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

Monique Lavoie

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (Biology)

GRADE / DEGREE

Department of Biology

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Characterisation of a Conserved Motif Found Within the D*lx cis*-regulatory Element URE2 and its Effect on Mouse URE2-*lacZ* Activity in Transgenic Mice

TITRE DE LA THÈSE / TITLE OF THESIS

M. Ekker

DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

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

L. Bonen

D. Johnson

B. Willmore

Gary W. Slater

Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Characterisation of a conserved motif found within the *Dlx cis*-regulatory element URE2 and its effect on mouse URE2-*lacZ* activity in transgenic mice.

Monique V. Lavoie

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies in partial fulfillment of the requirements for the Master of Science degree
Ottawa-Carleton Institute of Biology
University of Ottawa

Thèse soumise à la Faculté des Études Supérieures et Postdoctorales
En vue de l'obtention de la Maîtrise ès Sciences
L'institut de Biologie d'Ottawa-Carleton
Université d'Ottawa

Your file Votre référence
ISBN: 978-0-494-61341-2
Our file Notre référence
ISBN: 978-0-494-61341-2

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES 7

ABBREVIATIONS AND ACRONYMS 9

ABSTRACT 10

RÉSUMÉ 12

ACKNOWLEDGEMENTS 14

STATEMENT OF CONTRIBUTIONS 15

1. INTRODUCTION 16

1.1 Transcription factors are proteins that participate in DNA transcriptional regulation 16

1.2 cis-Regulatory elements also regulate DNA transcription 17

1.3 Conserved cis-regulatory elements can be indicators of important functions 18

1.4 The distal-less (Dll) gene is critical for proximodistal axis formation 19

1.5 Vertebrate Dlx genes are related in sequence to Dll 19

1.6 Nomenclature and evolution of Dlx genes 20

Figure 1.1: Organization of murine Dlx genes with corresponding known conserved cis-regulatory elements. 22

1.7 Dlx gene expression in branchial arch and ectoderm derived embryonic tissues 24

1.8 Dlx genes are expressed in the developing brain 25

1.9 GABAergic interneuronal differentiation and migration requires DLX 26

Figure 1.2. Spatiotemporal Dlx gene expression in the developing mouse forebrain at E12.5. 27

1.10 Dlx mutants have varying degrees of developmental defects 30

1.11 DLX transcription factors' gene targets 31

1.12 Known regulators of Dlx gene expression 33

1.13 cis-regulatory elements that control Dlx gene expression 33

Figure 1.3: lacZ reporter mice showing Dlx cis-regulatory elements' activity in whole embryos at E11.5 35

1.14 Identification of the URE2 enhancer in mouse and zebrafish 37

1.15 Identification of trans-activating factors binding to the URE2 enhancer 38

1.16 URE2 mutants have a shift in URE2 activity in the developing forebrain 38

Figure 1.4 Effect of the FP10 mutation on URE2 enhancer activity. 40

1.17 Initial identification of five candidate proteins for binding on base pairs 699-719 of URE2 42

1.17.1 Prime candidate: REST/NRSF 42

1.17.2 Revised list of candidates adds 4 more proteins from the Nuclear Factor I (NFI) family of transcription factors 43

Figure 1.5: Schematic of a conserved sequence found in the mouse URE2 element. 44

1.17.2.1 Candidate 2: Nuclear Factor I-A 46

1.17.2.2 Candidate 3: Nuclear Factor I-B 47

1.17.2.3 Candidate 4: Nuclear Factor I-C 47

1.17.2.4 Candidate 5: Nuclear Factor I-X 48

1.18 Objectives of this thesis 48

2. MATERIALS AND METHODS 51

2.1 Production and screening of transgenic mouse tissues 51

2.2 X-gal staining of lacZ transgenic whole embryos and forebrain tissues 52

2.3 Cryostat sectioning of mouse forebrain tissues 53

2.4 Preparation of double stranded oligonucleotides 53

2.5 Preparation of (mURE2699-719, mURE2699-719mut10 and consensus NRSE)-pGL3-Control reporter constructs 54

2.6 Preparation of the NFIB-pSPORT6 cDNA construct 56

2.7 Preparation of empty pSPORT6 vector 58

2.8 Cell culture 59

2.9 Transfections 60

2.10 Dual luciferase reporter gene assays 60

2. 11 Statistical analysis of dual luciferase reporter gene assay results 61

2.12 Primers used in all experiments 61

Figure 2.1 Map of the pGL3 vector used in dual luciferase assays. 64

Figure 2.2. Map of the pGL4 vector. 66

Table 2.1 Calculations for transfections done in test 1 (Figure 3.5) in HeLa and Neuro-2a cells. 68

Table 2.2 Calculations for transfections done in test 2 (Figure 3.6-3.8) in HeLa and Neuro-2a cells. 69

Table 2.3 Calculations for transfections done in test 3 (Figure 3.9-3.11) in HeLa and Neuro-2a cells. 70

2.12 MatInspector bioinformatical analysis. *71*

3. RESULTS *72*

3.1 mURE2lacZ and mURE2mut10lacZ reporter gene expression pattern results *72*

Figure 3.1: Comparison of mURE2lacZ and mURE2mut10lacZ reporter gene expression in the subpallial telencephalon of transgenic mice at E11.5. ..... *74*

Figure 3.2: Comparison of mURE2lacZ and mURE2mut10lacZ reporter gene expression in the subpalial telencephalon of transgenic mice at E13.5. ..... *77*

Figure 3.3: Comparison of mURE2lacZ and mURE2mut10lacZ reporter gene expression in the subpalial telencephalon of transgenic mice at E15.5. ..... *79*

3.2 Dual luciferase reporter assays *81*

Figure 3.4: Comparison of mURE2lacZ and mURE2mut10clacZ reporter gene expression in the telencephalon of transgenic mice at P0. ..... *82*

3.2.1 Effect of REST on the transcriptional activity of the mURE2nt699-719 sequence *85*

Table 2.1 Calculations for transfections done in test 1 (Figure 3.5) in HeLa and Neuro-2a cells. *87*

Figure 3.5: Effect of REST on the transcriptional activity of the mURE2nt699-719 sequence in HeLa and Neuro-2a cells *88*

3.2.2 An updated MatInspector database reveals that NFI proteins are also strong candidates for mURE2nt699-719 regulatory activity. *90*

3.2.2.1 Effect of NFI proteins on transcriptional activity of mURE2nt699-719 in HeLa and Neuro-2a cells *91*

3.2.3 Effects of a smaller amount of NFI cDNA on mURE2nt699-719 activity in Neuro-2a and HeLa cells. *94*

Table 2.2 Calculations for transfections done in test 2 (Figure 3.6-3.8) in HeLa and Neuro-2a cells. *95*

Figure 3.6: Effect of NFI factors on the transcriptional activity of the mURE2nt699-719 sequence in HeLa cells. *96*

Figure 3.7: Effect of NFI factors on the transcriptional activity of the mURE2nt699-719 sequence in Neuro-2a cells. *98*

Figure 3.8: Comparison of the effect of NFI factors on the transcriptional activity of the mURE2nt699-719 sequence in HeLa and Neuro-2a cells. *100*

Table 2.3 Calculations for transfections done in test 3 (Figure 3.9-3.11) in HeLa and Neuro-2a cells. *103*

Figure 3.9 Effect of NFI factors on the transcriptional activity of the mURE2nt699-719 sequence in HeLa cells. *104*

Figure 3.10 Effect of NFI factors on the transcriptional activity of the mURE2nt699-719 sequence in Neuro-2a cells. 106

3.2.4 Statistical analysis of raw data 108

Figure 3.11 Comparison of the effect of NFI factors on the transcriptional activity of the mURE2nt699-719 sequence in HeLa and Neuro-2a cells. 109

4.DISCUSSION AND CONCLUSIONS 112

4.1 Analysis of the effects of mutagenesis of mURE2nt699-719 on URE2 activity in the forebrain. 112

4.1.1 Effects of mutagenized mURE2nt699-719 on URE2 activity in E11.5 transgenic mice forebrain. 112

4.1.2 Effects of mutagenized mURE2nt699-719 on URE2 activity in E13.5 transgenic mice forebrains. 113

4.1.3 Effects of mutagenized mURE2nt699-719 on URE2 activity in E15.5 and P0 transgenic mice forebrains. 113

4.1.4 Overall effects of mutagenized mURE2nt699-719 on URE2 activity in the transgenic mouse forebrain. 114

4.2 Variability of lacZ profiles between transgenic mouse lines limits analysis 115

4.3 Luciferase assays support evidence of a transcription factor binding site on mURE2nt699-719 117

4.3.1 Luciferase activity reported does not support the hypothesis that REST is the repressor binding on mURE2nt699-719 117

4.3.2 Luciferase activity reported supports the hypothesis that NFI protein(s) are repressors that bind to mURE2nt699-719 117

4.3.2.1 Analysis of single transfections using 300ng of NFI cDNA 118

4.3.2.2 Analysis of co-transfections using 300ng of each NFI cDNA to form heterodimers 119

4.3.2.3 Analysis of single transfections using 200ng of candidate NFI cDNA 121

4.3.2.4 Analysis of co-transfections using 200ng of each candidate NFI cDNA to form heterodimers 121

4.3.3 Potential causes for variability in raw data of luciferase assays 123

4.4. Potential role of the URE2 enhancer element in the regulation of Dlx expression 123

4.5. Future directions 124

4.6 Conclusion 124

BIBLIOGRAPHY 127

APPENDICES 135

Appendix 1. Supplementary photos of coronal sections of forebrains in mURE2lacZ transgenic mice (Line 3300). *136*

Appendix 2. Supplementary photos of coronal sections of forebrains in mURE2mut10alacZ transgenic mice. *137*

Appendix 3. Supplementary photos of coronal sections of forebrains in mURE2mut10clacZ transgenic mice. *138*

Appendix 4. Raw data and statistical results for test 1 (Figure 3.5) using REST cDNA and various constructs. *139*

Appendix 5. Raw data and statistical results for co-transfections of 300ng NFI in HeLa. *141*

Appendix 6. Raw data and statistical results for co-transfections of 300ng NFI in Neuro-2a. *144*

Appendix 7. Raw data and statistical results for co-transfections of 200ng NFI in HeLa. *147*

Appendix 8. Raw data and statistical results for co-transfections of 200ng NFI in Neuro-2a. *151*

Appendix 9. Statistical analysis of effects of variables on results for all tests *156*

LIST OF FIGURES AND TABLES

Figure 1.1: Organization of murine Dlx genes with corresponding known conserved cis-regulatory elements. 22

Figure 1.2. Spatiotemporal Dlx gene expression in the developing mouse forebrain at E12.5. 27

Figure 1.3: lacZ reporter mice showing Dlx cis-regulatory elements' activity in whole embryos at E11.5. 35

Figure 1.4 Effect of the FP10 mutation on URE2 enhancer activity. 40

Figure 1.5: Schematic of a conserved sequence found in the mouse URE2 element. 44

Figure 2.1 Map of the pGL3 vector used in dual luciferase assays 64

Figure 2.2 Map of the pGL4 vector 66

Table 2.1 Calculations for transfections done in test 1 (Figure 3.5) in HeLa and Neuro-2a cells. 68

Table 2.2 Calculations for transfections done in test 2 (Figure 3.6-3.8) in HeLa and Neuro-2a cells. 69

Table 2.3 Calculations for transfections done in test 3 (Figure 3.9-3.11) in HeLa and Neuro-2a cells. 70

Figure 3.1: Comparison of mURE2lacZ and mURE2mut10lacZ reporter gene expression in the subpallial telencephalon of transgenic mice at E11.5. 74

Figure 3.2: Comparison of mURE2lacZ and mURE2mut10lacZ reporter gene expression in the subpalial telencephalon of transgenic mice at E13.5. 77

Figure 3.3: Comparison of mURE2lacZ and mURE2mut10lacZ reporter gene expression in the subpalial telencephalon of transgenic mice at E15.5. 79

Figure 3.4: Comparison of mURE2lacZ and mURE2mut10clacZ reporter gene expression in the telencephalon of transgenic mice at P0. 82

Table 2.1 Calculations for transfections done in test 1 (Figure 3.5) in HeLa and Neuro-2a cells. 87

Figure 3.5: Effect of REST on the transcriptional activity of the mURE2nt699-719 sequence in HeLa and Neuro-2a cells 88

Table 2.2 Calculations for transfections done in test 2 (Figure 3.6-3.8) in HeLa and Neuro-2a cells. 95

Figure 3.6: Effect of NFI factors on the transcriptional activity of the mURE2nt699-719 sequence in HeLa cells. 96

Figure 3.7: Effect of NFI factors on the transcriptional activity of the mURE2nt699-719 sequence in Neuro-2a cells. 98

Figure 3.8: Comparison of the effect of NFI factors on the transcriptional activity of the mURE2nt699-719 sequence in HeLa and Neuro-2a cells. ____ *100*

Table 2.3 Calculations for transfections done in test 3 (Figure 3.9-3.11) in HeLa and Neuro-2a cells. ____ *103*

Figure 3.9 Effect of NFI factors on the transcriptional activity of the mURE2nt699-719 sequence in HeLa cells. ____ *104*

Figure 3.10 Effect of NFI factors on the transcriptional activity of the mURE2nt699-719 sequence in Neuro-2a cells. ____ *106*

Figure 3.11 Comparison of the effect of NFI factors on the transcriptional activity of the mURE2nt699-719 sequence in HeLa cells. ____ *109*

ABBREVIATIONS AND ACRONYMS

bp: base pairs
E11:5: embryonic development, day 11.5 after conception
E13.5: embryonic development, day 13.5 after conception
E15.5: embryonic development, day 15.5 after conception
FP: area in DNAseI footprint protected by protein(s), numbered for easy reference
P0: postnatal day 0 (birth)
N-2a : Neuro-2a
NRSE: neural restrictive silencing element
NFIA: Nuclear factor 1-A
NFIB: Nuclear factor 1-B
NFIC: Nuclear factor 1-C
NFIX: Nuclear factor 1-X
nt: nucleotide(s)
m : mouse
REST: RE-1 silencing transcription factor

ABSTRACT

The *Distal-less*-like genes, commonly referred to as *Dlx* genes, encode homeodomain transcription factors essential for the establishment of GABAergic interneurons in the ventral forebrain as well as their migration to the cortex. The *cis*-regulatory element URE2 (enhancer) is located upstream of *Dlx1* and is thought to regulate the transcription of *Dlx1* and/or *Dlx2*. The objective of this study was to examine a short conserved sequence within the URE2 element, corresponding to nucleotides 699-719, and to evaluate its contribution to URE2 activity in transgenic mouse forebrain. My results showed that a mutant version of the URE2 element, in which nucleotides 699-719 were modified, was active in the cortex and the ventricular zone of forebrain of the transgenic mice produced with the mutant enhancer, where *Dlx1/Dlx2* are not usually expressed. This indicated a loss of binding by a protein with a transcriptional repressor function. Previous analyses suggested that mURE2nt699-719 would be a good candidate for a REST repressor binding site. However, we were not able to observe any transcriptional regulation by REST in co-transfection experiments performed in cultured cells. This, as well as revised bioinformatical information, led us to identify four new candidates possibly using mURE2nt699-719 as a binding site, namely the Nuclear-Factor-1 proteins: NFIA, NFIB, NFIC and NFIX. These genes are associated with transcriptional regulation, being repressors and/or activators of gene expression in the brain. We showed that NFIC was able to utilize the mURE2nt699-719 sequence to reduce the luciferase reporter expression in neuronal cells, which would explain the loss of repression in our transgenic mice at embryonic days E11.5 and E13.5. These preliminary results also allowed us to suggest that a heterodimer between NFIC/NFIB and NFIC/NFIX binds to the mURE2699-719 sequence at embryonic day E15.5 and postnatal day 0 (P0) of the transgenic

mice since the mutated enhancer's activity seems to indicate a loss of an activator at those stages. These results suggest that these factors could possibly be required for the regulation of the *Dlx* genes by URE2. These results will also help clarify the role of URE2 by identifying transcription factors that contribute to the genetic program involved in the cellular organization of the developing forebrain in vertebrates.

RÉSUMÉ

Les gènes *Distal-less-like*, appelés *Dlx*, codent pour des facteurs de transcription à boîte homéo essentiels pour l'établissement des interneurones GABAergiques dans le cerveau antérieur ventral ainsi que leur migration au cortex. L'élément agissant en *cis* (enhancer) URE2 est situé en amont du gène *Dlx1* et est soupçonné d'être l'un des éléments contrôlant la transcription des gènes *Dlx1* et/ou *Dlx2*. L'objectif de ce projet fut d'examiner une séquence spécifique à l'intérieur d'URE2, correspondant aux nucléotides 699-719 (mURE2nt699-719), et d'évaluer sa contribution à l'activité d'URE2 dans le cerveau de souris transgéniques. Mes résultats ont démontré qu'une version mutante de l'élément URE2, dans laquelle les nucléotides 699-719 ont été modifiés, était active dans le cortex et la zone ventriculaire des cerveaux de souris transgéniques mutantes où les gènes *Dlx1/Dlx2* ne sont pas habituellement exprimés, indiquant la perte de liaison d'une protéine de type répresseur. Des analyses antérieures proposaient que mURE2nt699-719 serait un bon candidat pour lier le répresseur REST. Toutefois, nous n'avons pas été capables d'observer de régulation transcriptionnelle par REST lors d'expériences de co-transfection dans des cellules en culture. Ceci, et des nouvelles informations nous ont emmené à l'identification de quatre nouveaux candidats se liant possiblement à la séquence mURE2nt699-719, soit les facteurs Nuclear-Factor-1: NFIA; NFIB; NFIC et NFIX. Ces gènes sont associés à la régulation transcriptionnelle, agissant soit comme répresseurs ou activateurs de l'expression génique dans le cerveau. Nous avons montré que NFIC était capable d'utiliser le site mURE2nt699-719 pour diminuer l'expression de la luciférase dans des cellules neuronales, ce qui expliquerait la perte de répression dans nos souris transgéniques aux jours 11.5 et 13.5 embryonnaires. Les résultats préliminaires nous ont aussi permis de suggérer qu'un

hétérodimère parmi NFIC/NFIB et NFIC/NFIX se lie à la séquence mURE2nt699-719 aux jours E15.5 et au jour de naissance (P0) des souris transgéniques puisque l'activité de l'enhancer muté semble indiquer la perte d'un activateur à ces stades. Ces résultats suggèrent que ces facteurs sont possiblement requis pour la régulation des *Dlx* par URE2. Ces résultats aideront à clarifier le rôle d'URE2 en identifiant des facteurs de transcription requis pour assurer un programme génétique impliqué dans l'organisation cellulaire du cerveau antérieur des vertébrés au cours du développement.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank Dr. Marc Ekker for allowing me to join his lab and learn more than I ever imagined I could. With his guidance, and especially his patience, I have accomplished something that five years ago I did not think was possible. I have the greatest respect for him as a supervisor and I would not hesitate in collaborating with him in the future. Thanks to all of the previous and current lab members for providing a creative and healthy work environment. To rise in the morning and want to go to work is something desired by many and I thank you guys for providing me with that.

Special thanks to Luc Poitras, Ph.D., for introducing me to the project and training me. Thank you for the many, many, times I asked for your advice and help. Your hard work, work ethic, and sense of humor really helped me during my project and I thank you sincerely. Thank you to Gary Hatch for also being a great help with general tasks and especially in helping me out with extra work near the end of this project; as well as the many occasions I asked for some assistance with difficult PCR's. Thanks also to Man Yu and Ryan MacDonald for their assistance and camaraderie throughout my entire project. I would like to thank Sylvie Emond, Tami Janveau and Angela Seguin for helping with the mouse work and training me to work with mice. Thanks also to the Biology graduate students and staff for their support and friendship throughout my Master's project, with some friendships extending back to when I was an undergraduate. Thank you and I value all of our memories!

Finally I would like to thank my family and family in-law, who have especially supported me emotionally throughout this endeavour. Special thanks to my husband, without whom this would not have been possible at all. Thanks for being an amazing spouse and statistician, helping me greatly with my finances and the statistical analyses for this project.

STATEMENT OF CONTRIBUTIONS

For the work presented in this thesis, the following people contributed their skills and time: PCR screening of transgenic mice, all *lacZ* staining, cryostat sectioning of embryonic tissue, subcloning of DNA constructs, cell culture, some co-transfections and the statistical analyses were done by the author; cell culture and most of the co-transfections were done by the author's summer student, Abbie Gagnon; injections of the mURE2mut10*lacZ* DNA constructs to create transgenic mouse lines were done by Adrianna Gambarotta; and finally, DNaseI footprint and initial MatInspector analysis were both done by Luc Poitras, Ph.D.

1. INTRODUCTION

1.1 Transcription factors are proteins that participate in DNA transcriptional regulation

In eukaryotes, DNA transcription is dependent on the recruitment of the preinitiation complex (PIC) in the promoter area of coding sequences, which the RNA Polymerase II is able to recognize and use as a transcription start site (TSS). Transcription factors are proteins that can attach themselves to promoters, *cis*-regulatory elements, and other proteins, which allow binding of, or blocks binding of, transcriptional machinery such as the PIC. These factors bind DNA in a sequence-specific manner, and thus only interact with the transcriptional machinery where required. When a transcription factor is no longer available to do its job, or is rendered non-functional, this often leads to disease (Kadonaga 2004).

The Distal-Less 2 (DLX2) homeobox transcription factor has been shown to be involved in a genetic cascade required for interneuronal differentiation and migration during early forebrain development (Anderson et al. 1997a). Homeobox transcription factors such as the DLX transcription factors share a conserved DNA binding domain called a homeodomain, which has been shown to be crucial to their transcriptional regulator activity. These transcription factors are crucial for gene regulation during early development since many genes turn on and off in a spatiotemporal-specific manner during that time. This type of complex control system requires many transcriptional regulators which carry a wide array of DNA- and protein-binding domains such as homeodomains, required for their regulatory activity.

Transcription factors recognize DNA sites that are very specific for their binding. Binding sites can vary from 6-8 base pairs (bp) to as long as 20-22 bp, and can be found throughout the genome. Transcription factor binding sites that do not represent the consensus of identified binding sites are also sometimes utilized, adding complexity to an already intricate system of transcriptional regulation (Mortazavi et al. 2006; Zhang et al. 2006).

1.2 *cis*-Regulatory elements also regulate DNA transcription

While transcription factors are able to initiate, repress or silence gene expression by binding to recognition sites in or nearby promoter regions to interact with transcriptional machinery, other *cis*-regulatory DNA sequences play a role in spatiotemporal-specific gene expression. Thus, *cis*-regulatory elements can be classified in two categories, the first category containing all elements that are defined as promoters and the cis-regulatory elements proximal to them; the second class of cis-regulatory elements contains the distal elements, which are classified according to their activity. This category includes enhancers, silencers, and insulators (Maston et al. 2006).

Enhancer elements are able to increase the rate of their target genes' expression in a manner that is independent of the way they are oriented, and have been shown to be active on genes that are up to hundreds of kilobase pairs away from them (Khoury and Gruss, 1983).

Proteins such as transcription factors can bind in a sequence specific manner to these *cis*-regulatory elements and create a signal for the transcriptional machinery. These elements are therefore just as important as transcription factors in the regulation of genes required in genetic cascades that are crucial during development.

1.3 Conserved *cis*-regulatory elements can be indicators of important functions

Some non-coding sequences are conserved among vertebrates and invertebrates, and have been shown to have *cis*-regulatory roles in transcriptional regulation (Kim and Pritchard, 2007). Some non-coding elements (NCE's) have been discovered that are shared by the human and pufferfish genomes, which shared a common ancestor over 450 million years ago. This is much longer than the 60 million years that separate the mouse and human from their common ancestor, and is another testament to their possible importance (Muller et al. 2002). When sequence conservation of 100% is present over 200bp of genomic DNA, it is labelled an ultraconserved sequence. One group of scientists identified almost 500 ultraconserved elements that were shared by the human, rat and mouse genomes, and may have had been evolutionarily conserved for the sake of their important roles for the ontogeny of mammals (Bejerano et al. 2004). Some ultraconserved non-coding elements in the human genome were found in "gene deserts", which are intergenic areas up to 100kb away from the nearest gene. The nearest genes were almost always involved in early development, making these particular conserved elements possible distal enhancers of these genes. It has also been suggested that highly conserved non-coding elements often play an essential role in gene regulation during development (Woolfe et al. 2005).

Conserved *cis*-regulatory elements often contain binding sites for multiple transcription factors, which have chromatin remodelling, transcriptional activation, repression or silencing roles. These elements provide an important contribution to the complex genomic instructions for developmental gene regulation (Woolfe et al. 2005).

1.4 The distal-less (*Dll*) gene is critical for proximodistal axis formation

The *distal-less* (*Dll*) gene is a member of the homeobox family of genes, where homeoboxes are present in each transcript, and once translated encode the homeodomain required for their DNA-binding activity.

Some of the first identified homeobox genes are known to be involved in the development of axis formation in embryos during development (McGinnis and Krumlauf 1992). Early development studies in *Drosophila* showed that embryos not only require proper dorsoventral, and anterior-posterior axis formation to develop the body segments correctly, but also a proximodistal axis for proper limb and antennae development. The *distal-less* (*Dll*) gene was discovered in the fruit fly when it was noted that distal limb development was significantly altered when *Dll* was mutated (Cohen et al. 1989a). It appeared then that *Dll* transcript abundance was inversely proportional to the severity of the decreased limb size and segment development and organization, suggesting that *Dll* was required for the proximodistal organization of the developing limb cells.

In a loss-of-function study done by the same group, a *Dll* knockout fly exhibited a lack of larval rudimentary limbs and antennae, which serve as the ears and nose of the fly, contributing to the embryonic lethality observed with the loss of *Dll* function (Cohen and Jürgens, 1989b).

1.5 Vertebrate *Dlx* genes are related in sequence to *Dll*

Genes related in sequence to the *Dll* gene have been identified in vertebrates, and have been named the *Distal-less-like* (*Dlx*) genes. These genes contain the homeodomain that shares sequence similarities with the *Dll* sequence. There have been six *Dlx* genes

identified in vertebrates, except for teleost fish such as the zebrafish and pufferfish, which have eight. In vertebrates they are organized as bigene clusters, where the genes *Dlx1* and *Dlx2* form the first pair; *Dlx3* and *Dlx4* are paired together; and *Dlx5* and *Dlx6* form the last pair (Figure1.1). These genes are also organized in an inverted manner and are convergently transcribed. Each of the clusters are on different chromosomes. The clusters that include *Dlx1/2* and *Dlx5/6* have been shown to be expressed in the forebrain during development, as is *Dll* in the fruitfly (Anderson et al. 1997a; Panganiban et al. 1997). *Dlx3* and *Dlx4* are expressed in structures such as the retina and olfactory placodes, the latter giving rise to the nose (Akimenko et al. 1994). The combined expression patterns of vertebrate *Dlx* genes are reminiscent of *Dll* expression in antennae and other sensory organs in *Drosophila*. Research has suggested that genome duplication events occurred after the echinoderm and the urochordate lineages split, as well as after the split of the vertebrates from the urochordates (Panganiban and Rubenstein, 2002; Stock et al. 1996). Based on the sequence similarity of the paralogs and orthologs' homeoboxes; gene expression profiles, and functional studies, researchers suspect that there were tandem and whole cluster duplications, followed by degeneration or loss-of-function of two copies of the *Dlx* genes, resulting in six *Dlx* genes in vertebrates.

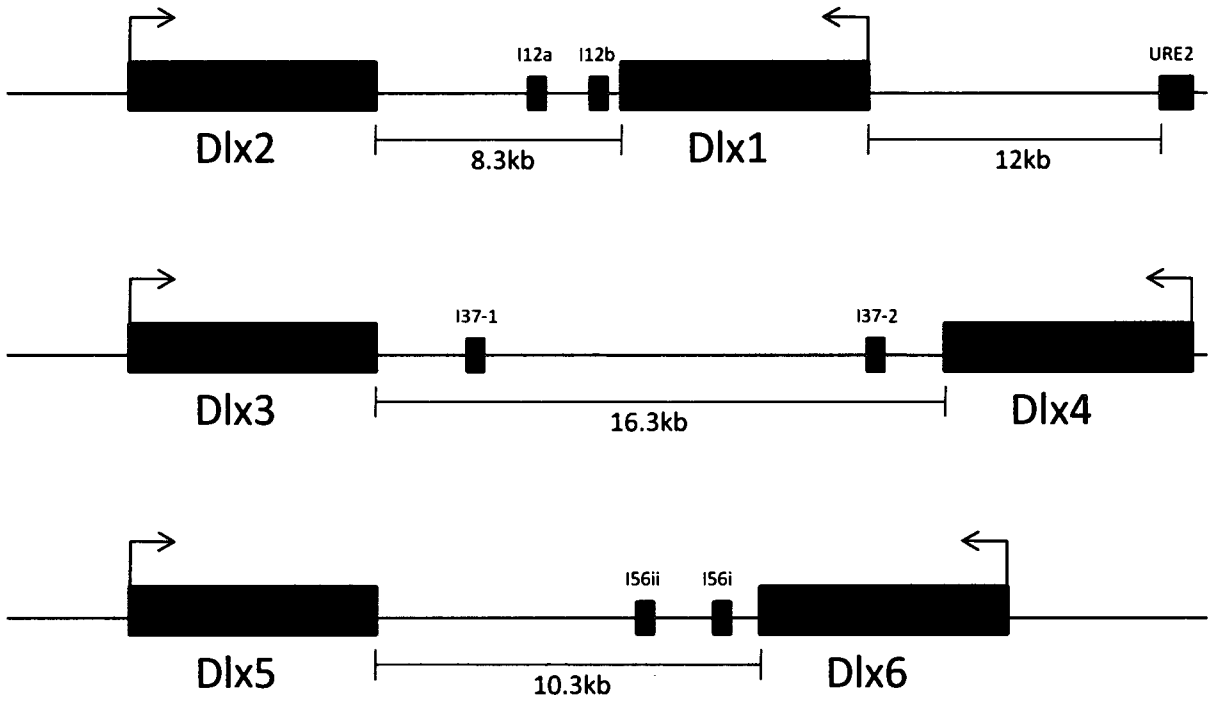
1.6 Nomenclature and evolution of *Dlx* genes

Since the *Dlx* genes all contain homeoboxes and that these share approximately 80% of the *Drosophila Dll* homeobox nucleotide sequence, it has been hypothesized that they derived from the ancestral *Dll* gene, of which there is only once copy in all invertebrates studied so far. The original function of *Dll* can only be firmly confirmed by looking at *Dll*

gene expression and functional characterization studies in more primitive taxa such as dogfish and gnathostomes like the lamprey. Studying genetics in these taxa is made difficult because of the levels of complexity for culturing and maintaining animals for a few generations, therefore researchers have not fully elucidated the role of the ancestral *Dll* gene yet. *Dlx* genes were named according to sequence similarities among the different genomes in which they were first identified. This is how they became known as *Dlx1* through 6, where all *Dlx1* genes identified in species investigated were more similar to each other in sequence than to other members of the *Dlx* family. Within the same genome there are also two clades of *Dlx* genes, where the *Dlx* genes forming the clades have more sequence similarity to each other than to the gene member with which they form a bigene cluster. For example, in *Dlx1/2* bigene cluster, the *Dlx1* gene is more similar in sequence to the *Dlx4* and *Dlx6*, and these form a first clade; while *Dlx2* is closer in sequence to *Dlx3* and *Dlx5* and these three form the second clade (for *Dlx* gene organization, see Figure 1.1).

