Carcinogenic Risk of Chemicals to Humans, vol. 32, 1983. p. 477. (1983).

Inafuku, M., Toda, T., Iwasaki, H. and Oku, H. (2004) Analysis of microsatellite instability and loss of heterozygosity in human aortic atherosclerotic lesions. *Rinsho Byori*, **52**, 961-965.

Institute, S. (1989.) SAS/STAT User's Guide, Version 6, fourth ed., SAS Institute, Cary, NC,.

Jackson, A.L. and Loeb, L.A. (2000) Microsatellite instability induced by hydrogen peroxide in Escherichia coli. *Mutat Res*, **447**, 187-198.

Jackson, M.A., Stack, H.F. and Waters, M.D. (1993) The genetic toxicology of putative nongenotoxic carcinogens. *Mutat Res*, **296**, 241-277.

Jeffreys, A.J., Neumann, R. and Wilson, V. (1990) Repeat unit sequence variation in minisatellites: a novel source of DNA polymorphism for studying variation and mutation by single molecule analysis. *Cell*, **60**, 473-485.

Jensen, M., Leffers, H., Petersen, J.H., Nyboe Andersen, A., Jorgensen, N., Carlsen, E., Jensen, T.K., Skakkebaek, N.E. and Rajpert-De Meyts, E. (2004) Frequent polymorphism of the mitochondrial DNA polymerase gamma gene (POLG) in patients with normal spermiograms and unexplained subfertility. *Hum Reprod*, **19**, 65-70.

Jeong, S.Y., Shin, K.H., Shin, J.H., Ku, J.L., Shin, Y.K., Park, S.Y., Kim, W.H. and Park, J.G. (2003) Microsatellite instability and mutations in DNA mismatch repair genes in sporadic colorectal cancers. *Dis Colon Rectum*, **46**, 1069-1077.

Jiricny, J. and Marra, G. (2003) DNA repair defects in colon cancer. *Curr Opin Genet Dev*, **13**, 61-69.

Kier, L.D., Brusick, D.J., Auletta, A.E., Von Halle, E.S., Brown, M.M., Simmon, V.F., Dunkel, V., McCann, J., Mortelmans, K. and et al. (1986) The Salmonella typhimurium/mammalian microsomal assay. A report of the U.S. Environmental Protection Agency Gene-Tox Program. *Mutat Res*, **168**, 69-240.

Klaunig, J.E. and Kamendulis, L.M. (2004) The role of oxidative stress in carcinogenesis. *Annu Rev Pharmacol Toxicol*, **44**, 239-267.

Kunz, D., Gerard, N.P. and Gerard, C. (1992) The human leukocyte platelet-activating factor receptor. cDNA cloning, cell surface expression, and construction of a novel epitope-bearing analog. *J Biol Chem*, **267**, 9101-9106.

Kyrtopoulos, S.A., Georgiadis, P., Autrup, H., Demopoulos, N.A., Farmer, P., Haugen, A., Katsouyanni, K., Lambert, B., Ovrebo, S., Sram, R., Stephanou, G. and Topinka, J. (2001) Biomarkers of genotoxicity of urban air pollution. Overview and descriptive data from a molecular epidemiology study on populations exposed to moderate-to-low levels of polycyclic aromatic hydrocarbons: the AULIS project. *Mutat Res*, **496**, 207-228.

Li, Y.C., Korol, A.B., Fahima, T. and Nevo, E. (2004) Microsatellites within genes: structure, function, and evolution. *Mol Biol Evol*, **21**, 991-1007.

Lin, C.T., Lyu, Y.L., Xiao, H., Lin, W.H. and Whang-Peng, J. (2001) Suppression of gene amplification and chromosomal DNA integration by the DNA mismatch repair system. *Nucleic Acids Res*, **29**, 3304-3310.

Lu, Z., Liu, D., Hornia, A., Devonish, W., Pagano, M. and Foster, D.A. (1998) Activation of protein kinase C triggers its ubiquitination and degradation. *Mol Cell Biol*, **18**, 839-845.

