A SNP associated with autism affects *Dlx5/Dlx6* regulation in the forebrain

Cindy Lesage-Pelletier

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

Statement of contribution

In chapter 3, «A SNP in an ultraconserved regulatory element affects *Dlx5/Dlx6* regulation in the forebrain», I carried out the following experiments. I cloned the I56i enhancer (wild type and SNP) in *luciferase* reporter plasmids. I carried out the Cotransfection and luciferase assays. I also performed electromobility shift assays experiments shown in Figure 3.6.

In chapter 4, « Phenotype analysis of mice carrying a variant I56 *Dlx* enhancer », I participated to all the phenotype characterization experiments but not in the production of the mice with a targeted enhancer replacement.

Abstract

Autism is a severe childhood neuropsychiatric condition characterized by impairments in socialization and communication, and by restricted and repetitive behaviours. Autism spectrum disorder (ASD) is a complex and largely unknown disease with a strong genetic basis, multiple genes involved and environmental factors determining its phenotype. Interestingly, the *DLX1/DLX2* and *DLX5/DLX6* bigene clusters are located in autism susceptibility loci and Dlx genes are involved in GABAergic interneurons differentiation and migration to the cortex during forebrain development. *Dlx* gene expression is controlled by different *cis*-regulatory elements. Of these, 4 are active in the forebrain, URE2, I12b, I56ii and I56i. In order to determine the role of the *DLX* genes in ASD, variants were found in gene exons and in *cis*-regulatory elements in autistic individuals. A single nucleotide polymorphism (SNP), a change of an adenine for a guanine, was identified in I56i enhancer. Finding a SNP in I56i was very surprising considering that it is located in a Dlx binding motif highly conserved among >40 species. We showed, using *in vitro* approaches, that the presence of this SNP affects the affinity of Dlx for their binding site and reduces the transcriptional activation of the enhancer. The SNP also affects activity of the I56i enhancer in transgenic mice. In order to determine the real impact of the SNP *in vivo*, mutant mice harboring the SNP in their I56i enhancer were produced. That involved the insertion of the I56i enhancer with the SNP, using homologous recombination in mouse embryonic stem cells to replace the wild type version of the enhancer. With these mutant mice, we demonstrated that, *in vivo,* this SNP reduces *Dlx5* and *Dlx6* expression in the forebrain. Furthermore, this decrease in *Dlx5/Dlx6* expression could affect the differentiation and/or migration of specific populations of inhibitory interneurons in the forebrain. No distinct

behavioural phenotypes were observed between wild type mice and those carrying the SNP, during social interaction and anxiety tests. Therefore, these results suggest that even a subtle change in a regulatory element can have an impact in the development of the forebrain and may even contribute to disorders such as autism.

Résumé

L'autisme est un trouble neuropsychiatrique infantile grave caractérisée par des déficiences dans la socialisation et la communication, avec des comportements limités et répétitifs. Les troubles du spectre autistique (TSA) sont complexes et encore inconnus avec une forte base héréditaire, plusieurs gènes impliqués et possiblement influencé par des facteurs environnementaux déterminant le phénotype. Il est intéressant de remarquer que deux paires de gènes *DLX1/DLX2* et *DLX5/DLX6* sont situés dans des loci de susceptibilité à l'autisme et sont impliqués dans la différenciation et la migration vers le cortex des interneurones GABAergiques, pendant le développement du cerveau antérieur. L'expression des gènes Dlx est contrôlée par différents éléments régulateurs, de ce nombre, quatre sont actifs dans le cerveau antérieur, soit URE2, I12b, I56ii et I56i. Afin de déterminer le rôle des gènes *DLX* dans les TSA, des polymorphismes ont été trouvés chez des patients autistes dans les exons des gènes et dans les éléments régulateurs. Un polymorphisme d'un seul nucléotide (SNP), un changement d'une adénine pour une guanine, a été identifié dans l'élément régulateur I56i. Trouver un SNP dans I56i était très surprenant puisqu'il se trouve dans un site de liaison de Dlx hautement conservé chez plus de 40 espèces. Nous avons montré, grâce à des expériences *in vitro*, que la présence de ce SNP affecte l'affinité des Dlx pour leur site de liaison et réduit l'activation transcriptionnelle de l'élément régulateur. Le SNP affecte l'activité de l'élément régulateur I56i dans les souris transgéniques. Afin de déterminer l'impact réel du SNP *in vivo*, des souris mutantes ayant le SNP dans leur élément régulateur I56i ont été produites. Cela implique l'insertion de l'élément régulateur I56i avec le SNP, en utilisant la recombinaison homologue, dans les cellules souches embryonnaires de souris afin de remplacer la version sauvage de I56i. Chez ces souris mutantes, nous avons démontré, *in*

vivo, que le SNP réduit l'expression de *Dlx5* et *Dlx6* dans le cerveau antérieur. Par ailleurs, cette diminution de l'expression de *Dlx5/Dlx6* pourrait affecter la différenciation et / ou la migration de populations spécifiques d'interneurones inhibiteurs dans le cerveau antérieur. Aucune différence comportementale n'a été observée entre les souris de type sauvage et celles qui ont le SNP, lors de tests d'interactions sociales et d'anxiété. Ces résultats suggèrent que même un changement subtil dans un élément régulateur peut avoir un impact dans le développement du cerveau antérieur et peut même contribuer à des troubles comme l'autisme.

Table of contents

List of figures

Abbreviations and acronyms

ASD: autism spectrum disorders

bp: base pairs

CGE: caudal ganglionic eminence

Dlx: distal-less homeobox

E11:5: embryonic development day 11.5 after conception

E13.5: embryonic development day 13.5 after conception

EMSA: electromobility shift assay

FP: area in DNAseI footprint protected by protein(s), numbered for easy reference

GABA: γ-aminobutyric acid

Gtf2i: general transcription factor 2I

LGE: lateral ganglionic eminence

MGE: medial ganglionic eminence

mI56i: mouse I56i enhancer

MZ: mantle zone

P0: postnatal day 0 (birth)

PDD: Pervasive Developmental Disorders

nt: nucleotide(s)

qPCR: real-time PCR

SNP: single nucleotide polymorphism

SVZ: subventricular zone

vI56i: variant form (with the SNP) of the I56i enhancer

VZ: ventricular zone

WBS: William-Beuren syndrom

Acknowledgements

I would first like to thank Dr. Marc Ekker for his trust, his guidance and his helpful comments on ma thesis. He has provided me with the opportunity to learn in his lab and conduct a project in which I felt involved from the very beginning.

My special thanks go to Dr. Luc Poitras for introducing me to the project and training me. He has given me valuable advice and without his knowledge and assistance this study would not have been successful. His sense of humour and his perpetual energy and enthusiasm in research motivated me to continue and always try to learn more.

I would like to show my gratitude to Dr. Mélanie Debiais-Thibaud who offered me support, advice and criticism. Thank you for using your precious time to read this thesis and to giving me constructive comments about it. But above all, thank you for your precious friendship.

I want to thank all other members of the Ekker lab, past and present, for their support and friendship throughout my Master's project, especially Sandra Noble, Jacob Pollack and Rafael Godoy.

I would like to thank, Mirela Hasu and Dr. Diane Lagacé for their great help with mouse behavioural work and for their helpful comments on the behavioural section of this thesis.

I am grateful to the members of my thesis advisory committee, Dr. Marie-Andrée Akimenko, Dr. Micheline Paulin-Levassseur and Dr. Jocelyn Côté for their useful advice and for providing interesting outside perspectives on my project. Furthermore, I would like to thank Dr. Claude Messier, Dr. Jocelyn Côté and Dr. Nicolas Corradi, who accepted, in the midst of all their work, to be members of my thesis examination board.

Finally, I would like to thank my family, especially my husband, without whose love, encouragement and support this thesis would not have evolved.

1 Introduction

1.1 Autism spectrum disorders (ASD)

Autism spectrum disorders, first described by Kanner in 1943, are neurobehavioural disorders that typically appear during the first three years of life. It is one of the most severe childhood neuropsychiatric conditions and it is characterized by an impairment of socialisation, verbal and nonverbal communication and range of interest, and restricted repetitive behaviours. ASD behaviours can range widely from mild to severe forms between individuals. ASD is one of the disorders classified as Pervasive Developmental Disorders (PDD) and includes Autistic Disorder (known as classical autism), Asperger's Disorder, and Pervasive Developmental Disorder-Not Otherwise Specified (PDD-NOS) (Freitag 2007). The prevalence of ASD was evaluated at approximately 2-6 cases per 1,000 and the ratio between boys and girls is about 4:1 (Fombonne, Quirke et al. 2009; Kim, Leventhal et al. 2011). Family and twins studies have shown that autism is mainly inherited, i.e in about 90% of the cases, with a prevalence of 8 to 9% higher in autistic siblings compared to the general population (Chudley, Gutierrez et al. 1998; Chakrabarti and Fombonne 2001; Newschaffer, Fallin et al. 2002). Furthermore, ASD is phenotypically and genetically heterogeneous, suggesting that autism is not a single-gene disorder. Genome-wide scans, cytogenetics and linkage analysis, and copy number variation (CNV) studies have shown several autism-susceptibility loci on almost all chromosomes, even though most of these loci were located on chromosomes 2q, 7q, 15q and 16p (Cook, Courchesne et al. 1998; Gillberg 1998; Schroer, Phelan et al. 1998; Risch, Spiker et al. 1999; Nurmi, Bradford et al. 2001; Anney, Klei et al. 2010). Since 75% of individuals with autism are males, it was suggested

that autism could be an X-linked disorder. However, linkage analyses have shown a modest linkage of the X chromosome to autism (Hallmayer, Hebert et al. 1996). Genes associated with ASD remain largely unknown, but several genes implicated in neurodevelopment are located in identified susceptibility loci, suggesting that mutation of these genes or of their regulatory elements could be linked to autism phenotype (Wassink, Piven et al. 2001; Jamain, Quach et al. 2003; Bernardet and Crusio 2006; Cheh, Millonig et al. 2006; Kuemerle, Gulden et al. 2007; Umeda, Takashima et al. 2010). One working hypothesis is therefore that ASD is a consequence of defects in nervous system development that can lead to cognitive, motor and intellectual disability (Rubenstein and Merzenich 2003)

1.1.1 Mouse models of ASD

To better understand the causes of autism and to further develop research in this field, animal models for autism are necessary, because there is a wide range of environmental and genetic factors involved in this neurodevelopmental disorder. Mice are the most commonly used vertebrate species. The use of mice is popular because they are readily available and inexpensive, they are small and easy to handle and they have a high reproductive rate. Mice are considered an excellent model for human genetic diseases since they have a genetic similarity >90% with human and they exhibit physical and behavioural aspects that may be comparable with those observed in humans (Church, Goodstadt et al. 2009; Ey, Leblond et al. 2011; Justice, Siracusa et al. 2011). In recent years, mouse models for autism have been developed and have allowed better modeling and understanding of the disease (Ey, Leblond et al. 2011). A good mouse model should have analogies to the phenotype observed in the human syndrome (face validity) and have the same biological dysfunction that causes the human disease, such as a gene mutation (construct validity) (Crawley 2007). Most models

developed so far have been focusing on the physiological and behavioural symptoms of autism (Moy and Nadler 2008; Ey, Leblond et al. 2011). These models are based on autism susceptibility loci determined by linkage studies in human populations, or on genetic mutations that lead to an abnormal synaptic homeostasis, due to problems with synaptic development, maturation or regulation or on prenatal environmental risk factors (Valproic acid, inflammatory agents), (Moy and Nadler 2008; Ey, Leblond et al. 2011). There are several mouse models that have been associated with ASD, but only a few are well studied and developed. Among them there are models for candidate genes that are related to synaptic development including neurexins (*Nrxn*), neuroligins (*Nlg),* shanks, reelin (*Reln*), integrins, cadherins and contactins (*Cntn*). The neuroligin family (*Nlgn1,2,3 and 4*) and neurexin1α (*Nrxn1α*), in particular, have been found to have an important role in excitatory and inhibitory synaptic contacts (Dean and Dresbach 2006; Moy and Nadler 2008). Nlgn, located in the postsynaptic region, is a synaptic cell adhesion protein that connects with presynaptic Nrxn1α, a neuronal cell surface protein. Both *Nlgn* genes and *Nrxn1α* gene are implicated in autism (Jamain, Quach et al. 2003; Szatmari, Paterson et al. 2007). Examination of *Nlgn* and *Nrxn* mutant lines indicated changes in some measures of synaptic function and displayed abnormal social and vocal behaviour (Missler, Zhang et al. 2003; Varoqueaux, Aramuni et al. 2006; Ey, Leblond et al. 2011). Specifically, *Nlgn4* mutant line showed a reduced reciprocal social interaction (Jamain, Quach et al. 2003; Jamain, Radyushkin et al. 2008), *Nlgn3* and *Nlgn4* mutant lines showed a lack of preference for social novelty in three-chambered social interaction test (Jamain, Radyushkin et al. 2008; Radyushkin, Hammerschmidt et al. 2009), *Nlgn3* and *Nlgn4* lines showed reduced ultrasonic vocalizations (Chadman, Gong et al. 2008; Jamain, Radyushkin et al. 2008) and finally *Nlgn1* and *Nrxn1α* showed impaired nest-building behaviour and repetitive self-grooming (Etherton, Blaiss et

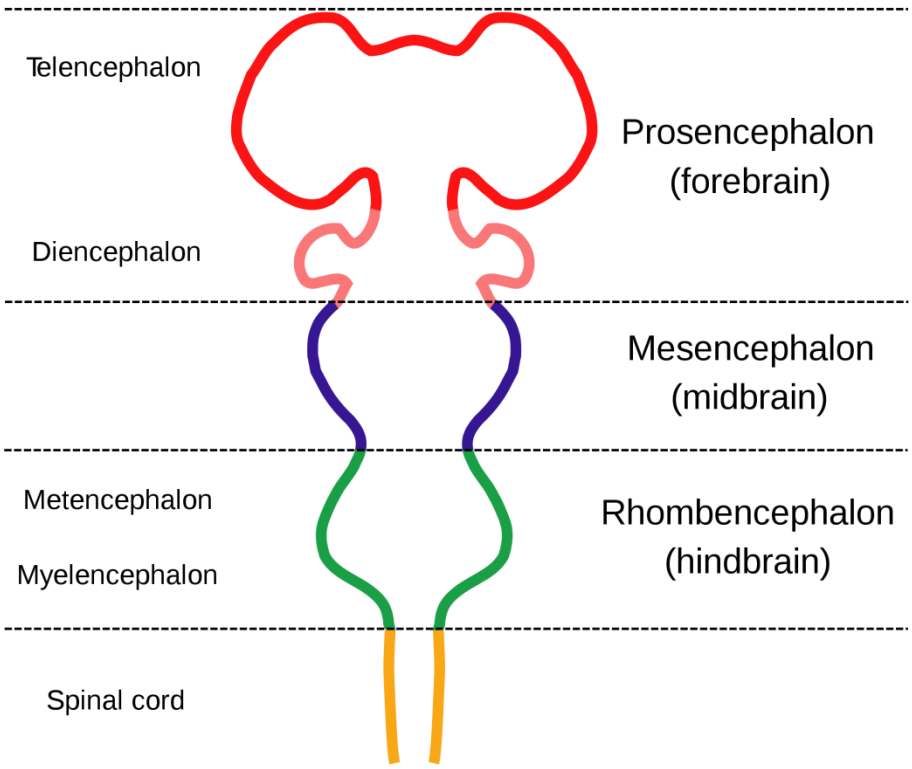
al. 2009; Blundell, Blaiss et al. 2010).Therefore those mouse models are relevant to characterize autism behaviour, with strong face validity (Silverman, Yang et al. 2010). There are also mouse models for candidates genes that are related to synaptic maturation and regulation such as methyl-CpG-binding protein 2 (*Mecp2*), fragile X mental retardation 1 (*Fmr1*), phosphatase and tensin homologue (*Pten*) known to influence neuronal migration and neurite extension and ubiquitin ligase E3A (*Ube3a*) that regulate the turnover of synaptic proteins (Moy and Nadler 2008; Silverman, Yang et al. 2010; Ey, Leblond et al. 2011). With these mouse models, it was possible to study different aspects of ASD behaviour. First, impairments in cognitive functions like anxiety, spatial learning, memory and social interaction are monitored in different models. Impairments in social interaction are the primary indicator of ASD, therefore the interactions between adult mice or with juveniles placed together in regular cages or specialized arenas provide the most detailed insights into reciprocal social interactions (Silverman, Yang et al. 2010). Second, physical disabilities including circadian rhythm, vision, olfaction, auditory startle, sensory-motor gating, seizure and locomotion, are observed. As a conclusion, even though mouse models for ASD have a different range of observable behaviours compared to human, they still represent the diversity of symptoms found in humans (Ey, Leblond et al. 2011).