This hypothesized evolutionary process is supported by findings showing that *Dlx* genes are linked with the *Hox* genes responsible for anterior-posterior and proximo-distal appendicular pattern specification. Additionally, the only copy of the *Dll* paralog, *ceh-43*, in the nematode *Caenorhabditis elegans* is linked to a single *Hox* cluster. In vertebrates, these genes are named *Hox* genes after the homeobox that they all share, and they were first discovered in *Drosophila*, where they are named *HOM-C* genes. Each *Dlx* bigene cluster is linked to a different *Hox* gene family member. The *Dlx1/2* cluster is linked to the *HoxD* cluster; *Dlx3/4* is linked to the *HoxA* cluster; the *Dlx5/6* genes are linked to the *HoxB* cluster; and no *Dlx* genes have been identified thus far has been linked to the *HoxC* cluster, which

Figure 1.1: Organization of murine *Dlx* genes with corresponding known conserved *cis*-regulatory elements. *Dlx* genes are represented by blue boxes, with arrows representing transcription start site and direction. Known intergenic conserved cis-regulatory elements are shown in red boxes. *Dlx1*, *Dlx4* and *Dlx6* genes form a clade within the *Dlx* family of genes due to their close sequence similarity, as do *Dlx2*, *Dlx3* and *Dlx5*. Figure is not to scale.


I12a
I12b
URE2
Dlx2
8.3kb
Dlx1
12kb
I37-1
I37-2
Dlx3
16.3kb
Dlx4
I56ii
I56i
Dlx5
10.3kb
Dlx6

could be the consequence of the loss of a bigene cluster, as previously mentioned (Krumlauf 1994; Stock et al. 1996).

1.7 *Dlx* gene expression in branchial arch and ectoderm derived embryonic tissues

Expression analysis of vertebrate *Dlx* genes provided the first indications as to their role in development. Their expression has been detected, among other areas, in cells of the branchial arches and tissues derived from ectoderm. All of the *Dlx* genes have been shown to be expressed in the apical ectodermal ridge (AER) of the developing limb bud, which reflects *Dll* expression and function in distal appendages and limb development (reviewed in (Panganiban and Rubenstein, 2002).

Dlx5 gene expression has been observed in developing osteoblasts, and they seem to have a role in the development of craniofacial structures, as seen by the lack of bones on the side and base of the skull in the *Dlx5* null allele mouse (Acampora et al. 1999).

Dlx4 genes are expressed in normal bone marrow, and are probably implicated in haematopoiesis, since cells lacking *Dlx4* function undergo increased apoptosis (Shimamoto et al. 1997; Shimamoto et al. 2000).

Dlx3 expression has been reported in the olfactory placodes, which develops into nasal structures, as mentioned previously (Akimenko et al. 1994). In zebrafish and chick embryos its expression has also been reported in the otic vesicle and other ectoderm derived structures that will form into inner ear and teeth structures later in development, making *Dlx3* important for nose, ear and teeth development (Ekker et al. 1992; Pera and Kessel, 1999). Since its transcripts have mostly been identified ectoderm derived tissues, this suggests that

Dlx3 is required primarily for these structures and not so much for the central nervous system.

Dlx1 and *Dlx2* genes are expressed in the entire first and second branchial arches in both mouse and zebrafish developing embryos. They share an expression pattern in the middle and distal sections of the arches with other *Dlx* genes, however, they are the only ones expressed proximally (Depew et al. 2002; Panganiban and Rubenstein, 2002; Qiu et al. 1997; Zerucha and Ekker, 2000a).

1.8 Dlx genes are expressed in the developing brain

Only four of the six known vertebrate *Dlx* genes are expressed in the developing brain. *Dlx1/2* and *Dlx5/6* are clusters for which there has been evidence of brain specific expression, and this was shown in the mouse and zebrafish (Bulfone et al. 1993; Liu et al. 1997). *Dlx1* and *Dlx2* are expressed with patterns that are almost identical rostrally, located in the proliferative area called the subventricular zone (SVZ) of the medial ganglionic eminence (MGE) and lateral ganglionic eminence (LGE), as well as in cells in the septum (SE) of the developing forebrain at E12.5. *Dlx1* and *Dlx2* genes are also expressed caudally in the diencephalon, notably in the Zona Limitans (ZL) and the Posterior Entopeduncular area (PEP).

Dlx5 is expressed in the SVZ along with *Dlx1* and *Dlx2*, where proliferation and differentiation of cells occurs. *Dlx5* is also expressed along with *Dlx6* in the mantle zone, the main site of *Dlx6* expression. Therefore, *Dlx1/2* are both expressed in the VZ of both the LGE and MGE; they are also expressed in the SVZ as is *Dlx5*, and *Dlx5/6* share a pattern of

expression in the mantle zone beneath the SVZ and VZ's of both the LGE and MGE at E12.5 (see Figure 1.2).

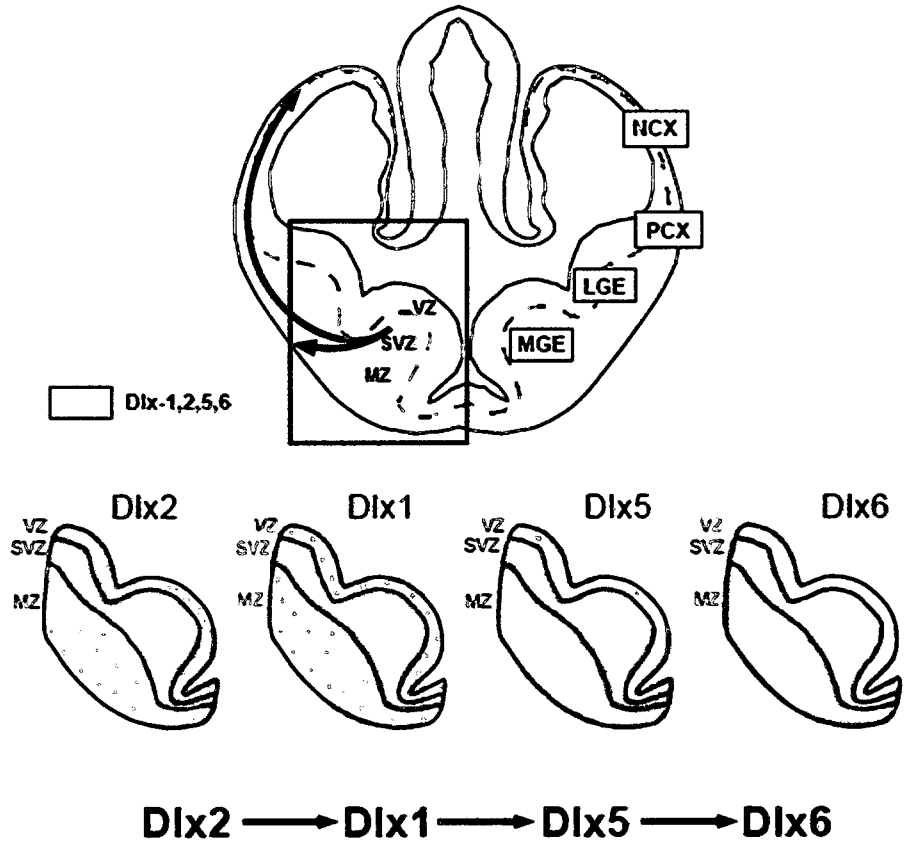
Further to having restricted spatial expression, *Dlx* gene transcription is also initiated in a temporally specific manner. The *Dlx2* gene is the first to be expressed, followed by *Dlx1*, *Dlx5* and then *Dlx6*. The later *Dlx5/6* expression coincides with cells that are more mature and migrating, therefore associating *Dlx1/2* gene expression with proliferating and differentiating cells of the VZ and SVZ, and Dlx5/6 with the more mature migrating interneuronal cells travelling through the mantle zone on their way to the neocortex via a tangential migration (Liu et al. 1997; Panganiban and Rubenstein, 2002).

1.9 GABAergic interneuronal differentiation and migration requires DLX

The *Dlx1/2* genes have been shown to be expressed by cells in the LGE and MGE between E11.5 and E15.5, which were already discussed as being proliferative zones of undifferentiated neural cells. *Dlx1/2* expression corresponds with *Gad1* (aka *Gad67*) and *Gad2* (aka *Gad65*) expression, which are genes that encode for the two-peptide glutamic acid decarboxylase (GAD), an enzyme that converts glutamate into γ-aminobutyric acid, or GABA (Stühmer et al. 2002). GABA is a neurotransmitter inhibitory in nature, which blocks signals in the neuronal signalling network, and required to maintain the balance of excitatory signals and blocking signals. DLX2 and DLX5, when ectopically expressed, are able to activate the expression of the genes required for GAD syntheis, and therefore able to induce the GABAergic phenotype in the cells affected (Stühmer et al. 2002).

The cells in the LGE and MGE go on to differentiate into GABAergic cortical interneurons, and these are classified into subtypes which express one or more peptides such

Figure 1.2. Spatiotemporal Dlx gene expression in the developing mouse forebrain at E12.5. The ventricular (VZ) and subventricular zones (SVZ) of the lateral and medial ganglionic emminences (LGE and MGE) are the sites where most cell proliferation occurs in the developing forebrain, correlating with Dlx1, Dlx2, Dlx5 and Dlx6 gene expression. Tangential migration of differentiating interneuronal precursor cells is depicted by arrows going from ventral to dorsal telencephalon. Intensity of Dlx-positive cells' signal is represented by solid color (more intense) and dots (less intense). NCX = neocortex; PCX = palliocortex. (Modified from Panganiban and Rubenstein, 2002)


NCX
PCX
LGE
MGE
VZ
SVZ
MZ
Dlx-1,2,5,6
Dlx2
Dlx1
Dlx5
Dlx6
Dlx2 → Dlx1 → Dlx5 → Dlx6

as Calbindin (CB), calretinin (CR), somatostatin (SOM), and neuropeptide-Y (NPY) reviewed in (Wonders and Anderson 2006). The cells secreting these peptides make up specific subpopulations of cortical interneurons depending on morphology, molecular markers, and physiological properties. Recent studies suggest that specific cortical GABAergic interneuron subpopulations may be produced depending on which enhancers are utilized to regulate *Dlx* genes at particular times and places during embryonic development (Ghanem et al. 2007; Ghanem et al. 2008; Marin and Rubenstein 2001; Potter et al. 2009).

One recent study has identified potential mechanisms behind the importance of the *Dlx* gene expression and the migration of GABAergic interneurons from the subpallium to the neocortex of the developing forebrain. Results indicate that when *Dlx1/2* genes are not expressed, the interneuron precursor cells in the VZ and SVZ have longer neuronal processes, which indicates that *Dlx1/2* expression is important to restrict their length (Cobos et al. 2007). Also, when *Dlx1/2* genes are not expressed, the interneuronal precursor cells undergo apoptosis, which confirmed previous reports of the importance of *Dlx* genes for cell survival in the MGE and LGE (Cobos et al. 2005). It is also of note that cells are able to tangentially migrate up to the neocortex, but have dendritic and axonal growth defects in these *Dlx1/2* mutants (Cobos et al. 2007). *Dlx1/2* genes were also shown to control the growth of neurite processes and promote cell migration through the repression of p21-activated serine/threonine kinase 3, also known as PAK3, which is a kinase responsible for controlling cytoskeleton actin and microtubule dynamics during cell motility, growth and transformation (Cobos et al. 2007).

1.10 *Dlx* mutants have varying degrees of developmental defects

Targeted loss of Dlx gene function was carried in the mouse and phenotypes include the malformation of branchial arch derived structures such as craniofacial bone structures. The *Dlx1*, *Dlx2* and *Dlx1/2* null mutant mice all exhibited defects in structures derived from the first and second branchial arch, although defects in the single mutant mice were not as severe as the double null mutants. Among the craniofacial bones affected in all three mutants were the ala temporalis, which is one of the two bones that make up the alisphenoid bone in the base of the skull, and is where all mutants had observed defects (Qiu et al. 1997). Other studies investigating the *Dlx1/2* double null mutant generated show an effect on upper molars, as they are lacking in these mutant newborn mice but are normal in $Dlx1^{-/-}$ and $Dlx2^{-/-}$ mice (Thomas et al. 1997).

The $Dlx1^{-/-}$ mutant mice survive until approximately P30, where as $Dlx1/2^{-/-}$ and $Dlx2^{-/-}$ mutants die a few hours after birth (Qiu et al. 1997).

$Dlx3^{-/-}$ null mutants die before birth due to the lack of vascularisation in the placenta, and are therefore difficult to analyse (Morasso et al. 1999).

$Dlx5^{-/-}$ null mutants have malformed ears, nose, and mandibles, probably due to their delayed osteogenesis, and die at birth (Acampora et al. 1999; Depew et al. 1999). The *Dlx5* null mice helped identify the link between *Dlx5* expression and its roles in development of all four branchial arches, since each of the four arches were affected in one way or another in the *Dlx5* null mouse.

Although limb development appears normal in *Dlx1/2* null mutants, *Dlx5/6* null mutants show an overall decrease in hindlimb size, split hindlimbs, kinked tails and exencephaly, where the brain is located outside of the skull (Robledo et al. 2002). *Dlx2/5*

null mutants also show abnormalities in limb development (Panganiban and Rubenstein, 2002).

Dlx mutants have also been identified in which brain structures are affected. *Dlx1* null mutants exhibit a decreased number of calretinin and somatostatin secreting subtypes of GABAergic interneuronal cells, and consequently showing reduced inhibitory postsynaptic current, resulting in mice displaying symptoms of epilepsy (Cobos et al. 2005). *Dlx1/2* double knockout experiments resulted in transgenic mice that exhibited a decrease in differentiated cortical interneurons and migration of these cells from the subpallium, resulting in a reduction in GABAergic interneurons in the cortex and an accumulation of undifferentiated cells in the LGE (Anderson et al. 1997a). This establishes *Dlx1/2* as critical genes in the GABAergic phenotype of cortical interneurons, which make up approximately 80% of cortical interneurons, or 20% of all cortical neurons, and are required to maintain a balanced network of synaptic conduction (Wonders and Anderson, 2006).

1.11 DLX transcription factors' gene targets

DLX proteins are able to recognize and bind to the consensus sequence (A/C/G)TAATT(G/A)(C/G). In addition to binding to their recognition sites, they can also form heterodimers with other homeobox transcription factors such as MSH-like proteins, called MSX in vertebrates, to perform their regulatory functions. The MSX/DLX heterodimers were previously shown to participate in the patterning of the anterior neural plate, as found in a study using *Xenopus* as a model (Feledy et al. 1999; Zhang et al. 1997). In *Xenopus*, *Dlx3* gene expression is high in the extreme anterior neural plate while *Msx1* gene expression is silent, and this is due to a gradient of *Dlx*-activating BMP signalling,

therefore providing a gradient of both homeobox genes' products. Consequently, by forming a gradient heterodimeric complexes with each other, activation of genes by DLX is not possible, as well as the gene repressing activity of MSX. Other studies looking at DLX transcription factors' function identified DLX5 as a transcriptional regulator of the *osteocalcin* (*OC*) gene, again implicating the *Dlx* genes in bone development (Ryoo et al. 1997). DLX binding sites have also been found in promoter areas of genes such as *neuropilin-2* (*nrp2*). Neuropilins such as the ones encoded by *nrp2* are transmembrane receptors that help control the repulsive action of class 3 semaphorin axon guiding signalling molecules, which inhibits neuronal migration. When the DLX1 and DLX2 proteins are bound to this promoter, there is repression of *nrp2*, making the DLX1/2 proteins direct repressors of *nrp2* and causing a non-tangential type of migration, which sees interneurons cross through the striatum, where they do not normally venture (Le et al. 2007; Tamagnone and Comoglio, 2000). Therefore, neuropilins have a central role in the tangential migration of interneurons from the ventral telencephalon to the more dorsal neocortex.

DLX proteins are also able to regulate the transcription of fellow *Dlx* gene family members. An example is the binding of DLX1 and/or DLX2 to the intergenic enhancers found between *Dlx5* and *Dlx6*. This is supported by the loss of the majority of *Dlx5/6* gene expression, especially in the SVZ, of the *Dlx1/2* double knockout mice (Anderson et al. 1997b; Zerucha et al. 2000b; Zhou et al. 2004). This makes the DLX proteins critical for their own expression and function in the differentiation and migration of GABAergic interneurons in the developing forebrain (Anderson et al. 1997a).

1.12 Known regulators of *Dlx* gene expression

To understand how the DLX transcription factors contribute to the complex genomic instructions of development, their transcriptional and post-translational regulators must be identified. So far, a limited number of proteins that control *Dlx* gene expression have been identified. These include the bone morphogenic proteins (BMP), as BMP-2 has been shown to transactivate *Dlx* genes such as *Dlx3* and *Dlx5*; and studies suggest that these interactions are important in the creation of the spatiotemporal gene expression pattern required for early embryogenesis (Feledy et al. 1999; Lee et al. 2003; Luo et al. 2001; Park and Morasso, 2002; Thomas et al. 2000).

Other known regulators of *Dlx* gene expression are the fibroblast growth factor-8 (FGF8) protein, shown to regulate the expression of *Dlx2* in the first branchial arch (Thomas et al. 2000); the pituitary homeobox factors (PITX) such as PITX2 are expressed in the same pattern as *Dlx2* during early development and are able to regulate the expression of *Dlx2* (Cox et al. 2002). The mammalian achate schute homolog 1 (MASH1) has also been shown bind to an enhancer of *Dlx1/2* and regulate its expression (Poitras et al. 2007). There has been recent evidence for a long polyadenylated non-coding RNA called *Evf2*, which is transcribed from the *Dlx5/6* intergenic area, and is able to regulate *Dlx5/6* expression by forming a complex with DLX2 on one of the intergenic enhancers and thereby activating *Dlx5/6* gene expression (Bond et al. 2009; Feng et al. 2006).

1.13 *cis*-regulatory elements that control *Dlx* gene expression

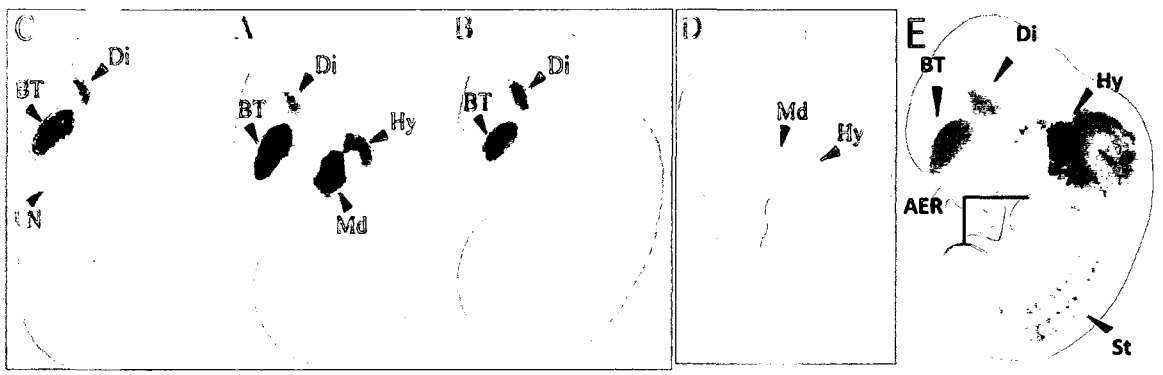
As mentioned before, the *Dlx* bigene clusters have similar gene expression profiles, which are allegedly due to shared regulatory elements, which would then target the gene expression to similar cells at similar times during development. To investigate this, research

has focused on conserved sequences between each gene bicluster, in the intergenic and surrounding areas of the *Dlx* genes. Conserved sequences were isolated and tested for regulatory activity, and at least six such elements have been identified.

The *Dlx* regulatory elements found thus far are the I12a, I12b, I56i, I56ii, the Upstream Regulatory Element 2 (URE2), and URE1 *cis*-regulatory elements (Ghanem et al. 2003; Ghanem et al. 2007; Zerucha et al. 2000b). These conserved elements display enhancer properties in studies where they targeted the transcription of a reporter gene in a spatiotemporal specific manner resembling various *Dlx* gene expression patterns (see Figure 1.3). The URE2 enhancer element is located approximately 12kb upstream of the mouse *Dlx1* and has been hypothesized to drive the transcription of the *Dlx1* and/or *Dlx2* gene(s), based on *lacZ* reporter gene and Cre-recombinase transgenic mice (Ghanem et al. 2007; Potter et al. 2009). URE1 targets expression of the *lacZ* reporter gene in mouse tissues other than the brain, such as the retina (Ghanem et al. 2007). Enhancers I12a and I12b are located in the *Dlx1* and *Dlx2* intergenic region, and are hypothesized to regulate both genes, also based on similar studies (Ghanem et al. 2003; Ghanem et al. 2007). Finally, I56i and I56ii are the conserved sequences found to be enhancers of the *Dlx5/6* cluster (Ghanem et al. 2003; Zerucha et al. 2000b).

The non-coding regulatory elements that regulate the *Dlx* genes may have a critical role in the development of the forebrain, and are therefore not expected to undergo rapid evolution (Drake et al. 2005). This is supported by a high degree of conservation among *Dlx* non-coding regulatory elements in mammals and teleost fish, which shared a common ancestor approximately 450 million years ago (see sequence alignments in

Figure 1.3: *lacZ* reporter mice showing *Dlx cis*-regulatory elements' activity in whole embryos at E11.5. Each intergenic cis-regulatory element is shown to target *lacZ* expression in the basal (ventral) telencephalon (BT) and diencephalon (Di) except for I12a. The I56i and I12a elements target reporter gene expression to the hyoid (2nd branchial arch, Hy) and the mandibular (1st branchial arch, Md) arches. URE2 targets reporter gene expression to the hyoid but not the mandibular arch, as well as the somites (St) and apical ectodermal ridge (AER). (A-D are adapted from Ghanem et al. 2003, E was taken by L. Poitras and edited by M. Lavoie).



I56i I56ii I12b I12a URE2

(Ghanem et al. 2003). It has been suggested that highly conserved non-coding elements frequently play an essential role in gene regulation during development (Woolfe et al. 2005).

1.14 Identification of the URE2 enhancer in mouse and zebrafish

The non-coding URE2 sequence upstream of the *Dlx1* gene was conserved in five species: the human, mouse, zebrafish and two pufferfish genomes, making it a good candidate for a *cis*-regulatory sequence (see supplemental data in (Ghanem et al. 2007). To test the function of the URE2 sequence, it was inserted in the p1230 vector, which contains a β-globin minimal promoter and the *lacZ* gene with a polyadenylation site. The multiple cloning site (MCS) for p1230 is that of the pBluescript SK, and the β-globin-*lacZ* cassette was inserted into the BamHI site of the MCS. To minimize a position effect due to the construct's random insertion into the mouse genome, multiple transgenic lines are generated when using this approach (Wilson et al., 1990).

In this experiment, the mouse URE2 sequence targeted *lacZ* gene expression in a manner similar to that of *Dlx1* and/or *Dlx2* expression patterns in mice (Ghanem et al. 2007). The creation of a transgenic Cre-recombinase mouse confirmed this, where the Cre gene, when under the control of URE2, was also able to very closely reproduce the expression pattern of *Dlx1* and *Dlx2* (Potter et al. 2009). Together, these data suggest that URE2 does indeed have a *cis*-regulatory role as an enhancer element which targets *Dlx1/2* genes, since it is able to spatiotemporally target two other genes to patterns very close to the *Dlx1/2* spatiotemporal expression pattern in mice.

1.15 Identification of trans-activating factors binding to the URE2 enhancer

As discussed before, TF binding sites are approximately 6-8bp, making a 1300pb non-coding sequences such as URE2 a candidate for many TF sites. To realize the full impact of URE2 on *Dlx1/2* gene expression, it is crucial to understand how URE2 functions, and a good place to start is to identify TF's that bind within its sequence and participate in *Dlx1/2* or other genes' transcription during embryonic development.

To test for transcription factors binding within the URE2 element, a DNAseI footprint was done previously in our laboratory (Poitras et al., unpublished observations). This type of experiment uses a radiolabeled DNA sequence digested with DNaseI to reveal a ladder-type pattern when put on a polyacrylamide gel and exposed on radiofilm. When the sequence is pre-incubated with proteins such as those from brain tissue nuclear extracts, the DNaseI enzyme used to cut the DNA probe is not be able to cut the sequence bound by protein(s) and this will result in a gap in the radioactive ladder-like pattern. This previous experiment revealed 13 areas where protein(s) from E11.5 wild type CD-1 background mice brain nuclear extracts were bound to the mouse URE2 sequence. One of the areas of the mouse URE2 sequence which was protected when fragmented by DNAseI was the segment from base pairs 699-719. The MatInspector bioinformatics program identified some potential transcription factor binding sites within this protected fragment and this is the basis for this thesis.

1.16 URE2 mutants have a shift in URE2 activity in the developing forebrain

To further understand the role of URE2, an attempt was made to identify potential TF binding sites. One area of URE2, corresponding to bp 699-719 (see above), identified by the

DNAseI footprinting analysis, was mutated and used to create transgenic mice in order to observe the mutated sequence's effect (see Figure 1.4). The mutated sequence was placed in the same *lacZ* reporter vector as described in section 1.14. This construct was named mURE2mut10-*lacZ*. A first line of transgenic mice containing mURE2mut10-lacZ existed at the onset of my thesis research. This nomenclature was established because mURE2 stands for the mouse sequence of URE2; mut10 is the name of the mutation given to this sequence since it originated from the footprint area 10 in the DNAseI footprint; and *lacZ* is the reporter gene that sits downstream of URE2.

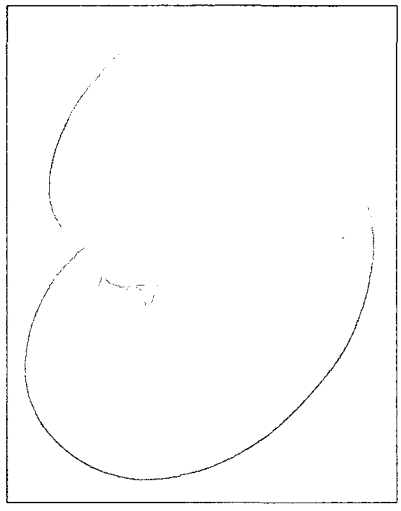
The mURE2mut10-*lacZ* mice demonstrated a *lacZ* reporter gene expression pattern different than that of the non-mutated mURE2-*lacZ* reporter gene expression pattern in mice. The mutant mice have a total of 6 base pairs mutated in the entire 1300bp enhancer, and display an overall dorsal shift in URE2 activity, seen by blue *lacZ* positive cells, in the telencephalon at E11.5. This suggested that perhaps a protein or complex of proteins with repressor activity were binding to this conserved site within URE2 in the brain at E11.5, which is the developmental tissue and stage tested in the DNAseI footprint. A bioinformatical analysis ensued in an attempt to identify the protein(s) using this sequence as a binding site.

As reviewed by Ooi and Wood (2007), REST is able to recruit chromatin remodelling factors such as its Co-Repressor (Co-R) at the C-terminal of the zinc finger protein, and mSin3 at its N-terminal. More elements join the complex upon the binding of REST to Co-R and mSin3. There is also research showing that a specific isoform of the

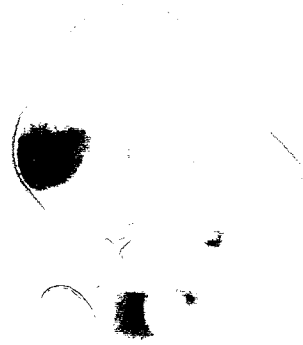
Figure 1.4 Effect of the FP10 mutation on URE2 enhancer activity. The mutant transgenic mice exhibit abnormal expression patterns of the *lacZ* reporter gene when compared to the non-mutated mURE2-*lacZ* activity seen in left-most panel. In right-most panel there is a less-restricted expression pattern in the forebrain resulting from the mutation of the conserved site. This site is suspected of being a REST repressor protein binding site called a Neural Restrictive Silencing Element (NRSE), or a Neural Factor-1 binding site (NFI). The middle panel, which is the transgenic mouse carrying a mutation in a suspected DLX binding site within URE2, displays a widespread loss of *lacZ* reporter gene expression. (L. Poitras, unpublished observations)

mURE2-*lacZ*

mURE2mut2b-*lacZ*
(DLX binding site?)



mURE2mut10-*lacZ*
(REST/NFI binding site?)



1.17 Initial identification of five candidate proteins for binding on base pairs 699-719 of URE2

1.17.1 Prime candidate: REST/NRSF

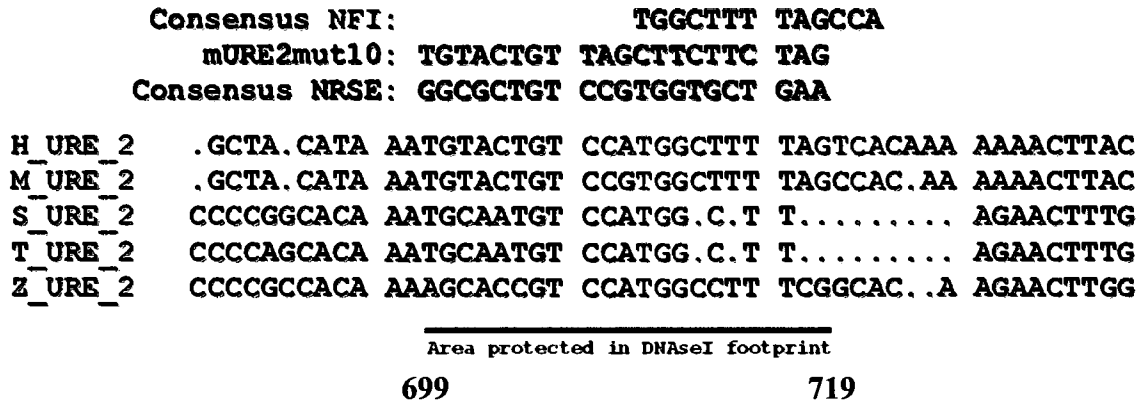
The MatInspector bioinformatics software used to identify possible transcription factors binding to the sequence bound by protein(s) within the URE2 enhancer originally identified the repressor protein REST as one candidate protein for this binding site. REST has been the focus of many reviews (Coulson 2005; Jones and Meech, 1999; Lunyak and Rosenfeld, 2005; Ooi and Wood, 2007), which identify it as a zinc finger protein that binds to Neural Restrictive Silencing Elements (NRSE), or Repressor Element 1 (RE-1) sites, these being widespread in the mammalian genome. The protein was discovered in 1995 as a non-neural repressor of a neural sodium channel gene (Mandel et al. 1992; Mandel et al. 1995), and functional analysis has provided evidence for REST's major role in the regulation of neurogenesis, among other roles (Ballas and Mandel, 2005; Kim et al. 2006; Kim et al. 2008; Li et al. 2008; Majumder 2006; Neumann et al. 2004; Reddy et al. 2009). The *rest* gene is ubiquitously expressed, and mutant mice depleted of REST were described as appearing normal until embryonic day 9.5 (E9.5), after which a rapid onset of apoptosis led to embryonic death by E11.5 (Jones and Meech, 1999). In the review by Jones and Meech (1999), the authors emphasize that REST represses neural gene expression in the brain, but does not constitute a general repressor within the brain, and is called upon in non-neural cells for neural gene silencing through a more permanent repression mechanism than the transient repression which occurs in the brain during neurogenesis. REST protein, REST4, is able to activate transcription in genes and compete for NRSE elements with full length REST

proteins, thereby silencing the silencing activity of the full length REST (Shimojo et al. 1999; Tabuchi et al. 2002). Thus, REST may act as a repressor or an activator, depending on the genomic context.