Mankodi, A., Takahashi, M.P., Jiang, H., Beck, C.L., Bowers, W.J., Moxley, R.T., Cannon, S.C. and Thornton, C.A. (2002) Expanded CUG repeats trigger aberrant splicing of ClC-1 chloride channel pre-mRNA and hyperexcitability of skeletal muscle in myotonic dystrophy. *Mol Cell*, **10**, 35-44.

Mao, L., Lee, D.J., Tockman, M.S., Erozan, Y.S., Askin, F. and Sidransky, D. (1994) Microsatellite alterations as clonal markers for the detection of human cancer. *Proc Natl Acad Sci U S A*, **91**, 9871-9875.

Marnett, L.J. (2000) Oxyradicals and DNA damage. *Carcinogenesis*, **21**, 361-370.

Marra, G., Iaccarino, I., Lettieri, T., Roscilli, G., Delmastro, P. and Jiricny, J. (1998) Mismatch repair deficiency associated with overexpression of the MSH3 gene. *Proc Natl Acad Sci U S A*, **95**, 8568-8573.

Masurekar, M., Krueuzer, K.N. and Ripley, L.S. (1991) The specificity of topoisomerase-mediated DNA cleavage defines acridine-induced frameshift specificity within a hotspot in bacteriophage T4. *Genetics*, **128**, 193.

McCaffrey, T.A., Du, B., Consigli, S., Szabo, P., Bray, P.J., Hartner, L., Weksler, B.B., Sanborn, T.A., Bergman, G. and Bush, H.L., Jr. (1997) Genomic instability in the type II TGF-beta1 receptor gene in atherosclerotic and restenotic vascular cells. *J Clin Invest*, **100**, 2182-2188.

McCord, J.M. and Edeas, M.A. (2005) SOD, oxidative stress and human pathologies: a brief history and a future vision. *Biomed Pharmacother*, **59**, 139-142.
Meehan, T. and Straub, K. (1979) Double-stranded DNA steroselectively binds benzo(a)pyrene diol epoxides. *Nature*, **277**, 410-412.
Miller, M.L., Vasunia, K., Talaska, G., Andringa, A., de Boer, J. and Dixon, K. (2000) The tumor promoter TPA enhances benzo[a]pyrene and benzo[a]pyrene diolepoxide mutagenesis in Big Blue mouse skin. *Environ Mol Mutagen*, **35**, 319-327.
Miniati, P., Sourvinos, G., Michalodimitrakis, M. and Spandidos, D.A. (2001) Loss of heterozygosity on chromosomes 1, 2, 8, 9 and 17 in cerebral atherosclerotic plaques. *Int J Biol Markers*, **16**, 167-171.
Mo, Q.H., Li, X.R., Li, C.F., He, Y.L. and Xu, X.M. (2005) A novel frameshift mutation (+G) at codons 15/16 in a beta0 thalassaemia gene results in a significant reduction of beta globin mRNA values. *J Clin Pathol*, **58**, 923-926.
Mondal, S. and Heidelberger, C. (1980) Inhibition of induced differentiation of C3H/10T 1/2 clone 8 mouse embryo cells by tumor promoters. *Cancer Res*, **40**, 334-338.
Mori, Y., Yin, J., Rashid, A., Leggett, B.A., Young, J., Simms, L., Kuehl, P.M., Langenberg, P., Meltzer, S.J. and Stine, O.C. (2001) Instabilotyping: comprehensive identification of frameshift mutations caused by coding region microsatellite instability. *Cancer Res*, **61**, 6046-6049.
Nagashima, T., Matsuda, H., Silva, D.G., Petrovsky, N., Konagaya, A., Schonbach, C., Kasukawa, T., Arakawa, T., Carninci, P., Kawai, J. and Hayashizaki, Y. (2004) FREP: a database of functional repeats in mouse cDNAs. *Nucleic Acids Res*, **32**, D471-475.
Neidle, S. and Abraham, Z. (1984) Structural and sequence-dependent aspects of drug intercalation into nucleic acids. *CRC Crit Rev Biochem*, **17**, 73-121.
OECD. (1997a) Test Guideline 471. Bacterial Reverse Mutation Test. In: OECD Guideline for Testing of Chemicals. Paris, Organization for Economic Cooperation & Development.
OECD. (1997b) Test Guideline 474. Mammalian Erythrocyte Micronucleus Test. In: OECD Guideline for Testing of Chemicals. Paris, Organization for Economic Cooperation & Development.
OECD. (1997a) Test Guideline 475. Mammalian Bone Marrow Chromosome Aberration Test. In: OECD Guideline for Testing of Chemicals. Paris, Organization for Economic Cooperation & Development.
OECD. (1997c) Test Guideline 471. Bacterial Reverse Mutation Test. In: OECD Guideline for Testing of Chemicals. Paris, Organization for Economic Cooperation & Development.
OECD. (1997d) Test Guideline 473. In Vitro Mammalian Chromosome Aberration Test. In: OECD Guideline for Testing of Chemicals. Paris, Organization for Economic Cooperation & Development.
OECD. (1997e) Test Guideline 476. In Vitro Mammalian Cell Gene Mutation Test. In: OECD Guideline for Testing of Chemicals. Paris, Organization for Economic.
Pamphlett, R. (2004) Somatic mutation: a cause of sporadic neurodegenerative diseases? *Med Hypotheses*, **62**, 679-682.