1.2 Development of the nervous system in vertebrate

The nervous system is a network of specialized cells, called the neurons, which transmit signals between different parts of the body. The nervous system can be divided in central nervous system (CNS) and peripheral nervous system (PNS). The central nervous system includes the brain, the spinal cord and the retina; whereas the peripheral nervous system consists of sensory neurons, ganglia, and nerves connecting between them and to the central nervous system. In vertebrate, the development of the central nervous system begins with the neural plate, consisting of a thin strip of ectodermal cells along the center of the dorsal surface of the embryo. As the development proceeds, the neural plate begins to fold and form the neural tube (future CNS) and to give rise to the neural crest above the neural tube (future PNS). The sequence of stages from neural plate to neural tube and neural crest is known as neurulation (Wolpert, Beddington et al. 1998). Most of the neural tube will give rise to the spinal cord. The rostral end of the neural tube will enlarge to form the three primary brain vesicles, the prosencephalon, the mesencephalon and the rhombencephalon (Fig1.1). Those three primary vesicles will be subdivided into five secondary vesicles. The rhombencephalon will divide into the metencephalon and the myelencephalon (Fig1.1), and that will differentiate into two major structures, respectively, the cerebellum and the medulla of the adult brain. This caudal region of the brain stem is called the hindbrain. The mesencephalon will remain as a single vesicle, the midbrain. Finally, the prosencephalon will divide into telencephalon and diencephalon (Fig1.1), and give rise to the forebrain. The diencephalon generates the thalamus and the hypothalamus in the mature brain. On either side of the diencephalon, two secondary vesicles also develop, called the optic vesicles. The optic vesicles will develop into the retinas. (Sanes D, Reh T.A et al. 2011). The progenitor cells of

the neural tube are known as neural stem cells. This represents the 1st stage of neural development, neurogenesis, when undifferentiated cells undergo mitotic divisions to produce either new stem cells or neuroblasts that will eventually differentiate into neurons (Purves, Augustine et al. 2008). The newly generated neurons migrate to different parts of the developing brain to self-organize into different brain structures. Various structures will emerge from the proliferation of the neurons of telencephalon: the cerebral cortex, the ventral telencephalon, and the olfactory bulb (Sanes, Reh et al. 2011). Once the neurons have reached their positions, they extend axons and dendrites, which allow them to communicate with other neurons via synapse and neurotransmitter. Synaptic communication between neurons allows the establishment of functional neural circuits that mediate sensory and motor processing, and underlie normal behaviour. It was proposed that some forms of autism were the result of an imbalance in the excitation/inhibition ratio of the cortex neurons during the development leading to an increase in the excitatory state of the brain (Rubenstein and Merzenich 2003).

Figure 1.1 : Diagram of the main subdivisions of the embryonic vertebrate brain during development. These regions will later differentiate into forebrain, midbrain and hindbrain structures.


Telencephalon
Diencephalon
Prosencephalon
(forebrain)
Mesencephalon
(midbrain)
Metencephalon
Myelencephalon
Rhombencephalon
(hindbrain)
Spinal cord

1.3 Development of the GABAergic interneurons in mammals

In mammals, the fully developed neocortex is composed of two major classes of neurons, projection neurons and local circuit neurons, also called interneurons. Projection neurons are excitatory neurons that contain the neurotransmitter glutamate while interneurons are inhibitory GABAergic neurons using γ-aminobutyric acid as a neurotransmitter (Benes and Berretta 2001). However, during embryonic development, GABAergic neurons will compose the first excitatory cortical network and then, after implementation of the projection neuron excitatory network, these GABAergic neurons will become inhibitory neurons (Ben-Ari 2002). There are distinct types of GABAergic interneurons in the developing and mature brain, each with different morphology, distribution or connectivity. Among these GABAergic interneurons subtypes there are basket cells (axo-somatic inhibitory synapse), chandeliers cells (axo-axonal inhibitory synapse), double bouquet cells and axon tuft cells (axo-dendritic inhibitory and disinhibitory synapse) (Benes and Berretta 2001). These distinct populations of GABAergic interneurons can be identified and characterized by the expression of different neurochemical markers, either calcium binding proteins such as parvalbumin, calretinin or calbindin, or neuropeptides such as somatostatin (Benes and Berretta 2001). Glutamic acid decarboxylase (GAD) is the enzyme that biosynthesizes the neurotransmitter GABA through decarboxylation of glutamic acid, and is therefore common to all GABAergic neurons. DLX transcription factors are important in the development of GABAergic interneurons since they regulate the expression of GAD genes (Stuhmer, Puelles et al. 2002). In mammals, GAD exists in two isoforms GAD_{67} and GAD_{65} with molecular weights of 67 and 65 kDa respectively. These isoforms are encoded by 2 different genes, *GAD1* and *2*, which are both expressed in the brain (Erlander and Tobin 1991). During

embryonic development, it was shown that most GABAergic interneurons in the neocortex originate from cell proliferation in the ventral telencephalon (lateral ganglionic eminence (LGE), medial ganglionic eminence (MGE), caudal ganglionic eminence (CGE), septum and anterior entopeduncular region), and these neuronal progenitors contribute to the neuronal diversity of distinct areas of the brain: cerebral cortex, olfactory bulb and hippocampus, through radial and tangential migrations (Fig1.3B). Within the cortex, interneurons integrate into the local circuitry while they develop mature and distinct morphological and physiological properties (Anderson, Eisenstat et al. 1997). As GABAergic interneurons display a variety of morphological and physiological behaviours that are crucial for both the development and the function of mature cortical circuits, it has been proposed that defects in cortical inhibition may be implicated in neurological and psychiatric disorders like epilepsy, depression, Rett syndrome, anxiety, schizophrenia and autistic disorders. (Benes and Berretta 2001; Rubenstein and Merzenich 2003; Cobos, Calcagnotto et al. 2005; Horike, Cai et al. 2005). This hypothesis is consistent with the fact that key regulators of GABAergic progenitor migration and differentiation, members of the *Dlx* homeobox gene family, are found in two autism-susceptibility loci on chromosome 2q and 7q. Specifically, *DLX1* and *DLX2* genes are located on chromosome 2q31.1 whereas *DLX5* and *DLX6* are found on chromosome 7q21.3. Until now, there are no observations that support a significant contribution of *DLX* genes in autism susceptibility. However, recent data obtained from autistic patients strongly implicate defects in *DLX* regulatory sequences in the development of ASD (Hamilton, Woo et al. 2005).

1.4 Distal-less-related homeobox (*Dlx*) genes

1.4.1 Homeobox genes and homeodomain proteins

Transcription factors are proteins that contain at least one DNA-binding domain (DBD), they can bind to specific DNA sequences and thereby control the transcription of genetic information. They can either up- or down-regulate the transcription of a gene by binding to sequences such as enhancers or promoters. Enhancers modulate the level of transcription which depends on the type of tissue, developmental stage, stage of the cell cycle, induction by hormones or other molecular signals, while a promoter is the regulatory part of gene that is responsible for the basal level of transcription (Alberts 2007). Transcription factors are classified on the basis of their DBD, such as the helix-turn-helix motif found in the homeobox transcription factors (Latchman 1997). The homeobox is a DNA sequence of about 180 bp that encodes a protein domain of 60 amino acids called the homeodomain and responsible for DNA binding. Homeobox genes encode transcription factors that play critical roles in the regulation of development of eukaryote organisms (Scott, Tamkun et al. 1989; Gehring 1994). Homeodomain proteins are found in all eukaryotic organisms studied so far (Alberts 2007). There are several families of homeobox genes including *Hox*, *Evx*, *Emx*, *Msx* and *Dlx* (Gehring 1993), to name only a few.

1.4.2 *Dlx* genes organization

Dlx genes encode a family of homeodomain-containing transcription factors related to the Drosophila's gene *dll* (*Drosophila distal-less*) (Panganiban and Rubenstein 2002). Vertebrates have more than one *Dlx* gene, whereas in drosophila and other protostomes a single *dll* gene has been identified. *Dlx* genes in vertebrates come from a tandem duplication event of an ancestral *Dlx* gene, followed by 2 (or 3 in teleost fish) genomic duplications and

the loss of a *Dlx* pair. Thus, there are six known *Dlx* genes in mammals and eight *dlx* genes in the zebrafish (a teleost fish) (Stock, Ellies et al. 1996; Panganiban and Rubenstein 2002). All *Dlx* genes have the same gene organization, three exons and two introns with their coding sequence on all three exons, and the homeobox region separated between exons 2 and 3 (Fig.1.2) (Liu, Ghattas et al. 1997). In mammals, *Dlx* genes are arranged in three pairs of genes in a tail to tail orientation which are also referred to as bigene clusters: *Dlx1* with *Dlx2*, *Dlx3* with *Dlx4*, and *Dlx5* with *Dlx6*. Each of these bigenes is found on different chromosomes and is located in synteny with a *Hox* gene cluster (Stock, Ellies et al. 1996; Panganiban and Rubenstein 2002).

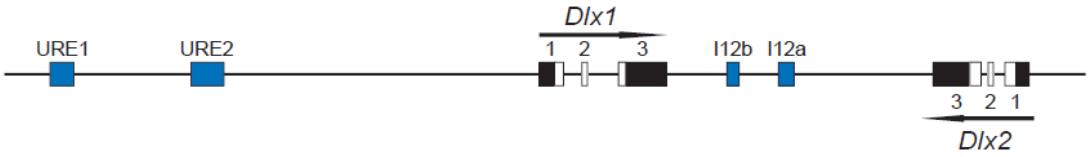
1.4.3 *Dlx* gene expression

1.4.3.1 Overview

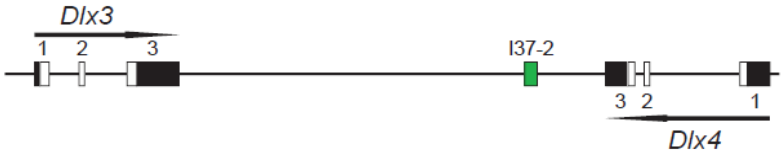
The expression of *Dlx* genes has mostly been studied during embryonic development (Zerucha and Ekker 2000; Panganiban and Rubenstein 2002). *Dlx* gene are expressed along the proximodistal axis in a subset of migratory neural crest cells that form the branchial arches (Panganiban and Rubenstein 2002; Kraus and Lufkin 2006). More specifically, *Dlx1* and *2* are expressed in the proximal, intermediate and distal regions of branchial arches, whereas *Dlx5* and *6* are expressed in intermediate and distal regions and, finally, *Dlx3* and *4* are only expressed in distal regions (Qiu, Bulfone et al. 1997; Depew, Lufkin et al. 2002). Therefore, *Dlx* genes demonstrate overlapping expression patterns in developing branchial arches (Acampora, Merlo et al. 1999; Depew, Liu et al. 1999; Zerucha and Ekker 2000; Panganiban and Rubenstein 2002). *Dlx* genes are also expressed in skeletal tissues and in developing teeth (Depew, Lufkin et al. 2002). Expression of *Dlx* genes, particularly *Dlx2, 3, 5* and *6*, is found in the sensory organs such as the olfactory organ, the otic vesicle and the

Figure 1.2 : *Dlx* gene organization. Exons are numbered (black = non-coding; white = coding). Regulatory sequences in blue (*Dlx1/Dlx2*), green (*Dlx3/Dlx4*) and red (*Dlx5/Dlx6*).

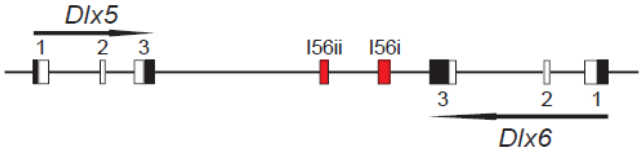
Dlx1/Dlx2



Dlx3/Dlx4



Dlx5/Dlx6

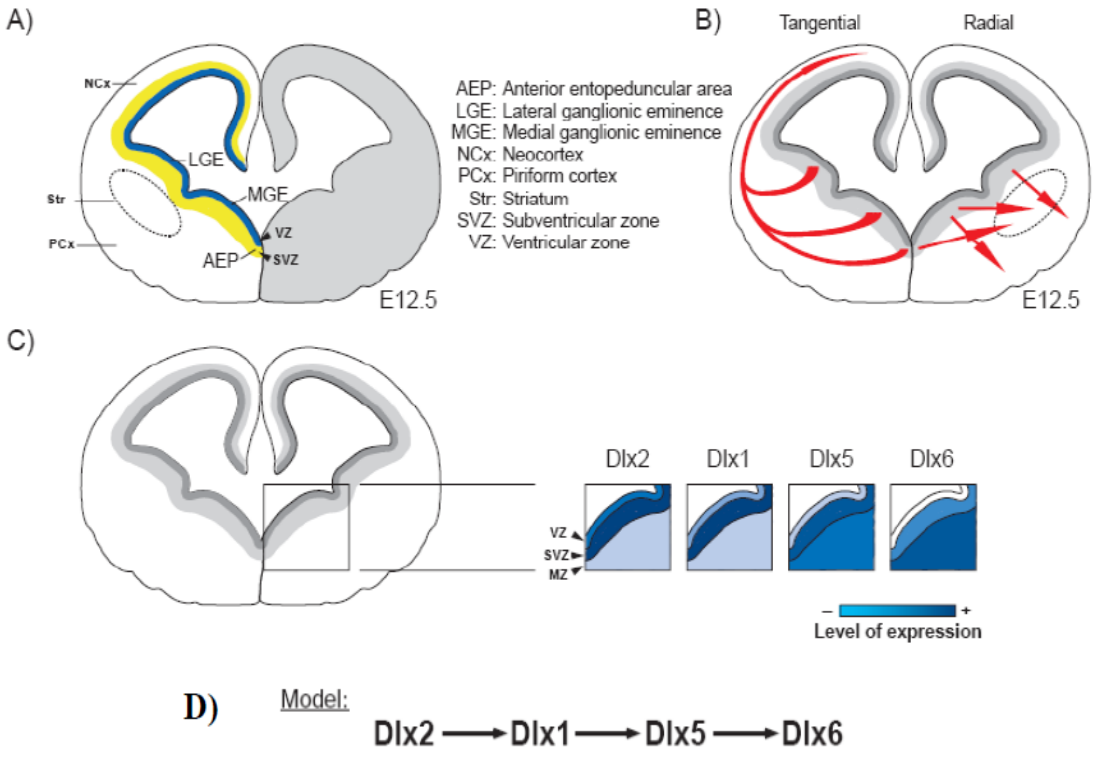


retina (Ekker, Akimenko et al. 1992; Depew, Liu et al. 1999). Finally, *Dlx* transcripts are found in the apical epidermal ridge of the developing forelimb and hindlimb (Dollé, Price et al. 1992; Morasso, Mahon et al. 1995).

1.4.3.2 Developing forebrain

In mouse, only four out of six *Dlx* genes, namely *Dlx1, 2, 4,* and *6,* are expressed during the development of the forebrain, in the ventral telencephalon and in the diencephalon (Eisenstat, Liu et al. 1999). Within the ventral telencephalon, *Dlx* genes are expressed in parts of the septum, in the lateral and medial ganglionic eminences (respectively MGE and LGE), in the preoptic area, in parts of the amygdala and finally in pallial interneurons (Wang, Lufkin et al. 2011). *Dlx* genes show sequential and overlapping spatial and temporal patterns of expression in the telencephalon (Fig1.3C). *Dlx2* is mainly expressed in the ventricular zone (VZ) and in the subventricular zone (SVZ) of the MGE and LGE. *Dlx1* expression is found in ventricular, subventricular (SVZ) and mantle zones (MZ) of both the MGE and LGE. *Dlx5* and *Dlx6* expression patterns are shown in the post-mitotic differentiating neurons of the subventricular and mantle zones (Panganiban and Rubenstein 2002). During development, neural cells move from the ventricular zone (their proliferation zone) through the subventricular and mantle zones and then towards the cortex (Marin and Rubenstein 2001). The spatially overlapping expression patterns therefore indicate a temporal sequence of *Dlx* gene expression: early activation of *Dlx*2 in the VZ, then *Dlx*1 when cells go through the SVZ, followed by late activation of *Dlx*5 and *Dlx*6 expression (Fig1.3D) (Panganiban and Rubenstein 2002). This suggests that different *Dlx* genes are necessary at different stages of GABAergic differentiation during development (Wang, Lufkin et al. 2011).