1.17.2 Revised list of candidates adds 4 more proteins from the Nuclear Factor I (NFI) family of transcription factors

After one year and a half of this project, the MatInspector software database was revised to include new information and new candidates were found in the 699-719 URE2 sequence that was protected by DNAseI footprinting (see Figure 1.5). These new candidates include the Nuclear Factor-I (NFI) family of genes which consists of four members in vertebrates. NFI genes are expressed in many tissues during development (Gronostajski 2000; Mason et al. 2009). The four genes are called Nuclear Factor-I-A (NFIA); Nuclear Factor-IB (NFIB); Nuclear FactorI-C (NFIC) and Nuclear Factor-IX (NFIX). The NFI binding sites are identical to the sequence found within the mouse URE2 with the exception of the last three base pairs, which were not protected in the DNAseI footprinting assay. These proteins can bind to NFI binding sites and form hetero- and homodimers with themselves and/or other members of the NFI family of proteins. They can also form complexes with other transcription factor such as the SP1 protein. These complexes have been shown to be involved in gene activation and/or repression. Their transactivation/repression domains are located in their C-termini and their N-termini are mostly used to bind to DNA and form hetero- or homodimers, though they have been reportedly used in transactivation of a specific promoter (Mason et al. 2009).

Figure 1.5: Schematic of a conserved sequence found in the mouse URE2 element. This sequence, which was found to be bound by protein in DNAseI footprinting analysis, resembles a Neural Restrictive Silencing Element (NRSE), which is a consensus binding site for the REST protein, and a half Nuclear Factor-1 binding site. The sequence shown above the bottom line was the area protected by DNAseI in a previous assay, and is 95% conserved between human (H) and mouse (M), where 20/21bp are the same. Above the sequence are the consensus NRSE site (GGC GCT GTC CGT GGT GCT GAA), the "mutant 10" version of the sequence (TGT ACT GT T AGC TTC TTC TAG), as well as the consensus NFI binding site (TGG CNN NNN GCC A). Residues in blue are areas critical for binding of the REST protein, and the nucleotides represented in green are the two areas critical for NFI binding. Z= *Danio rerio* (zebrafish), T= *Takifugu rubripes*, S= *Spheroides nephelus*.


Consensus NFI: TGGCTTT TAGCCA
mURE2mut10: TGTACTGT TAGCTTCTTC TAG
Consensus NRSE: GGCGCTGT CCGTGGTGCT GAA
H_URE_2 .GCTA.CATA AATGTACTGT CCATGGCTTT TAGTCACAAA AAAACTTAC
M_URE_2 .GCTA.CATA AATGTACTGT CCGTGGCTTT TAGCCAC.AA AAAACTTAC
S_URE_2 CCCCGGCACA AATGCAATGT CCATGG.C.T T..... AGAACTTTG
T_URE_2 CCCCAGCACA AATGCAATGT CCATGG.C.T T..... AGAACTTTG
Z_URE_2 CCCCGCCACA AAAGCACCGT CCATGGCCTT TCGGCAC..A AGAACTTGG
Area protected in DNAseI footprint
699
719

1.17.2.1 Candidate 2: Nuclear Factor I-A

Nfia is expressed in the developing forebrain (Plachez et al. 2008). When one group knocked out the *Nfia* gene in mouse, the result was a combination of postnatal neuroanatomical defects (das Neves et al. 1999; Shu et al. 2003). Such defects included the failure of the corpus callosum to develop, which is the element of the brain that connects the left and right hemispheres of the brain together. Another impact of the loss of *Nfia* was hydrocephaly, where the ventricles of the brain filled with cerebral spinal fluid (CSF). Even though the *Nfia* $^{-/-}$ knockout mice died postnatally, they did not show any morphological defects during development that would render the mutation lethal before P0, the day of birth. Also, only four out of 102 *Nfia*$^{-/-}$ homozygous mice were able to survive after birth, most of them dying after two weeks. NFIA is also required for the inhibition of the methyltransferase DNMT1 binding to the *gfap* promoter required for astrocyte differentiation from neuron precursor cells (Deneen et al. 2006; Miller and Gauthier, 2007).

Nfia gene expression has also been documented in the cortical cells after the formation of the preplate, which is an important cell layer for cortex organization (Xie et al. 2002). *Nfia* was also seen to be expressed in the posterior thalamus, the hypothalamus and the optic chiasm, as well as in the midline of the developing cerebral cortex (Plachez et al. 2008). It has been documented that the consensus recognition sites for NFIA and NFIC are identical and could vary between TGG(C/A)NNNNNGCCAA and TTGG(C/A)NNNNGCCAA. Also, it has been shown that NFIA is able to bind to half sites resembling TTGG(C/A) when there is no other half site nearby (Osada et al. 1996).

1.17.2.2 Candidate 3: Nuclear Factor I-B

The second candidate from this family is the NFIB protein. When the *Nfib* gene is knocked out in mice, pontine nuclei, which are precerebellar neurons, decrease in number (Steele-Perkins et al. 2005). The same mice had the same, but less severe gestational neuroanatomical defects as the *Nfia*$^{-/-}$ mice. Additionally, they died at birth with poorly developed lungs and an abnormally shaped hippocampus. Therefore these data suggest that *Nfia* and *Nfib* genes share a pattern of expression although *Nfib* gene expression is required for lung maturation as well as pontine nuclei and hippocampus development. NFIB has been known to form heterodimers with NFIC, thereby bringing the combined transactivation potential to somewhere in between NFIB's potential and NFIC's potential, which is the higher of the two (Chaudhry et al. 1998).

1.17.2.3 Candidate 4: Nuclear Factor I-C

The next member of the NFI family of genes that might bind within the conserved sequence within the URE2 enhancer is the NFIC transcription factor. The *Nfic* gene is not highly expressed in the developing brain as are the other gene family members, but a *Nfic*$^{-/-}$ knockout mouse reveals its important function in tooth and craniofacial skeleton development, since these mice lack both teeth and surrounding facial bones (Steele-Perkins et al. 2003). NFIC is able to activate gene expression via their promoters, but when combined with equal amounts of NFIX or NFIB, this transactivation is reduced by about half, which is somewhere above the potential for activation of NFIB or NFIX and that of NFIC (Chaudhry et al. 1998).

1.17.2.4 Candidate 5: Nuclear Factor I-X

Last but not least, the fourth member of the NFI multi-gene family is the NFIX protein. This transcription factor has been shown to be crucial for proper forebrain development, where *Nfix*$^{-/-}$ mice exhibited defects that were different than those observed in *Nfia-/-*, *Nfib-/-* and *Nfic-/-* knockout mice. Mostly, there was an increase in brain weight in the NFIX deficient mice, as well as a surplus of abnormally shaped *Pax6*- and *doublecortin*-positive cells in the lateral ventricles (Campbell et al. 2008; Driller et al. 2007). It was also observed that the entire brain, starting at the cingulated cortex, was expanded dorso-ventrally, as well as having an odd-shaped hippocampus. These defects are indicative of a crucial role in brain development. NFIX proteins are also transactivators of gene expression and have been shown to have reduced activity when dimerized with the NFIC protein (Chaudhry et al. 1998).

1.18 Objectives of this thesis

The goal of this research is to identify the cause of the abnormal *lacZ* reporter gene profile seen in mURE2mut10-*lacZ* transgenic mice when compared to non-mutated mURE2-*lacZ* profiles observed during embryonic brain development at E11.5. Thus, our aim was to characterize the URE2 activity as seen by *lacZ* gene expression during embryonic development and leading up to birth in forebrains of mURE2mut10-*lacZ* transgenic mice. Patterns of *lacZ* expression will be compared between the mURE2-*lacZ* non-mutated mouse URE2 transgenic mice and mURE2mut10-*lacZ* mice at E11.5, E13.5, E15.5 and P0. This will provide valuable information as to the developmental stages most impacted by the loss of the conserved site at base pairs 699-719 of the mouse URE2 sequence.

Areas of the brain in which *Dlx1/2* genes are expressed will be examined for potential abnormal enhancer activity at the four time points chosen. Since URE2 is known to enhance *Dlx1/2* transcription, any abnormal activity seen in its activity may reflect an impact on *Dlx1/2* gene expression. Therefore, the *lacZ* reporter mice will provide preliminary evidence of the possible impacts on *Dlx1/2* gene expression profiles when mutations are introduced in conserved areas possibly serving as transcription factor binding sites within the mouse URE2 sequence.

This project is also an attempt to identify the protein(s) binding to the base pairs 699-719 of the mouse URE2 sequence. This will be done by using the dual luciferase reporter assay, a system which is able to test sequences for repressor activity using transfections of the cultured cells, and measuring *Firefly* and *Renilla luciferase* light emission. The effects of candidate transcription factors on the mURE2nt699-719 sequence will be determined.

This work will elucidate some of the mechanisms of transcriptional control utilized by the URE2 enhancer element by identifying one or more of the transcription factors binding to its sequence between base pairs 699-719. The identified candidate proteins will be tested for their regulatory activity, and the effects of mutagenesis will be evaluated and analysed in transgenic mice. Only then will a clearer understanding of the genetic pathways which involve the DLX transcription factors be possible. The results obtained from this section of the project will provide information useful to investigators studying *Dlx1/2* transcriptional regulation. Also, this research will provide knowledge of potential and/or confirmed transcription factors involved in the genetic cascades important for normal forebrain development. Since no DLX-REST, DLX-NFI or DLX-REST-NFI complexes

have ever been documented, this research is novel in its nature and will provide new information in the context of *Dlx* gene regulation.

Lastly, this entire project is based on the hypothesis that the abnormal URE2 enhancer activity observed with a *lacZ* reporter gene system in E11.5 forebrains of mURE2mut10*lacZ* transgenic mice, carrying a mutated form of a putative transcription factor binding site, is due to the absence of an element with repressor activity, possibly one of the following five candidate proteins: NRSF/REST; Nuclear Factor-1A; Nuclear-Factor 1-B; Nuclear Factor 1-C and/or Nuclear Factor-1X.

2. MATERIALS AND METHODS

2.1 Production and screening of transgenic mouse tissues

The mouse URE2 sequence was subcloned into the p1229 vector, which contains a β-globin minimal promoter and the *lacZ* reporter gene. The same was done for the mURE2mut10*lacZ* mice. The vectors were sent to the Ottawa Health Research Institute, where they were injected into CD-1 strain pronuclear single-cell embryos using standard procedures. Tissue samples from the end of tail or ear of weaned mice, or extraembryonic tissue of embryos, were digested overnight at 55°C in 350µl of digestion buffer (50mM Tris-HCl pH 8.0, 100mM EDTA, 100mM NaCl and 1%SDS) and 35µl (350ug) of Proteinase K (10mg/ml). Genomic DNA was isolated using a salt and ethanol precipitation protocol, where digested tissue was precipitated with 0.05 volumes of 5M NaCl and then 0.8 volumes of room temperature isopropanol. Maximum speed centrifugation (Microlite, Thermo Electron Corporation) followed for 2 minutes at room temperature, and the resulting pellet was washed in cold 70% ethanol. After maximum centrifugation for 2 minutes at room temperature, the pellet was resuspended in room temperature TE buffer (100mM Tris-HCl pH 8.0 and 10mM EDTA).

Screening for mice carrying the transgene was done using PCR with primers used to amplify a fragment of the Fetal Hemoglobin (HbF) gene as a positive control (FH1: GAT CAT GAC CGC CGT AGG; FH2: CAT GAA CTT GTC CCA GGC TT) and also a fragment between the β -globin promoter and the *lacZ* gene (BgF: AGG GCA GAG CCA TCT ATT GC; lzR: CGC TCA TCC GCC ACA TAT CC). For the *lacZ* screening PCR, the following conditions were met: 1XPCR buffer (100mM TrisHCl pH8.3, 500mM KCl, 12mM $MgCl_2$), 0.2mM dNTPs (supplied by Invitrogen and diluted to stock solution of 2mM), 10-

100ng of template DNA, 0.2pM of each primer (Bg, Lz, FH1 and FH2), and 1 unit of Taq Polymerase. The cycle conditions were for one cycle at 95° for 5 minutes, 30 cycles of: (94°Cx30s, 55°Cx30s, 72°Cx30s), an extension cycle of 72°Cx5-8 minutes, and cooling at 4°C until the products were removed from the thermal cycler. The products were then run on a 1% agarose gel and visualized under U.V. light for screening of samples having two amplified products in one lane, resulting in a sample positive for the transgenic gene insertion of the *lacZ* gene.

2.2 X-gal staining of *lacZ* transgenic whole embryos and forebrain tissues

E11.5-12.5 embryos were washed in 1XPBS pH7.4 for 30 minutes, and fixed in D-FIX (0.2% gluteraldehyde, 1%formaldehyde, 0.02% Nonidet P-40 (NP-40) and 1XPBS pH7.4) for 30 minutes at 4°C. Whole embryos were subsequently washed 3X in 1XPBS pH7.4 with light agitation and stored in 1XPBS pH7.4 on agitator until overnight X-gal staining. For β-galactosidase activity analysis, X-gal staining was done overnight and in the dark at 32°C (1mg/mL X-gal, 5mM $K_3Fe(CN)_6$, 5mM $K_4Fe(CN)_6$, 2mM $MgCl_2$, 0.02% NP-40 and 1XPBS pH7.4). In the morning, embryos were retrieved from the culture dish, placed in a new dish, and washed 3X in PBS pH7.4 with light agitation. Embryos were finally stored in 1XPBS-10mMEDTA at 4°C until further use.

For E13.5-P0 embryos, embryos were washed in 1XPBS pH7.4 prior to fixing in 4% paraformaldehyde (PFA) for 4 hours at 4°C. Embryos were then washed three times in 1XPBS pH7.4 with light agitation. Following washes, the forebrains were dissected out and when required, they were re-fixed in 4% PFA for an hour. Forebrains were then washed in 1XPBS pH7.4 and stored in 1XPBS-10mMEDTA for storage at 4°C until further use.

2.3 Cryostat sectioning of mouse forebrain tissues

For E11.5 embryos, forebrains stored in 1XPBS-10mM EDTA were placed in warm 1.5% agarose/5% sucrose for positional fixing of the sample. Once solidified, cubes of agarose/sucrose containing the forebrain were cut out of the solidified agarose/sucrose media and placed in a 30% sucrose solution overnight for freezing protection and cryostat preparation of the samples. The following day the forebrain samples were mounted in Shandon Cryomatrix (Thermo Scientific) and placed in the Leica CM1850 chamber for at least 30 minutes to supercool the sample for sectioning. Sections of 25μm were taken and placed on Superfrost/Plus microscope slides (Fisher) and preserved with Aquatex (EMD Chemicals).

For embryos E12.5 and older, the samples stored in 1XPBS-10mM EDTA were placed in a 30% sucrose solution overnight for freezing protection and cryostat preparation of the samples. The following day they were placed directly on the mounting stage of the Leica CM1850 with Shandon Cryomatrix mounting media and were supercooled in the chamber for at least 30 minutes. Sections of 25μm were taken and placed on Superfrost/Plus microscope slides and preserved with Aquatex.

2.4 Preparation of double stranded oligonucleotides

Oligonucleotides (Invitrogen, see section 2.12 for sequences) were designed to contain overhangs that would fit in existing pGL3 *Asp*718 and *Bgl*II restriction sites. Upon receipt they were resuspended in sterile distilled water to achieve a final concentration of 50ng/μl. They were then annealed overnight by placing 100ng of each oligo (2μl) in sterile distilled water for a final concentration of 2.5ng/μl for each oligo. The oligos were then

placed at 80°C for 10 minutes, at which point the heat source was shut off, letting the oligos incubate overnight in a contained environment, and assuring a gradual heat loss and annealing process.

2.5 Preparation of (mURE2699-719, mURE2699-719mut10 and consensus NRSE)-pGL3-Control reporter constructs

For testing of sequences in dual luciferase reporter system, the following fragments were inserted in pGL3-Control using *Asp*718 (G˅GTACC) and *Bgl*II (A˅GATCT) restriction sites: mURE2699-719; its mutant form mURE2mut10 and the consensus NRSE fragment.

To start the subcloning process, 10µg of the pGL3-Control vector was opened, 5' overhang ends were generated by cutting with 10 units (1µl) of *Asp*718 (Roche Applied Science) and 10 units (1µl) of *Bgl*II (New England BioLabs) in 20µl of 10X Roche B Buffer (Roche Applied Science) (100mM NaCl, 10mM Tris-HCl, 5mM $MgCl_2$, 1mM B-mercaptoethanol, pH 8.0 at 37°C). Enzymatic reaction in a final volume of 200µl and at 37°C for an overnight period to ensure no partial digestion occurred. Calf Intestinal Alkaline Phosphatase (CiAP) was used for a 30 minute dephosphorylation process at 37°C, followed by a 10 minute heat inactivation of the CiAP at 65°C.

Gel purification of *Asp*718-pGL3-Control-*Bgl*II was then done using a QIAquick gel extraction kit (Qiagen) following manufacturer's protocols. Twenty ng of eluted *Asp*718-pGL3-Control-*Bgl*II was then ligated with each of the double stranded oligos previously described. Ligation mastermix consisted of 4 units of T4 Ligase (Invitrogen), 16µl of 5XLigase Buffer, 30.44µl of sterile and distilled water, and 4µl (20ng/µl) of the *Asp*718-pGL3-Control-*Bgl*II plasmid. The mastermix was divided into four equal parts and 20ng of

each double stranded oligo was introduced (mURE2699-719, mURE2699-719mut10 and consensus NRSE) for a final ligation reaction volume of 20µl. Ligation was conducted in a thermal cycler at 14°C for an overnight period.

Ligation products were transformed into 100µl of chemically competent XL-10 *Escherichia coli* cells by incubation on ice of both ligation product and cells for 60min, followed by a 42°C heat shock for 90s and 60min incubation at 37°C in prewarmed LB broth. After the final 37°C incubation, cells were centrifuged in the (Microlite, Thermo Electron Corporation) at 8000rpm for 1min and the supernatant was discarded except for ~50µl. The pellet was resuspended in this volume and plated on prewarmed LB-Ampicillin+ plates for growth of colonies.

The Ampicillin-resistant candidate clones were then picked from the plate and screened by colony PCR. To obtain a plasmid DNA sample, colonies were selected and picked from the main plate, streaked on a new plate identifying all selected clones, and the remaining sample was placed in a microfuge tube containing 30µl TE (100mM Tris-HCl pH 8.0 and 10mM EDTA). Samples were boiled for 5min and two PCR experiments were done: 1) was done on 3µl of cooled sample using a primer against the pGL3-Control backbone and one against the insert, and 2) was done on 3µl of cooled sample using two primers against the backbone (primers are listed in 2.12). The PCR reaction conditions were as follows for a 50µl reaction: 1XPCR buffer (100mM TrisHCl pH8.3, 500mM KCl, 12mM $MgCl_2$), 0.2mM dNTPs (Invitrogen), 10-100ng of template DNA, 0.2pM of each primer, and 1 unit of Taq Polymerase. The PCR cycle consisted of one cycle at 95° for 5 minutes, 30 cycles of: (94°Cx30s; annealing conditions varied between: 55°Cx30s for annealing RV3 and GL2 backbone primers, mURE2699-719 and mURE2699-719mut10 insert specific primers

and GL2 backbone primer, and 70°Cx30s for annealing of consensus NRSE insert primer and GL2 backbone primer), and 72°Cx30s), a final extension cycle of 72°Cx5-8 minutes was followed by cooling at 4°C. PCR products were visualized under U.V. light on a 2% agarose gel for the presence of an amplified insert and part of the backbone (330bp) or an amplified product containing no insert (300bp).

Positive clones were selected and grown in 2mL of LB at 37 °C overnight. Plasmid DNa was isolated using Wizard® Plus SV Minipreps DNA Purification System (Promega), and the DNA of two positive clones for all three constructs were sent for DNA sequencing at Ottawa Health Research Institute (OHRI, StemCore Laboratories at The Sprott Center for Stem Cell Research).

Following confirmation of positive clones, one clone per construct was chosen for growing in 50mL of LB at 37°C overnight. Plasmid DNA isolation was conducted using the Qiagen MIDI kit (Qiagen), and were stored at 4°C until required for luciferase reporter assays.

2.6 Preparation of the NFIB-pSPORT6 cDNA construct

To test the various sequences' regulatory activity in the pGL3 vector, cDNA of five candidate binding proteins were required. pcDNA1.1-REST cDNA was a generous gift from Dr. Paul Albert (Ottawa Health Research Institute), while the following were purchased from Open Biosystems: NFIA-pSPORT1 (Catalogue # MMM1013-98478788, clone # 30087137, accession # BC075072), NFIB-pSPORT6 (Catalogue #MMM1013-64884, clone # 4038233, accession # BC014290), NFIC-pSPORT6 (Catalogue # MMM1013-9200864, clone # 4987731, accession # BC055405), and NFIX-pSPORT6 (Catalogue # MMM1013-63211,

clone # 3491917, accession # BC003766). Because NFIA-pSPORT1 was not suitable for expression in cells, NFIA was removed from pSPORT1. A total of 10µg of NFIA-pSPORT1 was cut using *Eco*RI (Invitrogen) and *Not*I (Invitrogen), and the NFIA band was purified using QIAquick gel extraction kit (Qiagen) following the manufacturer's protocols. Following this, 20ng of purified NFIA were ligated as previously described (section 2.5) with the pSPORT6 vector which had been opened at *Eco*RI and *Not*I restriction sites, dephosphorylated (previously described in section 2.5) for 10 minutes, and purified with the QIAquick gel extraction kit as per the manufacturer's protocol. The ligation product was transformed in 100µl of chemically competent XL-10 *E.coli* cells as previously described (section 2.5). Colonies on LB-Amp+ plates were grown in 2mL of LB at 37°C overnight, and DNA was recovered using an alkaline lysis protocol. This protocol consisted of pelleting 1.5mL of bacterial culture at 8000 rpm for 5 minutes in the (Microlite, Thermo Electron Corporation), and discarding the supernatant. A volume of 150µl of Solution I were added (50mM Glucose, 10mM EDTA, 25mM Tris-HCl pH 8.0), as well as 0.05mg of RNAse A. Contents of tubes were then vortexed to resuspend pellet. A total of 300µl of Solution II were added (0.2M NaOH, 1% SDS) and samples were mixed by inversion 6X, followed by 5 min incubation at room temperature. A volume of 150µl of Solution III (3M KHAc pH 4.8) were added; samples were inverted 6X and incubated another 5min at room temperature. Maximum speed centrifugation ensued for 10min using the (Microlite, Thermo Electron Corporation), and the supernatant was transferred to a new tube, in which a total of 900µl of cold 99% ethanol was added. Incubation for 60min at -20°C followed, and then centrifugation at maximum speed for 5 minutes in the centrifuge (Microlite, Thermo Electron Corporation). Supernatant was discarded and the pellet washed with 500µl 70% cold

ethanol. Finally, centrifugation (Microlite, Thermo Electron Corporation) at top speed occurred for 5 minutes, followed by removal of ethanol and re-suspension of the pellet in 30µl sterile distilled water.

DNA preparations were screened by cutting NFIA-pSPORT6 positive clones with *Eco*RI and *Not*I to verify that the two bands representing each the insert and the vector were of the right size. One positive clone was regrown in 50mL LB and its DNA extracted with a Qiagen MIDI preparation kit as per the manufacturer's instructions. A sample was sent to OHRI for sequencing at the StemCore laboratories for confirmation.

2.7 Preparation of empty pSPORT6 vector

Because the cDNAs used in the dual luciferase reporter assays were contained in two different types of expression vectors, empty vectors of pSPORT6 and pcDNA1.1 were required for controls in single transfections when tests containing double and triple transfected cells were done at the same time. It was determined that pcDNA3 was a good substitute for a control since time was a limiting factor for the preparation of the empty pcDNA1.1 vector. To obtain empty pSPORT6, 5µg of the NFIB-pSPORT6 construct was cut with *Eco*RI and *Not*I in a final volume of 50µl. 5' Overhangs were filled using the Klenow large fragment of the DNA I polymerase (Invitrogen) for 30min at 37°C (50µl of template DNA, 10µl REACT2 buffer (Invitrogen), 10µl of 0.5mm dNTP's (Invitrogen) and water to complete to a final volume of 100µl). Following filling of ends, extraction and purification of the blunt *Eco*RI-pSPORT6-*Not*I construct was done using QIAquick gel extraction kit as per the manufacturer's protocol. A total of 20ng of eluted construct were self-ligated in a protocol described earlier in section 2.5. Ligation products were then

transformed in chemically competent XL-10 *E.coli* cells, as described in section 2.5. Colonies on LB-Amp+ plates were isolated by growing in 2mL LB overnight at 37°C, and their DNA extracted using the alkaline lysis protocol previously described in section 2.6. The resulting DNA samples were screened for the presence of empty pSPORT6 vector by cutting with *Sal*I (Invitrogen) and *Xho*I (Invitrogen). Confirmation of positive clone was obtained by sequencing of the DNA sample at OHRI at the StemCore laboratories.

2.8 Cell culture

Neuro-2a cells were a generous gift from Dr. Steffany Bennett (OHRI), and HeLa cells were a kind gift from Dr. Micheline Paulin-Levasseur (uOttawa). Cells were cultured in Dulbecco's Modified Medium (DMEM, Invitrogen) with 10% Fetal Bovine Serum (FBS, Invitrogen) and 1% penicillin and streptomycin (PS) mixture (Invitrogen). Cells were cultured at 37°C in 60mmX100mm tissue culture plates (Corning), and were subcultured in 1:3-1:10 ratios for HeLa or 1:10-1:20 for Neuro-2a, upon reaching approximately 50-70% confluence. The protocol used for passing the cells consisted of removing the growing medium from each 60mmX100mm plate of cultured cells to be subcultured, followed by an incubation of the cells for 5 minutes at 37°C with 1ml of 0.05% trypsin in each plate. A total of 9ml of growing media was then added to the 1ml of trypsin and the cells to obtain a volume of 10ml of trypsinized cells. For 1:10 passages, 1ml of this trypsinized cell mixture was added to 9ml of new growing media in a new tissue culture 60mmx100mm plate.

2.9 Transfections

When cells reached approximately 50-70% confluence, they were subcultured as described in section 2.8 into 24 well plates (500µl of trypsinized cells/well) for transfection 24hrs later. Each test transfection was done in triplicate. Calculations for transfections are as indicated in the Tables 2.1, 2.2 and 2.3. The pGL3-Control and pRenillaLuciferase (pRL) vectors were purchased from Promega and used in all transfections. All transfections were done with the SuperFect (Invitrogen) lipid-mediated reagent. The transfection protocol for each triplicate test sample consisted of incubating 6µl of the SuperFect reagent with 144µl of growing media (DMEM) containing no FBS, for a total reaction volume of 150µl. The incubation was done at room temperature for 5 minutes in a Biological Safety Cabinet (BSC). After the initial incubation, this mixture was then incubated with the test plasmid DNA which was already mixed with FBS-less DMEM in a total volume of 150ul. The total volume for this final 20 minute incubation was 300µl and this was again done at room temperature in the BSC. Following this 20 minute incubation, 90µl of the 300µl incubated SuperFect/DNA/DMEM solution was taken and pipetted into each well containing the cells to be transfected. The 24-well plate containing the cells and the transfection reactions was gently tapped and rotated to ensure even distribution of the transfection reagents, and incubated for 20h at 37°C in a growth chamber containing 5% CO_2.

2.10 Dual luciferase reporter gene assays

All luciferase assays were done using Dual-Luciferase® Reporter Assay System (Promega) as per the manufacturer's protocol. Transfected cells were washed 3X with 1XPBS pH7.4, and incubated at room temperature with a Passive Cell Lysis solution

prepared from the 5X stock solution provided by the manufacturer. This incubation was done at room temperature on a shaker for 30-45 minutes. The cell extracts were collected from each well and pipetted into sterile microfuge 1.5ml tubes. The lysates were then cleared by maximum centrifugation for 30s in the centrifuge (Microlite, Thermo Electron Corporation) and placed at -80°C until further use. *Firefly* and *Renilla luciferase* expression were measured in a luminometer (LMaxII 384, Molecular Devices). All assays were done on 20µl of the prepared cell lysates with 100µl of each Stop 'n Glo (diluted from the 50X stock solution provided with the kit) reagent, and Luciferase Assay Reagent (LAR).

2. 11 Statistical analysis of dual luciferase reporter gene assay results

All transfection experiments were organized by date, and all three readings relating to *Firefly* luciferase expression reported were averaged, as were the *Renilla* luciferase expression readings. The averaged *Firefly* luciferase expression reading was divided with the final *Renilla* luciferase expression reading. This was done for all tests, and all of these numbers were normalized to the empty pGL3 vector's averaged dual luciferase readings. This ensured that each triplicate was a sample (n=1). All tests had an n>2, therefore the mean and standard error were measured using the averaged readings and ratios, and multiple Student t-tests, ANOVA and Tukey's multiple comparisons analyses were performed on the raw data. Statistical analyses were done using R (R Foundation for Statistical Computing, version 2.9.2) for Windows. Significant p values were set at $p \leq 0.1$.

2.12 Primers used in all experiments

Purpose: screening of transgenic animal/tissue for presence of *lacZ* transgene:

Fetal haemoglobin gene specific primers:

FH1: GAT CAT GAC CGC CGT AGG

FH2: CAT GAA CTT GTC CCA GGC TT

β-globin minimal promoter and *lacZ* gene specific primers:

BgF: AGG GCA GAG CCA TCT ATT GC

LzR: CGC TCA TCC GCC ACA TAT CC

Purpose: Annealing of oligonucleotides for the creation of double stranded mURE2699-719, mURE2699-719mut10 and consensusNRSE, all containing extra nucleotides to form cohesive *Asp*718 and *Bgl*II cohesive ends, represented by underlined residues:

*Asp*718-mURE2699-719-*Bgl*IIa:

G TAC CTA AAA GCC ACG GAC AGT ACA

*Asp*718-mURE2699-719-*Bgl*IIb:

GAT TTC GGT GCC TGT CAT GTC TAG

Asp718-mURE2699-719mut10-*BglII*a:

G TAC CTA GAA GAA GCT TAC AGT ACA

Asp718-mURE2699-719mut10-*BglII*b:

GAT CTT CTT CGA ATG TAC ATG TCT AG

Asp718-ConsensusNRSE-*BglII*a:

GT ACC TTC AGC ACC ACG GAC AGC GCC A

Asp718-ConsensusNRSE-*BglII*b:

G AAG TCG TGG TGC CTG TCG CGG TCT AG

Purpose: screening of positive clones for subcloning of mURE2699-719, mURE2699-719mut10 and consensusNRSE in pGL3:

RV3: CTA GCA AAA TAG GCT GTC CC

GL2: CTT TAT GTT TTT GGC GTC TTC CA

mURE2699-719: CTA AAA GCC ACG GAC AGT ACA

mURE2699-719mut10: CTA GAA GAA GCT TAC AGT ACA

ConsensusNRSE: TTC AGC ACC ACG GAC AGC GCC

Figure 2.1 Map of the pGL3 vector used in dual luciferase assays. This vector was used to test the repressor activity of candidate transcription factors. Source: Promega Corporation.