Parfett, C.L. (2003) Combined effects of tumor promoters and serum on proliferin mRNA induction: a biomarker sensitive to saccharin, 2,3,7,8-TCDD, and other compounds at minimal concentrations promoting C3H/10T1/2 cell transformation. *J Toxicol Environ Health A*, **66**, 1943-1966.

Parfett, C.L. and Healy, C.M. (submitted) Tandemly repeated DNA sequence instabilities induced by two promoters of morphological transformation in vitro: a short-term response to
non-mutagenic agents in C3H/10T1/2 cells. *Mutagen.*

Park, B.K., Heo, M.Y., Park, H. and Kim, H.P. (2001) Inhibition of TPA-induced cyclooxygenase-2 expression and skin inflammation in mice by wogonin, a plant flavone from Scutellaria radix. *Eur J Pharmacol*, **425**, 153-157.

Pierce, A.J., Johnson, R.D., Thompson, L.H. and Jasin, M. (1999) XRCC3 promotes homology-directed repair of DNA damage in mammalian cells. *Genes Dev*, **13**, 2633-2638.

Prolla, T.A., Baker, S.M., Harris, A.C., Tsao, J.L., Yao, X., Bronner, C.E., Zheng, B., Gordon, M., Reneker, J., Arnheim, N., Shibata, D., Bradley, A. and Liskay, R.M. (1998) Tumour susceptibility and spontaneous mutation in mice deficient in Mlh1, Pms1 and Pms2 DNA mismatch repair. *Nat Genet*, **18**, 276-279.

Pryor, W.A., Stone, K., Zang, L.Y. and Bermudez, E. (1998) Fractionation of aqueous cigarette tar extracts: fractions that contain the tar radical cause DNA damage. *Chem Res Toxicol*, **11**, 441-448.

Radcliff, G. and Jaroszeski, M.J. (1998) Basics of flow cytometry. *Methods Mol Biol*, **91**, 1-24.

Reznikoff, C.A., Bertram, J.S., Brankow, D.W. and Heidelberger, C. (1973a) Quantitative and qualitative studies of chemical transformation of cloned C3H mouse embryo cells sensitive to postconfluence inhibition of cell division. *Cancer Res*, **33**, 3239-3249.

Reznikoff, C.A., Brankow, D.W. and Heidelberger, C. (1973b) Establishment and characterization of a cloned line of C3H mouse embryo cells sensitive to postconfluence inhibition of division. *Cancer Res*, **33**, 3231-3238.

Richard, G.F. and Paques, F. (2000) Mini- and microsatellite expansions: the recombination connection. *EMBO Rep*, **1**, 122-126.