Figure 1.3: Cross sections of a basal telencephalon from an E12.5 mouse embryo showing : (A) Structures of the telencephalon, (B) Migration of neurons during development following radial and tangential paths, (C) Expression domains and levels of *Dlx1*, *Dlx2*, *Dlx5* and *Dlx6* genes and (D) Proposed model for *Dlx* sequential activation in the mouse forebrain. (Adapted from (Panganiban and Rubenstein 2002)



D) Model:

Dlx2 → Dlx1 → Dlx5 → Dlx6

1.4.4 *Dlx* gene function

The function of Dlx transcription factors during forebrain development was mainly studied in the mouse through the production of targeted null mutations of the *Dlx* genes by homologous recombination. Only subtle forebrain phenotypes were observed in *Dlx* single mutants, suggesting that these transcription factors have both unique and redundant functions (Panganiban and Rubenstein 2002; Cobos, Calcagnotto et al. 2005; Wang, Lufkin et al. 2011). *Dlx1* single mutants show abnormal interneurons differentiation and partial loss of interneurons in the neocortex and hippocampus after birth, resulting in epilepsy later in life (Cobos, Calcagnotto et al. 2005; Mao, Page et al. 2009). Dlx1 has therefore been proposed to be responsible for the preservation of the functional longevity of interneurons in adult (Mao, Page et al. 2009). *Dlx2* and *Dlx5* single mutants both have reduced number of interneurons in the olfactory bulb (Zerucha and Ekker 2000; Panganiban and Rubenstein 2002; Wang, Lufkin et al. 2011). Loss of Dlx5 function also causes reduction of parvalbumin neocortical interneurons (Wang, Dye et al. 2010). More recently, it was shown that *Dlx6* single mutants have defects in the striatum and in part of the amygdala (Wang, Lufkin et al. 2011).

In the case of double mutants, the observed forebrain phenotypes are more severe. The *Dlx1/Dlx2* double mutants exhibit a major block in neurogenesis within the ventral telencephalon, affecting the development and migration of several types of GABAergic, dopaminergic, and cholinergic interneurons in the cortex (Anderson, Eisenstat et al. 1997; Anderson, Qiu et al. 1997; Panganiban and Rubenstein 2002). Also, Dlx1 and Dlx2 together control neuronal versus oligodendroglial cell fate acquisition (Petryniak, Potter et al. 2007). In the Dlx1/2 double mutant, the expression of *Dlx5*, *Dlx6* and *Gad1* is reduced (Anderson, Qiu et al. 1997; Wang, Lufkin et al. 2011). Finally, in the Dlx5/6 double mutant, the forebrain phenotype and therefore the function is not easy to determine due to exencephaly

in these animals. However, recently, a defect in the migration and differentiation of interneurons coming from the MGE was observed (Wang, Lufkin et al. 2011).

1.4.5 *Dlx* gene regulation

The expression of *Dlx* genes is thought to be controlled by *cis*-acting regulatory elements located upstream of these genes or in their intergenic region. *cis*-acting elements are enhancers that can be bound by transcription factors to enhance the transcription levels of genes in a gene cluster. After sequence comparisons and transgenic assays, three conserved non-coding elements were shown to be putative *cis*-acting regulatory elements in the *Dlx1/2* locus (Ghanem, Jarinova et al. 2003). Two of them, I12a and I12b, are located in the intergenic region between *Dlx1* and *Dlx2*. I12b can target reporter gene expression (*lacZ*) in both the forebrain and the branchial arches while I12a can only drive reporter gene expression in the branchial arches. A third putative enhancer, located 12 kb upstream of the *Dlx1* gene, was found and named Upstream Regulatory Element-2 (URE2). Transgenic experiments revealed that URE2 is able to target expression of the *lacZ* reporter gene in the forebrain (telencephalon and diencephalon) (Ghanem, Yu et al. 2007).

Concerning the *Dlx5-Dlx6* cluster, two putative enhancers have been identified: I56i and I56ii (Intergenic enhancer of *Dlx*5/*Dlx*6 i and ii). Sequence alignments revealed that these elements were highly conserved between species (Zerucha and Ekker 2000). In fact these two regulatory elements are part of a group of 481 elements, more than 200 base pairs long, that are perfectly conserved, i.e. display 100% identity without insertion or deletion between orthologous regions of human, rat and mouse (Bejerano, Pheasant et al. 2004). It was shown that both elements drive the expression of a *lacZ* reporter construct to the basal telencephalon and the diencephalon, with a pattern which is consistent with the endogenous *Dlx* expression

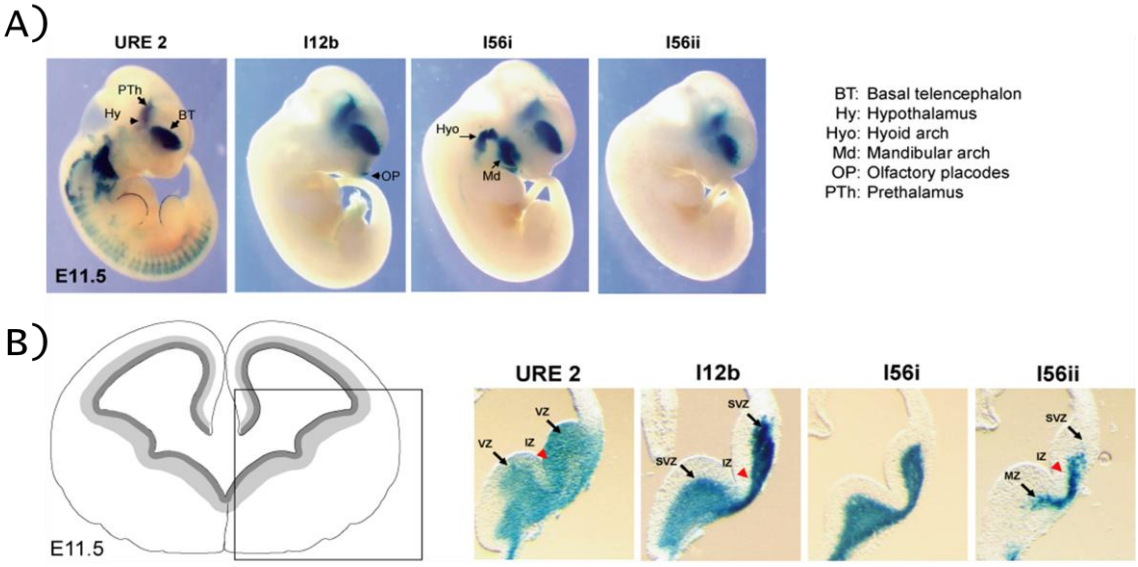
patterns. Branchial arch expression was also found in transgenic experiments using the I56i enhancer (Zerucha, Stuhmer et al. 2000; Park, Sperber et al. 2004).

To summarize, four conserved regions are able to drive gene expression in the forebrain: I12b, URE2, I56i and I56ii. The *lacZ* expression pattern driven by the putative forebrain enhancers reproduce *Dlx* expression patterns (Zerucha, Stuhmer et al. 2000). When these four putative enhancers were compared to each other within a given species, their sequences were found to be very divergent, suggesting that each enhancer could be responsive to different trans-acting factors (Zerucha, Stuhmer et al. 2000; Ghanem, Jarinova et al. 2003). The forebrain expression of *lacZ* activated by different *Dlx* putative enhancers is almost indistinguishable morphologically in whole-mount embryos (Fig1.4A). However, slight expression differences are detectable in the transverse sections of transgenic embryos at different stages of their development (Fig1.4B). It was proposed that these differences correspond to distinct regulatory mechanisms in precursors of different classes of cortical interneurons (Ghanem, Yu et al. 2007).

1.4.6 Identification of a SNP in the I56i forebrain enhancer

In order to be able to determine the role of DLX genes in ASD, genomic DNA of 161 individuals with autism (samples from the Autism Genetic Resource Exchange collection) and 58 non-autistic siblings was sequenced and compared to 188 samples of non-affected patients (from the NIGMS / NHGRI Polymorphism Discovery Resource) (Hamilton, Woo et al. 2005). Specifically, exons, exon-intron boundaries and regulatory elements (URE2, I12a, I12b, I56i and I56ii) of *DLX1, 2, 5* and *6* were examined in each individual. In total, thirty-three variants were identified. Of these, 31 variants are single nucleotide polymorphisms (SNPs) and two are insertion/deletion polymorphisms. Among these variants

Figure 1.4: Activity of conserved *Dlx* forebrain enhancers in E11.5 transgenic embryos. A) Whole embryos B) Transverse sections of the medial telencephalon. Adapted from (Ghanem, Yu et al. 2008)


A)
URE 2
I12b
I56i
I56ii
PTh
Hy
BT
OP
Hyo
Md
E11.5
BT: Basal telencephalon
Hy: Hypothalamus
Hyo: Hyoid arch
Md: Mandibular arch
OP: Olfactory placodes
PTh: Prethalamus
B)
E11.5
URE 2
I12b
I56i
I56ii
VZ
IZ
SVZ
MZ

7 are found in cis-regulatory elements, specifically, four SNPs are present in URE2, a deletion of a nucleotide is found in I12a, a SNP is observed in I12b and finally a SNP is identified in I56i (Hamilton, Woo et al. 2005). The SNPs found in the regulatory element I56i is a change of an adenine for a guanine and is found at a position being previously characterized as a binding site for homeodomain transcription factors. Moreover, it was shown that this binding site is a functional Dlx binding site (Zerucha, Stuhmer et al. 2000). Furthermore, finding a SNP in I56i enhancer was very surprising considering that this sequence is ultraconserved, therefore not susceptible to variation (Bejerano, Pheasant et al. 2004). Thus, this strongly suggests that the SNP could be causally related to the behavioural defects seen in the autistic individuals who carry it.

1.5 Hypothesis and objectives

The working hypothesis for this work was that a change in the sequence of the I56i enhancer corresponding to the SNP could change the expression of *Dlx5* and/or *Dlx6* in the forebrain and therefore may affect the differentiation and migration of GABAergic interneuron progenitors during development. This may contribute to explaining the neurodevelopmental problems observed in autistic patients.

Therefore, my first objective was to determine the impact of the SNP on the genetic interactions taking place at the I56i enhancer.

Specifically, I tested the impact of the SNP on the binding of Dlx transcription factors to the I56i enhancer as well as the enhancer activity of I56i in response to the Dlx2 or Dlx5 proteins. I also studied the implication of GTF2I in *Dlx5/Dlx6* gene regulation via the I56i enhancer.

My second objective was to perform the phenotypic and behavioural analyses of mice carrying a mutation in the I56i SNP enhancer.

To achieve that, I used mutant mice harbouring the SNP in the I56i regulatory element that was generated by a targeted replacement (Knock-in) of the wild-type enhancer using homologous recombination in embryonic stem cells. I characterized the impact of the mutation on *Dlx* expression in the forebrain during development. I also assessed behavioural skills of the I56i SNP mice using behavioural tests of sociability, infant and juvenile play and anxiety.

2 Materials and Methods

Since I participated only in some of the experiments of chapter three, this section describes the materials and methods that are relevant to the experiments that I have done.

2.1 Transient transfection assays

2.1.1 Reporter plasmids.

10µg of the pGL4.24 vector (Promega) were digested with 10 units of *Hin*dIII (Invitrogen) and *Bgl*II (Invitrogen) at 37°C for 2 hours followed by dephosphorylation for 30 minutes at 37°C with Calf Intestinal Alkaline Phosphatase (CiAP). Then, the CiAP was heat inactivated 10 minutes at 65°C. The DNA was purified by gel with a gel extraction kit QIAquick (Qiagen), that purified PCR products >100 bp using a bind-wash-elute procedure. Ten µg of the wild type and SNP variant versions of the mouse I56i enhancer were subcloned in the vector pDrive (Qiagen) were digested with 10 units of restriction enzymes *Hin*dIII (Invitrogen) and 10 units of *Bam*H1 (Invitrogen) at 37°C for 2 hours. Then, DNA was purified by gel with the QIAquick extraction kit. Then, 7ng of each of these inserts were inserted upstream of a minimal promoter driving expression of the synthetic *firefly luciferase (luc2)* gene in 20ng of the vector pGL4,24. Ligation was performed using 1 unit of T4 Ligase (Invitrogen). Ligation was conducted in a thermal cycler at 14°C for an overnight period. The two resulting constructs were transformed into 100µl of chemically competent XL-10 *Escherichia coli* cells by incubation on ice of both ligation product and cells for 1 hour, followed by a 42°C heat shock for 90s and a 60 minutes incubation at 37°C in pre-warmed LB broth. After the final 37°C incubation, cells were centrifuged in a Microlite centrifuge, (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. Positive clones were selected and grown in 2mL of LB at 37°C overnight. Plasmid DNA was isolated using the Wizard Plus SV Minipreps DNA Purification System (Promega), a silica membrane-based system with a yield of 1.8µg/mL, and the DNA was sent for DNA sequencing at the 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 growth in 50mL of LB media at 37°C overnight. Plasmid DNA isolation was conducted using the Qiagen MIDI kit (Qiagen). The kits deliver DNA yields of up to 100 µg. Lysate clearing and isopropanol precipitation were achieved by centrifugation. Plasmid DNA was stored at 4°C until required for luciferase reporter assays.

2.1.2 Expression vectors

A zebrafish *dlx2a* cDNA (845bp) encompassing the full coding sequence was amplified by PCR and cloned into the *EcoRI* site of the pcDNA3-FLAG expression vector. The construction of the DLX2A homeodomain mutant was obtained by changing essential amino acids of the third loop of the homeodomain. Therefore, Trp170, Phe171, Gln172, Asn173, Arg174 and Arg175 were changed to glycine using overlapping PCR as described in (Poitras, Ghanem et al. 2007). The zebrafish *dlx5a* cDNA (1.2kb) was amplified by PCR and cloned into the *Bgl*II/*Xho*I sites of the pTL2 expression vector. The rat Gtf2i cDNA (accession # BC085815) subcloned in the expression vector pExpress-1 was purchased from Open Biosystem.

2.1.3 Cotransfection and luciferase assays.

The cotransfection experiments were performed in the NIH 3T3 mouse fibroblast cell line. The cells were grown at 37 °C in DMEM high glucose 1X (4.5 g / L D-glucose, L-glutamine and 110 mg / L sodium pyruvate) (GIBCO) enriched with 10% FBS (Fetal bovine Serum) (GIBCO ©) and 1% Penicillin / streptomycin (GIBCO). The cells were seeded in a 24-well plate, 24 hours before transfection at a density of 80,000 cells per well. A total of 705 ng of DNA per well was used in each transfection. This included 5 ng of pRL-CMV (*Renilla* luciferase vector) from Promega, an internal control for transfection efficiency, 500 ng of reporter plasmid (constructs described above) and 200 ng of *Dlx* expression plasmid (as described above). Lipofectamine reagent (Invitrogen) were used to introduce plasmids into cultured cells by lipofection. The cells were harvested 24 hours after transfection and lysed in the presence of "passive lysis buffer" (Promega). Luciferase tests were performed in accordance with the protocol of Dual-Luciferase Reporter Assay System (Promega) and using the solutions provided by the company. The firefly luciferase reporter is measured first by adding Luciferase Assay Reagent II (LAR II) to generate a luminescent signal lasting at least one minute. After quantifying the firefly luminescence, this reaction is quenched, and the *Renilla* luciferase reaction is initiated simultaneously by adding Stop & Glo® Reagent to the same sample. Both reporters yield linear assays with attomole ($<10^{-18}$) sensitivities. All luminescent signal were obtained from a *Molecular Device* LMAXII (384) luminometer and Softmax Pro 4.7.1 software. Luciferase activity was standardized to *Renilla* luciferase levels to compensate for variations in transfection efficiency. Experiments were performed in triplicate and repeated a minimum of three times.

2.1.4 Statistics

Statistical significance (*P* values < 0.05) was determined using paired Student's *t* test. All quantitative data were presented as the mean ± S.E.M.

2.2 Electromobility shift assays (EMSA)

Gel mobility shift assays were performed using double-stranded oligonucleotides corresponding to the putative protein binding site. This double-stranded oligonucleotide was obtained by the annealing of two complementary 21 nucleotide oligonucleotides (mI56i FP3.for and mI56i FP3.rev). They were annealed overnight by placing 100ng of each oligonucleotide in sterile distilled water for a final concentration of 2.5ng/µl for each oligonucleotide. The oligonucleotides were then placed at 80°C for 10 minutes, at which point the heat source was shut off, letting the oligonucleotides incubate overnight in a contained environment, and assuring a gradual heat loss and annealing process. They were end-labeling using [^{32}P] γ-ATP (Amersham) and T4 Polynucleotide kinase (New England Biolabs).