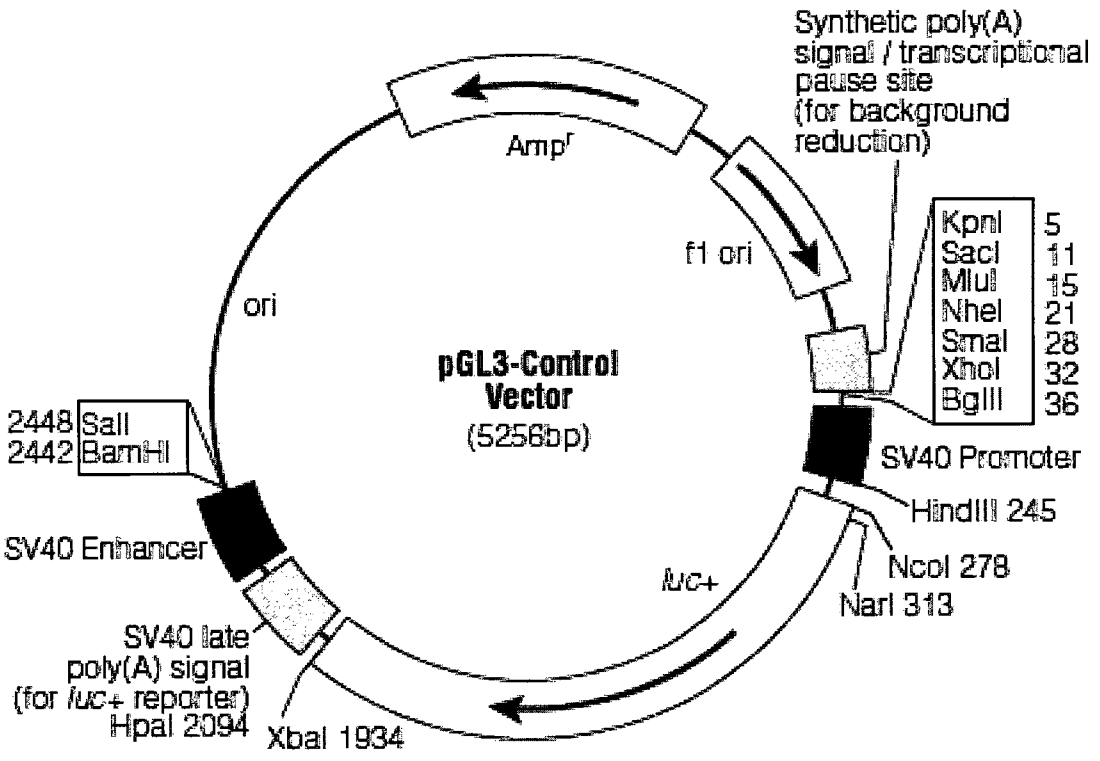

Synthetic poly(A)
signal / transcriptional
pause site
(for background
reduction)
Amp^r
f1 ori
ori
pGL3-Control
Vector
(5256bp)
KpnI 5
SacI 11
MluI 15
NheI 21
SmaI 28
XhoI 32
BglII 36
SV40 Promoter
HindIII 245
NcoI 278
NarI 313
luc+
2448 SalI
2442 BamHI
SV40 Enhancer
SV40 late
poly(A) signal
(for luc+ reporter)
HpaI 2094
XbaI 1934

Figure 2.2. Map of the pGL4 vector. This vector should be used as per future directions of this thesis, to test the activating potential of candidate transcription factors. Source: Promega Corporation.

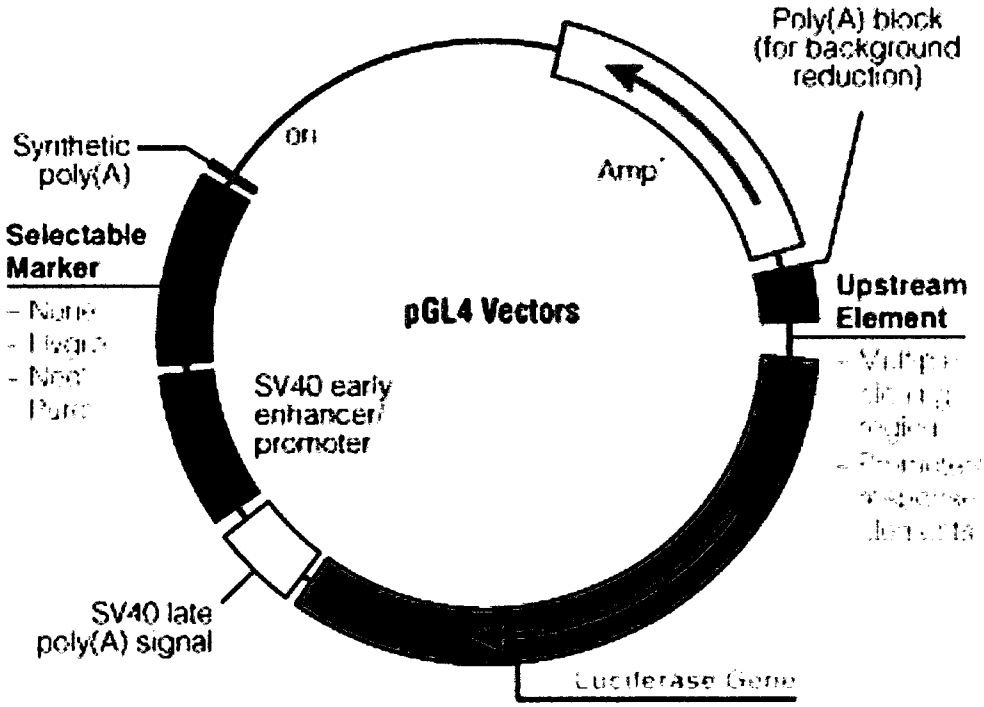

Poly(A) block
(for background
reduction)
Synthetic
poly(A)
ori
Amp
Selectable
Marker
pGL4 Vectors
Upstream
Element
SV40 early
enhancer/
promoter
SV40 late
poly(A) signal
Luciferase Gene

Table 2.1 DNA amounts for transfections done in test 1 (Figure 3.5) in HeLa and Neuro-2a cells. All amounts are in nanograms.

Test 1 (Figure 3.5)	Lane 1	Lane 2	Lane 3	Lane 4	Lane 5
pGL3-Control	400				
pRL	5	5	5	5	5
mURE2699-719-pGL3-Control		400	400		
ConsensusNRSE-pGL3-Control				400	400
pcDNA1.1-REST cDNA			400		400
pcDNA3	800	800	400	800	400

Table 2.2 DNA amounts for transfections done in test 2 (Figure 3.6-3.8) in HeLa and Neuro-2a cells. All amounts are in nanogams.

Test 2 (Figs 3.6-3.8)	1	2	3	4	5	6	7	8	9	10	11	12
pGL3-Control	300											
pRL	5	5	5	5	5	5	5	5	5	5	5	5
mURE2699-719-pGL3-Control		300	300	300	300	300	300	300	300	300	300	300
pSPORT6-NFIA			300				300	300	300			
pSPORT6-NFIB				300			300			300	300	
pSPORT6-NFIC					300			300		300		300
pSPORT6-NFIX						300			300		300	300
pSPORT6	600	600	300	300	300	300						

Table 2.3 DNA amounts for transfections done in test 3 (Figure 3.9-3.11) in HeLa and Neuro-2a cells. All amounts are in nanograms.

Test 3 (Figs 3.9-3.11)	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
pGL3-Control	200											
pRL	5	5	5	5	5	5	5	5	5	5	5	5
mURE2699-719-pGL3-Control		200	200	200	200	200	200	200	200	200	200	200
pSPORT6-NFIA			200				200	200	200			
pSPORT6-NFIB				200			200			200	200	
pSPORT6-NFIC					200			200		200		200
pSPORT6-NFIX						200			200		200	200
pcDNA3	200	200	200	200	200	200	200	200	200	200	200	200
pSPORT6	400	400	200	200	200	200						

2.12 MatInspector bioinformatical analysis.

To identify potential transcription factor binding sites within the mouse URE2 sequence, it was entered in the MatInspector program (Genomatix software), and a readout of potential transcription factor binding sites was produced (for mouse URE2 sequence see Ghanem, et al. 2007).

3. RESULTS

3.1 mURE2*lacZ* and mURE2mut10*lacZ* reporter gene expression pattern results

It is possible to characterize the function of candidate *cis*-regulatory elements by placing them upstream of a minimal promoter and a reporter gene such as *lacZ* in a DNA construct. The entire construct is then injected in fertilized eggs from wild type female mice. This method results in a random insertion of the DNA construct, and renders possible the visualization of the candidate sequence's regulatory activity on the reporter gene in transgenic mouse tissues. The reporter gene expression pattern which results from this approach is usually, at least part of the expression pattern exhibited by the endogenous gene which the candidate *cis*-regulatory element normally targets.

The URE2 element was discovered in an attempt to search for other *cis*-regulatory elements regulating *Dlx* gene expression other than the I12a, I12b, I56i and I56ii elements. The mURE2-*lacZ* transgenic mouse was created when the activity of conserved *cis*-regulatory elements regulating *Dlx* gene expression in vertebrates was investigated (Ghanem 2007). It was determined that the URE2 element was able to target *lacZ* expression in domains of the forebrain which showed a similar pattern to the endogenous *Dlx1/2* expression profile. The URE2 element from the mouse was also able to drive *lacZ* expression in the apical ectodermal ridge (AER) and the hyoid arch of whole embryos, two of the areas where *Dlx* genes are known to be expressed at E11.5 and E12.5. It was therefore determined that URE2 was probably an enhancer of *Dlx1/2*.

To test the function of the small conserved sequence at nucleotides 699-719 within the URE2 element, a mutant construct was created that changed the sequence in a way that largely differs from a putative REST binding site, and named mURE2mut10*lacZ*. This

construct was injected into wild type CD-1 background fertilized eggs, and the resulting *lacZ* expression profile was compared to the expression pattern seen in the mURE2*lacZ* mice. To eliminate any positional effects due to the genomic area in which the transgene was integrated, many injections were done and four independent lines were successfully generated, which were referred to "mut10a-d".

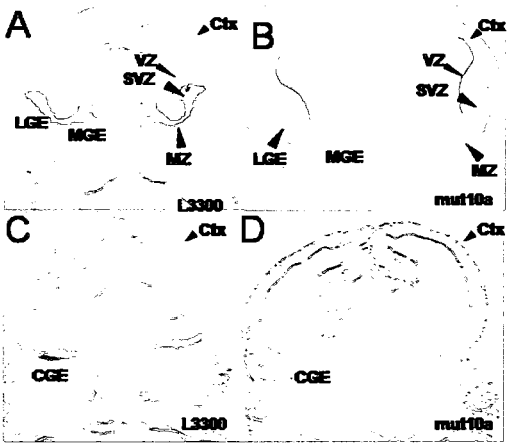
All four lines of transgenic mice were bred and the dams were euthanized during gestation at E11.5, E13.5 and E15.5 for harvesting of the embryos. Pups at postnatal day 0 (P0) were also euthanized and their brains harvested. To study the URE2 activity at these time points, coronal sections of the forebrain of embryos and pups were obtained for both the mURE2*lacZ* and mURE2mut10*lacZ* lines. Sections were stained with an X-gal based reagent for visualization of the *lacZ* activity, and this was done during a16hr incubation on many separate occasions throughout the two years of my project.

The mURE2*lacZ* transgenic mouse forebrain coronal sections gathered for my project revealed *lacZ* positive cells in the SVZ of the LGE, medialMGE (mMGE) and caudal ganglionic eminence (CGE) of the ventral telencephalon at E11.5 (Figure 3.1, A). The *lacZ* signal in the SVZ of the MGE was stronger in the medial sections. There were also some chains of *lacZ* positive cells arranged radially and extending from the SVZ to the VZ of the LGE. This is most likely due to the fact that the ganglionic eminences are cell proliferation zones, and the differentiation and migration taken by these differentiating cells are dependent on *Dlx* gene expression and the enhancer elements used for *Dlx* gene regulation. In the mutant lines, there is a reduced expression in the MGE and a shift in the expression profile which extends past the LGE-neocortex limit and into the neocortex (Figure 3.1, B). When looking at the *lacZ* profile in more caudal sections of the forebrain, there is an overall

Figure 3.1: Comparison of mURE2*lacZ* and mURE2mut10*lacZ* reporter gene expression in the subpallial telencephalon of transgenic mice at E11.5. Coronal sections are shown of the medial and caudal forebrain at E11.5, and show the *lacZ* expression profile under the control of the mouse URE2 enhancer. Sections describe the pattern in medial levels (A,B) and caudal levels (C,D) of mURE2*lacZ* mice (A,C) and the mutant lines mURE2mut10*lacZ* (B,D). Results are shown for the L3300 line of the mURE2*lacZ* mice, and the mURE2mut10a-*lacZ* mutant line.

E11.5

mURE2-lacZ mURE2mut10-lacZ



LGE, lateral ganglionic eminence; MGE, medial ganglionic eminence; CGE, caudal ganglionic eminence; VZ, ventricular zone; SVZ, subventricular zone; MZ, mantle zone; Ctx, cerebral cortex.

reduced expression in the caudal ganglionic eminence (CGE) (Figure 3.1, D) than in the mURE2*lacZ* line (Figure 3.1, C). At E11.5 in the mURE2mut10a-*lacZ* line, the extension of the *lacZ* expression into the cortex could be seen in all lines early on in the project except for the mURE2mut10b-*lacZ* (see Appendix 2). The mURE2mut10c-*lacZ* line appears to have lost this extended expression profile as seen in the pictures in Appendix 3, and this will be discussed in the next chapter.

At E13.5, the URE2 element is mainly active in the SVZ of the LGE, with a weaker activity in the MGE and mantle zone (Figure 3.2). There are also a few cells arranged radially extending from the SVZ to the VZ that are *lacZ* positive in the rostral LGE, suggesting that the cells are also clonally related as seen in the E11.5 sections. The mutant URE2 lines depict a pattern in which there is no such SVZ-restricted pattern, and the *lacZ* expression is strong in the VZ of the LGE. This is true along the entire anterior-posterior axis. The expression pattern in the MGE appears weaker than in the LGE but to the same degree in both mutant and non-mutant lines.

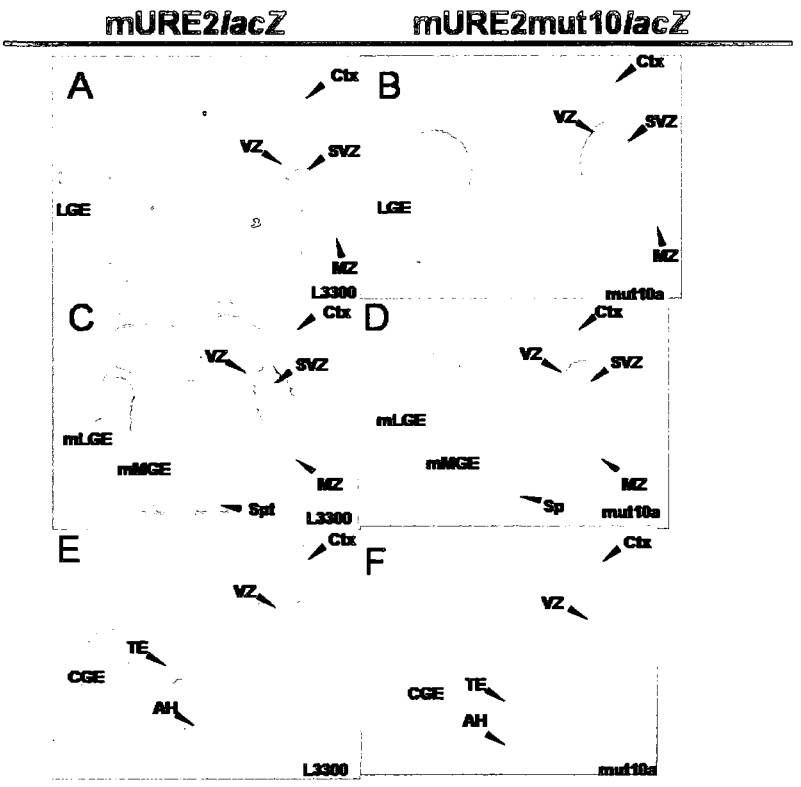
At E13.5, the *lacZ* expression was observed in the VZ of the LGE of the mURE2mut10a-*lacZ* and mut10c-*lacZ* lines. This was not observed in the mURE2*lacZ* line. We have, at this time, obtained E13.5 data for these lines only.

The E15.5 time point exhibited an expression profile similar to that seen in the non-mutated transgenic line. This was seen in both mURE2mut10a-*lacZ* and mURE2mut10c- *lacZ* lines (Figure 3.3), and these two lines were the only ones from which information was available for E15.5.

On postnatal day 0, P0, the mURE2-*lacZ* transgenic mice express *lacZ* in cells restricted to the ventral telencephalon, with a strong signal in the LGE where cell

Figure 3.2: Comparison of mURE2*lacZ* and mURE2mut10*lacZ* reporter gene expression in the subpalial telencephalon of transgenic mice at E13.5. Coronal sections are shown of the medial and caudal forebrain at E13.5, and show the *lacZ* expression profile under the control of the mouse URE2 enhancer. Sections describe the pattern in medial levels (A,B,C,D) and caudal levels (E,F) of mURE2*lacZ* mice (A,C,E) and the mutant lines mURE2mut10*lacZ* (B,D,F). Results are shown for the L3300 line of the mURE2*lacZ* mice, and the mURE2mut10a-*lacZ* mutant lines though similar results were obtained in mURE2mut10c-*lacZ* mice.

E13.5

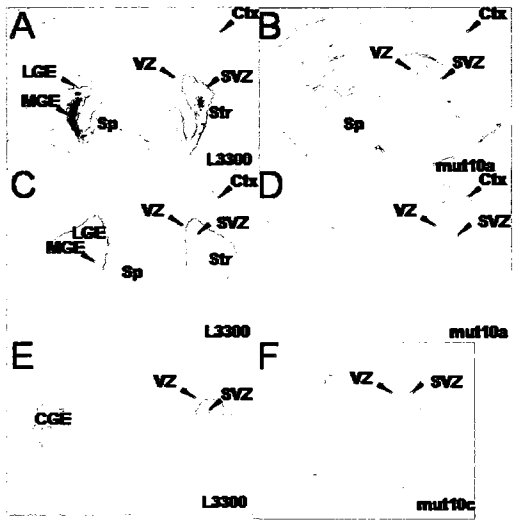


(m)LGE, (medial) lateral ganglionic eminence; (m)MGE, (medial) medial ganglionic eminence;CGE, caudal ganglionic eminence; VZ, ventricular zone; SVZ, subventricular zone; MZ, mantle zone; Ctx, cerebral cortex; Spt, septum; TE, thalamic eminence; AH, anterior hypothalamic area.

Figure 3.3: Comparison of mURE2*lacZ* and mURE2mut10*lacZ* reporter gene expression in the subpalial telencephalon of transgenic mice at E15.5. Coronal sections are shown of the medial and caudal forebrain at E15.5, and show the *lacZ* expression profile under the control of the mouse URE2 enhancer. Sections describe the pattern in medial levels (A,B,C,D) and caudal levels (E,F) of mURE2*lacZ* mice (A,C,E) and the mutant lines mURE2mut10*lacZ* (B,D,F). Results are shown for the L3300 line of the mURE2*lacZ* mice, and the mURE2mut10a-*lacZ* mutant lines though similar results were obtained in mURE2mut10c-*lacZ* mice.

E15.5

mURE2-*lacZ* **mURE2mut10-*lacZ***



LGE, lateral ganglionic eminence; MGE, medial ganglionic eminence; CGE, caudal ganglionic eminence; VZ, ventricular zone; SVZ, subventricular zone; MZ, mantle zone; Sp, septum; Str, developing striatum; Ctx, cerebral cortex.

proliferation is still occurring (Figure 3.4). The mutant lines express the same *lacZ* pattern as the mURE2*lacZ* line, but in an overall decreased manner. This seems to be the only difference between the non-mutated and the mutant URE2 element's activity in the forebrain along the rostral-caudal axis. This was confirmed in the mURE2mut10a-*lacZ* and mURE2mut10c-*lacZ* lines.

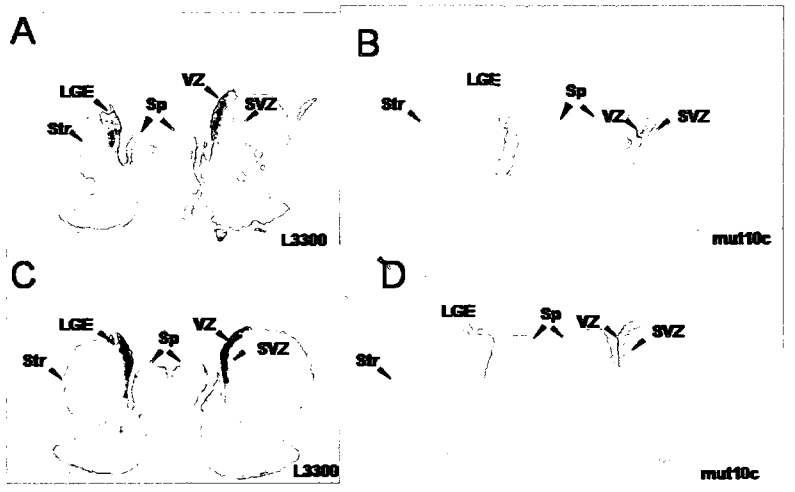
As seen in panels A and B of Figure 3.1, one of the most affected areas in the transgenic mice carrying the mutant URE2 is the entire medial MGE at E11.5, where there is a remarkable decrease in *lacZ* expression when compared to the mURE2*lacZ* transgenic mice. The CGE is also affected at E11.5, having lost almost all of the *lacZ* signal in the cells expressing *lacZ* in the mURE2*lacZ* mouse line. The mutation in URE2 also seemed to have an effect at the E11.5 time point on the SVZ-restricted *lacZ* expression pattern observed in mURE2*lacZ* mice brains, as seen in panels B and D of Figure 3.1, as well as in the same panels of Figure 3.2. The decrease in *lacZ* expression in the CGE and the MGE at this time point suggests that the mutation in the URE2 sequence is able to prevent the URE2 element from targeting gene expression in these areas.

3.2 Dual luciferase reporter assays

The luciferase reporter assay system can report the regulatory activity of DNA sequences placed upstream or downstream of a luciferase gene, which results in an activation, repression or no change in its expression. The luciferase transcripts produced are translated to form a protein capable of emitting light when it is subjected to a chemical reaction in a luminometer. This system utilizes a number of different vectors, which have the luciferase gene in a DNA construct, with and without elements that affect its rate of

Figure 3.4: Comparison of mURE2lacZ and mURE2mut10clacZ reporter gene expression in the telencephalon of transgenic mice at P0. Coronal sections are shown of the medial forebrain on the first day of birth, and show the *lacZ* expression profile under the control of the mouse URE2 enhancer. Sections describe the pattern in medial levels of mURE2*lacZ* mice (A,C) and the mutant line mURE2mut10c-*lacZ* (B,D). Results are shown for L3300 line of the mURE2*lacZ* mice, and the mURE2mut10a-*lacZ* mutant lines though similar results were obtained in mURE2mut10c-*lacZ* mice.

P0



LGE, lateral ganglionic eminence; VZ, ventricular zone; SVZ, subventricular zone; Sp, septum; Str, Striatum.

transcription such as a promoter and an enhancer. For example, the pGL3-Control vector used in experiments reported in this section is usually used in luciferase assays in which sequences with suspected repressor activity are tested. The luciferase gene in this vector is under the control of an SV-40 promoter and an SV-40 enhancer. Therefore, when this vector is transfected into cells, it is able to produce *Firefly luciferase*. A pRL-CMV plasmid is also always co-transfected in these assays to control for possible variations in transfection efficiency. The pRL-CMV plasmid consists of a CMV promoter and a *Renilla luciferase* gene and is used for normalization of experimental *Firefly luciferase* readings. When both pGL3 and pRL-CMV vectors are transfected in mammalian cells, the two luciferase readings are normalized and this raw data is used for statistical analysis and determination of the overall effect on the luciferase gene.

The luciferase activity can be reduced when a reporter plasmid containing a test regulatory sequence upstream of the luciferase gene is bound to proteins with repressor activity. When co-transfected with cDNAs that, once in the cell, are translated into the candidate proteins, it is possible to identify which proteins have regulatory functions associated with the test sequence. To test the regulatory function of the mouse URE2 sequence at nucleotides 699-719, the cDNAs of candidate proteins were transfected to observe which ones had a repressor function. A bioinformatical analysis using MatInspector software revealed that possible transcription factors binding to this site were the REST protein and four possible factors from the NFI family of proteins, more specifically the NFIA, NFIB, NFIC and NFIX factors.

To test the candidate cDNAs, the mURE2nt699-719 sequence was inserted into the pGL3 plasmid. As previously mentioned in Section 1.15 of this thesis, the work done in

transgenic mice and with MatInspector software revealed that one of the candidate proteins binding to the test site was the neural repressor protein REST. This protein is actively transcribed in epithelial cells where the protein is thought to be required to repress neural genes, but is not always required in neural cells where neural genes are expressed (Ooi and Wood, 2007). Consequently, REST protein levels are higher in epithelial cells than in neural cells, and transfection experiments of all candidate cDNAs were conducted in both types of cell lines.

The epithelial cell line chosen for the luciferase assays was HeLa, which was reported as having high levels of REST (Schoenherr and Anderson, 1995). The neural cell line chosen for this study was Neuro-2a, which is reported as having very low and undetectable levels of REST (Palm, K. et al., 1999).

3.2.1 Effect of REST on the transcriptional activity of the mURE2nt699-719 sequence

The mURE2nt699-719 sequence was tested to see if it was indeed a binding site for the REST repressor protein in two cell types where endogenous levels of REST were already known. I also co-transfected the pcDNA1.1-REST expression vector to observe if further repression could be achieved. The results were then compared with those obtained with a consensus REST binding site, called an NRSE. All tests were done by transfecting 400ng of each construct as well as 5ng of pRL-CMV (described in section 3.2), and cells received the same amount of DNA for each transfection (see Table 2.1).

Results in Figure 3.5 suggest that endogenous proteins in HeLa are able to identify the consensus NRSE and repress the expression of the luciferase gene by 67% (pGL3-Control fold + SEM = 1.00; consensusNRSE-pGL3-Control fold + SEM = 0.33). This

repression was due to the binding proteins that are repressors in nature and endogenous to HeLa cells; and are likely the REST proteins. However, in Neuro-2a cells, where there are lower endogenous levels of REST, there was an activation of approximately 68% in luciferase gene expression (fold + SEM = 1.68).

To observe if further repression can be conferred by exogenous REST proteins in HeLa cells, the pcDNA1.1-REST expression vector was co-transfected with the consensusNRSE-pGL3- plasmid. Statistically, there was no further reduction in luciferase expression following this co-transfection in HeLa (fold+ SEM = 0.27). When exogenous REST was introduced in Neuro-2a cells, there was also no statistically significant difference between this result and the results obtained with only the consensusNSRSE-pGL3 plasmid (fold + SEM = 1.25). These results were surprising in the Neuro-2a cells since a consensus NRSE was present in the reporter plasmid and REST cDNA was provided to the cell for the production of the recombinant REST protein. This will be discussed in more detail in the chapter 4.

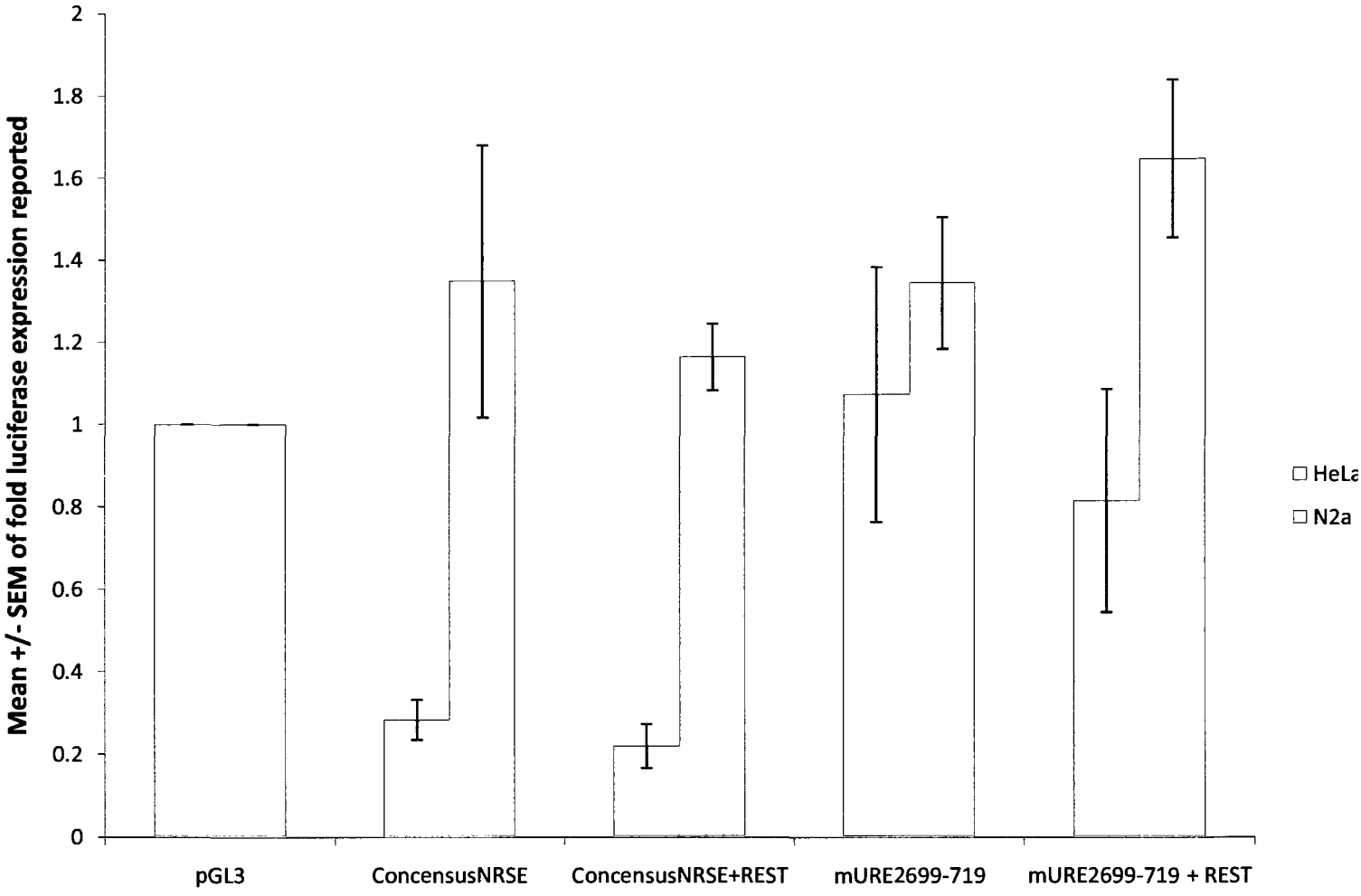
Results show that the mURE2nt699-719 sequence, when transfected without the REST cDNA (Figure 3.5), is not able to change the baseline luciferase levels generated by the empty pGL3 reporter plasmid in HeLa cells (fold + SEM = 1 vs. 1.06). The same result was observed when 400ng of pcDNA1.1- REST cDNA was co-transfected with the mURE2nt699-719 sequence (fold + SEM = 1.05). These results suggest that endogenous REST and/or other proteins with repressor activity in HeLa do not bind to this sequence.

The same figure shows that in Neuro-2a cells, the mURE2nt699-719 construct generated luciferase expression levels approximately 50% higher than levels generated by the control pGL3 reporter plasmid (fold + SEM = 1.5). When Neuro-2a cells were co-

Table 2.1 DNA amounts for transfections done in test 1 (Figure 3.5) in HeLa and Neuro-2a cells. All amounts are in nanograms.

Test 1 (Figure 3.5)	Lane 1	Lane 2	Lane 3	Lane 4	Lane 5
pGL3-Control	400				
pRL	5	5	5	5	5
mURE2699-719-pGL3-Control		400	400		
ConsensusNRSE-pGL3-Control				400	400
pcDNA1.1-REST cDNA			400		400
pcDNA3	800	800	400	800	400

Figure 3.5: Effect of REST on the transcriptional activity of the mURE2nt699-719 sequence in HeLa and Neuro-2a cells. Results are represented as means +/- SEM of averaged triplicates (n=6 in HeLa; n=3 in N-2a) which were normalized to the empty pGL3 vector. A total of 400ng of pGL3-Control, ConsensusNRSE-pGL3-Control or mURE2nt699-719-pGL3-Control constructs were transfected in HeLa and Neuro-2a cells. 400ng of pcDNA1.1-REST was co-transfected with mURE2nt699-719-pGL3-Control as well as with ConsensusNRSE-pGL3-Control.