Ross, J.S., Stagliano, N.E., Donovan, M.J., Breitbart, R.E. and Ginsburg, G.S. (2001) Atherosclerosis and cancer: common molecular pathways of disease development and progression. *Ann N Y Acad Sci*, **947**, 271-292; discussion 292-273.

Ruggiero, T., Olivero, M., Follenzi, A., Naldini, L., Calogero, R. and Di Renzo, M.F. (2003) Deletion in a (T)8 microsatellite abrogates expression regulation by 3'-UTR. *Nucleic Acids Res*, **31**, 6561-6569.

Shen, Z., Wells, R.L. and Elkind, M.M. (1994) Enhanced cytochrome P450 (Cyp1b1) expression, aryl hydrocarbon hydroxylase activity, cytotoxicity, and transformation of C3H 10T1/2 cells by dimethylbenz(a)anthracene in conditioned medium. *Cancer Res*, **54**, 4052-4056.

Shibuya, T. and Morimoto, K. (1993) A review of the genotoxicity of 1-ethyl-1-nitrosourea. *Mutat Res*, **297**, 3-38.

Sia, E.A., Jinks-Robertson, S. and Petes, T.D. (1997) Genetic control of microsatellite stability. *Mutat Res*, **383**, 61-70.

Siddique, Y.H. and Afzal, M. (2005) In vivo induction of sister chromatid exchanges (SCEs) and chromosomal aberrations (CAs) by lynestrenol. *Indian J Exp Biol*, **43**, 291-293.

Sies, H., Stahl, W. and Sevanian, A. (2005) Nutritional, dietary and postprandial oxidative stress. *J Nutr*, **135**, 969-972.

Slebos, R.J., Oh, D.S., Umbach, D.M. and Taylor, J.A. (2002) Mutations in tetranucleotide repeats following DNA damage depend on repeat sequence and carcinogenic agent. *Cancer Res*, **62**, 6052-6060.

Slebos, R.J. and Taylor, J.A. (2001) A novel host cell reactivation assay to assess homologous recombination capacity in human cancer cell lines. *Biochem Biophys Res Commun*, **281**, 212-219.

Snyder, R.D. (2000) Use of catalytic topoisomerase II inhibitors to probe mechanisms of chemical-induced clastogenicity in Chinese hamster V79 cells. *Environ Mol Mutagen*, **35**, 13-21.

Somers, C.M., Yauk, C.L., White, P.A., Parfett, C.L. and Quinn, J.S. (2002) Air pollution induces heritable DNA mutations. *Proc Natl Acad Sci U S A*, **99**, 15904-15907.

Squadrito, G.L., Cueto, R., Dellinger, B. and Pryor, W.A. (2001) Quinoid redox cycling as a mechanism for sustained free radical generation by inhaled airborne particulate matter. *Free Radic Biol Med*, **31**, 1132-1138.

Streisinger, G., Okada, Y., Emrich, J., Newton, J., Tsugita, A., Terzaghi, E. and Inouye, M. (1966) Frameshift mutations and the genetic code. This paper is dedicated to Professor Theodosius Dobzhansky on the occasion of his 66th birthday. *Cold Spring Harb Symp Quant Biol*, **31**, 77-84.

Subramanian, S., Madgula, V.M., George, R., Mishra, R.K., Pandit, M.W., Kumar, C.S. and Singh, L. (2003) Triplet repeats in human genome: distribution and their association with genes and other genomic regions. *Bioinformatics*, **19**, 549-552.

Sun, Y.J. and Han, X. (2004) Dynamic behavior of fragile X full mutations in cultured female fetal fibroblasts. *Acta Pharmacol Sin*, **25**, 973-976.

Tautz, D. and Schlotterer. (1994) Simple sequences. *Curr Opin Genet Dev*, **4**, 832-837.

Taylor, S.M. and Jones, P.A. (1979) Multiple new phenotypes induced in 10T1/2 and 3T3 cells treated with 5-azacytidine. *Cell*, **17**, 771-779.

Tippin, B., Kobayashi, S., Bertram, J.G. and Goodman, M.F. (2004) To slip or skip, visualizing frameshift mutation dynamics for error-prone DNA polymerases. *J Biol Chem*, **279**, 45360-45368.