For each binding reaction, 1 ng of labeled DNA was incubated with recombinant Flag-Dlx2 (TNT Quick Coupled Transcription/Translation Kit, Promega, used to obtain *in vitro* translation from a gene construct), GST-Dlx5 proteins or with nuclear extracts in reaction buffer (7 mM Tris pH 7.5, 81 mM NaCl, 2.75 mM dithiothreitol, 5 mM $MgCl_2$, 0.05% NP-40, 1 mg/mL bovine serum albumin, 25 µg/mL poly(dIdC), 10% glycerol) at room temperature for 30 minutes. Reactions were loaded on a polyacrylamide gel (6% acrylamide, 0.16% bis-acrylamide) in 0.5×TBE. The gel was run at 80 volts in 0.5×TBE and dried. Radioactive bands were visualized by phosphorimaging (Molecular Dynamics) as previously described in (Poitras, Ghanem et al. 2007).

The EMSA with Flag-Dlx2 was carried using double-stranded ^{32}P-labelled oligonucleotides, previously described, with increasing amounts of Flag-Dlx2 protein. For the supershift assay, the unpurified Flag-Dlx2 TNT reaction was preincubated overnight at 4°C with 1 µg of anti-Flag antibody (Sigma).

The competition assay with recombinant Dlx5 protein was carried out using double-stranded ^{32}P-labelled oligonucleotides with fixed amounts of GST-zDlx5a protein. This double-stranded oligonucleotide corresponds to a triplicate of the (ATAATT) core of the identified putative FP3 binding site. It was obtained by the annealing of two complementary oligonucleotides (mI56i FP3X3.for and mI56i FP3X3.rev), as previously described. These binding reactions were competed with increasing amounts (1, 3 and 10 ng) of the unlabeled double-stranded wt oligonucleotide competitor or the variant version (vI56i FP3X3.for and vI56i FP3X3.rev). Quantification of the percentage of the remaining complexes was achieved by using the ImageJ software (http://rsbweb.nih.gov/ij/index.html).

The primers used to do the double-stranded oligonucleotides were:

mI56i FP3.for 5'- AAATGCAGCCATAATTAGAGT-3'

mI56i FP3.rev 5'- ACTCTAATTATGGCTGCATTT-3'

mI56i FP3X3.for 5'-CCCCATAATTCCCCATAATTCCCCATAATTCCCC-3'

mI56i FP3X3.rev 5'-GGGGAATTATGGGGAATTATGGGGAATTATGGGG-3'

vI56i FP3X3.for 5'-CCCCGTAATTCCCCGTAATTCCCCGTAATTCCCC-3'

vI56i FP3X3.rev 5'-GGGGAATTACGGGGAATTACGGGGAATTACGGGG-3'

2.3 "Knock-in" mice

Knock-in mice were generated by homologous recombination in 129Sv embryonic stem (ES) cells. A modified enhancer-Neomycin resistance cassette was inserted in place of the wild type I56i enhancer present on a Bacterial Artificial Chromosome (BAC) containing the *Dlx5/Dlx6* locus. This was accomplished by homologous recombination in bacteria as described in (Zhang, Buchholz et al. 1998; Lee, Yu et al. 2001). The recombined BAC clone was verified for sequence accuracy at the site of recombination and then was electroporated into ES cells. This part of the work was done at the Transgenic Core Facility of the McGill Cancer Center. Clones positive for neomycin were selected with G418, and 100 ES cell colonies were screened for the correct gene using quantitative real-time PCR as described by (Valenzuela, Murphy et al. 2003). The A to G mutation was verified by sequencing the genomic DNA of positive ES clones. One clone was correctly targeted and used to generate the "knock-in" mice. ES cells were then injected into C57BL/6 blastocysts to make chimeras, at the Transgenic Core Facility of the McGill Cancer Center. Chimeric founders were mated with wild-type C57BL/6 mice. Genotyped (as described below in section 2.3.3) progeny that were positive for germline transmission were bred. First generation offspring that inherited the targeted allele with neomycin were subsequently mated with C57BL/6 mice. I started my project with these heterozygous offspring. Heterozygous offspring within each mating scheme were subsequently bred to yield homozygous mice.

Mice were group-housed in standard mouse cages in a room with a 12 h light-dark cycle and *ad libitum* access to food and water and all animal experiments were approved by the Standard Operating Procedures and Guidelines of the uOttawa Animal Care Committee. The

mice were anaesthetized with CO_2, followed by a vertebral dislocation. All researchers interacting with these mice obtained the required National Institute for Animal User Care (NIAUT) training.

2.4 Production and screening of mutant mouse tissues

2.4.1 Genomic DNA extraction

Tissue samples from the tail end of weaned mice, or extraembryonic tissue of embryos, were digested overnight at 55°C in 500µl of digestion buffer (50mM Tris-HCl pH 8.0, 100mM EDTA, 100mM NaCl and 1%SDS) and 50µl (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) was carried out 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).

2.4.2 Genotyping

Genotyping of mice carrying zero, one or two SNP alleles of the I56i enhancer was done using PCR with primers used to amplify the neomycin gene (Neo.for and Neo.rev) or primers used to amplify the wild type allele only (I56iscreen.for and I561screen.rev, and primers used, in the same mixed as Neo primers, to amplify a fragment of the Fetal Hemoglobin (HbF) gene as a positive control (FH1 and FH2). The I56i screen.for primer was located directly on the I56i enhancer and the I56isceen.rev was located 383bp upstream.

However in the mutant allele, the neomycin cassette was inserted between the two primers giving a distance too large (2.3kb) to be amplified by PCR under the conditions used. The following conditions were used in PCR experiments : 1XPCR buffer (100mm TrisHCl pH8.3, 500mM KCl, 12mM MgCl2), 0.2mM dNTPs (supplied by Invitrogen and diluted to stock solution of 2mM), 10-100ng of template DNA, 2pmol of each primer (I56iscreen.for and I561Screen.Rev, Neo.for and Neo.Rev, FH1 and FH2), and 1 unit of Taq Polymerase. The cycle conditions were: one cycle at 95° for 4 minutes, 30 cycles of: (94°Cx60s, 58°Cx60s, 72°Cx120s), an extension cycle of 72°C of 5 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. Therefore, in wild-type mice, no Neo band was seen but there is an I56i band (383bp). Heterozygous mice had a band for Neo (542bp) and I56iscreen (383bp), while mice homozygous for the vI56i had a Neo band (542bp) and no I56iscreen band. The following primers were used to genotype the mice

Neo.for 5'-TCTCCTGCCGAGAAAGTAT-3'

Neo.rev 5'-CAACAGATGGCTGGCAACTA-3'

I56iscreen.for 5'-GCTTCAAATTGGATGGCACT-3'

I56iscreen.rev 5'-TACAGACCTGGGCATCCTTC-3'

FH1 5'-GATCATGACCGCCGTAGG-3'

FH2 5'- CATGAACTTGTCCCAGGCTT-3'

2.5 Quantitative real-time PCR (qPCR)

2.5.1 RNA extract preparation

RNA extracts were obtained from brains collected from wild type mice, heterozygous mice or vI56i mutant mice at E11.5, E13.5 and P0. The embryonic brains were quickly removed and stabilized in 1mL of RNAlater: RNA stabilization reagent (Qiagen) to prevent degradation of RNA by RNases. For P0 brain, part of the cortex was manually dissected immediately after the whole brain collection and also stabilized in RNAlater. All samples were stored at 4°C until used. Total RNA was isolated by extraction with the commercially available QIAGEN RNAeasy kit. The kit used guanidine-isothiocyanate lysis with silica-membrane purification. The residual amount of genomic DNA remaining was removed by treating the extracted RNA with DNaseI (Deoxiribonuclease I; Roche). RNA concentration and purity were evaluated by spectrophotometry. The absence of RNA degradation was confirmed by agarose gel electrophoresis. To maintain the integrity of the RNA until cDNA preparation, the collection tubes were frozen in liquid nitrogen and placed in a -80oC freezer.

2.5.2 cDNA Synthesis

Two μg of total RNA were reverse-transcribed to cDNA using the First-Strand cDNA synthesis using SuperScript II reverse transcriptase (RT) (Invitrogen) protocol. Subsequently, 50 ng of random primer 6 (New England Biolabs) and dNTPs were used to synthesize the cDNA. RNaseOUT was used to remove any remaining RNase and 5X first-strand buffer was used to stabilize the mixture. Finally, SuperScript II RT enzyme was added to the mixture and incubated in a 42oC water bath to produce the final cDNA sample. cDNA samples were stored at -20°C until use.

2.5.3 qPCR

qPCR was conducted using SYBR green with dissociative curves and ROX as a reference dye. SYBR green is a fluorescent dye that binds to double-stranded DNA during PCR amplification. In addition to primers specific to the gene under study (*Dlx1, Dlx2, Dlx5,* or *Dlx6*), PCR amplification was carried out in parallel with a reference house-keeping gene (*GAPDH*). qPCR was performed using the Agilent technologies Stratagene Mx3000P instrument according to the following reaction conditions: 95°C for 10 minutes, followed by 40 cycles of 95°C for 30 seconds, 55°C for 1 minutes and 72°C for 1min and finally 1 minutes at 95°C, 30 seconds at 55°C and 30 seconds at 95°C. Blank controls were performed on each PCR plate. The CT (Cycle Threshold) was automatically calculated and used to quantify the starting copy number of the target mRNA. The amounts of target RNA copies were normalized to the endogenous reference gene, *GAPDH.* Relative expressions of each of the four Dlx genes were determined for each genotype using the Livak method of data interpretation (Livak and Schmittgen 2001). The primers used in qPCR experiments were:

qPCR.GAPDH.for 5'-TGCACCACCAACTGCTTAGC-3'

qPCR.GAPDH.rev 5'-GGCATGGACTGTGGTCATGAG-3'

qPCR.Dlx1.for 5'-CAGTTGCAGGCTTTGAACC-3'

qPCR.Dlx1.rev 5'-ACTTGGAGCGTTTGTTCTGG-3'

qPCR.Dlx2.for 5'-GCCTCACCCAAACTCAGG-3'

qPCR.Dlx2.rev 5'-GCCGCTTTTCCACATCTTC-3',

qPCR.Dlx5.for 5'-CGACTTCCAAGCTCCGTTC-3'

qPCR.Dlx5.rev 5'-TTCTTTCTCTGGCTGGCTG-3'

qPCR.Dlx6.for 5'-CGGACCATTTATTCCAGCC-3'

qPCR.Dlx5.rev 5'-CGCTTATTCTGAAACCATATC-3'

2.5.4 Statistics

Differences in transcript expression were statistically evaluated by Student's t-test. A *P* value < 0.05 was considered significant. All results were expressed as mean ± standard error (SE). All qPCR experiments were done in duplicata.

2.6 Cryostat sectioning of mouse forebrain tissues

Embryonic brains were collected from wild type mice and vI56i mutant mice at E11.5 and E13.5 developmental stages. The brains were quickly removed and fixed in 4% paraformaldehyde (PFA) in PBS 1X overnight at 4°C. They were washed 3 times in 1X PBS and were placed in a 30% sucrose solution overnight for freezing protection and cryostat preparation of the samples. The following day, brain samples were embedded in OCT (VWR International LTD, Poole, England) and immediately frozen at -80°C. Brains were placed in the Leica CM1850 chamber for at least 30 minutes to "supercool" the sample for sectioning. 20µm sections were taken and placed on Superfrost/Plus microscope slides (Fisher) and kept at -80°C until use.

2.7 *In situ* hybridization

2.7.1 Riboprobe

An antisense mRNA probes was labeled with digoxygenin-11-UTP (Roche) and synthesized from the following cDNA clones: mouse *Dlx-5* (Liu, Ghattas et al. 1997). The pKS vector that contained the full-length *Dlx5* cDNA was linearized with *Sal*I and the antisense riboprobe was generated using T7 RNA polymerase (Roche)

2.7.2 Hybridization

The following *in situ* hybridization protocol was adapted from (Smith, Zhang et al. 2008). Frozen slides were thawed for 1 hr at room temperature (RT). Edges were sealed using a pap-pen to keep liquid on the slide. Probes were diluted in hybridization buffer (1× Salt solution [1× Salt: 0.2 M NaCl, 10 mM Tris HCl, 5 mM NaH_2PO_4, 5 mM Na_2HPO_4, 1 mM Tris-base, 5 mM EDTA], 50% deionized formamide, 10× dextran sulphate, 1 mg/ml yeast tRNA, and 1× Denhardts) at a concentration of 1:400. The probe mix was denatured at 70°C for 5–10 min. Slides were placed in a sealed Tupperware box lined with paper towel and Whatman paper soaked with 1× Salt solution in 50% formamide to make a humidification chamber. Three hundred µL of the probe mix were added to each slide that was then covered with a coverslip and placed in an oven at 65°C overnight. For washes, slides were transferred to a Coplin jar and washed 2× 30 min at 65°C in solution A (1× SSC, 50% formamide, and 0.1% Tween 20), followed by 2× 30 min in 1× TBST at room temperature (RT). Slides were then blocked in 10% calf serum in 1× TBST (blocking solution) for >1 hr at RT. For incubation with the antibody, the slides were placed back in a sealed Tupperware box with water-soaked paper towels at the bottom.Three hundred µL of 1:2,000 dilution of anti-DIG AP Fab fragment antibody (Roche) in blocking solution were added to each slide that was then covered with a coverslip and incubated overnight at 4°C. Slides were placed in Coplin jars and washed 4–5× for 20 min in 1× TBST at RT. Slides were then equilibrated for 2× 10 min in NTMT staining buffer (100 mM NaCl, 100 mM Tris HCL pH 9.5, 50 mM $MgCl_2$, and 0.1% Tween 20). Each slide was stained in the dark and was covered with 400 µl of NTMT containing 1.4 µl of BCIP and 2.7 µl of NBT. Staining was performed for 2 hrs. The reaction was stopped by 2–4× 10-min washes in PBST followed by fixation in 4% PFA

overnight at 4°C. Slides were subsequently washed in distilled water and allowed to dry before mounting with coverslips using Aquatex (EMD Chemicals).

2.8 Behavioural tests

2.8.1 Overview

All the following behavioural procedures were adapted from the Animal Care and Veterinary Service (ACVS) standard operating procedure of the behavioural core facility of the faculty of medicine at the University of Ottawa. Behavioural tests were initiated when the mice were 8–11 weeks of age. A cohort of 12 males (4 wild type and 8 vI56i mutant) and a cohort of 16 females (3 wild type and 13 vI56i mutant) were used. All studies were performed by observers who were blinded to the genotype of the mice. Tests were performed in the following sequence: Elevated plus maze, Beambreak, Adult social interaction and Juvenile interaction. Animals were tested during the light phase of the light-dark cycle, except for Beambreak test that was done overnight. Mice were allowed to habituate in the testing room for 1 h with a background noise level set up for 70dB. Prior to the Day of Testing, the tails of the test mice were marked with a Sharpie marker to easily and rapidly identify them. We employed an inter-test rest period of approximately 3-4 days. In experiments where the same testing apparatus was repeatedly utilized for multiple mice, the surface was wiped with an antiseptic solution to eliminate any possible odour cues left by previous subjects. All experiments were performed in accordance with ACVS of university of Ottawa and follow the standard protocol for the behaviour testing.

2.8.2 Elevated plus maze

The elevated plus maze (Hogg 1996), to evaluate anxiety-like behaviour, has two arms, measuring ~5 cm wide and ~60 cm long. The arms are crossed perpendicularly and are open where they meet. One arm (which has two sides) has open platforms; the other arm (which has two sides) has walls (~14.5 cm high) and is enclosed. The plus maze was raised 1 meter above the floor. Mice were tested with the lights fully on to increase the anxiety of being on the open arm. In each trial, mice were placed in the center of the maze and were allowed to freely explore the maze for 10 minutes. Their location within the maze was tracked by videotracking (Ethovision software from Noldus) and analyzed the number of times the animal enters each of the arms and the time spent in each arm.

2.8.3 Beam Break

Locomotor activity is tracked by invisible infrared light beams that do not interfere with the normal mouse behaviour and are located on a metal frame (from automated Accuscan equipment) that surrounds regular mouse cage. The mice were individually placed into test cages that contain just enough clean bedding to cover the floor without interfering with the sensors. Each metal frame was equipped with infrared receptors and emitters that provide a series of one-dimensional infrared beams that, when broken by the animals' movement, allows for the measurement of locomotor activity. Locomotor activity was monitored continuously over a 15hrs period during the night and a tally was recorded every 5 min using micromax analyzer software from Accuscan.

2.8.4 Adult social interaction

The social interaction task consisted of two 5 min long trials performed in darkness under red light with a background noise. An infrared-sensitive videocamera was used for tracking mice

during darkness. In the first trial, the mouse is placed in the corner of the open field box measuring ~45 cm long on each side and ~45cm high and containing a 5.5 X 9.6 cm wire mesh rectangular cage placed along the center of one wall. The mouse is left in the box for 5 min to explore and get familiar with the environment and then removed from the box. During the second trial, the mouse is placed back into the same arena, which now possesses an unfamiliar social target mouse of the same gender, age and strain, within the wire mesh cage. Videotracking (Ethovision from Noldus) was employed to measure the time spent by the mouse interacting with the test mouse in the interaction zone around the wire mesh cage. The interaction zone was defined as a 15x30 cm rectangular area surrounding the wire mesh.