Constructs tested (400ng) with and without pcDNA1.1-REST cDNA (400ng) in HeLa and Neuro-2a cells

transfected with the pcDNA1.1-REST expression vector, there was a further increase in luciferase activity of 34%, resulting in a total of 84% more luciferase activity compared to the empty pGL3 control (fold + SEM = 1.84). Even though the results show an increase of 34% more luciferase activity following the transfection of REST cDNA, this was not a statistically significant increase compared to the luciferase levels generated with the mURE2nt699-719 sequence and no REST cDNA.

The increase in luciferase activity in Neuro-2a cells transfected with the mURE2nt699-719 construct compared to those transfected with the empty pGL3 plasmid is statistically significant (fold + SEM = 1.68 vs 1).

These results suggest that there are no proteins in HeLa or Neuro-2a cells that are capable of binding to mURE2nt699-719 and repress luciferase expression. However, a 68% increase in luciferase expression was observed when the mURE2nt699-719 construct was transfected in Neuro-2a cells, suggesting that a protein present in these cells was able to recognize this site and activate the luciferase gene in a neural setting.

3.2.2 An updated MatInspector database reveals that NFI proteins are also strong candidates for mURE2nt699-719 regulatory activity.

After the REST protein was identified as a potential candidate for binding to the mURE2nt699-719 sequence, the public database used by the program MatInspector was updated to include another possible transcription factor binding site that overlaps partially with the putative REST binding site. With the updated information, the mURE2nt699-719 sequence was now also a strong candidate for a Nuclear Factor-1 (NFI) binding site. This candidate could be one of the four members of the NFI family of transcription factors. These

factors, NFIA; NFIB; NFIC and NFIX, are proteins that result from four differential splicing events from the same gene. Another indicator that mURE2nt699-719 was a possible NFI binding site was the fact that two NFI binding sites were found in close proximity to each other according to the MatInspector analysis, and these factors are known to form dimers. This led us to test whether or not the NFI family of transcription factors could use the mURE2nt699-719 sequence for regulation of the luciferase gene in the pGL3-Control vector.

3.2.2.1 Effect of NFI proteins on transcriptional activity of mURE2nt699-719 in HeLa and Neuro-2a cells

The next step of this project therefore consisted in testing the four Nuclear Factor-1 candidates to measure their impact on reporter gene expression driven by the mURE2nt699-719 sequence. Each NFI cDNA that was not already in the pSPORT6 expression vector was inserted in it, and 300 ng of each of the four NFI-pSPORT6 cDNA's was co-transfected with the mURE2nt699-719-pGL3 reporter plasmid. Total amounts of transfected DNA were kept constant in all transfections.

In HeLa cells, all NFI proteins were able to cause a significant reduction in luciferase expression from the mURE2nt699-719-pGL3 plasmid except for NFIA (Figure 3.6). The NFI protein with the most potent repressor activity in this particular assay was NFIB, as it was able to repress luciferase gene expression by 54% (fold + SEM = 0.58). The NFIC protein was able to reduce it by 46% (fold + SEM = 0.68), whereas the NFIX factor was able to repress it by 33% (fold + SEM = 0.84).

This set of results also shows that the mURE2nt699-719 sequence was able to behave differently than in the assay where 400ng of it was transfected in HeLa cells (Figure

3.5). In the experiment shown in Fig. 3.6, 300ng of the mURE2nt699-719 -pGL3 plasmid was unable to repress the luciferase gene with endogenous HeLa proteins when compared to readings generated from the empty pGL3-Control vector (fold + SEM =1.25 vs 1). In the previous section (Figure 3.5), 400ng of mURE2nt699-719 in the reporter plasmid was able to produce a repression in luciferase through endogenous proteins in HeLa. This difference in results will be discussed in more detail in Chapter 4.

NFI proteins are not only able to form homodimers, but they can also form heterodimers with other NFI family members. For this reason, all possible combinations of heterodimers were tested for their repressor activity on mURE2nt699-719 in HeLa cells.

Thus, the NFIA/NFIB combination was able to repress the luciferase gene expression by approximately 42% (fold + SEM = 0.73). Similarly, the NFIA/NFIC complex was able to reduce the luciferase gene expression by 32% (fold + SEM = 0.85). Lastly, the strongest repression occurs when the NFIA protein is paired with the NFIX member of the NFI family, as it is able to repress luciferase gene expression by 55% (fold + SEM = 0.56). All of the combinations that included NFIA were able to repress the luciferase gene even though NFIA was not able to do so alone (fold + SEM = 1.11 vs 1.25). These results indicate that NFIA is not able to function as a repressor unless it is co-translated with 300ng of cDNA of another NFI family member to form a repressor complex. The three remaining possible heterodimer complexes include the NFIB/NFIC complex, where a reduction of approximately 38% occurs when they are co-transfected together (fold + SEM = 0.77). The NFIB protein is able to form a potent repressor complex with NFIX, as that combination was able to reduce luciferase gene expression by 61% (fold + SEM = 0.49). Finally, when NFIC and NFIX are co-

transfected, they are able to form a heterodimer which is not able to reduce luciferase gene expression (fold + SEM = 1.21).

When looking at the combinatorial effects reported above, it is important to look at how these complexes may be interacting together. For example the NFIA and NFIX heterodimer results suggest that synergistic activity is occurring between those two NFI factors in HeLa cells, as their combined repression effect exceeds the sum the repression effects obtained with each of the two NFI. This is the only example of synergy observed in HeLa cells.

Overall, my results suggest that the most potent repressors of luciferase in HeLa cells using the mURE2699-719-pGL3-reporter plasmid were the NFIA/NFIX and NFIB/NFIX heterodimers.

In Neuro-2a cells (Figure 3.7), the data show that the mURE2nt699-719 sequence was not able to change the luciferase levels compared to the empty pGL3 plasmid when exposed only to endogenous Neuro-2a proteins (fold + SEM = 1.03). When each of the four NFI cDNA's were transfected, transactivation was observed in all four cases. NFIC had the least activating potential (fold + SEM = 1.11), while it was able to reduce the luciferase activity by 46% in HeLa cells. NFIA was the other NFI factor capable of only slightly elevating luciferase activity (fold + SEM = 1.64) in Neuro-2a cells, while it had a similar effect in HeLa cells (fold + SEM = 1.11). The NFIB and NFIX factors were able to both transactivate the luciferase gene by approximately 3 fold (fold + SEM = 3.34 and 3.6 respectively), whereas they were both able to repress the luciferase gene in HeLa (fold + SEM = 0.58 and 0.85 respectively).

Next, heterodimer combinations were tested for repression or activation potential in Neuro-2a cells transfected with the mURE2nt699-719 sequence in the pGL3 reporter plasmid. The results showed that all possible NFI protein pairs, when transfected together, showed no repressive effects in Neuro-2a cells. However, the combination with the least activation potential was the NFIA/NFIC heterodimer.

When comparing the various single and combinatorial transfections of NFI cDNAs' effects between the two cell types used, there are noted trends but the standard error bars hamper the analysis, rendering the p values not significant. The only tests where there were statistical significances were in the single NFIB and the NFIB/NFIC heterodimer transfections. NFIB is able to increase luciferase expression 3.25 fold in Neuro-2a cells, but repress it by 48% in HeLa cells. The NFIB/NFIC heterodimer was able to transactivate the luciferase gene by 4.5 fold in Neuro-2a cells, but instead produced a 61% reduction in HeLa cells. The rest of the results are statistically non-significant, but show a trend where NFI factors are activators of gene expression using mURE2699-719 as a test sequence in Neuro-2a cells, but are all repressors in a HeLa environment with the exception of NFIA.

3.2.3 Effects of a smaller amount of NFI cDNA on mURE2nt699-719 activity in Neuro-2a and HeLa cells.

Given that the pcDNA1.1-REST cDNA had no significant effect on the regulation of the luciferase gene found in the pGL3-Control vector using the mURE2699-719 as a regulatory test sequence, the possibility that REST and NFI heterodimers formed at the site still existed. Therefore, all REST and NFI combinations were tested for repressor

Table 2.2 DNA amounts for transfections done in test 2 (Figure 3.6-3.8) in HeLa and Neuro-2a cells. All amounts are in nanogams.

Test 2 (Figs 3.6-3.8)	1	2	3	4	5	6	7	8	9	10	11	12
pGL3-Control	300											
pRL	5	5	5	5	5	5	5	5	5	5	5	5
mURE2699-719-pGL3-Control		300	300	300	300	300	300	300	300	300	300	300
pSPORT6-NFIA			300				300	300	300			
pSPORT6-NFIB				300			300			300	300	
pSPORT6-NFIC					300			300		300		300
pSPORT6-NFIX						300			300		300	300
pSPORT6	600	600	300	300	300	300						

Figure 3.6: Effect of NFI factors on the transcriptional activity of the mURE2nt699-719 sequence in HeLa cells. Results are represented by means +/- SEM (n=4) of fold luciferase expression normalized to averaged triplicate readings of luciferase generated from the empty pGL3-Control vector. A total of 300ng of pGL3-Control and/or mURE2699-719-pGL3-Control were tested with and without 300ng of each of the NFI factors, either in single transfection or double transfections. * $p \leq 0.05$ when compared to luciferase generated from empty pGL3 vector.

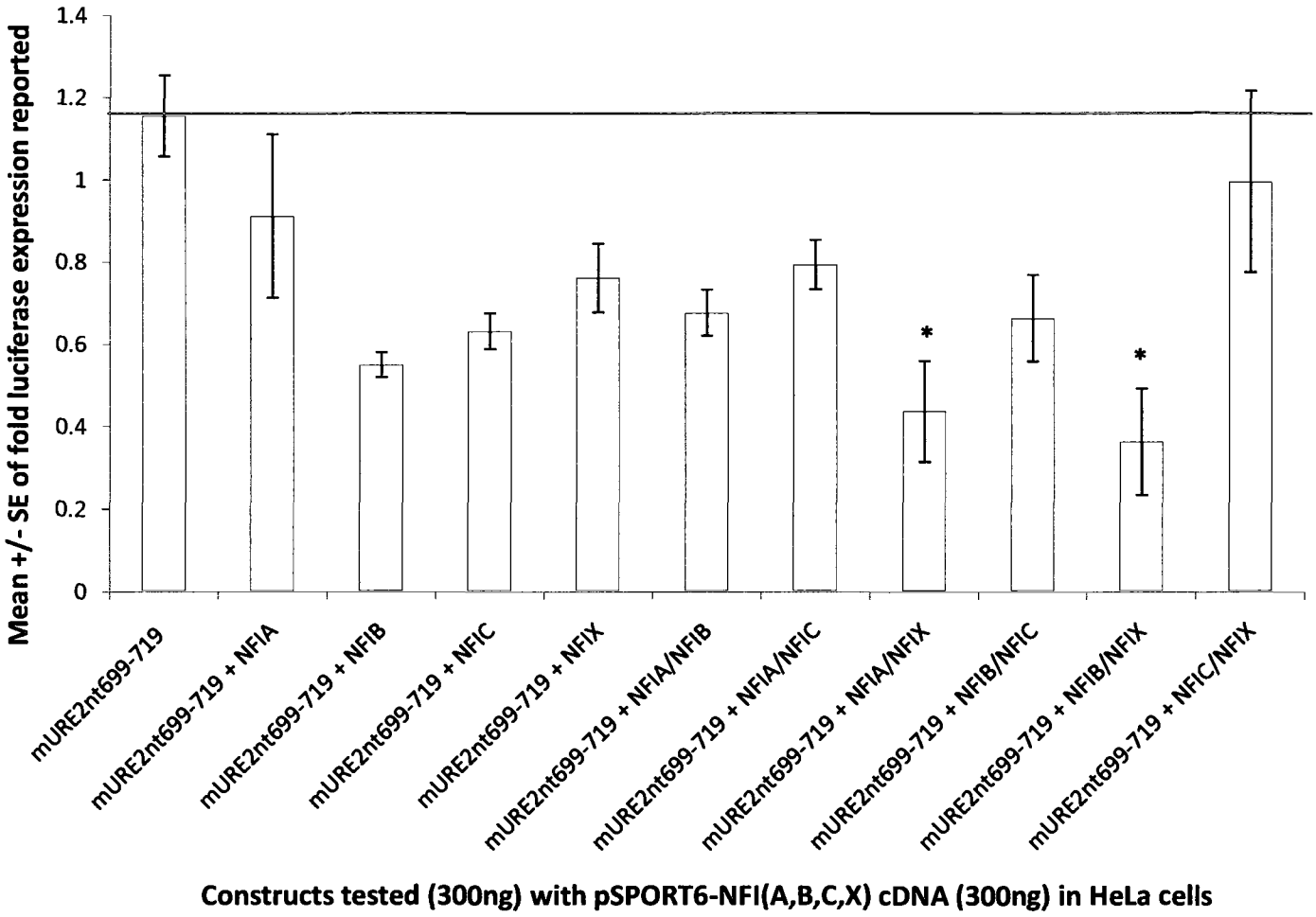

Mean +/- SE of fold luciferase expression reported
1.4
1.2
1
0.8
0.6
0.4
0.2
0
*
*
mURE2nt699-719
mURE2nt699-719 + NFIA
mURE2nt699-719 + NFIB
mURE2nt699-719 + NFIC
mURE2nt699-719 + NFIX
mURE2nt699-719 + NFIA/NFIB
mURE2nt699-719 + NFIA/NFIC
mURE2nt699-719 + NFIA/NFIX
mURE2nt699-719 + NFIB/NFIC
mURE2nt699-719 + NFIB/NFIX
mURE2nt699-719 + NFIC/NFIX
Constructs tested (300ng) with pSPORT6-NFI(A,B,C,X) cDNA (300ng) in HeLa cells

Figure 3.7: Effect of NFI factors on the transcriptional activity of the mURE2nt699-719 sequence in Neuro-2a cells. Results are represented by means +/- SEM (n=4) of fold luciferase expression normalized to averaged triplicate readings of luciferase generated from the empty pGL3-Control vector. A total of 300ng of pGL3-Control and/or mURE2699-719-pGL3-Control were tested with and without 300ng of each of the NFI factors, either in single transfection or double transfections.

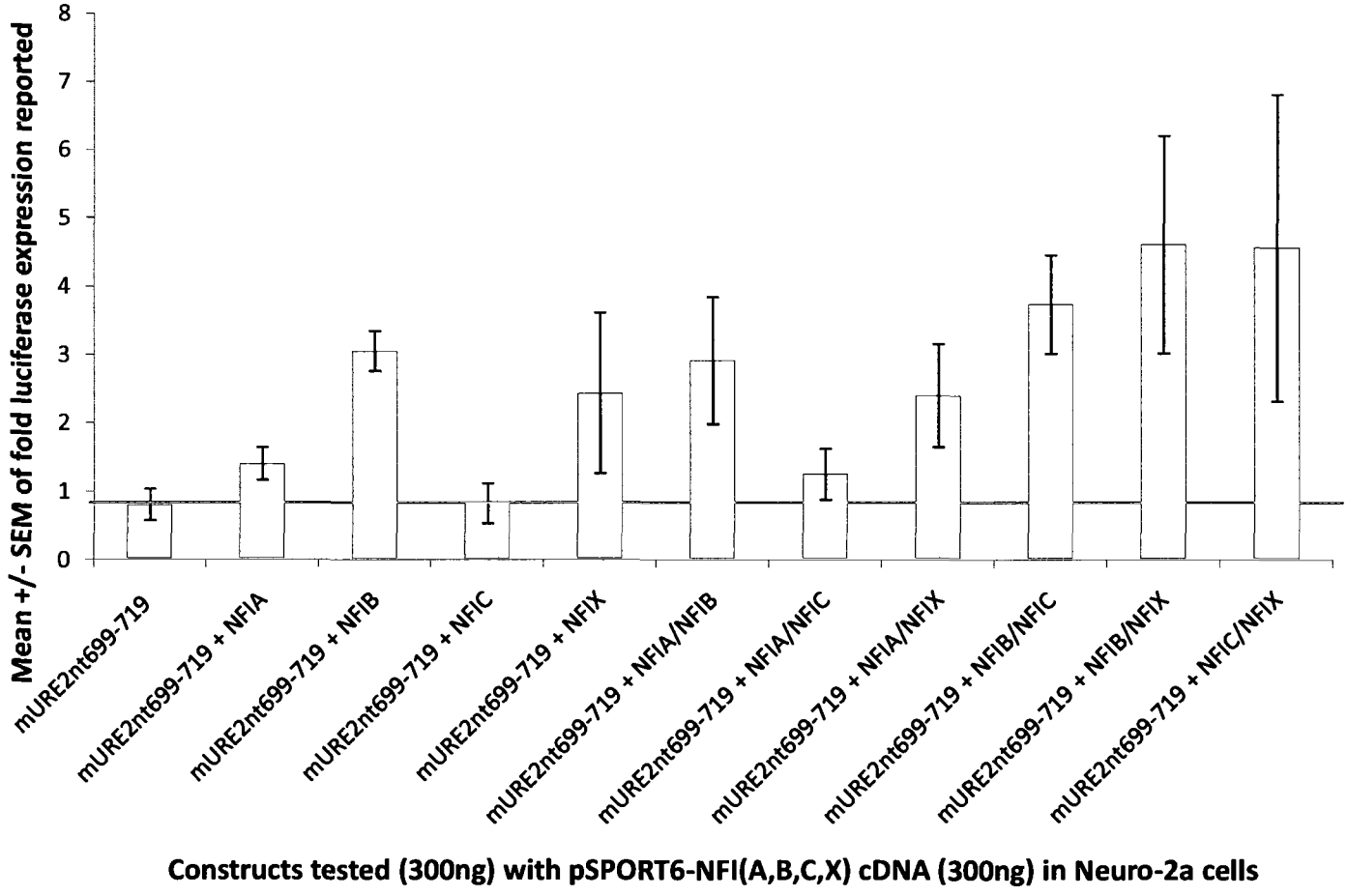
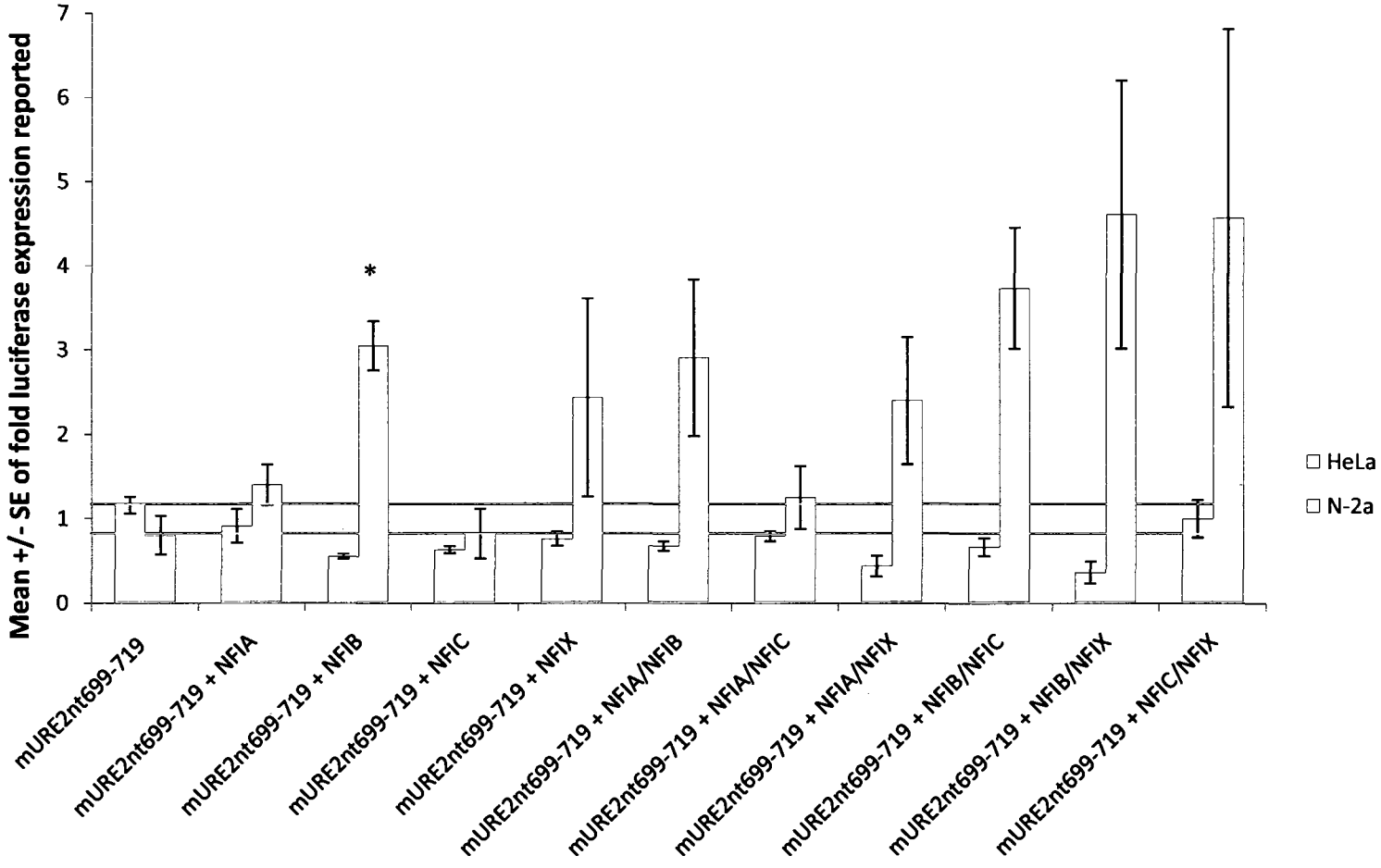

Mean +/- SEM of fold luciferase expression reported
0
1
2
3
4
5
6
7
8
mURE2nt699-719
mURE2nt699-719 + NFIA
mURE2nt699-719 + NFIB
mURE2nt699-719 + NFIC
mURE2nt699-719 + NFIX
mURE2nt699-719 + NFIA/NFIB
mURE2nt699-719 + NFIA/NFIC
mURE2nt699-719 + NFIA/NFIX
mURE2nt699-719 + NFIB/NFIC
mURE2nt699-719 + NFIB/NFIX
mURE2nt699-719 + NFIC/NFIX
Constructs tested (300ng) with pSPORT6-NFI(A,B,C,X) cDNA (300ng) in Neuro-2a cells

Figure 3.8: Comparison of the effect of NFI factors on the transcriptional activity of the mURE2nt699-719 sequence in HeLa and Neuro-2a cells. Results are represented by means +/- SEM (n=4 for each cell type) of fold luciferase expression normalized to averaged triplicate readings of luciferase generated from the empty pGL3-Control vector. A total of 300ng of pGL3-Control and/or mURE2699-719-pGL3-Control were tested with and without 300ng of each of the NFI factors, either in single transfection or double transfections. Multiple t-test comparisons were conducted between Neuro-2a and HeLa mean luciferase activity folds to measure which NFI factors had statistically different impacts between the two cell types: * $p \leq 0.05$ for Student's t-tests comparing HeLa results to N-2a results.



Constructs tested (300ng) with and without NFI(A,B,C,X)-pSPORT6 cDNA (300ng) in HeLa and Neuro-2a cells

activity. Again, the pcDNA1.1-REST cDNA had no significant effect on luciferase expression and will therefore not be described here. However, this experiment provided useful data for the NFI transfections of 200ng and will be the focus of this section.

Results from the transfections done in HeLa cells (Figure 3.9) show that 200ng of the NFIA plasmid is able to recognize the mURE2nt699-719 sequence and repress luciferase activity by 42% (mURE2nt699-719 fold + SEM = 0.85; NFIA fold + SEM = 0.49). This was a substantial decrease in activity compared to the previous assay in which 300ng of NFIA was only able to reduce the luciferase activity by 11% in HeLa (fold + SEM = 1.11 Figure 3.6). The only other factor which showed repressive activity when transfected alone was NFIX, as it was able to reduce the luciferase gene expression by 27% (fold + SEM = 0.62). The NFIB transfection was not able to produce such strong results, with a 5% reduction in activity (fold + SEM = 0.81). Finally, NFIC was the only NFI factor with a transactivator function, as it increased the luciferase activity by 32% (fold + SEM = 1.12).

When combined with other NFI factors, NFIA was again the only factor able to repress the luciferase gene in transfections done in HeLa cells, and this repression only occurred when it was paired with NFIB (fold + SEM = 0.66). The remaining NFI combinations did not show any repressive effects, but showed some activating effects. All of the pairs were able to transactivate luciferase activity by 22% (fold + SEM = 1.21) and the NFIC/NFIX combination increased luciferase activity by 43% (fold + SEM = 1.21).

Results in Neuro-2a cells (Figure 3.10) show that the NFIC factor is able to repress luciferase expression by 46 % (mURE2699-719 fold + SEM = 1.27; NFIC fold + SEM = 0.69). The NFIA/NFIC heterodimer was also able to repress, though not as well as NFIC alone (NFIA/NFIC fold + SEM = 1.01), resulting in a decrease of luciferase expression by

Table 2.3 DNA amounts for transfections done in test 3 (Figure 3.9-3.11) in HeLa and Neuro-2a cells. All amounts are in nanograms.

Test 3 (Figs 3.9-3.11)	1	2	3	4	5	6	7	8	9	10	11	12
pGL3-Control	200											
pRL	5	5	5	5	5	5	5	5	5	5	5	5
mURE2699-719-pGL3-Control		200	200	200	200	200	200	200	200	200	200	200
pSPORT6-NFIA			200				200	200	200			
pSPORT6-NFIB				200			200			200	200	
pSPORT6-NFIC					200			200		200		200
pSPORT6-NFIX						200			200		200	200
pcDNA3	200	200	200	200	200	200	200	200	200	200	200	200
pSPORT6	400	400	200	200	200	200						

Figure 3.9 Effect of NFI factors on the transcriptional activity of the mURE2nt699-719 sequence in HeLa cells. Results are represented by means +/- SEM (n=3) of fold luciferase expression normalized to averaged triplicate readings of luciferase generated from the empty pGL3-Control vector. A total of 200ng of pGL3-Control and/or mURE2699-719-pGL3-Control were tested with and without 200ng of each of the NFI factors, either in single transfection or double transfections.

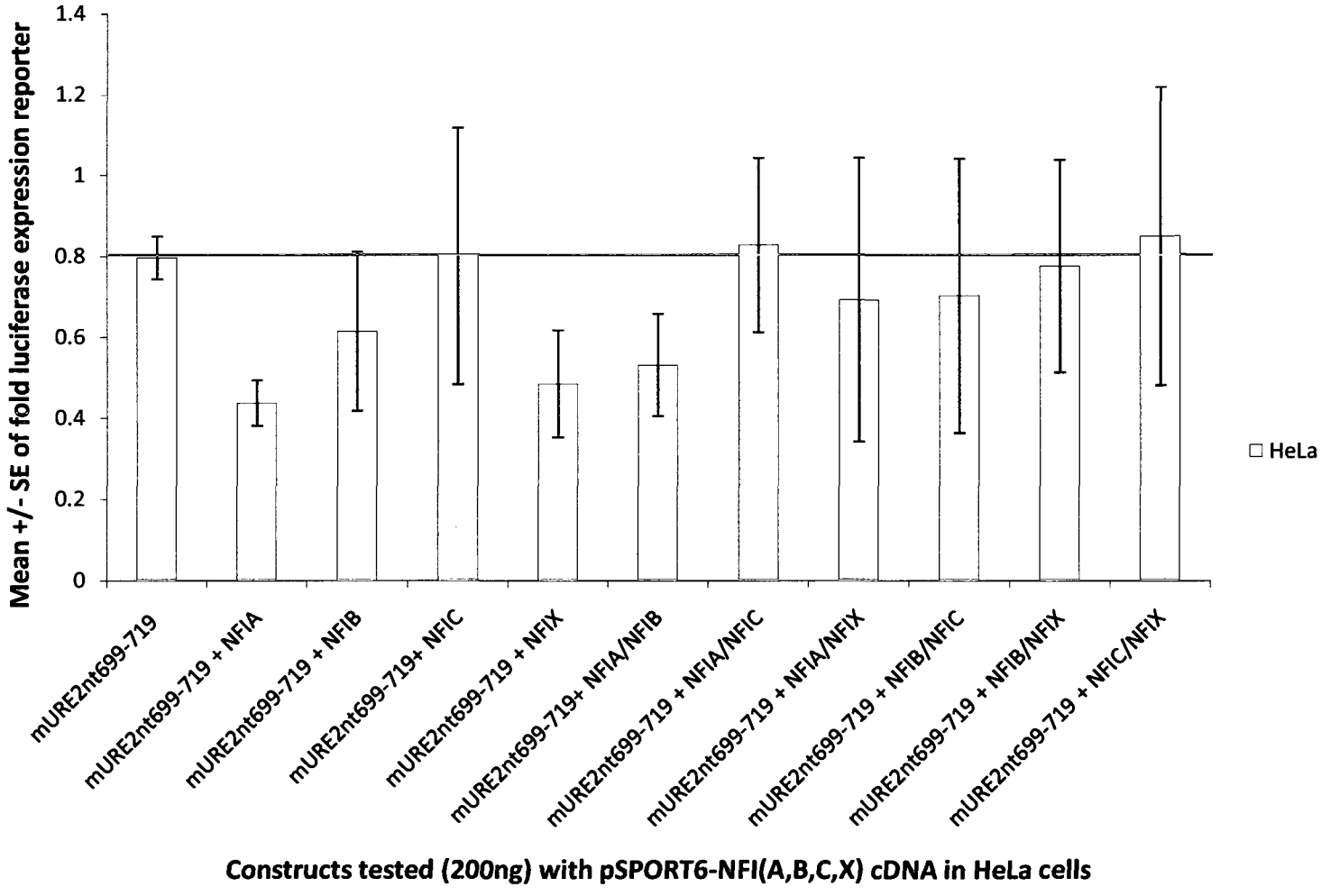
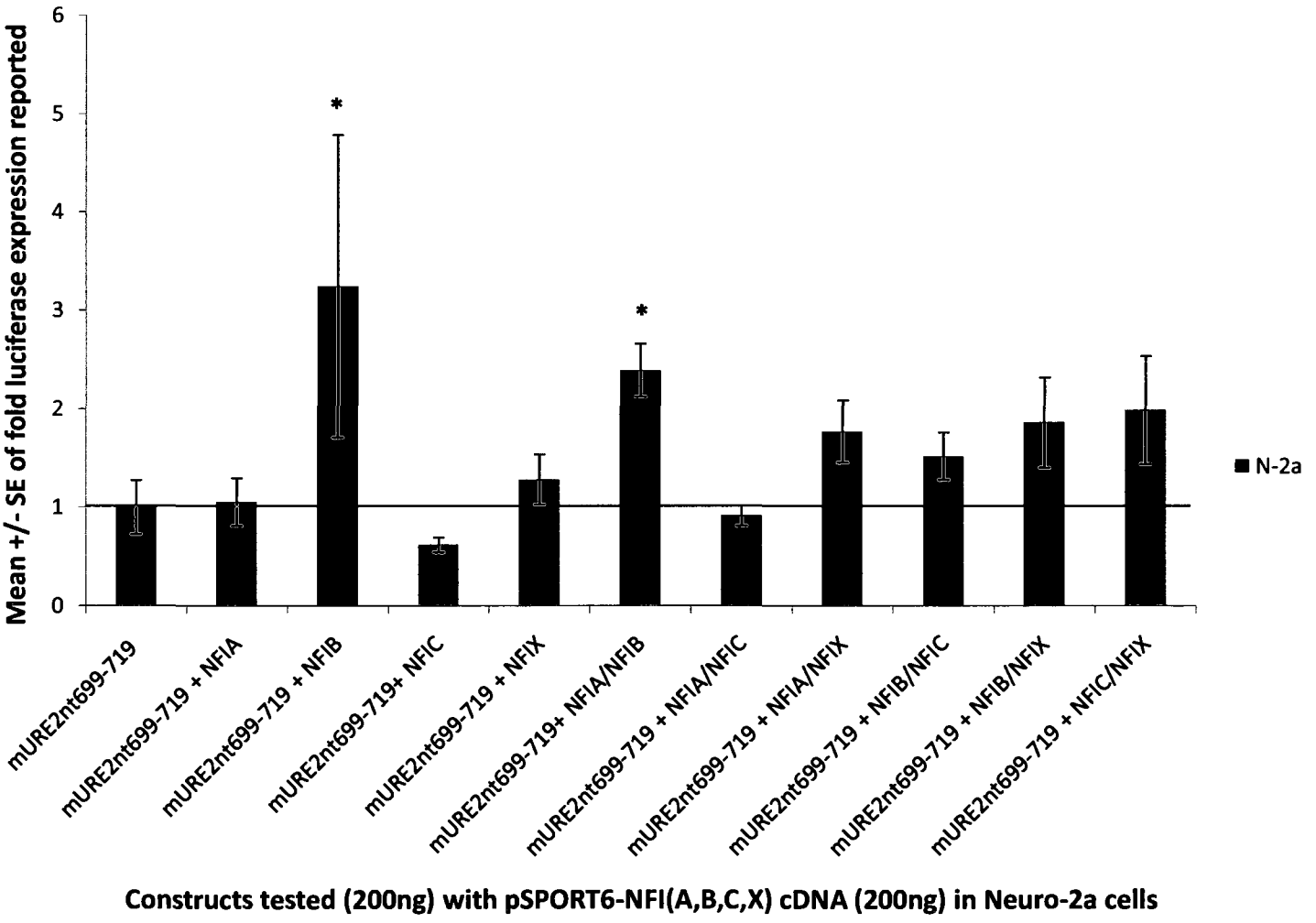

Mean +/- SE of fold luciferase expression reporter
1.4
1.2
1
0.8
0.6
0.4
0.2
0
mURE2nt699-719
mURE2nt699-719 + NFIA
mURE2nt699-719 + NFIB
mURE2nt699-719+ NFIC
mURE2nt699-719 + NFIX
mURE2nt699-719+ NFIA/NFIB
mURE2nt699-719 + NFIA/NFIC
mURE2nt699-719 + NFIA/NFIX
mURE2nt699-719 + NFIB/NFIC
mURE2nt699-719 + NFIB/NFIX
mURE2nt699-719 + NFIC/NFIX
HeLa
Constructs tested (200ng) with pSPORT6-NFI(A,B,C,X) cDNA in HeLa cells

Figure 3.10 Effect of NFI factors on the transcriptional activity of the mURE2nt699-719 sequence in Neuro-2a cells. Results are represented by means +/- SEM (n=4) of fold luciferase expression normalized to averaged triplicate readings of luciferase generated from the empty pGL3-Control vector. A total of 200ng of pGL3-Control and/or mURE2699-719-pGL3-Control were tested with and without 200ng of each of the NFI factors, either in single transfection or double transfections. * $p \leq 0.05$ when compared to luciferase generated from empty pGL3 vector.