Toth, G., Gaspari, Z. and Jurka, J. (2000) Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Res*, **10**, 967-981.

Trifunovic, A., Wredenberg, A., Falkenberg, M., Spelbrink, J.N., Rovio, A.T., Bruder, C.E., Bohlooly, Y.M., Gidlof, S., Oldfors, A., Wibom, R., Tornell, J., Jacobs, H.T. and Larsson, N.G. (2004) Premature ageing in mice expressing defective mitochondrial DNA polymerase. *Nature*, **429**, 417-423.

Trosic, I., Busljeta, I. and Modlic, B. (2004) Investigation of the genotoxic effect of microwave irradiation in rat bone marrow cells: in vivo exposure. *Mutagenesis*, **19**, 361-364.

Tucker, J.D., Auletta, A., Cimino, M.C., Dearfield, K.L., Jacobson-Kram, D., Tice, R.R. and Carrano, A.V. (1993) Sister-chromatid exchange: second report of the Gene-Tox Program. *Mutat Res*, **297**, 101-180.

Vergnaud, G. and Denoeud, F. (2000) Minisatellites: mutability and genome architecture. *Genome Res*, **10**, 899-907.

Vo, A.T., Zhu, F., Wu, X., Yuan, F., Gao, Y., Gu, L., Li, G.M., Lee, T.H. and Her, C. (2005) hMRE11 deficiency leads to microsatellite instability and defective DNA mismatch repair. *EMBO Rep.*

Watson, D.E., Cunningham, M.L. and Tindall, K.R. (1998) Spontaneous and ENU-induced mutation spectra at the cII locus in Big Blue Rat2 embryonic fibroblasts. *Mutagenesis*, **13**, 487-497.

Yauk, C. (1998) Monitoring for induced heritable mutations in natural populations: application of minisatellite DNA screening. *Mutat Res*, **411**, 1-10.

Yauk, C.L. (2004) Advances in the application of germline tandem repeat instability for in situ monitoring. *Mutat Res*, **566**, 169-182.

Yauk, C.L., Dubrova, Y.E., Grant, G.R. and Jeffreys, A.J. (2002) A novel single molecule analysis of spontaneous and radiation-induced mutation at a mouse tandem repeat locus. *Mutat Res*, **500**, 147-156.

Youn, C.K., Cho, H.J., Kim, S.H., Kim, H.B., Kim, M.H., Chang, I.Y., Lee, J.S., Chung, M.H., Hahm, K.S. and You, H.J. (2005) Bcl-2 expression suppresses mismatch repair activity through inhibition of E2F transcriptional activity. *Nat Cell Biol*, **7**, 137-147.

Young, L.C., Thulien, K.J., Campbell, M.R., Tron, V.A. and Andrew, S.E. (2004) DNA mismatch repair proteins promote apoptosis and suppress tumorigenesis in response to UVB irradiation: an in vivo study. *Carcinogenesis*, **25**, 1821-1827.

Yuryev, A., Patturajan, M., Litingtung, Y., Joshi, R.V., Gentile, C., Gebara, M. and Corden, J.L. (1996) The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine-rich proteins. *Proc Natl Acad Sci U S A*, **93**, 6975-6980.

Zhang, X.B., Urlando, C., Tao, K.S. and Heddle, J.A. (1995) Factors affecting somatic mutation frequencies in vivo. *Mutat Res*, **338**, 189-201.

Zhu, Y., Bye, S., Stambrook, P.J. and Tischfield, J.A. (1994) Single-base deletion induced by benzo[a]pyrene diol epoxide at the adenine phosphoribosyltransferase locus in human fibrosarcoma cell lines. *Mutat Res*, **321**, 73-79.

Zhuang, P., Kolbanovskiy, A., Amin, S. and Geacintov, N.E. (2001) Base sequence dependence of in vitro translesional DNA replication past a bulky lesion catalyzed by the exo- Klenow fragment of Pol I. *Biochemistry*, **40**, 6660-6669.