2.8.5 Juvenile interaction

The juvenile interaction test involved 2 mice, a test animal and a juvenile of the same sex. Test mice were allowed to habituate in the testing room for 15 minutes prior the test. Then, the test mouse and the juvenile were placed in a novel cage together at the same time, and allowed to explore for 2 minutes. Total duration of social interactions is scored as cumulative seconds spent by the test mouse in sniffing the nose, anogenital and other body regions of the juvenile. Three days later, the same two mice were placed together again in a new cages for another 2 minutes. The time the test mouse spends interacting with the juvenile was again recorded. The interaction time from first exposure to the juvenile and the second exposure to same juvenile are compared. All the trials were performed in darkness under red light with a background noise. Silent stop watches were used to record the 2 minutes intervals and the cumulative interaction time.

2.8.6 Statistics

Data are presented as the mean ± SEM. Data were analyzed with a two ways ANOVA or repeated measure ANOVA, using SPSS software version 19. All statistical tests were performed at a significance level of 0.05. When there was no significant sex effect, it is noted in text, but data is not shown in graph format.

3 A SNP in an ultraconserved regulatory element affects *Dlx5/Dlx6* regulation in the forebrain

Luc Poitras[1], Man Yu[1], Cindy Lesage-Pelletier[1], Ryan B. MacDonald[1], Jean-Philippe Gagné[2], Gary Hatch[1], Isabelle Kelly[3], Steven P. Hamilton[4, 5], John L.R. Rubenstein[5, 6], Guy G. Poirier[2, 3] and Marc Ekker[1,]

[1] Center for Advanced Research in Environmental Genomics (CAREG), Department of Biology, University of Ottawa, Ottawa, Ontario, Canada.

[2] Centre de Recherche du CHUQ-CHUL, Faculty of Medicine, Laval University, Québec, Canada.

[3] Proteomics Platform of the Québec Genomic Center, Centre de Recherche du CHUQ-CHUL, Laval University, Québec Canada.

[4] Department of Psychiatry, University of California, San Francisco, CA, USA

[5] Institute for Human Genetics, University of California, San Francisco, CA, USA

[6] Nina Ireland Laboratory, University of California, San Francisco, CA, USA

Summary

Dlx homeobox genes play a crucial role in the migration and differentiation of subpallial precursor cells that will give rise to various subtypes of gamma-amino-butyric-acid (GABA)-expressing neurons of the forebrain, including local-circuit cortical interneurons. Aberrant development of GABAergic interneurons has been linked to several neurodevelopmental disorders including epilepsy, schizophrenia, Rett syndrome and autism. Here, we report that a single nucleotide polymorphism (SNP) found in an autistic proband falls within a functional protein binding site in an ultraconserved *cis*-regulatory element. This element, I56i, is involved in regulating *Dlx5/Dlx6* homeobox gene expression in the developing forebrain. We show that the SNP results in lower I56i activity, predominantly in the medial and caudal ganglionic eminences and in streams of neurons tangentially migrating to the cortex. Reduced activity is also observed in GABAergic interneurons of the adult somatosensory cortex. The SNP affects the affinity of Dlx proteins for their binding site *in vitro* and reduces the transcriptional activation of the enhancer by Dlx proteins. Affinity purification using I56i sequences leads to the identification of a novel regulator of *Dlx* expression, the General Transcription Factor 2-I (Gtf2i), which is among the genes most often deleted in Williams-Beuren syndrome (WBS), a neurodevelopmental disorder. This study illustrates the clear functional consequences of a single nucleotide variation in an ultraconserved non-coding sequence in the context of developmental abnormalities associated with disease.

3.1 Introduction

During embryonic development, two major types of neurons will migrate and populate the cortex, the excitatory glutamate-releasing neurons and the inhibitory γ–Amino-Butyric Acid (GABA)-releasing interneurons. It has been proposed that some forms of autism are the results of an imbalance in the excitation/inhibition ratio of the cortical neurons leading to an increase in the excitatory state of the brain (Rubenstein and Merzenich 2003). This excitatory state could be achieved by a reduction in the inhibitory signaling or an increase in the excitatory system.

The differentiation of multipotent neuronal progenitors into various neuronal cell types requires complex genetic regulatory mechanisms. For instance, the differentiation and proper migration of the GABAergic interneurons to the cortex is controlled to a large extent by homeodomain-containing transcription factors of the *Dlx* gene family (Anderson, Eisenstat et al. 1997; Anderson, Qiu et al. 1997). Four *Dlx* genes, *Dlx1*, *Dlx2*, *Dlx5*, and *Dlx6* are expressed in two distinct forebrain domains, the ventral telencephalon and parts of the diencephalon (Robinson, Wray et al. 1991; Bulfone, Puelles et al. 1993). Their expression in the ventral telencephalon is restricted to the differentiating GABAergic projection neurons and interneurons, which will later migrate to the cortex and olfactory bulb (Anderson, Eisenstat et al. 1997; Stuhmer, Anderson et al. 2002; Stuhmer, Puelles et al. 2002).

The *Dlx* genes are organized in convergently transcribed bigene clusters and have highly similar patterns of expression possibly due to shared *cis*-acting regulatory elements (CREs) (Porteus, Bulfone et al. 1991; Price, Lemaistre et al. 1991; Robinson, Wray et al. 1991; Robinson and Mahon 1994; Simeone, Acampora et al. 1994; Scherer, Heng et al. 1995; Nakamura, Stock et al. 1996; Stock, Ellies et al. 1996). Through phylogenetic footprinting

and transgenic analysis, two CREs were discovered in the *Dlx5*/*Dlx6* intergenic region, called I56i and I56ii (Zerucha, Stuhmer et al. 2000). In reporter assays, both elements were active in the telencephalon and diencephalon. In-depth analysis of their enhancer activities has shown that these individual enhancers may participate in *Dlx* regulation during the development of specific populations of GABAergic interneurons (Ghanem, Yu et al. 2007; Ghanem, Yu et al. 2008).

The *DLX* genes have been associated with human autism spectrum disorders (ASDs) (Hamilton, Woo et al. 2005). Two *DLX* bigene clusters, *DLX1/DLX2* and *DLX5/DLX6*, are found in autism susceptibility loci on chromosomes 2q31.1 and 7q21.3. Two recent linkage studies have provided evidence suggesting that these two loci can be linked to ASDs (Liu, Novosedlik et al. 2009; Nakashima, Yamagata et al. 2009). Liu and collaborators showed that two polymorphisms found in the *DLX1/DLX2* loci could potentially increase susceptibility, or cause, autism in cohorts of multiplex families where there was a greater genetic component for these conditions (Liu, Novosedlik et al. 2009). In a search for autism genomic variants in the *DLX1/2* and *DLX5/6* loci, we identified thirty-one single nucleotide polymorphisms (SNPs) and two insertion/deletion polymorphisms. One adenine to guanine SNP was found at position 182 of I56i (Hamilton, Woo et al. 2005). The location of the SNP coincides with one of the two putative DLX binding sites (Feledy, Morasso et al. 1999) within I56i (Zerucha, Stuhmer et al. 2000). This region of I56i also codes for the second exon of the long polyadenylated non-coding RNA (lpncRNA) *Evf2* (Feng, Bi et al. 2006), which was recently shown to contribute to the mechanisms by which the *Dlx5/Dlx6* bigene cluster is regulated (Bond, Vangompel et al. 2009). Furthermore, the I56i CRE was identified as one of 481 ultraconserved elements of the mouse genome (i.e., 100% identity with no insertion or deletions on DNA fragments longer than 200bp between orthologous regions of

the human, rat and mouse genomes) (Bejerano, Pheasant et al. 2004). The 182-SNP was found to be part of a completely conserved 8 bp motif based on the sequence conservation between forty vertebrate genomes (Fig 3.1). Therefore, the unexpected identification of this SNP in I56i enhancer suggested it could have functional consequences for *Dlx5/Dlx6* regulation during development.

Here, we present evidence suggesting that the I56i SNP affects the transcriptional regulation of the *Dlx5/Dlx6* bigene cluster through the I56i CRE. Enhancer activity is impaired by the I56i SNP leading to a decrease in the expression of a reporter gene in the developing telencephalon as well as in different subpopulations of GABAergic neurons in the adult cortex. Our *in vivo* and *in vitro* assays strongly suggest that Dlx proteins have a reduced affinity for the variant site. Affinity purification confirm that Dlx proteins are bound at or near the SNP sequence and also lead to the identification of Gtf2i as a new regulator of the *Dlx5/Dlx6* bigene cluster. We propose that the impaired I56i enhancer activity resulting from the 182-SNP could has impacts on *Dlx* function and therefore the development of cortical interneuron development and could constitute one contributing factor to developmental

Figure 3.1 : The SNP found in I56i is located in a highly conserved sequence. Forty vertebrate genome sequences in the vicinity of the targeted variant (shown in bold) were aligned. Non-consensus bases are underlined, and inserted bases in tenrec and medaka are indicated as gaps in all other species. This region corresponds to bases 96,641,417-96,641,443 of chromosome 7 from the February 2009 human genome assembly (GRCh37).

```
Human           ttaaatgcagcc**a**taattaga-gtaa-tt
Chimp           ttaaatgcagcc**a**taattaga-gtaa-tt
Orangutan       ttaaatgcagcc**a**taattaga-gtaa-tt
Rhesus          ttaaatgcagcc**a**taattaga-gtaa-tt
Marmoset        ttaaatgcagcc**a**taattaga-gtaa-tt
Tarsier         ttaaatgcagcc**a**taattaga-gtaa-tt
Mouse lemur     ttaaatgcagcc**a**taattaga-gtaa-tt
Bushbaby        ttaaatgcagcc**a**taattaga-gtaa-tt
Mouse           ttaaatgcagcc**a**taattaga-gtaa-tt
Rat             ttaaatgcagcc**a**taattaga-gtaa-tt
Kangaroo rat    ctaaatgcagcc**a**taattaga-gtaa-tt
Guinea Pig      ttaaatgcagcc**a**taattaga-gtaa-tt
Squirrel        ttaaatgcagcc**a**taattaga-gtaa-tt
Rabbit          ttaaatgcagcc**a**taattaga-gtaa-tt
Pika            ttaaatgcagcc**a**taattaga-gtaa-tt
Alpaca          ttaaatgcagcc**a**taattaga-gtaa-tt
Dolphin         ttaaatgcagcc**a**taattaga-gtaa-tt
Cow             ttaaatgcagcc**a**taattaga-gtaa-tt
Horse           ttaaatgcagcc**a**taattaga-gtaa-tt
Cat             ttaaatgcagcc**a**taattaga-gtaa-tt
Dog             ttaaatgcagcc**a**taattaga-gtaa-tt
Microbat        ttaaatgcagcc**a**taattaga-gtaa-tt
Megabat         ttaaatgcagcc**a**taattaga-gtaa-tt
Hedgehog        ttaaatgcagcc**a**taattaga-gtaa-tt
Elephant        ttaaatgcagcc**a**taattaga-gtaa-tt
Rock hyrax      ttaaatgcagcc**a**taattaga-gtaa-tt
Tenrec          ttaaatgcagcg**a**taattagaggtaa-gt
Armadillo       ttaaatgcagcc**a**taattaga-gtaa-tt
Sloth           ttaaatgcagcc**a**taattaga-gtaa-tt
Opossum         ttaaatgcagcc**a**taattaga-gtaa-tt
Platypus        ttaaatgcagcc**a**taattaga-gtaa-tt
Chicken         ttaaatgcagcc**a**taattaga-gtaa-tt
Zebra finch     ttaaatgcagcc**a**taattaga-gtaa-tt
Lizard          ttaaatgcagcc**a**taattaga-gtaa-tt
X. tropicalis   ttaaatgcagcc**a**taattaga-gtaa-tt
Tetraodon       ataaatgcaggc**a**taattagg-gtaa-tt
Fugu            ataaatgcaggc**a**taattagg-gtaa-tt
Stickleback     ataaatgcaggc**a**taattagg-gtaa-tt
Medaka          ataaatgcaggc**a**taattagg-gtaagtt
Zebrafish       ataaatgcagac**a**taattagg-gtaa-tt
```

3.2 Materials and Methods

3.2.1 Mutagenesis of mouse I56i.

The 182 SNP-containing I56i enhancer was generated by PCR using the overlapping fragment technique. A first fragment was amplified from the mouse I56i enhancer using the following oligonucleotides I56i-1.for and I56i-316.rev (all oligonucleotide sequences are listed in the **Table S1**). A second overlapping fragment was generated using an oligonucleotide containing the desired mutation I56i SNP in combination with the I56i-436 primer. After purification by agarose gel electrophoresis, the two fragments were used as templates for a final PCR reaction with I56i-1.for and I56i-436.rev. A mix (1:3) of *Taq* and *Pfx* DNA polymerase (Invitrogen) was used to avoid unwanted mutations. Finally, the mutant enhancer was purified on gel, TA-cloned and sequenced.

3.2.2 Transgenic experiments.

The I56i and vI56i enhancers were subcloned into the *Xho*I site of the p1230 vector (Yee and Rigby 1993) that contains a human β-*GLOBIN* minimal promoter and a *lacZ* reporter gene. Transgenic animals were produced and analyzed as previously described (Zerucha, Stuhmer et al. 2000). Two transgenic lines and three primary transgenic embryos were obtained for the vI56i-*lacZ* construct. These transgenic embryos were compared to two mouse transgenic lines obtained with the I56i-*lacZ* construct (Zerucha, Stuhmer et al. 2000; Ghanem, Jarinova et al. 2003; Ghanem, Yu et al. 2007).

3.2.3 Histology and double imunohistochemistry.

Fixation of E13.5 embryonic telencephalon was carried out overnight in 4% PFA in 1xPBS. For adult brains, mice (P30) were deeply anesthetized and underwent intracardiac perfusion

with 10% saline solution followed by 4% PFA in 1×PBS. The adult brains were removed and post-fixed in 4% PFA for 2h at room temperature. After fixation, all brains were cryoprotected by immersion in 30% sucrose, frozen in OCT (Tissue-Tek) on dry ice, and stored at −80°C until use. For single and double immunohistochemistry, E13.5 brains were cut at 60 μm and adult brains at 40-45 μm on the cryostat (Leica CM3050 S). Double immunostaining was performed as previously described (Ghanem, Yu et al. 2007). An antibody from Abnova was used for Gtf2i immunohistochemistry (IHC-00273).

3.2.4 Nuclear extract preparation.

Nuclear extracts were prepared from the brains of E13.5 mouse embryos (1ml of tissue). Preparation of nuclear extracts was carried out as described by Sambrook et al. (Sambrook, Fritsch et al. 1989). In addition, the protein extract was dialyzed against the resuspension buffer using a Spectra/Por dialysis cuvette (VWR) overnight at 4^{o}C. Aliquots of 50 μL were frozen in liquid nitrogen and stored at -80^{o}C.

3.2.5 DNase I footprinting analysis.

A portion of the I56i enhancer corresponding to nucleotides 187-316 was amplified by PCR using the following primers: zI56i-187 and zI56i-316. PCR fragments were TA-cloned in the pDrive vector (Qiagen). This fragment was recovered by digestion with *Eco*RI and *Kpn*I followed by gel purification (QIAquick gel purification). Directional labeling of each fragment was done by 5' end fill-in using the large fragment of DNA polymerase I (Invitrogen). Footprint reactions were carried out as previously described (Poitras, Ghanem et al. 2007).

3.2.6 Protein-DNA complexes purification

Nuclear protein extracts from E13.5 mouse embryos were incubated with 2 μg of a double-stranded biotinylated oligonucleotide encompassing the two Dlx binding sites (See mI56i FP3-FP5. For and .Rev in **Table S1**) in 1X EMSA binding buffer (7 mM Tris pH 7.5, 81 mM NaCl, 2.75 mM dithiothreitol, 5 mM $MgCl_2$, 0.05% NP-40, 1 mg/mL bovine serum albumin, 25 μg/mL poly(dIdC), 10% glycerol) at room temperature for 60 minutes. After 30 minutes of incubation, streptavidin-coated agarose (Sigma, #S1638) was added to the reaction. Protein-DNA complexes coupled to the streptavidin sepharose beads were washed with 1X EMSA binding buffer. Elution of the protein components of the complexes was carried out by boiling the reaction mixture in Laemmli loading buffer (final concentration of 2% SDS, 10% glycerol, 5% 2-mercaptoethanol, 0.002% bromophenol blue and 0.0625M Tris-HCl). Removal of the sepharose beads was achieved by filtering the boiled reaction through a SPIN-X® centrifuge tube filter (Corning).