Mean +/- SE of fold luciferase expression reported
6
5
4
3
2
1
0
*
*
N-2a
mURE2nt699-719
mURE2nt699-719 + NFIA
mURE2nt699-719 + NFIB
mURE2nt699-719+ NFIC
mURE2nt699-719 + NFIX
mURE2nt699-719+ NFIA/NFIB
mURE2nt699-719 + NFIA/NFIC
mURE2nt699-719 + NFIA/NFIX
mURE2nt699-719 + NFIB/NFIC
mURE2nt699-719 + NFIB/NFIX
mURE2nt699-719 + NFIC/NFIX
Constructs tested (200ng) with pSPORT6-NFI(A,B,C,X) cDNA (200ng) in Neuro-2a cells

about 20%. NFIA had no significant increase or decrease in luciferase expression, and the rest of the tests resulted in transactivation of the luciferase gene via the mURE2699-719 sequence.

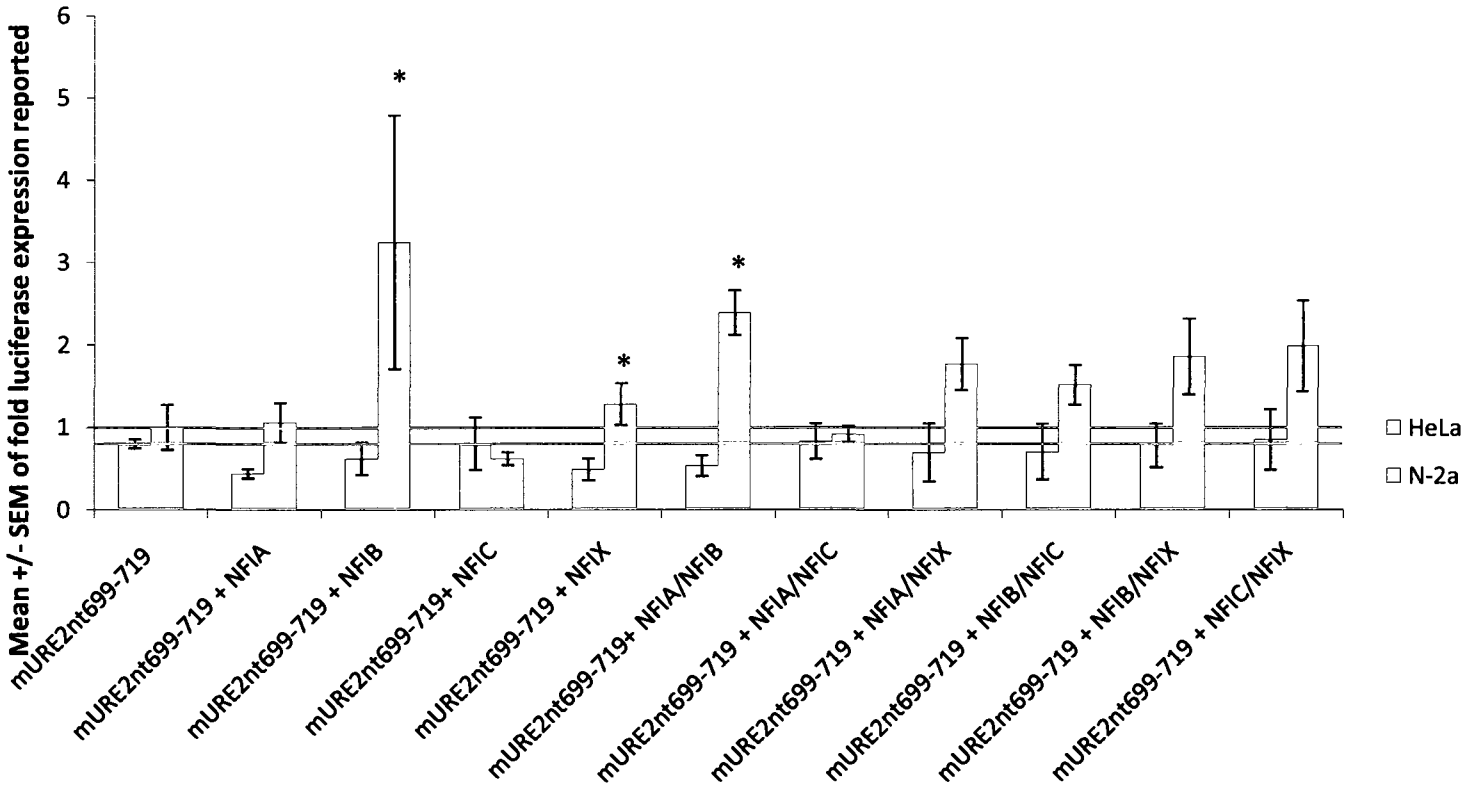
The comparison of mURE2699-719 activity in HeLa and Neuro-2a cells (Figure 3.11) when transfected with various NFI factors alone or in combinations was verified using multiple Student t-tests (see Appendix) between HeLa and Neuro-2a fold luciferase observed in all tests. Results show that overall, the mURE2nt699-719 sequence alone was not able to generate statistically different levels of luciferase activity between HeLa and Neuro-2a cells. However, most transfections were able to provide a different response according to cell type.

To sum up my findings, there is an indication that in transfections of 200ng of candidate cDNA; in HeLa cells NFIA is the strongest repressor when transfected alone, and that NFIX is able to moderately reduce the luciferase activity generated by the mURE2nt699-719 reporter plasmid. In Neuro-2a cells, the only factor showing repressive activity was NFIC, and only two heterodimers were able to repress gene expression in either cell line, namely the NFIA/NFIB combination in HeLa and the NFIA/NFIC complex in Neuro-2a.

3.2.4 Statistical analysis of raw data

A Tukey multiple comparisons test (see Appendix for raw data and results) was also conducted on all tests in order to see whether or not the NFI factors and REST protein elicited significant responses in the tests. The p value assigned as significant for all tests was of $p \leq 0.05$, though there were many values situated between the less stringent p value of $p =\leq 0.1$ and $p=0.05$. It was in using Tukey's test that it was determined that the REST co-transfection did not statistically change the luciferase gene expression in all tests in which it

Figure 3.11 Comparison of the effect of NFI factors on the transcriptional activity of the mURE2nt699-719 sequence in HeLa and Neuro-2a cells. Results are represented by means +/- SEM (n=4 for Neuro-2a and n=3 for Hela) of fold luciferase expression normalized to averaged triplicate readings of luciferase generated from the empty pGL3-Control vector. A total of 200ng of pGL3-Control and/or mURE2699-719-pGL3-Control were tested with and without 200ng of each of the NFI factors, either in single transfection or double transfections. Multiple t-test comparisons were conducted between Neuro-2a and HeLa mean luciferase activity folds to measure which NFI factors had statistically different impacts between the two cell types: $*p \leq 0.05$ for Student's t-tests comparing HeLa results to N-2a results.



Constructs tested (200ng) with pSPORT6-NFI(A,B,C,X) cDNA(200ng) in HeLa and Neuro-2a cells

was used. This statistical test also showed that pooled results using NFI co-transfection test results had distinct responses in HeLa and Neuro-2a tests, where most NFI factors were repressors in HeLa and activators in Neuro-2a, therefore the cell lines used were able to show overall different responses by most proteins tested in each cell type. The 3-factor ANOVA test (see Appendix for raw data and results) and Student t-tests (not shown) also showed that controls such as the mURE2699-719-pGL3-Control test with no cDNA did not have any variability between or within cell types. However, multiple Student t-tests were able to show that some of the results showing that NFI factors repressing gene expression in HeLa but activating gene expression in Neuro-2a were statistically significant, as seen by asterisks in Figures 3.8 and 3.11 (see Appendix 8).

4.DISCUSSION AND CONCLUSIONS

4.1 Analysis of the effects of mutagenesis of mURE2nt699-719 on URE2 activity in the forebrain.

4.1.1 Effects of mutagenized mURE2nt699-719 on URE2 activity in E11.5 transgenic mice forebrain.

At E11.5 of embryonic development, the observed differences in URE2 activity between the mURE2*lacZ* and the mURE2mut10a-*lacZ* transgenic mouse lines show an extension of the *lacZ* expression profile in the VZ and into the cerebral cortex. This result suggests that the mutation in URE2 is preventing a protein with repressor activity of binding to the URE2 sequence and assuring the SVZ-restricted expression as well as the restriction of the *lacZ* expression to the LGE-cortex border.

However, the mURE2mut10c-*lacZ* transgenic line is contradictory to the result seen in mURE2mut10a*lacZ*, showing that URE2 is active mainly in the SVZ of the LGE and the MGE, with no extension of *lacZ* expression into the cortex. This needs to be further investigated by doing more injections in order to obtain primary embryos These would not result in established transgenic mouse lines, but they would help identify the dominant *lacZ* pattern and we would then be able to see if the mURE2mut10a*lacZ* or the mURE2mut10c*lacZ* line is the one showing what is probably the true effect of the mutation in URE2.

4.1.2 Effects of mutagenized mURE2nt699-719 on URE2 activity in E13.5 transgenic mice forebrains.

The results obtained at E13.5 show that the *lacZ* signal is strong in the VZ of the LGE in the mURE2mut10*lacZ* lines where it is not present in the mURE2*lacZ* lines. Again, this suggests that a repressor is lacking in this tissue at this particular time point. This extension of the *lacZ* expression profile was seen throughout the entire anterior-posterior axis, suggesting that the VZ of the entire forebrain is affected by the mutation introduced in URE2. This time point is important because this is the time during development when the *Dlx* genes are co-expressed with the *gad65* and *gad67* genes, ensuring that the glutamic acid decarboxylase enzyme is produced by the developing interneurons. This enzyme is required to form GABA, which is crucial in the interneurons' signal inhibition function. This inhibition is required for proper forebrain signaling networks which when altered, results in disorders such as epilepsy. Since both of the mURE2mut10a*lacZ* and mURE2mut10c*lacZ* lines were able to produce this pattern, then it is possible to say that the mutation introduced in URE2 is most likely responsible for this VZ expression of *lacZ*. This suggests that there is no longer a SVZ-specific control of the URE2 activity in the LGE, and that a protein with repressor activity is probably not able to bind to URE2 at nt 699-719 at this time point in these mice.

4.1.3 Effects of mutagenized mURE2nt699-719 on URE2 activity in E15.5 and P0 transgenic mice forebrains.

For the tissues obtained at E15.5 and P0, there seemed to be an overall decrease in the strength of the *lacZ* expression in the lines carrying the mutant enhancer, but the profile

remained similar between the mURE2*lacZ* and mURE2mut10*lacZ* lines. Because the mURE2*lacZ* mouse brains sections were not always stained for *lacZ* expression at the same time as the specimen of the mutant lines for these time points, it is not possible to conclude that the decreased *lacZ* signal is definitely due to the mutation or if it's just due to variability in the staining experiment. One way to determine what caused the *lacZ* pattern to be similar but decreased in color intensity would be to slice brains from the different time points and store them until enough mURE2*lacZ* and mURE2lacZmut10*lacZ* brains are obtained for the *lacZ* staining procedure. They would then be treated at the same time and the signal could be analysed further than the current analysis.

4.1.4 Overall effects of mutagenized mURE2nt699-719 on URE2 activity in the transgenic mouse forebrain.

If the β-galactosidase signal is indeed truly decreased in the mutant lines at E15.5 and P0 as discussed in the previous section, this could cause the main take home message of these experiments to be interpreted in three ways.

The first explanation would be that protein(s) binding to the mutated site could have a transcriptional repressor role only in early embryogenesis such as E11.5 and E13.5 and that the similar but less intense *lacZ* profiles at E15.5 and P0 are due to protocol inconsistencies. This would explain the URE2 activity in the VZ of the LGE in mutants at E11.5 and E13.5, as well as the lack of striking differences except for the less intense URE2 activity between mURE2*lacZ* and mURE2mut10*lacZ* mice forebrains at E15.5 and P0.

The second interpretation possible is that the protein in question could have a repressor role at E11.5 and E13.5, but that at E15.5 and P0 the protein acts as an activator of

gene expression. This would also explain the URE2 activity seen in the VZ of the LGE since a repressor would keep URE2 active only in the SVZ of the LGE, and it also could explain the overall decreased *lacZ* signal in the forebrain at E15.5 and P0 since this protein would be required to increase the rate of transcription of the genes under the control of URE2 at those times. A protein such as this could have both transcriptional repressor and activator domains, such as the BRX2/PITX1 or RUNX1 homeobox transcription factors (Lopez et al. 2000; Wotton et al. 1994; Lutterbach et al. 2000). These types of repressors/activators have different functions which are determined by the identity of the co-factors they are bound to at their binding sites. Alternatively, the protein which utilizes this sequence could repress genes at E11.5 and E13.5 but passively activate gene expression by inhibiting the binding of a repressor at E15.5 and P0.

Finally, there could be more than one protein binding to this site at the times studied, with one transcriptional repressor binding to the mURE2nt699-719 sequence at E11.5 and E13.5, while a transcriptional activator binds to it at E15.5 and P0.

4.2 Variability of *lacZ* profiles between transgenic mouse lines limits analysis

When comparing the *lacZ* expression between mURE2*lacZ* and mURE2mut10*lacZ* transgenic mouse lines, there seemed to be some stronger *lacZ* signals in some of the lines compared to others. This could be due to the number of copies of the transgene inserted in the transgenic mouse genome, and could be addressed by conducting a quantitative PCR (q-PCR) reaction to obtain an approximation of the number of *lacZ* transgene transcripts in all of the mouse lines generated.

By doing this type of experiment it would assure a better representation of URE2 activity in all *lacZ* transgenic mice forebrains analysed. Knowing that a certain line contains twice or three times the number of *lacZ* copies could possibly explain some differences observed in previous analyses which were difficult to explain.

There is also another explanation for the variability between transgenic mouse lines, since there is the potential for a positional effect during the DNA construct's integration into the mouse genome. By inserting in areas not prone to DNA expression such as heterochromatin, the transgene will not be expressed by the mouse but will be amplified when screening by PCR. There is no reason to believe that this is happening in lines for which results were presented in this thesis. Whether or not a gene also integrates in an area under the influence of potent regulatory elements will also either hinder or increase its expression, and this could explain some of the differences seen between transgenic lines created using the same DNA construct. To reduce the complexity that positional effects may pose to these types of analyses is to create many established transgenic mouse lines or create inject eggs with the goal of harvesting and studying many primary embryos. This will ensure that profiles which are the minority of all *lacZ* profiles observed are not included in analyses. One could also inject constructs of the test *cis*-regulatory sequence upstream of a reporter gene such as done in this project, but using insulator sequences which flank the cis-regulatory element and the reporter gene. This would ensure that the reporter gene expression is not affected by neighbouring regulatory elements, and only of the test cis-regulatory sequence.

4.3 Luciferase assays support evidence of a transcription factor binding site on mURE2nt699-719

4.3.1 Luciferase activity reported does not support the hypothesis that REST is the repressor binding on mURE2nt699-719

Based on the results shown in Figure 3.5, the REST protein does not use the mURE2nt699-719 sequence for its binding site in order to carry out its repressor function. This was observed in both neural (N-2a) and epithelial (HeLa) cell lines. This is supported by the statistical analysis conducted on the raw data used for Figure 3.5. The Student t-test was not able to find that data obtained with the REST cDNA transfections were at all different from the data obtained without it. However, a Western blot will have to be done in the future to confirm the endogenous REST levels in these cells as well as to confirm that pcDNA1.1-REST cDNA transfections were successful in all cell types used. With the Western blot, it will also be possible to see if the pcDNA1.1-REST plasmid was indeed producing a REST mRNA which was then translated into REST protein in the cell lines used. Post-translational modifications to REST proteins produced by REST cDNA transfections could also be the cause of non-functional REST. This should also be investigated. Only after these types of analyses will we be able to say for certain that REST does not bind to this sequence within URE2.

4.3.2 Luciferase activity reported supports the hypothesis that NFI protein(s) are repressors that bind to mURE2nt699-719

4.3.2.1 Analysis of single transfections using 300ng of NFI cDNA

Based on the results shown in Figures 3.6-3.8, some NFI proteins are able to repress gene expression via the mURE2nt699-719 in HeLa cells when 300ng of their cDNA is transfected into these cells. The NFI factor with the most potent repressor activity was NFIB, repressing luciferase gene expression by 54%; NFIC was able to reduce it by 46%; and the NFIX factor was able to repress it by 33%. The statistically significant results, however did not include the NFIB repression ($p=0.12$, see Appendix 5).

All four NFI proteins, when transfected alone in Neuro-2a cells were not able to repress the luciferase gene (Figure 3.7). Out of the three NFI factors able to repress the luciferase gene in HeLa cells, namely NFIB, NFIC and NFIX, only two were able to transactivate the luciferase gene in Neuro-2a. NFIC had a minimal impact on the change of luciferase activity, whereas NFIB and NFIX were able to increase luciferase activity. However, though the trend indicates that the NFI factors mainly increases the luciferase activity, these increases were found to be statistically non-significant (see Appendix 6).

Together, these results allow us to deduce that these NFI factors are required to repress gene expression in non-neural cells such as HeLa and activate gene expression at different levels in neural cells such as Neuro-2a. The results in HeLa agree with previous studies in which NFIC and NFIX were able to repress genes in HeLa (Chaudhry et al., 1999), but no such gene repression in HeLa by NFIB could be found in the literature.

4.3.2.2 Analysis of co-transfections using 300ng of each NFI cDNA to form heterodimers

When looking for heterodimers using the mURE2nt699-719 site as a potential binding site, results from HeLa (Figure 3.6) show that the NFIA/NFIC combination was able to repress the luciferase gene expression by 32%, the NFIA/NFIB by 42%, while the NFIA/NFIX complex was able to reduce it by 55%. The NFIA/NFIX heterodimer result suggests that synergistic activity is occurring between those two NFI factors in HeLa cells, since the sum of both of their activities is smaller than their activity when they are part of a heterodimer. This is the only example of synergy observed in HeLa cells. There are other repressing heterodimers in HeLa, including the NFIB/NFIC complex, where a reduction of approximately 38% occurs when they are co-transfected together, and most importantly the NFIB/NFIX complex, which is able to reduce luciferase gene expression by 61%. The heterodimers mentioned in this paragraph with repressor functions could not be found in supporting literature, therefore this work is novel in showing they are able to repress the luciferase gene in HeLa as complexes.

Overall, these results suggest that the most potent repressors of luciferase in HeLa cells using the mURE2699-719-pGL3-reporter plasmid were the NFIA/NFIX and NFIB/NFIX heterodimers. These also happen to be the only statistically significant differences between the luciferase activity generated by mURE2nt699-719 and co-transfections involving 300ng of NFI cDNA's (see Appendix 5).

The complexes are therefore capable of repression by binding to mURE2nt699-719 in epithelial cells such as HeLa, but the repressor activity that was to be characterized in this

project occurred in cells of the forebrain during development, which is why Neuro-2a cells were also chosen for investigation. Originally we were testing to find a potential repressor of neural genes sometimes required in neural cells, but potentially always active in non-neural cells, such as the REST repressor.

As in the single transfections, no NFI heterodimers were able to repress the luciferase gene using mURE2nt699-719 as a regulatory sequence in Neuro-2a. This was also supported by statistical analysis, where none of the data were statistically different, though the trend seemed to show that some of the pairs are able to transactive gene expression (see Appendix 6).

To identify which factors are truly activators of gene expression, the transfection experiments would have to be redone using the pGL4 vector instead of the pGL3 vector. Because the pGL3 vector is designed to test sequences with repressor functions, it does not give a true idea of how much activation occurs when it does occur while using this system since there is a basal level of luciferase expression due to the nature of the elements within the pGL3 vector. The pGL4 vector is ideal for testing transcriptional activating regulatory elements since it contains an SV-40 enhancer element in the plasmid, which is downstream of the luciferase gene. The test sequence is inserted upstream of the luciferase gene in the pGL4 vector, and will enable the promoter-like gene activation effects of the sequence when it is exposed to various proteins already found in the cells used, and/or those resulting from cDNA transfections. This pGL4 system is much more sensitive to gene activation than the pGL3 system, and all candidates suspected of having gene expression activator roles should be retested with pGL4. We could then draw stronger conclusions as to which factor has which regulatory role in which cell line.

4.3.2.3 Analysis of single transfections using 200ng of candidate NFI cDNA

When 200ng of candidate NFI cDNA's are transfected in HeLa cells (Figure 3.9), NFIA is the strongest repressor when transfected alone, being able to repress the luciferase gene by 42%. The NFIX factor is also able to moderately reduce the luciferase activity generated by the mURE2nt699-719 reporter plasmid, reducing luciferase activity by 27%. These trends are however not statistically significant (see Appendix 7). Additionally, no support for this role for NFIA or NFIX in HeLa could be found in the literature.

The results for Neuro-2a transfections of single NFI factors showed that only NFIC was able to repress the luciferase gene when 200ng of its cDNA was co-transfected with the luciferase pGL3 plasmid (Figure 3.10). This factor was able to repress the luciferase activity by 46%. However, this could not be supported by the statistical analysis, the only statistically significant result being the one showing that NFIB is an activator of gene expression (see Appendix 8). The literature also only supports NFIC as a repressor in epithelial breast cancer cells, not in neural cells (Eeckhoute et al. 2006).

4.3.2.4 Analysis of co-transfections using 200ng of each candidate NFI cDNA to form heterodimers

In HeLa cells, the only complex able to change the luciferase activity in this experiment was the NFIA/NFIB complex, as it was able to reduce the luciferase activity by 23%. However, this was not found to be statistically significant and was also not supported by any previous studies showing such results.

In Neuro-2a cells, the NFIA/NFIC heterodimer was able to reduce the luciferase activity by 21%. All of the other NFI factors formed heterodimers in Neuro-2a cells which

were shown to only have transactivational or no effect on luciferase gene expression. These results suggest that further studies need to be done using the NFIA/NFIC complex to investigate if they indeed bind to the mURE2nt699-719 sequence in neural cells. Though the trend indicates this complex may have repressor functions while others have activator roles, the only statistically significant differences in luciferase activity found were between the luciferase activity seen in the mURE2nt699-719 transfection and the luciferase activity resulting from the transfection of NFIA/NFIB, which happens to transactivate the luciferase gene in the pGL3 plasmid (see Appendix 8). No information on NFI proteins' regulatory activity in Neuro-2a cells could be found in the literature, making all of these findings novel.

Together, the data for NFI heterodimers' effects on luciferase activity suggest that most heterodimers capable of repression in HeLa cells are probably required for repression of genes in non-neural cells but lose their repressor functions in neural cells such as Neuro-2a. For example, the transfection of 300ng of NFIB/NFIX cDNA is able to cause a transactivation of the luciferase gene by 6.2 fold in Neuro-2a cells, but instead produce a 61% reduction in HeLa cells. Even though the results are statistically non-significant, they do show a trend where NFI factors are activators of gene expression when using mURE2699-719 as binding site in Neuro-2a cells while in HeLa cells, they mostly use it as binding site for repressor functions, with the exception of NFIA.

So far, the literature, when present, has been quite contradictory and misleading on the roles of NFI proteins in various cell lines. For example, it has been documented that the NFIC/NFIX heterodimer is a transactivating complex in epithelial JEG-3 cells (Chaudhry et al. 1998), where the results in my project indicate that this complex is capable of transactivating gene expression in neural cells only and not in epithelial HeLa cells. This is

an example of how these factors can have cell- and promoter-specific roles, therefore it is important to contribute to the database of knowledge accumulated for NFI factors' effects on cell specific gene regulation.

4.3.3 Potential causes for variability in raw data of luciferase assays

Since the mURE2nt699-719 sequence, when inserted in the pGL3 plasmid and transfected in the same cell line but in different amounts, was able to produce different results, there was a possibility that the effects observed were due to the amount of DNA transfected in the cells. The effects due to this variable, namely the amount of cDNA transfected, were shown by a statistical analysis to almost be statistically significant, but not so (see Appendix 8). Therefore, an experiment using different amounts of cDNA in one cell line should be done in the future to optimize the transfections.

4.4. Potential role of the URE2 enhancer element in the regulation of *Dlx* expression

Three scenarios were identified as possibly happening in the mutant transgenic lines. The first possibility is a protein recognizing mURE2nt699-719 that would have a repressor role at E11.5 and E13.5 and the overall decrease in *lacZ* expression in E15.5 and P0 was due to protocol inconsistencies. The second possibility suggests that the protein involved could have a transcriptional repressor role at E11.5 and E13.5, but then change functions and become a transcriptional activator at E15.5 and P0. And finally, the third possibility stipulates that more than one protein could bind to the mURE2nt699-719 sequence, one of them being a transcriptional repressor while the other would have a transcriptional activator role.

Based on the transfection results of this project, it could be possible that the second scenario mentioned in the above paragraph is at play in the mURE2*lacZ* transgenic mice forebrains studied. It is possible that the protein acting as a transcriptional repressor at E11.5 and E13.5 in the mURE2*lacZ* transgenic mouse forebrain is NFIC, as it was the only repressor in Neuro-2a cells, reducing the luciferase activity by 46% in Neuro-2a cells. Later on at E15.5 and P0, it is possible that the NFIC protein dimerizes with either NFIB or NFIX as both of these heterodimers were able to increase the luciferase activity in Neuro-2a.

4.5. Future directions

To identify which scenario of the three mentioned is occurring in the mouse, the best experiment that could identify the protein present on this site at E11.5, E13, E15.5 and P0 time points would be a Chromatin Immunoprecipitation Assay followed by sequencing of binding sites that were isolated at each time point. This is possible since antibodies for the REST and Nuclear Factor-1 proteins are currently available on the market for exactly this type of assay, therefore we could find out quickly if one or more of these five candidates utilize(s) the mURE2nt699-719 binding site at the time points mentioned.

Another assay to be done would be the Electrophoretic Mobility Shift Assay using recombinant REST and/or recombinant Nuclear Factor-1 proteins to see if they can bind to the mURE2699-719. Supershifts using REST and NFI antibodies could also be attempted.

4.6 Conclusion

A bioinformatical analysis of the mouse URE2 enhancer sequence was conducted for comparison of the conserved sequence found (and protected by a DNaseI footprinting assay) to known transcription factor binding sites. Originally, only one candidate protein was

suspected of binding in the mouse URE2's sequence from base pairs 699-719; however, with revised data the MatInspector program identified four additional candidate proteins.

The goal of this project was to identify the protein(s) bound to this 20bp segment, as well as to elaborate on this protein's role in *Dlx1/2* regulation via the URE2 enhancer. Additionally, because of the close proximity of the predicted binding sites of the five candidates tested, namely the four NFI proteins and the REST protein, it is possible that the protected area as per the DNAseI footprinting analysis represents a complex of proteins, or hetero- and/or homodimer(s). The NFI genes are able to form many isoforms, and many of which are known to form heterodimers, making this analysis novel in the context of *Dlx* gene expression and/or function, but complex. No association has ever been made between REST and NFI proteins, however they are both known to be able to dimerize with the ubiquitously expressed trans-activator SP3 zinc finger protein (Kim 2006; Luciakova et al., 2008). Also, no DLX proteins have ever been identified to bind to either REST or NFI family of proteins, though the *nfic*$^{-/-}$ and *Dlx1/2* null mice both have severe defects in their teeth, reflecting their roles in molar development (Qiu 1997; Steele-Perkins 2003).

While results were not entirely clear on which candidate protein(s) represented the transcription factor binding to mURE2699-719, they did provide us with a better idea of where to take the project in the future. For example, further research should be limited to cell culture and transfections using a neural cell line only and the NFI factors since the REST protein did not seem to have any repressor role in a neural setting. Since the transgenic mouse results indicate that there is a possible repression occurring in the forebrain at E11.5 and E13.5, and perhaps activation at E15.5 and P0, this should be an indication to continue

investigating the NFI factors because according to my results, they can fulfill these two roles in a neural setting.

Finally, the URE2 enhancer is a complex regulatory sequence. It is necessary to further study this sequence and its role in *Dlx* regulation as it may be acting on other neighboring genes and interacting with other important proteins not previously identified. By establishing which transcription factors are involved in the interneuronal specification and their eventual migration to the developing cortex, such as in this project, we can identify the transcription factors for which their expression is necessary for correct and specific cell type functions in the brain. By establishing all regulators of *Dlx* genes as well as DLX targets, our work will not only be able to provide a genetic map which will be useful for understanding interneuronal specification and migration, but we will be participating in the deciphering of the complex puzzle which is the entire set of molecular cues required for proper cell organization in the developing forebrain of vertebrates. This type of work benefits the scientific community by shedding light on how brain disorders involving interneuronal inhibitory signalling, such as epilepsy and autism, are caused, and can help find new treatments and cures for such disorders.

BIBLIOGRAPHY

References

Acampora D., Merlo GR., et al. 1999. Craniofacial, vestibular and bone defects in mice lacking the distal-less-related gene Dlx5. Development 126(17):3795.

Akimenko MA., Ekker M., et al. 1994. Combinatorial expression of three zebrafish genes related to distal-less: Part of a homeobox gene code for the head. The Journal of Neuroscience 14(6):3475.