3.2.7 Sample preparation for mass spectrometry analysis

Eluted proteins were resolved using 4-12% Criterion™ XT Bis-Tris gradient gel (Bio-Rad) and stained with Sypro Ruby (Invitrogen) according to the manufacturer's instructions. Images were acquired on a Geliance CCD-based bioimaging system (PerkinElmer).

3.2.8 LC-MS/MS analysis

The entire protein profile on SDS-PAGE was sliced from the gel into 25 bands using a gel excision Lanepicker™ (The Gel Company). In-gel protein digestion was performed on a MassPrep™ liquid handling station (Micromass) according to the manufacturer's specifications and using sequencing-grade modified trypsin (Promega). Peptide extracts were dried out using a SpeedVac™ .

Peptide extracts were separated by online reversed-phase (RP) nanoscale capillary liquid chromatography (nanoLC) and analyzed by electrospray mass spectrometry (ES MS/MS). The experiments were performed on a Thermo Surveyor MS pump connected to a LTQ linear ion trap mass spectrometer (Thermo Electron, San Jose, CA) equipped with a nanoelectrospray ion source (Thermo Electron, San Jose, CA). Peptide separation took place within a PicoFrit column BioBasic C18, 10 cm x 0.075 mm internal diameter (New Objective, Woburn, MA) with a linear gradient from 2 to 50% solvent B (acetonitrile, 0.1% formic acid) in 30 min, at 200 nL/min. Mass spectra were acquired using data dependent acquisition mode (Xcalibure software, version 2.0). Each full-scan mass spectrum (400 to 2000 m/z) was followed by collision-induced dissociation of the seven most intense ions. The dynamic exclusion function was enabled (30 sec exclusion), and the relative collisional fragmentation energy was set to 35%.

3.2.9 Interpretation of Tandem Mass Spectrometry Spectra

All MS/MS samples were analyzed using Mascot (Matrix Science, London, UK; version 2.2.0). Mascot was set up to search against rodent Uniref_100 protein database [TaxID: 9989] assuming a digestion with trypsin. Fragment and parent ion mass tolerance were respectively of 0.5 Da and 2.0 Da. Iodoacetamide derivative of cysteine was specified as a fixed modification and oxidation of methionine was specified as variable modification. Two missed cleavages were allowed.

3.2.10 Criteria for protein identification

Scaffold (version 02_02_03, Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications. Peptide identifications were accepted if

they could be established at greater than 95.0% probability as specified by the Peptide Prophet algorithm (Keller, Nesvizhskii et al. 2002). Protein probabilities were assigned by the Protein Prophet algorithm (Nesvizhskii, Keller et al. 2003). Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Using these stringent identification parameters, the rate of false positive identifications is less than 1%.

3.2.11 Co-transfection experiments.

A zebrafish *dlx2a* cDNA (845bp) encompassing the full coding sequence was amplified by PCR and cloned into the *EcoRI* site of the pcDNA3-FLAG expression vector. The construction of the DLX2A homeodomain mutant was previously described in (Poitras, Ghanem et al. 2007). The zebrafish *dlx5a* cDNA (1.2kb) was amplified by PCR and cloned into the *Bgl*II/*Xho*I sites of the pTL2 expression vector. The rat Gtf2i cDNA (accession # BC085815) subcloned in the expression vector pExpress-1 was purchased from Open Biosystem.

The wild-type and 182-SNP variant versions of the mouse I56i enhancer were PCR-amplified and inserted upstream of a minimal promoter driving expression of the synthetic *firefly luciferase* (*luc2*) gene (pGL4.23 vector from Promega). Cell culture and transfection were carried out as previously described (Zerucha, Stuhmer et al. 2000). Luciferase assays were performed according to the manufacturer's protocol.

3.2.12 Gel mobility shift assay.

Gel mobility shift assays were performed using double-stranded oligonucleotides corresponding to the identified putative binding site: complementary 21bp oligonucleotides (mI56i FP3.for and mI56i FP3.rev) were annealed and end-labeled using [^{32}P] γ-ATP

(Amersham) and T4 Polynucleotide kinase (New England Biolabs). For each binding reaction, 1 ng of labeled DNA was incubated with recombinant Flag-Dlx2 (TNT Quick Coupled Transcription/Translation Kit, Promega), GST-Dlx proteins or with nuclear extracts in reaction buffer. Electromobility shift assays were carried out as previously described (Poitras, Ghanem et al. 2007).

In the competition assay using nuclear extract, increasing amounts of unlabeled competitor (0, 5, 20, 50 and 100 ng) were added to the binding reaction. As a competitor, an unlabeled double stranded wt oligonucleotide competitor or a same-length version containing the 182-SNP (see Fig3.2A) were used.

The competition assay with recombinant Dlx proteins was carried out using double-stranded ^{32}P-labelled oligonucleotides corresponding to a triplicate of the core FP3 site (ATAATT), I56i FP3 X3, with fixed amounts of GST-zdlx5a protein. These binding reactions were competed with increasing amounts (1, 3 and 10 ng) of the unlabeled double-stranded wt oligonucleotide competitor (I56i FP3 X3) or the variant version, vI56i FP3 X3. Quantification of the percentage of the remaining complexes was achieved by using the ImageJ software (http://rsbweb.nih.gov/ij/index.html).

For the supershift assay, the Flag-Dlx2 unpurified TNT reaction was preincubated overnight at 4°C with 1 μg of anti-Flag antibody (F7425, Sigma).

3.2.13 Western blots

Proteins from the previously described E13.5 nuclear extract were separated on an 8% SDS acrylamide gel and electrophoretically transferred to Hybond-C membrane (GE healthcare) in 25 mM Tris, 192 mM glycine, and 20% methanol. Membranes were blocked in PBS-T (145 mM NaCl, 10 mM Na-phosphate buffer, pH 7.4, 0.2% Tween-20) containing 2.5% dry

milk. The membranes were then incubated with an anti-Gtf2i (Abnova, IHC-00273)(1/1000) in PBS-T containing 1% dry milk overnight at 4^{o}C. After washing with PBS-T/1% dry milk, blots were incubated with horseradish peroxidase-conjugated goat anti-rabbit IgG (1/5000) in PBS-T/1% dry milk. The blots were developed by using Immun-Star HRP chemiluminescence substrate kit.

3.2.14 Statistical Analysis

Statistical significance (*P* values < 0.05) presented in Fig 3.6C and 3.6D were determined using paired Student's *t* test. All quantitative data were presented as the mean ± S.E.M.

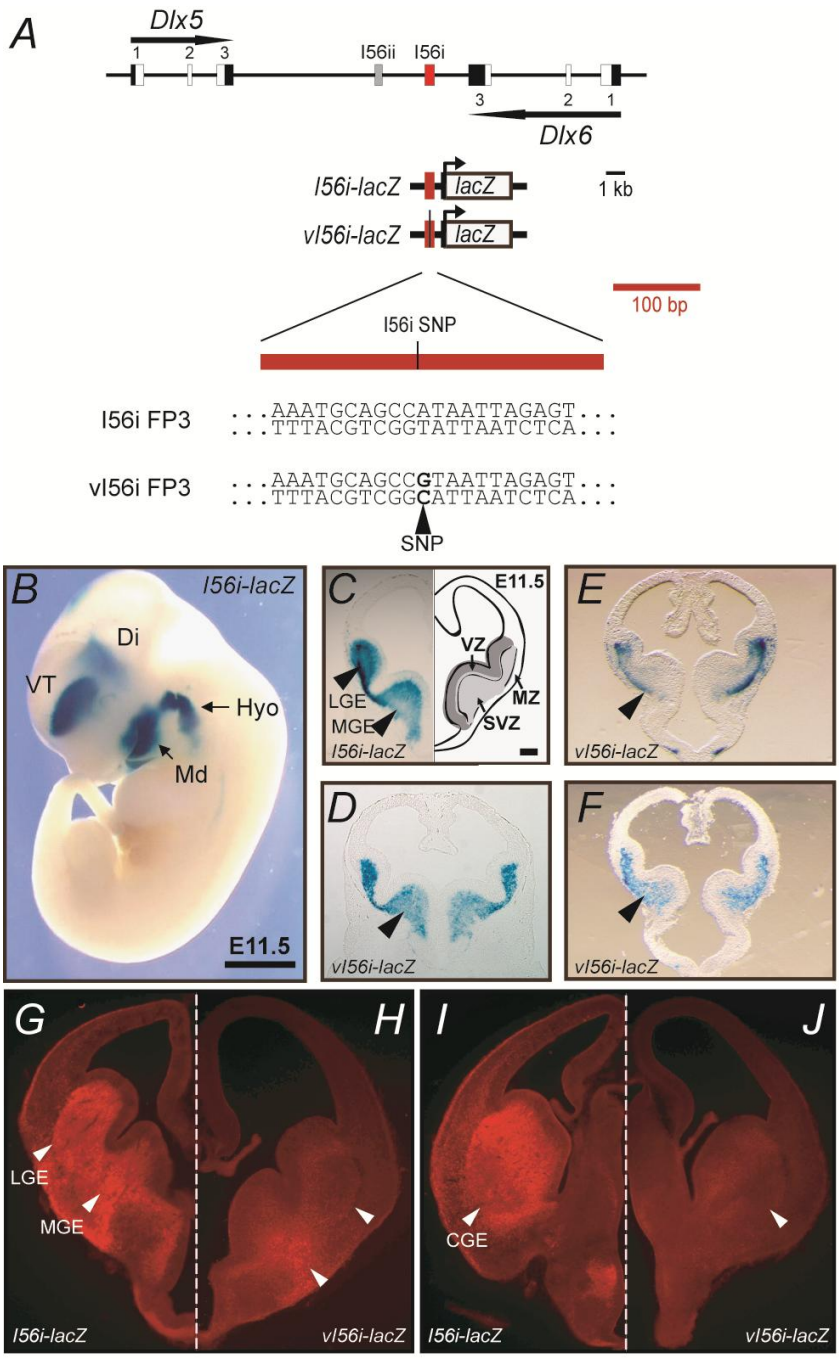
3.3 Results

3.3.1 A SNP affects the activity of the I56i enhancer.

We previously identified seven variants in CREs of *Dlx* genes from a collection of DNA samples from autistic probands (Hamilton, Woo et al. 2005). Using sequence conservation analysis in orthologous sequences (Fig. S1) and DNaseI footprinting to localize binding of trans-acting ((Poitras, Ghanem et al. 2007) data not shown, see also Fig. 3.4A), we found that position 182 in I56i (base 96,641,429 on chromosome 7, Fig. 3.2A), changing an adenine for a guanine, was the only sequence variant predicted to have a major impact on CRE function. By comparing the nucleotide sequence of the I56i enhancer from forty vertebrate genomes, we found that this adenine was part of a completely conserved 8 bp motif (**A**TAATTAG) (Fig. 3.1).

To evaluate potential effects of the 182-SNP on the *in vivo* enhancer activity of I56i, we produced transgenic mouse embryos with reporter constructs that contained either the wild-type I56i or its corresponding 182-SNP-containing variant (hereafter referred to as vI56i) in combination with a β-*GLOBIN* minimal promoter-*lacZ* cassette (Yee and Rigby 1993). As previously reported, the I56i enhancer targets the expression of the reporter cassette to the developing forebrain and branchial arches (Zerucha, Stuhmer et al. 2000; Ghanem, Jarinova et al. 2003; Ghanem, Yu et al. 2007)(Fig. 3.2B). In the telencephalon of E11.5 embryos, I56i targets *lacZ* expression to the subventricular (SVZ) and mantle zone (MZ) of the medial and lateral ganglionic eminences (MGE and LGE) (Fig. 3.2C). By contrast, *lacZ* expression driven by vI56i-*lacZ* was reduced in all aforementioned regions of the forebrain. A more pronounced effect was observed in the MGE (Fig. 3.2D-F). The reductions in transgene expression varied when comparing embryos from two independent transgenic lines and three

Figure 3.2 : The 182-SNP affects the I56i enhancer activity in the developing forebrain. (A) *Dlx5/Dlx6* bigene cluster with exons in black and white (coding), the enhancers I56i in red and I56ii in grey. The I56i enhancer was enlarged to show the position of the 182-SNP variant identified by Hamilton and collaborators (Hamilton, Woo et al. 2005). Sequence of the 182-SNP variant (vI56i) and wild-type I56i. The rest of the mouse and human I56i sequences are virtually identical (Zerucha et al., 2000). (B) At E11.5, the mouse I56i enhancer drives *lacZ* reporter expression to the ventral telencephalon (VT), diencephalon (Di) and, to the hyoid (Hyo) and mandibular (Md) arches. Transverse sections through the telencephalon of *I56i-lacZ* (C, G, I) and vI56i-*lacZ* (D, E, F, H and J) embryos at E11.5 (C-F) and E13.5 (G-J). (A-F) show β-galactosidase assay staining and G-J show immunohistochemistry for β-galactosidase. (E-F) are sections from primary transgenic embryos. Black arrowheads in (C, D, E, F) and white arrowheads in (I, J) indicate reduced vI56i enhancer activity in the MGE at E11.5 and in the MGE, LGE and CGE at E13.5, respectively. The asymmetric staining observed in E-F could be attributed to tilted sectioning and not the result of an asymmetric expression of *lacZ* transgene similar to the transgenic experiments reported by Lettice and collaborators on point mutation in the ZRS enhancer of the *Shh* gene (Lettice, Hill et al. 2008). Scale bars: B, 1mm; C-H, 500μm.

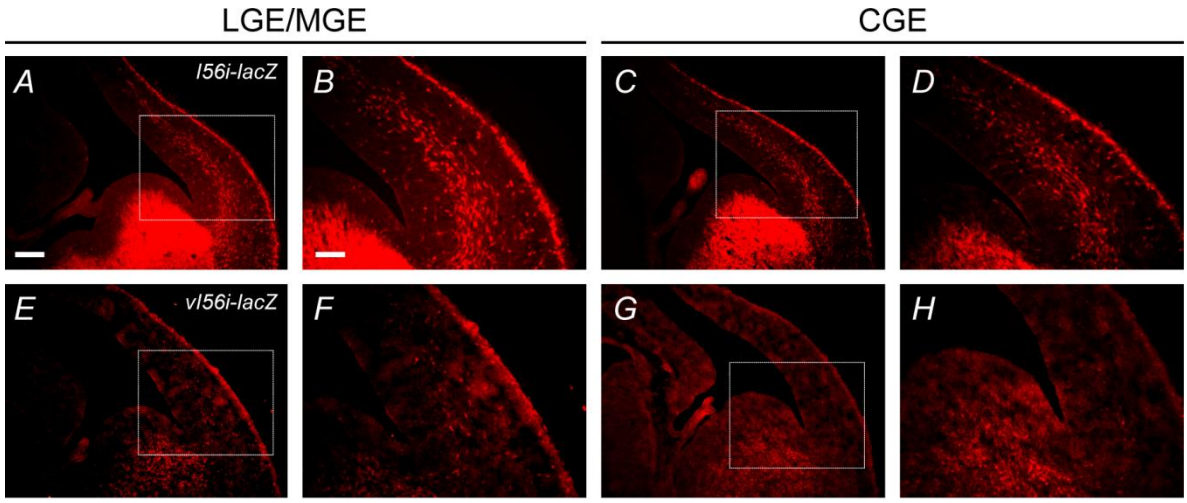

A
Dlx5
1 2 3
I56ii I56i
3 2 1
Dlx6
I56i-lacZ
lacZ
1 kb
vI56i-lacZ
lacZ
100 bp
I56i SNP
I56i FP3 ...AAATGCAGCCATAATTAGAGT...
...TTTACGTCGGTATTAATCTCA...
vI56i FP3 ...AAATGCAGCCGTAATTAGAGT...
...TTTACGTCGGCATTAATCTCA...
SNP
B
I56i-lacZ
Di
VT
Hyo
Md
E11.5
C
E11.5
VZ
MZ
SVZ
LGE
MGE
I56i-lacZ
E
vI56i-lacZ
D
vI56i-lacZ
F
vI56i-lacZ
G
H
LGE
MGE
I56i-lacZ
vI56i-lacZ
I
J
CGE
I56i-lacZ
vI56i-lacZ

primary transgenic animals, ranging from slight reductions to an almost total absence of transgene expression (Fig. 3.2C-F and data not shown). As we never observed such variations in activity in at least 9 primary transgenic embryos and 4 established lines expressing I56i-*lacZ* (Zerucha et al., 2000 and data not shown), it is most unlikely that the decreases we observed with vI56i-*lacZ* are attributable to integration or to copy number effects.