Anderson SA, Eisenstat DD, et al. 1997a. Interneuron migration from basal forebrain to neocortex: Dependence on dlx genes. Science 278(5337):474.

Anderson SA., Qiu M., et al. 1997b. Mutations of the homeobox genes dlx-1 and dlx-2 disrupt the striatal subventricular zone and differentiation of late born striatal neurons. Neuron 19(1):27.

Ballas N., and Mandel G. 2005. The many faces of REST oversee epigenetic programming of neuronal genes. Curr Opin Neurobiol 15(5):500.

Bejerano G., Pheasant M., et al. 2004. Ultraconserved elements in the human genome. Science 304(5675):1321.

Bond AM., Vangompel MJ., et al. 2009. Balanced gene regulation by an embryonic brain ncRNA is critical for adult hippocampal GABA circuitry. Nat Neurosci 12(8):1020-7.

Bulfone A., Puelles L., et al. 1993. Spatially restricted expression of dlx-1, dlx-2 (tes-1), gbx-2, and wnt-3 in the embryonic day 12.5 mouse forebrain defines potential transverse and longitudinal segmental boundaries. The Journal of Neuroscience 13(7):3155.

Campbell CE., Piper M., et al. 2008. The transcription factor nfix is essential for normal brain development. BMC Dev Biol 8:52.

Chaudhry AZ., Vitullo AD., et al. 1998. Nuclear factor I (NFI) isoforms differentially activate simple versus complex NFI-responsive promoters. J Biol Chem 273(29):18538-46.

Chaudhry AZ., Vitullo AD., and Gronostajski RM. 1999. Nuclear factor I-mediated repression of the mouse mammary tumor virus promoter is abrogated by the coactivators p300/CBP and SRC-1. J Biol Chem 274(11)7072-7081.

Cobos I., Borello U., et al. 2007. Dlx transcription factors promote migration through repression of axon and dendrite growth. Neuron 54(6):873.

Cobos I., Calcagnotto, ME., et al. 2005. Mice lacking Dlx1 show subtype-specific loss of interneurons, reduced inhibition and epilepsy. Nat Neurosci 8(8):1059.

Cohen S., Bronner M., et al. 1989a. Distal-less encodes a homeodomain protein required for limb formation in drosophila. Nature 338:432.

Cohen S., and Jürgens G. 1989b. Distal pattern formation in drosophila: Cell autonomous requirement for distal-less gene activity in. EMBO J 8(7):2045.

Coulson JM. 2005. Transcriptional regulation: Cancer, neurons and the REST. Current Biology 15(17):665.

Cox CJ., Espinoza HM., et al. 2002. Differential regulation of gene expression by PITX2 isoforms. J. Biol. Chem. 277(28):25001.

das Neves L., Duchala CS., et al. 1999. Disruption of the murine nuclear factor I-A gene (nfia) results in perinatal lethality, hydrocephalus, and agenesis of the corpus callosum. Proc Natl Acad Sci U S A 96(21):11946-51.

Deneen B., Ho R., et al. 2006. The transcription factor NFIA controls the onset of gliogenesis in the developing spinal cord. Neuron 52(6):953.

Depew MJ., Lufkin, T., et al. 2002. Specification of jaw subdivisions by dlx genes. Science 298(5592):381.

Depew MJ., Liu JK., et al. 1999. Dlx5 regulates regional development of the branchial arches and sensory capsules. Development 126(17):3831.

Drake JA., Bird C., et al. 2005. Conserved noncoding sequences are selectively constrained and not mutation cold spots. Nat Genet 38(2):223.

Driller K., Pagenstecher A., et al. 2007. Nuclear factor I X deficiency causes brain malformation and severe skeletal defects. Mol Cell Biol 27(10):3855-67.

Eeckhoute, JR., Carroll J., et al. 2006. A cell-type-specific transcriptional network required for estrogen regulation of cyclin D1 and cell cycle progression in breast cancer. Genes & Dev. 20:2513-2526.

Ekker M., Akimenko MA., et al. 1992. Regional expression of three homeobox transcripts in the inner ear of zebrafish embryos. Neuron 9(1):27.

Feledy JA., Beanan MJ., et al. 1999. Inhibitory patterning of the anterior neural plate in xenopus by homeodomain factors Dlx3 and Msx1. Dev Biol 212(2):455.

Feng J., Bi C., et al. 2006. The evf-2 noncoding RNA is transcribed from the dlx-5/6 ultraconserved region and functions as a dlx-2 transcriptional coactivator. Genes & Dev.

20(11):1470.

Ghanem N., Yu M., et al. 2008. Characterization of a distinct subpopulation of striatal projection neurons expressing the dlx genes in the basal ganglia through the activity of the I56ii enhancer. Dev Biol 322(2):415.

Ghanem N., Yu, M., et al. 2007. Distinct cis-regulatory elements from the Dlx1/Dlx2 locus mark different progenitor cell populations in the ganglionic eminences and different subtypes of adult cortical interneurons. J. Neurosci. 27(19):5012.

Ghanem N., Jarinova O., et al. 2003. Regulatory roles of conserved intergenic domains in vertebrate dlx bigene clusters. Genome Res 13(4):533.

Gronostajski RM. 2000. Roles of the NFI/CTF gene family in transcription and development. Gene 249(1-2):31.

Jones FS., and Meech R. 1999. Knockout of REST/NRSF shows that the protein is a potent repressor of neuronally expressed genes in non-neural tissues. [Cambridge, UK]:Co. of Biologists.

Kadonaga. 2004. Regulation of RNA polymerase II transcription by sequence-specific DNA binding factors. Cell 116(2):247.

Khoury G., and Gruss P. 1983. Enhancer elements. Cell 33(2):313.

Kim CS., Choi HS., et al. 2006. Evidence of the neuron-restrictive silencer factor (NRSF) interaction with Sp3 and its synergic repression to the mu opioid receptor (MOR) gene. Nucleic Acids Res 34(22):6392.

Kim SY., and Pritchard JK. 2007. Adaptive evolution of conserved noncoding elements in mammals. PLOS Genetics 3(9):1572.

Kim CS., Hwang CK., et al. 2008. Novel function of neuron-restrictive silencer factor (NRSF) for posttranscriptional regulation. Biochim Biophys Acta 1783(10):1835-46.

Krumlauf R. 1994. Hox genes in vertebrate development. Cell 78(2):191.

Le TN., Du G., et al. 2007. Dlx homeobox genes promote cortical interneuron migration from the basal forebrain by direct repression of the semaphorin receptor neuropilin-2. J. Biol. Chem. 282(26):19071.

Lee MH., Kwon TG., et al. 2003. BMP-2-induced osterix expression is mediated by Dlx5 but is independent of Runx2. Biochem Biophys Res Commun 309(3):689.

Li Y., Liu Q., et al. 2008. Regulatory role of neuron-restrictive silencing factor in the specific expression of cocaine- and amphetamine-regulated transcript gene. J Neurochem

106(3):1314-24.

Liu JK., Ghattas I., et al. 1997. Dlx genes encode DNA-binding proteins that are expressed in an overlapping and sequential pattern during basal ganglia differentiation. Dev Dyn 210(4):498-512.

Lopez, S., Island ML., et al. 2000. Repression of Virus-Induced Interferon A Promoters by Homeodomain Transcription Factor Ptx1. Mol. Cell. Biol. 20(20):7527-7540,

Luciakova K., Kollarovic G., et al. 2008. Growth-dependent repression of human adenine nucleotide translocator-2 (ANT2) transcription: Evidence for the participation of smad and sp family proteins in the NF1-dependent repressor complex. Biochem J 412(1):123-30.

Lunyak VV., and Rosenfeld MG. 2005. No rest for REST: REST/NRSF regulation of neurogenesis. Cell 121(4):499.

Lutterbach, B., Westendorf, J.J., et al. 2000. A mechanism of repression by acute myeloid leukemia-1, the target of multiple chromosomal translocations in acute leukemia. J. Biol. Chem. 275:651–656.

Luo T., Matsuo-Takasaki M., et al. 2001. Differential regulation of dlx gene expression by a BMP morphogenetic gradient. Int J Dev Biol 45(4):681.

Majumder S. 2006. REST in good times and bad: Roles in tumor suppressor and oncogenic activities. Cell Cycle 5(17):1929.

Mandel G., Tsay HJ., et al. 1992. Silencing the type II sodium channel gene: A model for neural-specific gene regulation. Neuron 9(1):37.

Mandel G., Kraner SD., et al. 1995. REST: A mammalian silencer protein that restricts sodium channel gene expression to neurons. Cell 80(6):949.

Marin O., and Rubenstein JL. 2001. A long, remarkable journey: Tangential migration in the telencephalon. Nat Rev Neurosci 2(11):780-90.

Mason S., Piper M., et al. 2009. Nuclear factor one transcription factors in CNS development. Mol Neurobiol 39(1):10-23.

Maston. 2006. Transcriptional regulatory elements in the human genome. Annual Review of Genomics and Human Genetics 7:29.

McGinnis W, and Krumlauf R. 1992. Homeobox genes and axial patterning. Cell 68(2):283.

Miller FD, and Gauthier AS. 2007. Timing is everything: Making neurons versus glia in the developing cortex. Neuron 54(3):357.

Morasso MI, Grinberg A., et al. 1999. Placental failure in mice lacking the homeobox gene dlx 3. Proc Natl Acad Sci U S A 96(1):162.

Mortazavi A., Thompson ECL., et al. 2006. Comparative genomics modeling of the NRSF/REST repressor network: From single conserved sites to genome-wide repertoire. Genome Res 16(10):1208.

Muller F., Blader P., et al. 2002. Search for enhancers: Teleost models in comparative genomic and transgenic analysis of cis regulatory elements. BioEssays 24(6):564.

Neumann SB., Seitz R., et al. 2004. Relaxation of glycine receptor and onconeural gene transcription control in NRSF deficient small cell lung cancer cell lines. Mol Brain Res 120(2):173.

Ooi L., and Wood IC. 2007. Chromatin crosstalk in development and disease: Lessons from REST. Nature Reviews Genetics 8(7):544.

Osada S., Daimon S., et al. 1996. Identification of DNA binding-site preferences for nuclear factor IA. FEBS Lett 390(1):44.

Palm, K., Belluardo, N., et al. 1998. Neuronal expression of zinc finger transcription factor REST/NRSF/XBR gene. J. Neurosci 18(4):1280–1296.

Panganiban G., and Rubenstein JLR. 2002. Developmental functions of the distal-less/Dlx homeobox genes. Development 129(19):4371.

Panganiban G., Irvine SM., et al. 1997. The origin and evolution of animal appendages. Proc Natl Acad Sci U S A 94(10):5162.

Park GT., and Morasso MI. 2002. Bone morphogenetic protein-2 (BMP-2) transactivates Dlx3 through Smad1 and Smad4: Alternative mode for Dlx3 induction in mouse keratinocytes. Nucleic Acids Res 30(2):515.

Pera E., and Kessel M. 1999. Expression of DLX3 in chick embryos. Mech Dev 89(1-2):189.

Plachez C., Lindwall C., et al. 2008. Nuclear factor I gene expression in the developing forebrain. J Comp Neurol 508(3):385-401.

Poitras L., Ghanem N., et al. 2007. The proneural determinant MASH1 regulates forebrain Dlx1/2 expression through the I12b intergenic enhancer. Development 134(9):1755-65.

Potter GB., Petryniak MA., et al. 2009. Generation of cre-transgenic mice using Dlx1/Dlx2 enhancers and their characterization in GABAergic interneurons. Mol Cell Neurosci

40(2):167.

Qiu M., Bulfone A., et al. 1997. Role of the dlx homeobox genes in proximodistal patterning of the branchial arches: Mutations of dlx-1, dlx-2, and dlx-1 and-2 alter morphogenesis of proximal skeletal and soft tissue structures derived from the first and second arches. Dev Biol 185(2):165.

Reddy BY., Greco SJ., et al. 2009. RE-1-silencing transcription factor shows tumor-suppressor functions and negatively regulates the oncogenic TAC1 in breast cancer cells. Proc Natl Acad Sci U S A 106(11):4408-13.

Robledo RF., Rajan L., et al. 2002. The Dlx5 and Dlx6 homeobox genes are essential for craniofacial, axial, and appendicular skeletal development. Genes & Dev 16(9):1089.

Ryoo HM., Hoffmann HM., et al. 1997. Stage-specific expression of dlx-5 during osteoblast differentiation: Involvement in regulation of osteocalcin gene expression. Molecular Endocrinology 11(11):1681.

Schoenherr CJ., and Anderson DJ. 1995. The neuron-restrictive silencer factor (NRSF): a coordinate repressor of multiple neuron-specific genes. Science. 267(5202)1360-63.

Shimamoto T., Ohyashiki K., et al. 2000. Overexpression of the homeobox gene DLX-7 inhibits apoptosis by induced expression of intercellular adhesion molecule-1. Exp Hematol 28(4):433.

Shimamoto T., Nakamura S., et al. 1997. Inhibition of DLX-7 homeobox gene causes decreased expression of GATA-1 and c-myc genes and apoptosis. Proc Natl Acad Sci U S A 94(7):3245.

Shimojo M., Paquette AJ., et al. 1999. Protein kinase A regulates cholinergic gene expression in PC12 cells: REST4 silences the silencing activity of neuron-restrictive silencer factor/REST. Mol Cell Biol 19(10):6788.

Shu T., Butz KG., et al. 2003. Abnormal development of forebrain midline glia and commissural projections in nfia knock-out mice. J Neurosci 23(1):203-12.

Steele-Perkins G., Plachez C., et al. 2005. The transcription factor gene nfib is essential for both lung maturation and brain development. Mol Cell Biol 25(2):685.

Steele-Perkins G., Butz KG., et al. 2003. Essential role for NFI-C/CTF transcription-replication factor in tooth root development. Mol Cell Biol 23(3):1075.

Stock DW., Ellies DL., et al. 1996. The evolution of the vertebrate dlx gene family. Proc Natl Acad Sci U S A 93(20):10858.

Stühmer T., Anderson SA., et al. 2002. Ectopic expression of the dlx genes induces glutamic acid decarboxylase and dlx expression. Development 129(1):245.

Tabuchi A., Yamada T., et al. 2002. REST4-mediated modulation of REST/NRSF-silencing function during BDNF gene promoter activation. Biochem Biophys Res Commun 290(1):415.

Tamagnone T., and Comoglio PM. 2000. Signalling by semaphorin receptors: Cell guidance and beyond. Trends Cell Biol 10(9):377.

Thomas BL., Liu JK., et al. 2000. Independent regulation of Dlx2 expression in the epithelium and mesenchyme of the first branchial arch. Development 127(2):217.

Thomas BL., Tucker AS., et al. 1997. Role of dlx-1 and dlx-2 genes in patterning of the murine dentition. Development 124(23):4811.

Wilson, C., Bellen, HJ., and Gehring, WJ. 1990. Position effects on eukaryotic gene expression. Annu. Rev. Cell Biol. (6)679-714.

Wonders CP., and Anderson SA. 2006. The origin and specification of cortical interneurons. Nat Rev Neurosci 7(9):687-96.

Woolfe A., Goodson M., et al. 2005. Highly conserved non-coding sequences are associated with vertebrate development. PLoS Biology 3(1):e7.

Wotton, D., Ghysdael, J., et al. 1994. Cooperative binding of Ets-1 and core binding factor to DNA. Mol. Cell. Biol. 14:840–850.

Xie Y., Skinner E., et al. 2002. Influence of the embryonic preplate on the organization of the cerebral cortex: A targeted ablation model. The Journal of Neuroscience 22(20):8981.

Zerucha T., and Ekker M. 2000a. Distal-less-related homeobox genes of vertebrates: Evolution, function, and regulation. Biochem. Cell Biol 78(5):593.

Zerucha T., Stühmer T., et al. 2000b. A highly conserved enhancer in the Dlx5/Dlx6 intergenic region is the site of cross-regulatory interactions between dlx genes in the embryonic forebrain. J. Neurosci 20(2):709.

Zhang C., Xuan Z., et al. 2006. A clustering property of highly-degenerate transcription factor binding sites in the mammalian genome. Nucleic Acids Res 34(8):2238.

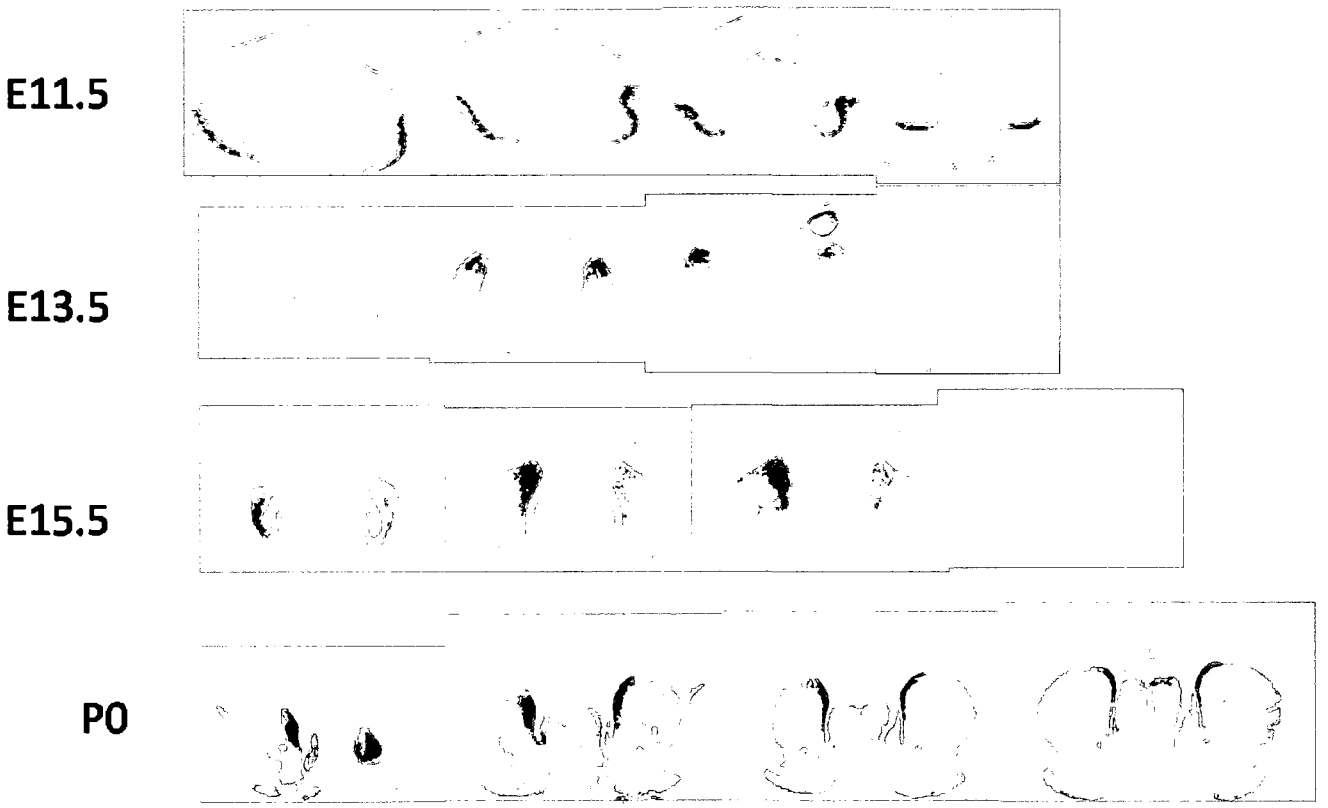
Zhang H., Hu G., et al. 1997. Heterodimerization of msx and dlx homeoproteins results in functional antagonism. Mol Cell Biol 17(5):2920.

Zhou QP., Le TN., et al. 2004. Identification of a direct dlx homeodomain target in the developing mouse forebrain and retina by optimization of chromatin

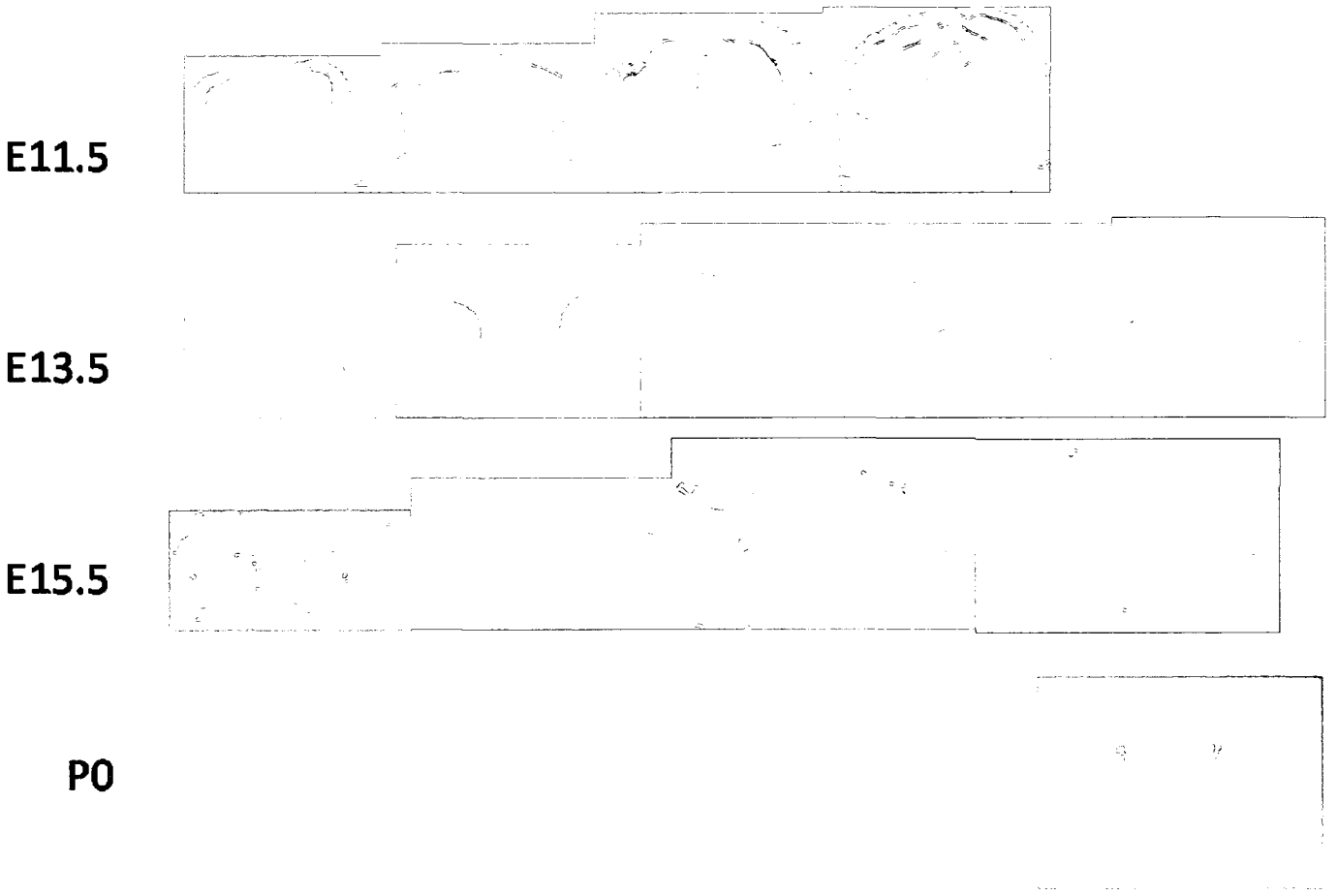
immunoprecipitation. Nucleic Acids Res 32(3):884.

APPENDICES

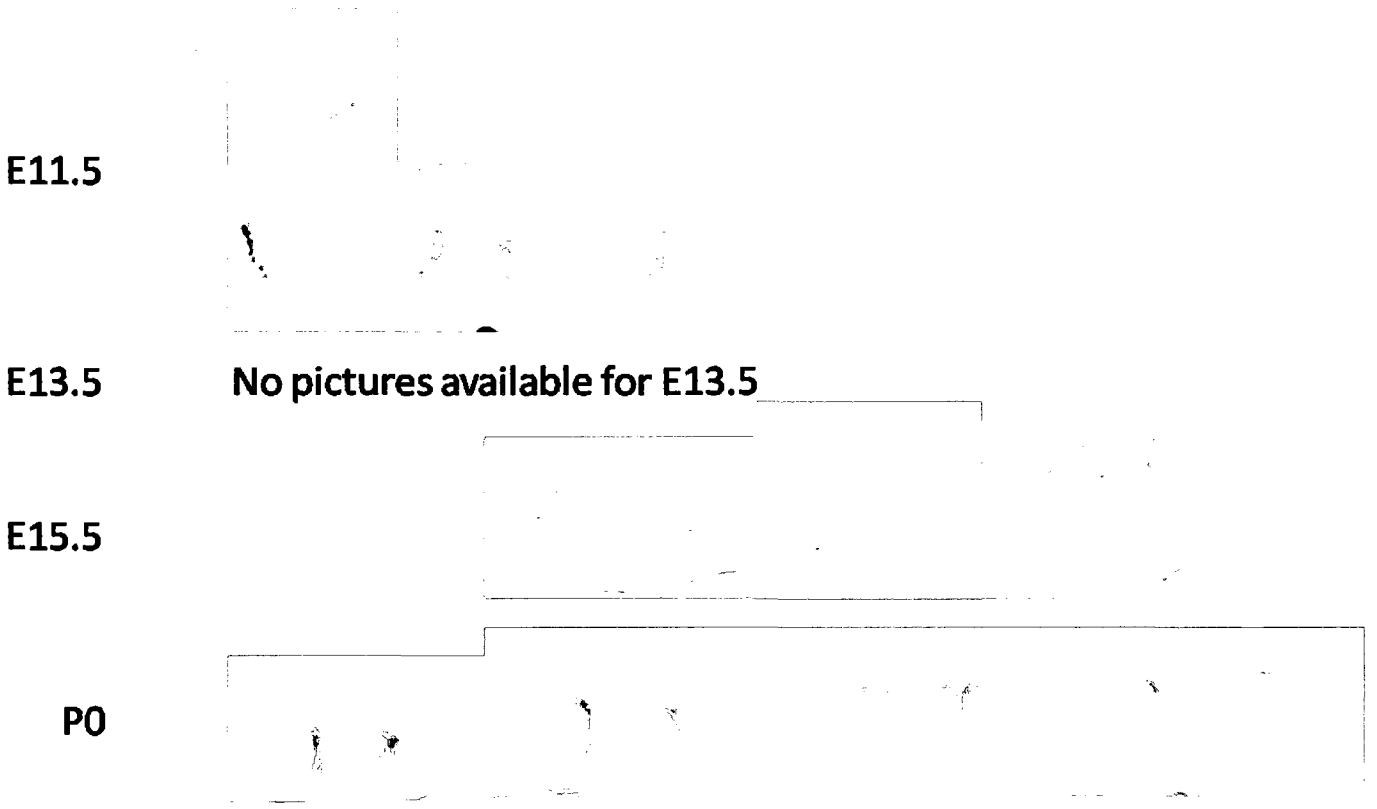
Appendix 1. Supplementary photos of coronal sections of forebrains in mURE2*lacZ* transgenic mice (Line 3300).



Appendix 2. Supplementary photos of coronal sections of forebrains in mURE2mut10a*lacZ* transgenic mice.



Appendix 3. Supplementary photos of coronal sections of forebrains in mURE2mut10c*lacZ* transgenic mice.



Appendix 4. Raw data and statistical results for test 1 (Figure 3.5) using REST cDNA and various constructs.

Date	Test	Cell type	Construct	cDNAr	Fold
30-Jul-09	test1	Hela	Concensus NRSE	no REST	0.151719928
30-Jul-09	test1	Hela	Concensus NRSE	REST	0.087496338
30-Jul-09	test1	N-2a	Concensus NRSE	no REST	0.717686596
30-Jul-09	test1	N-2a	Concensus NRSE	REST	1.150363355
30-Jul-09	test1	Hela	mURE2699-719	no REST	0.473724827
30-Jul-09	test1	Hela	mURE2699-719	REST	0.262735469
30-Jul-09	test1	N-2a	mURE2699-719	no REST	0.960898184
30-Jul-09	test1	N-2a	mURE2699-719	REST	1.130176529
30-Jul-09	test1	Hela	mURE2699-719mut10	no REST	0.332215019
30-Jul-09	test1	Hela	mURE2699-719mut10	REST	0.23034517
30-Jul-09	test1	N-2a	mURE2699-719mut10	no REST	0.725104989
30-Jul-09	test1	N-2a	mURE2699-719mut10	REST	1.17644181
31-Jul-09	test1	Hela	Concensus NRSE	no REST	0.275117852
31-Jul-09	test1	Hela	Concensus NRSE	REST	0.212300001
31-Jul-09	test1	Hela	mURE2699-719	no REST	1.397723894
31-Jul-09	test1	Hela	mURE2699-719	REST	0.587043028
31-Jul-09	test1	Hela	mURE2699-719mut10	no REST	1.380453671
31-Jul-09	test1	Hela	mURE2699-719mut10	REST	1.025528396
02-Jul-09	test1	Hela	Concensus NRSE	no REST	0.236571177
02-Jul-09	test1	Hela	Concensus NRSE	REST	0.21245932
02-Jul-09	test1	Hela	Concensus NRSE	no REST	0.456619943
02-Jul-09	test1	Hela	Concensus NRSE	REST	0.403428952
02-Jul-09	test1	N-2a	Concensus NRSE	no REST	1.841161573
02-Jul-09	test1	N-2a	Concensus NRSE	REST	1.310651259
02-Jul-09	test1	N-2a	Concensus NRSE	no REST	1.487633604
02-Jul-09	test1	N-2a	Concensus NRSE	REST	1.029775789
02-Jul-09	test1	Hela	mURE2699-719	no REST	0.639789288
02-Jul-09	test1	Hela	mURE2699-719	REST	0.685617893
02-Jul-09	test1	Hela	mURE2699-719	no REST	1.485020311
02-Jul-09	test1	Hela	mURE2699-719	REST	1.743859499
02-Jul-09	test1	Hela	mURE2699-719	no REST	2.1829435
02-Jul-09	test1	Hela	mURE2699-719	REST	1.357214166
02-Jul-09	test1	N-2a	mURE2699-719	no REST	1.653990739
02-Jul-09	test1	N-2a	mURE2699-719	REST	1.988582184
02-Jul-09	test1	N-2a	mURE2699-719	no REST	1.202692328
02-Jul-09	test1	N-2a	mURE2699-719	REST	1.586333074
02-Jul-09	test1	N-2a	mURE2699-719	no REST	1.561927196
02-Jul-09	test1	N-2a	mURE2699-719	REST	1.881826905
19-Aug-09	test1	Hela	Concensus NRSE	no REST	0.295508274
19-Aug-09	test1	Hela	Concensus NRSE	REST	0.18115942
19-Aug-09	test1	Hela	mURE2699-719	no REST	0.2596054
19-Aug-09	test1	Hela	mURE2699-719	REST	0.252278646
19-Aug-09	test1	Hela	mURE2699-719mut10	no REST	0.283370619
19-Aug-09	test1	Hela	mURE2699-719mut10	REST	0.311996779

Analysis of Variance Table (1 way ANOVA or Student t-test)

Response: Fold

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
cDNAr	1	0.0324	0.0324	0.084	0.7734
Residuals	42	16.1991	0.3857		

Appendix 5. Raw data and statistical results for co-transfections of 300ng NFI in HeLa.