The activity of the variant enhancer was more severely affected at E13.5 than at E11.5 (Fig. 3.2G-H). Moreover, almost no vI56i-*lacZ* -expressing cells could be detected in the SVZ and MZ of the caudal ganglionic eminence (CGE) (Fig. 3.2I-J). Reductions in enhancer activity were also observed in more rostral sections of the telencephalon (data not shown).

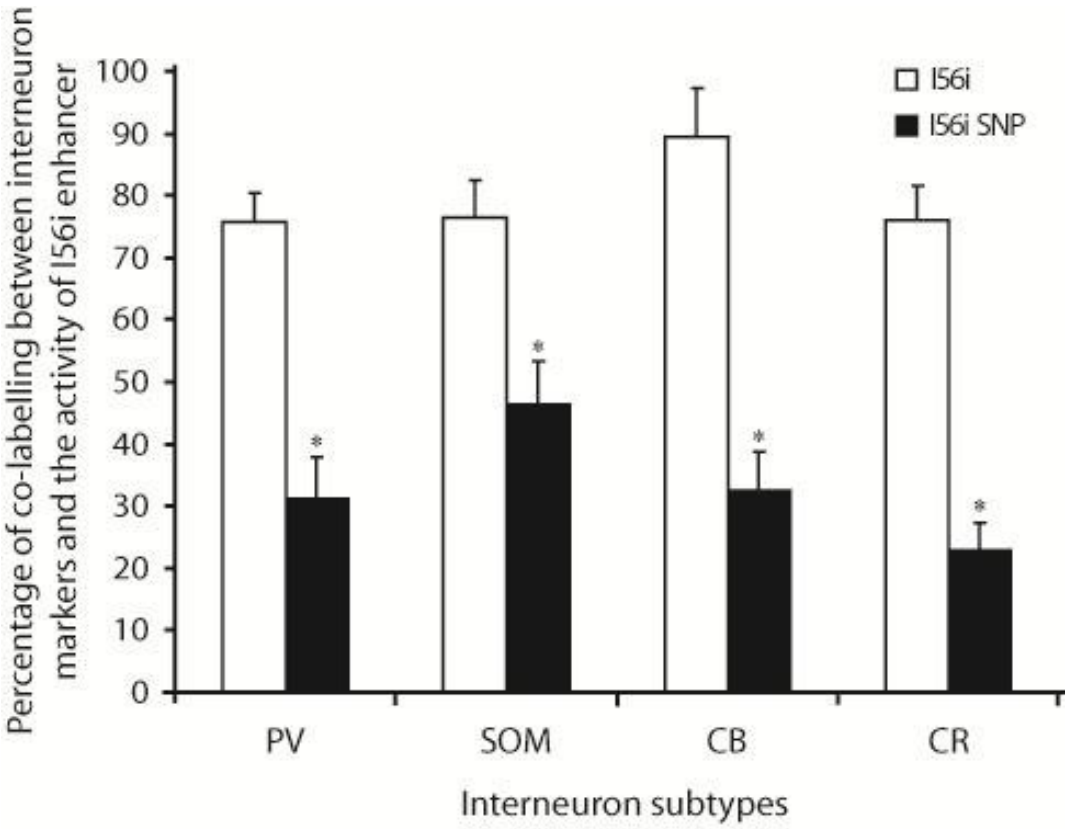
We have previously shown that the transcriptional activity of the I56i enhancer could be observed in GABAergic interneurons migrating tangentially from the ventral telencephalon to the cortex (Ghanem, Yu et al. 2007). Hence, we evaluated the activity of the vI56i-*lacZ* transgene in migrating interneurons at E13.5. In the brains of I56i-*lacZ* embryos, numerous *lacZ*-expressing neurons, organized in streams, were observed migrating towards the pallium at the levels of the mid-telencephalon (LGE/MGE) and caudal telencephalon (CGE) (Fig. 3.3A-D). By contrast, only a few scattered migrating cells expressed the vI56i-*lacZ* transgene at both levels of the embryonic telencephalon (Fig. 3.3E-H).

Figure 3.3: I56i enhancer activity is impaired by the 182-SNP in migrating neuronal precursors at E13.5. Immunolocalization of β-galactosidase (shown in red) in *I56i-lacZ* (A-D) and *vI56i-lacZ* (E-H) transgenic brains at the level of the LGE/MGE (A, B, E, F) and CGE (C, D, G, H). Panels B, D, F, H are higher-magnification pictures of boxes shown in panels A, C, E, G. Numerous *lacZ*+ neurons were observed migrating from the MGE and CGE toward the developing cortex in *I56i-lacZ* brains (A-D), whereas, only few scattered *lacZ*-expressing cells were detected in the migrating streams leaving both eminences from the *vI56i-lacZ* brains (E-H). Scale bars = 500μm in A, C, E and G; 250μm in B, D, F and H.


LGE/MGE
CGE
A
l56i-lacZ
B
C
D
E
vl56i-lacZ
F
G
H

To determine whether the effect of the 182-SNP on I56i enhancer activity persists after birth, we performed double immunohistochemistry on the somatosensory cortex of P35 transgenic mice using cortical interneuron markers. We co-labeled *lacZ*-expressing cells with markers for four majorsubtypes of interneurons: the calcium-binding proteins calbindin, calretinin and parvalbumin (CB, CR and PV) as well as the neuropeptide somatostatin (SOM). These four markers were specifically chosen on the basis of previous findings showing that most CB, PV and SOM+ interneurons mainly originate from the MGE whereas the CGE is one major origin of CR+ interneurons (Wonders and Anderson 2006; Fogarty, Grist et al. 2007; Wonders, Taylor et al. 2008). We observed a significant decrease in co-labeling in the somatosensory cortex of the vI56i-*lacZ* mice as compared to the I56i-*lacZ* mice (Fig. 3.4). The mutant vI56i-*lacZ* transgene was expressed in 32.4% of CB+, 22.8% of CR+, 31.2% of PV+ and 46.5% of SOM+ interneurons. In contrast, the wild type I56i-*lacZ* was found in 89.5 % of CB+, 75.9% of CR+, 75.8% of PV+ and 76.4% of SOM+ interneurons. Taken together, these results suggest that the 182-SNP has important functional consequences on I56i enhancer activity at both embryonic and postnatal stages of brain development. Therefore, we propose that a mouse homozygous for the 182-SNP mutation may present reduced levels of *Dlx5/Dlx6* expression, and this may in turn affect the development of cortical interneurons

Figure 3.4: I56i enhancer activity is impaired by the 182-SNP in the adult cortex. Quantification of the percentage of *lacZ*-expressing cells that co-express various GABAergic interneuron subtypes (PV, parvalbumin; SOM, somatostatin; CB, calbindin; CR, calretinin) in the adult somatosensory cortex at P35. Histograms show that the contribution of vI56i-*lacZ* expressing precursors to adult cortical interneuron populations expressing PV, SOM, CB and CR was significantly reduced when compared to that of I56i-*lacZ* expressing cells (Ghanem, Yu et al. 2007). All data are presented as mean ± SEM. *: $P < 0.001$.

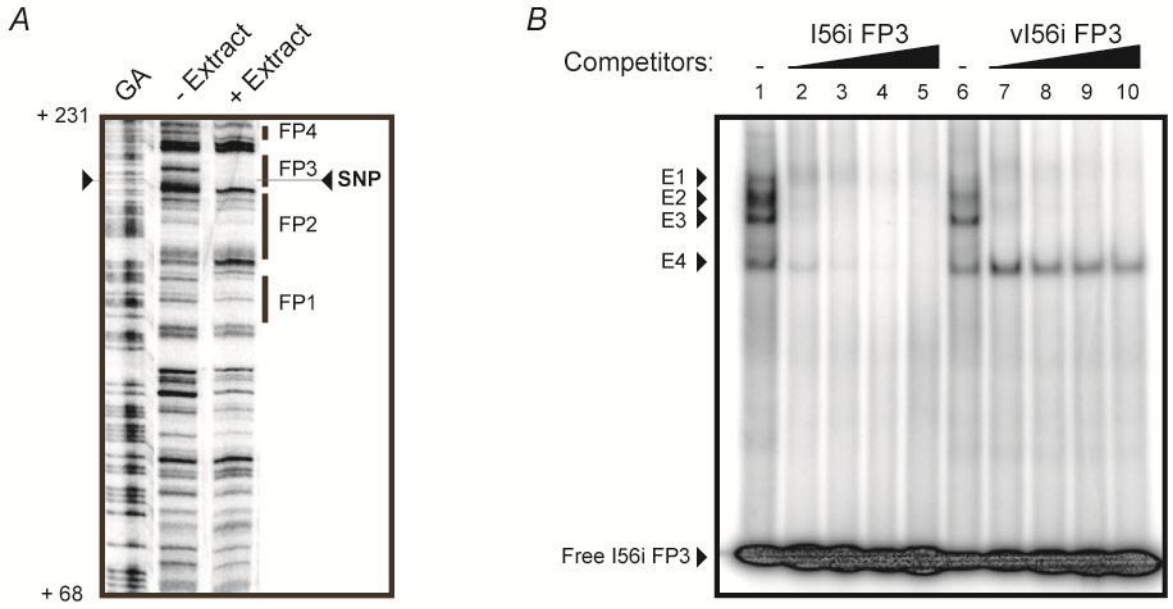

Percentage of co-labelling between interneuron markers and the activity of I56i enhancer
100
90
80
70
60
50
40
30
20
10
0
I56i
I56i SNP
*
*
*
*
PV
SOM
CB
CR
Interneuron subtypes

3.3.2 Binding of a forebrain transcription factor is affected by the presence of the 182-SNP.

To characterize the DNA-protein interactions taking place on the I56i enhancer, we performed a DNaseI footprinting analysis using mouse E13.5 telencephalon nuclear protein extract. Seven DNA-protein complexes were observed, four of which are shown in Fig. 3.5A. The 182-SNP is located within the boundary of DNA-protein complex #3 (FP3). To examine the effects of the 182-SNP on the binding of forebrain transcription factors to the I56i enhancer, a competitive electrophoretic-mobility shift assay (EMSA) was performed. When a 21 bp double-stranded oligonucleotide corresponding to the FP3 binding site (Fig. 3.2A) was incubated with the E13.5 telencephalon nuclear extract, four DNA-protein complexes were observed (Fig. 3.5B, E1-E4). As increasing amounts of the unlabeled I56i-FP3 oligonucleotide were added to the reaction, all four complexes were abolished. When the unlabeled vI56i-FP3 oligonucleotide (Fig. 3.2A) was used as a competitor, three DNA-protein complexes were efficiently competed. However, the E4 complex was unaffected by the vI56i-FP3 oligonucleotide. Interestingly, the molecular weight of the E3 and E4 complexes were shown to be comparable to the complexes produced by a recombinant Dlx2 (Fig. 3.6A). This suggests that the 182-SNP is interfering, *in vitro*, with the formation of a DNA-protein complex, which either includes Dlx protein(s) or proteins that are similar in size.

To identify the potential factors that could bind to the I56i at and near the FP3 binding site, we carried out an affinity purification assay using a DNA fragment encompassing the two Dlx binding sites of the I56i enhancer and nuclear proteins prepared from E13.5 mouse embryonic telencephalon. Nuclear proteins that were bound to the affinity matrix were identified by mass spectrometry (Fig. S2).

Figure 3.5: Binding of a forebrain transcription factor to I56i is affected by the 182-SNP. (A) The 182-SNP is located within a nuclear protein binding site (FP3) in the I56i enhancer as shown by DNaseI footprinting analysis. A DNA fragment corresponding to the I56i enhancer (nt 1-276) was incubated with a nuclear protein extract from E13.5 mouse forebrain (+ Extract). Four regions of the enhancer (FP1 to FP4) were bound by nuclear proteins. Protected areas (footprints) are represented by thick black bars. Protein-DNA interactions were mapped on the mouse I56i enhancer sequence using Maxam-Gilbert Guanine/Adenine chemical sequencing reaction (GA). The 182-SNP was pinpointed within the protected area FP3. Numbers on the left side of each panel indicate the relative positions on the I56i sequence. (B) A competitive EMSA on I56i FP3 was performed using E13.5 embryonic forebrain nuclear extract. In the presence of nuclear extract, four protein-DNA complexes are observed (lane 1 and 6, complexes E1 to E4). These four complexes were competed by increasing amounts of the unlabeled I56i FP3 oligonucleotide (lane 2 to 5, I56i FP3). In competition using the SNP-containing vI56i FP3 oligonucleotide (see Fig. 1A for sequence), protein-DNA complexes E1 to E3 were competed whereas the lower mobility complex E4 was not affected (lane 7 to 10, vI56i FP3).


A
GA
- Extract
+ Extract
+ 231
+ 68
FP4
FP3
SNP
FP2
FP1
B
Competitors:
I56i FP3
vI56i FP3
-
1 2 3 4 5
-
6 7 8 9 10
E1
E2
E3
E4
Free I56i FP3

Consistent with previous observations that Dlx2 is able to bind and activate transcription through the I56i enhancer (Zerucha, Stuhmer et al. 2000) (Zhou, Le et al. 2004), Dlx proteins were identified in the list of purified proteins. However, the most abundant protein isolated from the purification was the General Transcription Factor-II (Gtf2i) with forty-nine distinct peptides identified by mass spectrometry (Fig. S2). Interestingly, a potential Gtf2i binding site (Chimge, Makeyev et al. 2008) is located halfway between the two Dlx binding sites of the sequence used for affinity purification (Fig. S1).

3.3.3 Transcriptional activity of I56i is affected by the presence of the 182-SNP.

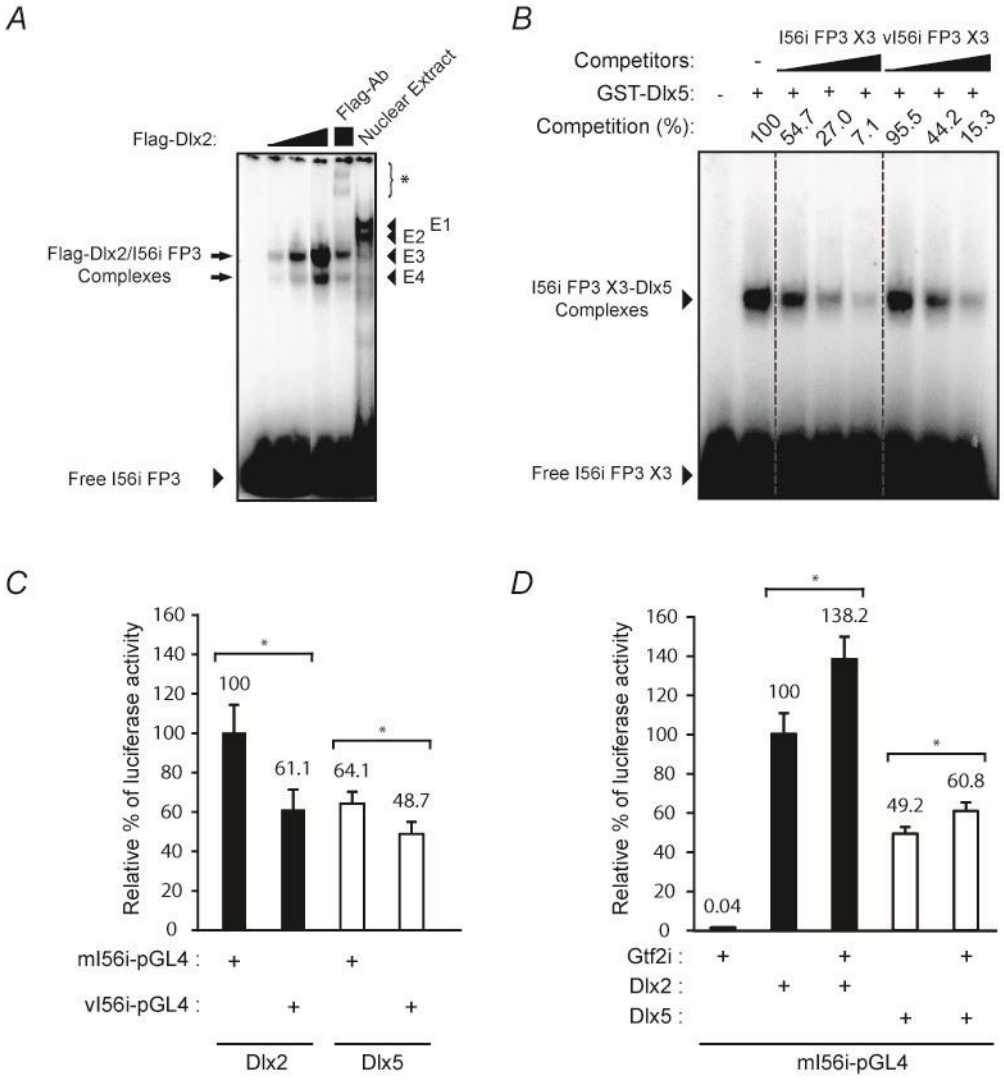
We first evaluated the affinity of the Dlx proteins for the FP3 binding site in its wild type or 182-SNP-containing versions by performing a competitive EMSA. The Dlx5-I56i FP3 X3 complexes were competed more efficiently by the unlabeled I56i FP3 X3 than by the variant competitor (vI56i FP3 X3), suggesting that the Dlx proteins possess more affinity for the wt FP3 binding site than by the variant sequence (Fig. 3.6B). Next, we examined the effects of the 182-SNP on the transcriptional activity of the I56i enhancer in response to the Dlx2 or Dlx5 proteins using luciferase reporter assays following co-transfections in murine fibroblast NIH3T3 cells. In the presence of the 182-SNP, the response of the I56i enhancer to Dlx2 and Dlx5 decreased by 39% and 24%, respectively (Fig. 3.6C).