Date	Test	Cell type	Construct	cDNAr	cDNAn	Amount	Fold
08-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	no	300	0.827586
08-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	NFIA	300	1.353888
08-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	NFIB	300	0.446707
08-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	NFIC	300	0.562771
08-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	NFIX	300	0.495238
08-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	NFIA + NFIB	300	0.808552
08-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	NFIA + NFIC	300	0.586665
08-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	NFIA + NFIX	300	0.551007
08-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	NFIB + NFIC	300	0.915892
08-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	NFIB + NFIX	300	0.047767
08-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	NFIC + NFIX	300	1.363699
17-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	no	300	1.073695
17-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	NFIA	300	0.674308
17-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	NFIB	300	0.482856
17-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	NFIC	300	0.712739
17-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	NFIX	300	0.528715
17-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	NFIA + NFIB	300	0.619262
17-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	NFIA + NFIC	300	0.690413
17-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	NFIA + NFIX	300	0.141048
17-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	NFIB + NFIC	300	0.562198
17-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	NFIB + NFIX	300	0.481256
17-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	NFIC + NFIX	300	0.598719
27-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	no	300	1.155857
27-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	NFIA	300	0.911645
27-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	NFIB	300	0.549788
27-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	NFIC	300	0.631443
27-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	NFIX	300	0.761066
27-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	NFIA + NFIB	300	0.676611
27-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	NFIA + NFIC	300	0.793455
27-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	NFIA + NFIX	300	0.437327
27-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	NFIB + NFIC	300	0.662982
27-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	NFIB + NFIX	300	0.362877
27-Jul-09	test2	HeLa	mURE2NRSE699-719	no REST	NFIC + NFIX	300	0.995116

```
> TukeyHSD(fit) # where fit comes from aov()
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = Fold ~ cDNAn, data = data, na.action = na.omit)
$cDNAn
```

	diff	lwr	upr	p adj
NFIA + NFIB-NFIA	-0.278472536	-0.88318646	0.326241393	0.8459232
NFIA + NFIC-NFIA	-0.289769757	-0.89448369	0.314944172	0.8139832
NFIA + NFIX-NFIA	-0.60348657	-1.2082005	0.001227359	0.0507648
NFIB-NFIA	-0.486830287	-1.09154422	0.117883642	0.1933031
NFIB + NFIC-NFIA	-0.266256593	-0.87097052	0.338457336	0.8769921
NFIB + NFIX-NFIA	-0.682646975	-1.2873609	-0.077933046	0.018483
NFIC-NFIA	-0.344296448	-0.94901038	0.260417481	0.629073
NFIC + NFIX-NFIA	0.005897279	-0.59881665	0.610611208	1
NFIX-NFIA	-0.384941033	-0.98965496	0.219772896	0.4811637
no-NFIA	0.039099059	-0.56561487	0.643812988	1
NFIA + NFIC-NFIA + NFIB	-0.011297221	-0.61601115	0.593416708	1
NFIA + NFIX-NFIA + NFIB	-0.325014034	-0.92972796	0.279699895	0.698582
NFIB-NFIA + NFIB	-0.208357751	-0.81307168	0.396356178	0.9711898
NFIB + NFIC-NFIA + NFIB	0.012215943	-0.59249799	0.616929872	1
NFIB + NFIX-NFIA + NFIB	-0.404174439	-1.00888837	0.20053949	0.4150843
NFIC-NFIA + NFIB	-0.065823912	-0.67053784	0.538890017	0.9999984
NFIC + NFIX-NFIA + NFIB	0.284369815	-0.32034411	0.889083744	0.8296154
NFIX-NFIA + NFIB	-0.106468497	-0.71118243	0.498245432	0.9998631
no-NFIA + NFIB	0.317571595	-0.28714233	0.922285524	0.7245244
NFIA + NFIX-NFIA + NFIC	-0.313716813	-0.91843074	0.290997116	0.7376883
NFIB-NFIA + NFIC	-0.197060531	-0.80177446	0.407653398	0.9802382
NFIB + NFIC-NFIA + NFIC	0.023513164	-0.58120076	0.628227093	1
NFIB + NFIX-NFIA + NFIC	-0.392877219	-0.99759115	0.21183671	0.4534113
NFIC-NFIA + NFIC	-0.054526691	-0.65924062	0.550187238	0.9999997
NFIC + NFIX-NFIA + NFIC	0.295667035	-0.30904689	0.900380964	0.7961935
NFIX-NFIA + NFIC	-0.095171277	-0.69988521	0.509542652	0.9999504
no-NFIA + NFIC	0.328868816	-0.27584511	0.933582745	0.6849111
NFIB-NFIA + NFIX	0.116656283	-0.48805765	0.721370212	0.9996928
NFIB + NFIC-NFIA + NFIX	0.337229977	-0.26748395	0.941943906	0.6548293
NFIB + NFIX-NFIA + NFIX	-0.079160405	-0.68387433	0.525553524	0.999991
NFIC-NFIA + NFIX	0.259190122	-0.34552381	0.863904051	0.8932068
NFIC + NFIX-NFIA + NFIX	0.609383849	0.00466992	1.214097778	0.0471862
NFIX-NFIA + NFIX	0.218545537	-0.38616839	0.823259466	0.9607236
no-NFIA + NFIX	0.642585629	0.0378717	1.247299558	0.0310449

NFIB + NFIC-NFIB	0.220573695	-0.38414023	0.825287624	0.9583562
NFIB + NFIX-NFIB	-0.195816688	-0.80053062	0.408897241	0.9810852
NFIC-NFIB	0.142533839	-0.46218009	0.747247768	0.9983159
NFIC + NFIX-NFIB	0.492727566	-0.11198636	1.097441495	0.181791
NFIX-NFIB	0.101889254	-0.50282467	0.706603183	0.9999078
no-NFIB	0.525929346	-0.07878458	1.130643275	0.126868
NFIB + NFIX-NFIB + NFIC	-0.416390383	-1.02110431	0.188323546	0.3754785
NFIC-NFIB + NFIC	-0.078039855	-0.68275378	0.526674074	0.9999921
NFIC + NFIX-NFIB + NFIC	0.272153871	-0.33256006	0.8768678	0.8624635
NFIX-NFIB + NFIC	-0.118684441	-0.72339837	0.486029488	0.9996429
no-NFIB + NFIC	0.305355652	-0.29935828	0.910069581	0.7654856
NFIC-NFIB + NFIX	0.338350527	-0.2663634	0.943064456	0.6507616
NFIC + NFIX-NFIB + NFIX	0.688544254	0.08383033	1.293258183	0.0171071
NFIX-NFIB + NFIX	0.297705942	-0.30700799	0.902419871	0.7898782
no-NFIB + NFIX	0.721746034	0.11703211	1.326459963	0.0110238
NFIC + NFIX-NFIC	0.350193727	-0.2545202	0.954907656	0.6074413
NFIX-NFIC	-0.040644585	-0.64535851	0.564069344	1
no-NFIC	0.383395507	-0.22131842	0.988109436	0.4866368
NFIX-NFIC + NFIX	-0.390838312	-0.99555224	0.213875617	0.4604816
no-NFIC + NFIX	0.03320178	-0.57151215	0.637915709	1
no-NFIX	0.424040092	-0.18067384	1.028754021	0.351768

Appendix 6. Raw data and statistical results for co-transfections of 300ng NFI in Neuro-2a.

Date	Test	Cell type	Construct	cDNAr	cDNAn	Amount	Fold
08-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	no	300	1.448162
08-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	NFIA	300	2.111156
08-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	NFIB	300	3.772398
08-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	NFIC	300	1.619716
08-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	NFIX	300	4.777013
08-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	NFIA + NFIB	300	4.763421
08-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	NFIA + NFIC	300	2.157052
08-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	NFIA + NFIX	300	4.334569
08-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	NFIB + NFIC	300	5.68852
08-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	NFIB + NFIX	300	4.54736
08-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	NFIC + NFIX	300	3.45839
09-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	no	300	0.494276
09-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	NFIA	300	1.223718
09-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	NFIB	300	3.226504
09-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	NFIC	300	0.40419
09-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	NFIX	300	1.362076
09-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	NFIA + NFIB	300	2.094818
09-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	NFIA + NFIC	300	0.660084
09-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	NFIA + NFIX	300	1.332741
09-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	NFIB + NFIC	300	2.249937
09-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	NFIB + NFIX	300	1.879477
09-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	NFIC + NFIX	300	1.342065
17-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	no	300	0.469995
17-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	NFIA	300	1.118428
17-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	NFIB	300	2.763061
17-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	NFIC	300	0.357692
17-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	NFIX	300	1.157815
17-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	NFIA + NFIB	300	1.868407
17-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	NFIA + NFIC	300	0.617195
17-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	NFIA + NFIX	300	1.082444
17-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	NFIB + NFIC	300	3.697781
17-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	NFIB + NFIX	300	7.391573
17-Jul-09	test2	N-2a	mURE2NRSE699-719	no REST	NFIC + NFIX	300	8.877412

```
> TukeyHSD(fit) # where fit comes from aov()
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = Fold ~ cDNAn, data = data, na.action = na.omit)
$cDNAn
```

	diff	lwr	upr	p adj
NFIA + NFIB-NFIA	1.42444782	-3.988383	6.837278	0.9959243
NFIA + NFIC-NFIA	-0.3396571	-5.752488	5.073173	1
NFIA + NFIX-NFIA	0.7654839	-4.647347	6.178314	0.9999815
NFIB-NFIA	1.76955365	-3.643277	7.182384	0.9797974
NFIB + NFIC-NFIA	2.39431201	-3.018519	7.807143	0.8740294
NFIB + NFIX-NFIA	3.1217026	-2.291128	8.534533	0.6127358
NFIC-NFIA	-0.69056786	-6.103398	4.722263	0.999993
NFIC + NFIX-NFIA	3.07485475	-2.337976	8.487685	0.6319191
NFIX-NFIA	0.94786744	-4.464963	6.360698	0.9998696
no-NFIA	-0.68028976	-6.09312	4.732541	0.9999939
NFIA + NFIC-NFIA + NFIB	-1.76410492	-7.176935	3.648726	0.9802224
NFIA + NFIX-NFIA + NFIB	-0.65896392	-6.071794	4.753867	0.9999955
NFIB-NFIA + NFIB	0.34510583	-5.067725	5.757936	1
NFIB + NFIC-NFIA + NFIB	0.96986419	-4.442966	6.382695	0.99984
NFIB + NFIX-NFIA + NFIB	1.69725478	-3.715576	7.110085	0.9849062
NFIC-NFIA + NFIB	-2.11501568	-7.527846	3.297815	0.9365719
NFIC + NFIX-NFIA + NFIB	1.65040693	-3.762424	7.063237	0.9876451
NFIX-NFIA + NFIB	-0.47658038	-5.889411	4.93625	0.9999998
no-NFIA + NFIB	-2.10473758	-7.517568	3.308093	0.9383775
NFIA + NFIX-NFIA + NFIC	1.10514099	-4.30769	6.517972	0.9994972
NFIB-NFIA + NFIC	2.10921074	-3.30362	7.522041	0.9375959
NFIB + NFIC-NFIA + NFIC	2.73396911	-2.678861	8.1468	0.7652263
NFIB + NFIX-NFIA + NFIC	3.4613597	-1.951471	8.87419	0.4749664
NFIC-NFIA + NFIC	-0.35091076	-5.763741	5.06192	1
NFIC + NFIX-NFIA + NFIC	3.41451184	-1.998319	8.827342	0.4935032
NFIX-NFIA + NFIC	1.28752454	-4.125306	6.700355	0.9981852
no-NFIA + NFIC	-0.34063266	-5.753463	5.072198	1
NFIB-NFIA + NFIX	1.00406975	-4.408761	6.4169	0.9997823
NFIB + NFIC-NFIA + NFIX	1.62882812	-3.784002	7.041659	0.9887692
NFIB + NFIX-NFIA + NFIX	2.35621871	-3.056612	7.769049	0.8840902
NFIC-NFIA + NFIX	-1.45605176	-6.868882	3.956779	0.9951651
NFIC + NFIX-NFIA + NFIX	2.30937085	-3.10346	7.722201	0.8958079
NFIX-NFIA + NFIX	0.18238354	-5.230447	5.595214	1
no-NFIA + NFIX	-1.44577365	-6.858604	3.967057	0.9954236
NFIB + NFIC-NFIB	0.62475837	-4.788072	6.037589	0.9999973
NFIB + NFIX-NFIB	1.35214896	-4.060682	6.764979	0.9973022

NFIC-NFIB	-2.46012151	-7.872952	2.952709	0.8555493
NFIC + NFIX-NFIB	1.3053011	-4.107529	6.718132	0.9979705
NFIX-NFIB	-0.82168621	-6.234517	4.591144	0.9999644
no-NFIB	-2.4498434	-7.862674	2.962987	0.8585256
NFIB + NFIX-NFIB + NFIC	0.72739059	-4.68544	6.140221	0.9999885
NFIC-NFIB + NFIC	-3.08487987	-8.49771	2.327951	0.6278209
NFIC + NFIX-NFIB + NFIC	0.68054273	-4.732288	6.093373	0.9999939
NFIX-NFIB + NFIC	-1.44644457	-6.859275	3.966386	0.995407
no-NFIB + NFIC	-3.07460177	-8.487432	2.338229	0.6320225
NFIC-NFIB + NFIX	-3.81227046	-9.225101	1.60056	0.346133
NFIC + NFIX-NFIB + NFIX	-0.04684785	-5.459678	5.365983	1
NFIX-NFIB + NFIX	-2.17383516	-7.586666	3.238995	0.9255691
no-NFIB + NFIX	-3.80199236	-9.214823	1.610838	0.3496026
NFIC + NFIX-NFIC	3.76542261	-1.647408	9.178253	0.362112
NFIX-NFIC	1.6384353	-3.774395	7.051266	0.988279
no-NFIC	0.0102781	-5.402552	5.423109	1
NFIX-NFIC + NFIX	-2.12698731	-7.539818	3.285843	0.9344252
no-NFIC + NFIX	-3.7551445	-9.167975	1.657686	0.3656734
no-NFIX	-1.6281572	-7.040988	3.784673	0.9888028

Appendix 7. Raw data and statistical results for co-transfections of 200ng NFI in HeLa.

Date	Test	Cell type	Construct	cDNAr	cDNAn	Amount	Fold
10-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	no	200	0.771765
10-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	no	200	0.485867
10-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	NFIA	200	0.326805
10-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	NFIA	200	0.416288
10-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	NFIB	200	0.407149
10-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	NFIB	200	0.428571
10-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	NFIC	200	0.445021
10-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	NFIC	200	0.365189
10-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	NFIX	200	0.297756
10-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	NFIX	200	0.41319
10-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	NFIA+NFIB	200	0.422906
10-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	NFIA+NFIB	200	0.334529
10-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	NFIA+NFIC	200	0.589514
10-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	NFIA+NFIC	200	0.388784
10-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	NFIA+NFIX	200	0.294095
10-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	NFIA+NFIX	200	0.319435
10-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	NFIB+NFIC	200	0.380537
10-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	NFIB+NFIC	200	0.346289
10-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	no	200	0.716309
10-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	NFIB+NFIX	200	0.692056
10-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	NFIB+NFIX	200	0.601321
10-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	NFIC+NFIX	200	0.566547
10-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	NFIC+NFIX	200	0.79153
31-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	no	200	0.84538
31-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	no	200	0.855672
31-Aug-	test3	HeLa	mURE2NRSE699-719	no REST	NFIA	200	0.477465

09							
31-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	NFIA	200	0.526422
31-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	NFIB	200	0.428779
31-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	NFIB	200	0.746109
31-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	NFIC	200	0.521878
31-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	NFIC	200	0.528431
31-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	NFIX	200	0.416754
31-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	NFIX	200	0.428984
31-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	NFIA+NFIB	200	0.387449
31-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	NFIA+NFIB	200	0.452875
31-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	NFIA+NFIC	200	0.633934
31-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	NFIA+NFIC	200	0.623425
31-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	NFIA+NFIX	200	0.391667
31-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	NFIA+NFIX	200	0.410246
31-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	NFIB+NFIC	200	0.346552
31-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	NFIB+NFIC	200	0.461553
31-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	NFIB+NFIX	200	0.367801
31-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	NFIB+NFIX	200	0.33086
31-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	NFIC+NFIX	200	0.4014
31-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	NFIC+NFIX	200	0.632148
31-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	no	200	0.907884
31-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	no	200	1.344368
31-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	NFIA	200	0.506521
31-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	NFIA	200	0.520164
31-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	NFIB	200	1.006477
31-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	NFIB	200	1.47764
31-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	NFIC	200	1.433406
31-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	NFIC	200	0.873382

31-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	NFIX	200	0.739868
31-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	NFIX	200	0.745679
31-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	NFIA+NFIB	200	0.782723
31-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	NFIA+NFIB	200	1.231263
31-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	NFIA+NFIC	200	1.258087
31-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	NFIA+NFIC	200	1.70634
31-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	NFIA+NFIX	200	1.392014
31-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	NFIA+NFIX	200	1.222822
31-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	NFIB+NFIC	200	1.379692
31-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	NFIB+NFIC	200	1.732393
31-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	NFIB+NFIX	200	1.267438
31-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	NFIB+NFIX	200	1.450679
31-Aug-09	test3	HeLa	mURE2NRSE699-719	no REST	NFIC+NFIX	200	1.581157
31-Aug-09	test3	HeLa	mURE2NRSE699-719	REST	NFIC+NFIX	200	2.66829

```
> TukeyHSD(fit) # where fit comes from aov()
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = Fold ~ cDNAr + cDNAn + cDNAr:cDNAn, data = data, na.action = na.omit)
$cDNAn
```

	diff	lwr	upr	p adj
NFIA+NFIB-NFIA	0.13967982	-0.8522296	1.1315893	0.9999923
NFIA+NFIC-NFIA	0.40440303	-0.5875064	1.3963125	0.9466566
NFIA+NFIX-NFIA	0.20943561	-0.7824738	1.2013451	0.9996813
NFIB-NFIA	0.28684334	-0.7050661	1.2787528	0.9954838
NFIB+NFIC-NFIA	0.31222499	-0.6796845	1.3041344	0.991268
NFIB+NFIX-NFIA	0.3227482	-0.6691613	1.3146576	0.98879
NFIC-NFIA	0.23227364	-0.7596358	1.2241831	0.999214
NFIC+NFIX-NFIA	0.64456769	-0.3473418	1.6364771	0.5150913
NFIX-NFIA	0.04476079	-0.9471487	1.0366702	1
no-NFIA	0.39131952	-0.5645083	1.3471474	0.945238
NFIA+NFIC-NFIA+NFIB	0.26472321	-0.7271862	1.2566327	0.9976386
NFIA+NFIX-NFIA+NFIB	0.06975579	-0.9221537	1.0616652	1
NFIB-NFIA+NFIB	0.14716352	-0.8447459	1.139073	0.9999874

NFIB+NFIC-NFIA+NFIB	0.17254518	-0.8193643	1.1644546	0.9999446
NFIB+NFIX-NFIA+NFIB	0.18306838	-0.8088411	1.1749778	0.9999048
NFIC-NFIA+NFIB	0.09259383	-0.8993156	1.0845033	0.9999999
NFIC+NFIX-NFIA+NFIB	0.50488787	-0.4870216	1.4967973	0.8148137
NFIX-NFIA+NFIB	-0.09491902	-1.0868285	0.8969904	0.9999998
no-NFIA+NFIB	0.25163971	-0.7041882	1.2074676	0.9978894
NFIA+NFIX-NFIA+NFIC	-0.19496742	-1.1868769	0.796942	0.9998317
NFIB-NFIA+NFIC	-0.11755969	-1.1094691	0.8743498	0.9999985
NFIB+NFIC-NFIA+NFIC	-0.09217803	-1.0840875	0.8997314	0.9999999
NFIB+NFIX-NFIA+NFIC	-0.08165483	-1.0735643	0.9102546	1
NFIC-NFIA+NFIC	-0.17212938	-1.1640388	0.8197801	0.9999458
NFIC+NFIX-NFIA+NFIC	0.24016466	-0.7517448	1.2320741	0.9989541
NFIX-NFIA+NFIC	-0.35964223	-1.3515517	0.6322672	0.9755211
no-NFIA+NFIC	-0.0130835	-0.9689114	0.9427444	1
NFIB-NFIA+NFIX	0.07740773	-0.9145017	1.0693172	1
NFIB+NFIC-NFIA+NFIX	0.10278938	-0.8891201	1.0946988	0.9999996
NFIB+NFIX-NFIA+NFIX	0.11331259	-0.8785969	1.105222	0.999999
NFIC-NFIA+NFIX	0.02283803	-0.9690714	1.0147475	1
NFIC+NFIX-NFIA+NFIX	0.43513208	-0.5567774	1.4270415	0.9166665
NFIX-NFIA+NFIX	-0.16467482	-1.1565843	0.8272346	0.999964
no-NFIA+NFIX	0.18188391	-0.773944	1.1377118	0.9998743
NFIB+NFIC-NFIB	0.02538166	-0.9665278	1.0172911	1
NFIB+NFIX-NFIB	0.03590486	-0.9560046	1.0278143	1
NFIC-NFIB	-0.05456969	-1.0464791	0.9373398	1
NFIC+NFIX-NFIB	0.35772435	-0.6341851	1.3496338	0.9764162
NFIX-NFIB	-0.24208254	-1.233992	0.7498269	0.998881
no-NFIB	0.10447619	-0.8513517	1.060304	0.9999993
NFIB+NFIX-NFIB+NFIC	0.0105232	-0.9813862	1.0024327	1
NFIC-NFIB+NFIC	-0.07995135	-1.0718608	0.9119581	1
NFIC+NFIX-NFIB+NFIC	0.3323427	-0.6595668	1.3242521	0.9860764
NFIX-NFIB+NFIC	-0.2674642	-1.2593736	0.7244452	0.9974305
no-NFIB+NFIC	0.07909453	-0.8767333	1.0349224	1
NFIC-NFIB+NFIX	-0.09047456	-1.082384	0.9014349	0.9999999
NFIC+NFIX-NFIB+NFIX	0.32181949	-0.67009	1.3137289	0.9890288
NFIX-NFIB+NFIX	-0.27798741	-1.2698969	0.713922	0.9964854
no-NFIB+NFIX	0.06857132	-0.8872565	1.0243992	1
NFIC+NFIX-NFIC	0.41229405	-0.5796154	1.4042035	0.9398031
NFIX-NFIC	-0.18751285	-1.1794223	0.8043966	0.9998816
no-NFIC	0.15904588	-0.796782	1.1148737	0.9999632
NFIX-NFIC+NFIX	-0.5998069	-1.5917163	0.3921026	0.617167
no-NFIC+NFIX	-0.25324817	-1.209076	0.7025797	0.9977756
no-NFIX	0.34655873	-0.6092691	1.3023866	0.9755217

Appendix 8. Raw data and statistical results for co-transfections of 200ng NFI in Neuro-2a.

Date	Test	Cell type	Construct	cDNAr	cDNAn	Amount	Fold
06-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	no	200	0.598519
06-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	no	200	0.481248
06-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIA	200	0.473285
06-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIA	200	0.732887
06-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIB	200	1.11029
06-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIB	200	1.329956
06-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIC	200	0.407221
06-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIC	200	0.437061
06-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIX	200	0.569042
06-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIX	200	0.848709
06-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIA+NFIB	200	1.535163
06-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIA+NFIC	200	0.674839
06-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIA+NFIX	200	0.97499
06-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIA+NFIX	200	1.143854
06-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIB+NFIC	200	0.858836
06-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIB+NFIC	200	1.235098
06-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIB+NFIX	200	1.027577
06-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIB+NFIX	200	1.532269
06-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIC+NFIX	200	0.950056
06-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIC+NFIX	200	0.800408
18-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	no	200	0.978423
18-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	no	200	0.632376
18-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIA	200	1.447293
18-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIA	200	0.690446
18-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIB	200	7.81676

18-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIB	200	2.204266
18-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIC	200	0.670339
18-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIC	200	0.74663
18-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIX	200	1.422441
18-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIX	200	1.261708
18-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIA+NFIB	200	2.603695
18-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIA+NFIB	200	5.323126
18-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIA+NFIC	200	1.116769
18-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIA+NFIX	200	2.22628
18-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIA+NFIX	200	2.044749
18-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIB+NFIC	200	1.656491
18-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIB+NFIC	200	2.887512
18-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIB+NFIX	200	3.119222
18-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIB+NFIX	200	2.753183
18-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIC+NFIX	200	1.439951
18-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIC+NFIX	200	1.992178
18-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	no	200	2.057194
18-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIA+NFIB	200	3.159758
18-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIA+NFIB	200	2.951068
18-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIA+NFIC	200	0.83077
18-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIA+NFIC	200	1.408029
20-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	no	200	0.723685
20-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	no	200	0.745672
20-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIA	200	1.454389
20-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIA	200	1.47483
20-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIB	200	2.102384
20-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIB	200	4.081409
20-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIC	200	0.76484

20-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIC	200	1.012242
20-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIX	200	1.77744
20-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIX	200	1.957236
20-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIA+NFIB	200	2.500911
20-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIA+NFIB	200	2.87758
20-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIA+NFIC	200	0.668757
20-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIA+NFIC	200	0.819219
20-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIA+NFIX	200	2.318777
20-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIA+NFIX	200	1.398213
20-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIB+NFIC	200	1.544562
20-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIB+NFIC	200	2.150633
20-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIB+NFIX	200	1.37876
20-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIB+NFIX	200	1.636087
20-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIC+NFIX	200	0.924925
20-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIC+NFIX	200	1.074491
26-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	no	200	0.625199
26-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	no	200	0.888045
26-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIA	200	0.835352
26-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIA	200	1.117348
26-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIB	200	1.948672
26-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIB	200	2.698162
26-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIC	200	0.620986
26-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIC	200	0.821789
26-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIX	200	1.342356
26-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIX	200	1.35465
26-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIA+NFIB	200	2.138965
26-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIA+NFIB	200	3.629602
26-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIA+NFIC	200	1.028512

26-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIA+NFIC	200	1.565074
26-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIA+NFIX	200	1.541778
26-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIA+NFIX	200	2.007271
26-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIB+NFIC	200	2.002465
26-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIB+NFIC	200	2.435901
26-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIB+NFIX	200	1.900275
26-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIB+NFIX	200	2.691161
26-Aug-09	test3	N-2a	mURE2NRSE699-719	no REST	NFIC+NFIX	200	3.497155
26-Aug-09	test3	N-2a	mURE2NRSE699-719	REST	NFIC+NFIX	200	4.263846

> TukeyHSD(fit) # where fit comes from aov()

Tukey multiple comparisons of means

95% family-wise confidence level

Fit: aov(formula = Fold ~ cDNAr + cDNAn + cDNAr:cDNAn, data = data, na.action = na.omit)

$cDNAn

	diff	lwr	upr	p adj
NFIA+NFIB-NFIA	1.949839	0.399502	3.500176	0.003728
NFIA+NFIC-NFIA	-0.01423	-1.60952	1.581051	1
NFIA+NFIX-NFIA	0.67876	-0.91652	2.274044	0.939815
NFIB-NFIA	1.883259	0.287975	3.478543	0.008574
NFIB+NFIC-NFIA	0.818209	-0.77708	2.413492	0.826087
NFIB+NFIX-NFIA	0.976588	-0.6187	2.571872	0.623187
NFIC-NFIA	-0.34309	-1.93837	1.252193	0.99971
NFIC+NFIX-NFIA	0.839648	-0.75564	2.434931	0.802369
NFIX-NFIA	0.288469	-1.30681	1.883753	0.99994
no-NFIA	-0.16011	-1.71044	1.390231	1
NFIA+NFIC-NFIA+NFIB	-1.96407	-3.51441	-0.41373	0.003367
NFIA+NFIX-NFIA+NFIB	-1.27108	-2.82142	0.279259	0.208429
NFIB-NFIA+NFIB	-0.06658	-1.61692	1.483757	1
NFIB+NFIC-NFIA+NFIB	-1.13163	-2.68197	0.418707	0.36318
NFIB+NFIX-NFIA+NFIB	-0.97325	-2.52359	0.577086	0.587667
NFIC-NFIA+NFIB	-2.29293	-3.84327	-0.74259	0.000281
NFIC+NFIX-NFIA+NFIB	-1.11019	-2.66053	0.440146	0.391377
NFIX-NFIA+NFIB	-1.66137	-3.21171	-0.11103	0.025709

no-NFIA+NFIB	-2.10995	-3.61399	-0.6059	0.000703
NFIA+NFIX-NFIA+NFIC	0.692993	-0.90229	2.288277	0.931525
NFIB-NFIA+NFIC	1.897491	0.302208	3.492775	0.007804
NFIB+NFIC-NFIA+NFIC	0.832441	-0.76284	2.427725	0.810509
NFIB+NFIX-NFIA+NFIC	0.99082	-0.60446	2.586104	0.602905
NFIC-NFIA+NFIC	-0.32886	-1.92414	1.266426	0.999801
NFIC+NFIX-NFIA+NFIC	0.85388	-0.7414	2.449164	0.785815
NFIX-NFIA+NFIC	0.302702	-1.29258	1.897985	0.999906
no-NFIA+NFIC	-0.14587	-1.69621	1.404463	1
NFIB-NFIA+NFIX	1.204499	-0.39079	2.799782	0.314728
NFIB+NFIC-NFIA+NFIX	0.139448	-1.45584	1.734732	1
NFIB+NFIX-NFIA+NFIX	0.297828	-1.29746	1.893111	0.999919
NFIC-NFIA+NFIX	-1.02185	-2.61713	0.573433	0.558426
NFIC+NFIX-NFIA+NFIX	0.160887	-1.4344	1.756171	1
NFIX-NFIA+NFIX	-0.39029	-1.98557	1.204993	0.999102
no-NFIA+NFIX	-0.83887	-2.3892	0.71147	0.774664
NFIB+NFIC-NFIB	-1.06505	-2.66033	0.530234	0.496806
NFIB+NFIX-NFIB	-0.90667	-2.50195	0.688613	0.719518
NFIC-NFIB	-2.22635	-3.82163	-0.63107	0.000767
NFIC+NFIX-NFIB	-1.04361	-2.6389	0.551673	0.527267
NFIX-NFIB	-1.59479	-3.19007	0.000494	0.050139
no-NFIB	-2.04337	-3.5937	-0.49303	0.001892
NFIB+NFIX-NFIB+NFIC	0.158379	-1.4369	1.753663	1
NFIC-NFIB+NFIC	-1.1613	-2.75658	0.433985	0.36713
NFIC+NFIX-NFIB+NFIC	0.021439	-1.57385	1.616723	1
NFIX-NFIB+NFIC	-0.52974	-2.12502	1.065544	0.989267
no-NFIB+NFIC	-0.97832	-2.52865	0.572022	0.580196
NFIC-NFIB+NFIX	-1.31968	-2.91496	0.275606	0.198161
NFIC+NFIX-NFIB+NFIX	-0.13694	-1.73222	1.458343	1
NFIX-NFIB+NFIX	-0.68812	-2.2834	0.907165	0.934446
no-NFIB+NFIX	-1.13669	-2.68703	0.413643	0.35667
NFIC+NFIX-NFIC	1.182738	-0.41255	2.778022	0.340584
NFIX-NFIC	0.631559	-0.96372	2.226843	0.962308
no-NFIC	0.182984	-1.36735	1.733321	0.999999
NFIX-NFIC+NFIX	-0.55118	-2.14646	1.044105	0.985564
no-NFIC+NFIX	-0.99975	-2.55009	0.550583	0.548549
no-NFIX	-0.44858	-1.99891	1.101762	0.996351

Appendix 9. Statistical analysis of effects of variables on results for all tests

Analysis of Variance Table

Response: Fold

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
type	1	79.942	79.942	95.6081	< 2.2e-16	***
cDNAn	16	64.753	4.047	4.8401	2.94E-08	***
amount	1	1.366	1.366	1.6331	0.2028	
type:cDNAn	16	64.567	4.035	4.8263	3.14E-08	***
type:amount	1	1.224	1.224	1.4636	0.2279	
cDNAn:amount	4	2.064	0.516	0.6171	0.6508	
Residuals	191	159.704	0.836			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1