To test a possible involvement of Gtf2i in *Dlx5/Dlx6* gene regulation via the I56i enhancer, additional co-transfections were performed in NIH3T3 cells. Gtf2i and the Dlx proteins were shown to synergize in activating the expression of the reporter *luc2* (Fig. 3.6D**)**. The transfection of a Gtf2i expressing plasmid *per se* did not stimulate the transcriptional activity (0.04% relative to Dlx2 activation taken as a 100% reference). However, when co-transfected with the Dlx2 or Dlx5 expression plasmids, the transcriptional activity increased

to 138.2- and 60.8%, respectively. These activations represent an increase of 38% (Dlx2) and 24% (Dlx5) when compared to the transfection reactions in which only Dlx2 or Dlx5 plasmids were transfected (100- and 49.2-%, respectively).

These results indicate that the 182-SNP interferes with the activation of the I56i enhancer by Dlx proteins and suggest a role for interactions between Dlx and Gtf2i proteins in *Dlx* regulation via the I56i enhancer.

Figure 3.6: Transcriptional activation of I56i by Dlx and Gtf2i proteins is impaired by the 182- SNP. (A) The molecular weight of the protein component(s) of the E4 complex is similar to Dlx2. Comparison of an electro-mobility shift assays (EMSA) using the I56i FP3 oligonucleotide with the E13.5 embryonic forebrain nuclear extract (Nuclear extract) or an increasing amounts of recombinant Flag-tagged Dlx2 (Flag-Dlx2). Preincubation of the recombinant Dlx2 with an anti-Flag antibody leads to the formation of a high molecular weight complex (asterisk, Flag-Ab lane). (B) Affinity of Dlx proteins for the FP3 site is affected by the 182-SNP. EMSA was performed using the I56i X3 oligonucleotide and a fixed amount of recombinant Dlx5. The resulting complexes (shown by the black arrowhead) were competed more efficiently by the unlabeled wt I56i X3 (increasing amounts in lane 3 to 5) than by the variant competitor (vI56i X3, lane 6 to 8). Quantification numbers (Competition) shown in this figure correspond the percentage of remaining complexes in respect to the control reaction (without competitor, 100%). (C) Replacement of I56i by its 182-SNP containing variant (vI56i-pGL4.23) impaired the transcription activation by Dlx2 or Dlx5. (D) Dlx2 and Dlx5 were shown to synergize with Gtf2i in activating the transcription of a *luc2* reporter construct containing the I56i enhancer. Values shown in C and D represent the mean of relative luciferase activity obtained from five (C) or six (D) independent experiments $^{\pm}$ S.E.M. Relative luciferase activity were normalized to Dlx2 + mI56i-pGL4 (100%) in both graphs. * $P < 0.05$.


A
Flag-Dlx2:
Flag-Ab
Nuclear Extract
*
E1
E2
E3
E4
Flag-Dlx2/I56i FP3
Complexes
Free I56i FP3
B
I56i FP3 X3
vI56i FP3 X3
Competitors:
GST-Dlx5:
Competition (%):
100
54.7
27.0
7.1
95.5
44.2
15.3
I56i FP3 X3-Dlx5
Complexes
Free I56i FP3 X3
C
Relative % of luciferase activity
100
61.1
64.1
48.7
mI56i-pGL4 :
vI56i-pGL4 :
Dlx2
Dlx5
D
Relative % of luciferase activity
0.04
100
138.2
49.2
60.8
Gtf2i :
Dlx2 :
Dlx5 :
mI56i-pGL4

3.4 Discussion

In this report, we demonstrated that a single nucleotide polymorphism, originally isolated in a screen of autistic probands and located in an ultraconserved *Dlx5/Dlx6* CRE (I56i), had functional consequences on the activity of the enhancer throughout forebrain development, continuing into adulthood. The identification of this functionally important SNP in the I56i enhancer provides us with an entry point to address the relationship between *Dlx* gene regulation in the telencephalon, GABAergic interneuron development, and, potentially, the etiology of ASD.

Since this element was identified as one of 481 ultraconserved DNA segments amongorthologous regions of the human, rat and mouse genomes, the I56i enhancer is considered as one region of the genome not susceptible to variation (Bejerano, Pheasant et al. 2004). Furthermore, the SNP was found in a region of I56i that is identical in forty vertebrates species surveyed (Fig. 3.1). Such sequence conservation made this SNP an unexpected finding and, thus, suggestive of a causal relationship. DNaseI footprinting analysis showed that the 182-SNP was located within the boundary of a DNA-protein interaction, the FP3 binding site. This site was previously identified as one of two Dlx binding sites present on the I56i enhancer. These Dlx binding sites play crucial roles in the auto- or cross regulation mechanisms of the *Dlx* bigene (Zerucha, Stuhmer et al. 2000). Thus, the presence of the 182-SNP within one of these important binding sites may, by itself, have functional consequences on the regulatory cascades controlling *Dlx* expression.

The hypothesis that the 182-SNP would affect auto- or cross regulatory mechanisms at I56i was supported by our EMSA and co-transfection experiments. Dlx affinity for the FP3 sequence was impaired by the presence of the 182-SNP as well as the transcriptional activity

of I56i in response to Dlx2 or Dlx5. In co-transfection assays, we further provided evidence that the 182-SNP affected the I56i enhancer activity induced by the Dlx proteins.

Our EMSA analysis also showed that binding of one factor to I56i was specifically affected by the SNP. We have shown that this factor was similar in size to the Dlx proteins. Complexes formed by the FP3 oligonucleotide with the recombinant Dlx2 migrated in similar way to the E3 and E4 nuclear complexes. A potential explanation for the observation of two complexes could be that Dlx2 binds to the FP3 site as a monomer and/or as a homodimers. It has been demonstrated that Dlx proteins can bind DNA as hetero- and homodimers (Zhang, Hu et al. 1997). Supershift experiments on these complexes with a number of anti-Dlx antibodies were unconclusive (data not shown). However, our affinity purification experiment enabled us to find a new regulator of the *Dlx* genes, General Transcription Factor 2-I (Gtf2i). In co-transfection assays, this factor was shown to synergize with the Dlx proteins in activating transcription through I56i and the overall activity of I56i in the presence of a Dlx-Gtf2i combination was affected by the SNP. The mechanism by which Gtf2i interacts with the Dlx on the I56i enhancer is still unknown. Possibilities include: 1) Gtf2i interacts directly, or through an intermediate factor, with Dlx proteins, independent of Gtf2i's ability to bind to DNA. However, attempts at co-immunoprecipitating Gtf2i and Dlx2 *in vitro* have not been successful (data not shown). 2) Gtf2i binds the I56i sequence, possibly at the putative binding site located between the two Dlx binding sites.

The identification of Gtf2i as a potential regulator of *Dlx* expression is particularly relevant because Gtf2i is a candidate gene for Williams-Beuren syndrome (WBS), a neurodevelopmental disorder. A hemizygous deletion of a cluster of 26 genes including three members of the GTF2 family on chromosome 7 (7q11.23) was observed in a majority of the WBS cases. Children with WBS display distinctive facial features, mental disability,

and growth retardation. They also share a unique set of personality traits, which include over-friendliness, abnormal social behavior (outwardly social) and an increased trust in strangers (reviewed in (Francke 1999)). Although WBS children show high sociability, almost half of them display socio-communicative deficits characteristic of ASD (Klein-Tasman, Mervis et al. 2007; Lincoln, Searcy et al. 2007). Moreover, the genetic analysis of an autistic child exhibiting the cognitive–behavioral profile (CBP) associated with WBS revealed an atypical deletion of the WBS critical region (Edelmann, Prosnitz et al. 2007). Although the authors of this study proposed that the WBS CBP was due to the deletion of the member of the GTF2 family and that the symptoms of autism were due to the deletion of another gene outside the typical WBS deletion, one cannot exclude that the mental retardation and autistic traits observed in this patient were due to a dosage effect of *Gtf2i* gene expression. Finally, a case in which an autistic individual with a 7q11.23 duplication was recently reported (Van der Aa, Rooms et al.). Altogether, there is accumulating evidence suggesting that members of the *Gtf2i* gene family play an important role in WBS even though it still remains unclear whether the *Gtf2i* genes can be accounted for socio-communicative deficits (characteristic of ASD) seen in WBS patients.

Dlx genes are key members of genetic cascades responsible for the establishment of a diversity of neuronal types in the central nervous system, such as the GABAergic interneurons. In transgenic animals, the 182-SNP affects I56i enhancer activity *in vivo*, at three key time points during GABAergic interneuron development: their early differentiation, their migration to the pallium and their differentiated phase in the somatosensory cortex. From these results, we can extrapolate that the presence of the 182-SNP in I56i may lead to a decrease in *Dlx5/Dlx6* expression, which in turn will probably affect the early process during of GABAergic interneuron differentiation as Dlx proteins

have been shown to regulate *Gad* expression (Long, Cobos et al. 2009; Long, Swan et al. 2009). As for the impact of the SNP on precursor migration, it was previously shown that *neuropilin-2* (*Nrp-2*) is ectopically expressed when *Dlx1* and *Dlx2* are inactivated (Le, Du et al. 2007). Nrp-2 is a transmembrane protein implicated in repulsive axon guidance and, therefore, it was proposed that down-regulation of *Nrp-2* by Dlx proteins could be a mechanism by which the migration of GABAergic interneurons was facilitated. We predict that reduced levels of Dlx5/Dlx6 proteins caused by the SNP could interfere with Nrp2 regulation, potentially leading to the impaired migration of interneuron progenitors to the cortex.

The impact of altered *Dlx* regulation in GABAergic interneurons of the newborn or adult mice remains to be established. In mice, the brain of *Dlx1* null mutants displays no overt phenotype at birth but later in life, they develop epileptic seizures (Cobos, Calcagnotto et al. 2005). *Dlx5*/6$^{+/-}$ mice, whose cortex appears normal, have spontaneous electrographic seizures, and reduced power of gamma oscillations suggesting defects in cortical interneurons; in addition, there is evidence for defective development of parvalbumin interneurons in *Dlx5*/6$^{-/-}$ and *Dlx5*$^{-/-}$ mutants (Wang, Dye et al. 2010). Thus, Dlx proteins could be exerting a specific role in the adult brain or subtle alterations that occur during development could have delayed consequences. In both cases, a decrease in I56i activity in mature neurons could potentially lead to dysfunctions of neural networks.

The 182-SNP is a rare polymorphism with a frequency of 0.3% amongst the autistic probands (one affected individual from 161 probands screened) (Hamilton, Woo et al. 2005). This SNP was paternally transmitted and was absent from an affected sibling (Richler, Reichert et al. 2006). This situation is not exceptional considering that siblings, who share half the genetic risk, often display different autistic phenotype (Skuse 2007). It was proposed

that one or several independent risk factors can elevate the chance of displaying the full autistic phenotype. Furthermore, using a broader concept of autism, two independent studies have shown that an increased proportion of unaffected siblings were reported to have subsyndromal autistic impairments (Bishop, Maybery et al. 2006; Constantino, Lajonchere et al. 2006), supporting the notion of independent risk factors or genetic susceptibility factors.

Paternal imprinting of the *Dlx5* locus was recently documented (Okita, Meguro et al. 2003; Horike, Cai et al. 2005) and would have phenotypic implications for a SNP in I56i, depending on the parent of origin. Horike and colleagues demonstrated that the loss of *Dlx5* imprinting, through *Mecp2* inactivation, leads to biallelic expression of *Dlx5* resulting in increased *Dlx5* expression. Since Mecp2 mutations are associated with Rett syndrome, it was suggested that misregulation of several genes, including *Dlx5*, resulting from the loss of their imprinting, may contribute to the etiology of this syndrome. Imprinting of *Dlx5* was also observed in human and swine loci (Okita, Meguro et al. 2003; Cheng, Zhang et al. 2008). However, paternal imprinting of *Dlx5* was challenged by other studies addressing biallelic expression of the gene (Kimura, Kazuki et al. 2004; Schule, Li et al. 2007; Bischoff, Tsai et al. 2009). In the presence of imprinting, one copy of the *Dlx5/Dlx6* locus will be silenced. Therefore, the presence of the SNP on the remaining allele will result in a decrease in *Dlx5/Dlx6* expression, which will be significantly higher (double) when compared to the situation without imprinting on *Dlx5/Dlx6* locus.

The reason for the exceptional sequence conservation of some non-coding regulatory elements in vertebrate genomes is still puzzling. While it is not likely due to chance that these regions have been maintained for hundreds of millions of years, it is also not clear what their function is, given the lack of phenotype shown when some of the ultraconserved regions were deleted (Ahituv, Zhu et al. 2007). Our study provides an example of the

consequences that a single nucleotide change could have on the function of an ultraconserved sequence.

3.5 Acknowledgments

We would like to thank Mélanie Debiais-Thibaud and Fabien Avaron for helpful discussions. We are grateful to Adriana Gambarotta for technical assistance in the production of transgenic mice. This work was supported by grants MOP14460 and MOP-74564 from the Canadian Institutes of Health Research (CIHR). JLRR's work was supported by the research grants from: Nina Ireland, the Althea Foundation, Cure Autism Now and R37 MH049428. GGP work was supported by grants from Genome Canada and Génome Québec. GGP is a recipient of a Canadian Research Chair in Proteomics.

3.6 Author contribution

L.P. performed experiments in **Figures 3.2-6** and contributed to the writing of the manuscript. M.Y. carried out the immunolocalisation experiments presented in **Figure 3.2-4**. C.L.P. carried out transfections and EMSA experiments shown in **Figure 3.6**. R.M. participated in the EMSA experiments shown in **Figure 3.6.** P.H., I.K. and G.P. were involved in the purification and identification of candidate proteins by mass spectrometry. G.H. participated in the design of the in vivo experiments. S.H. and J.L.R. identified the single nucleotide polymorphism presented in this study and contributed the writing of this manuscript. M.E. participated in the design of the experiments and contributed to the writing of the manuscript.

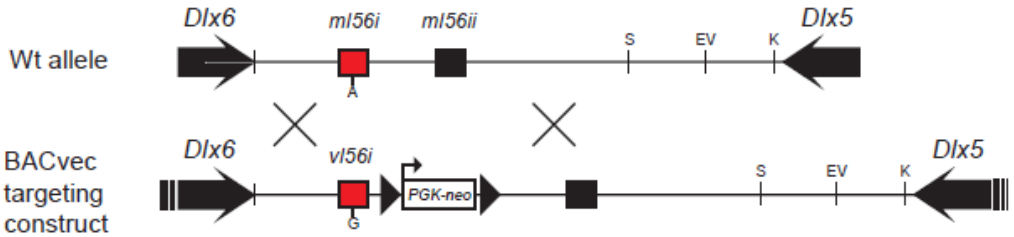
4 Phenotype analysis of mice carrying a variant I56 Dlx enhancer

4.1 Overview

To analyze the potential effects of the SNP found in the I56i enhancer of some autistic probands, I used mice in which the wild type I56i enhancer has been replaced with the SNP-containing allele. Replacement was carried out by homologous recombination in embryonic stem cells of the SV129 mouse strain using a Bacterial Artificial Chromosome (BAC) construct. A cassette including the modified enhancer (vI56i) and the neomycin resistance gene was inserted in place of the wild type I56i enhancer in the BAC containing the mouse *Dlx5/Dlx6* locus (Fig. 4.1). This was accomplished by homologous recombination in bacteria as described in (Zhang, Buchholz et al. 1998; Lee, Yu et al. 2001). The modified BAC clone was electroporated into mouse embryonic stem cells. Following G-418 selection, 100 resistant clones were tested for the proper recombination event using quantitative real-time PCR as described in (Valenzuela, Murphy et al. 2003). To facilitate detection of the SNP allele by PCR, a small deletion outside of the I56i enhancer was engineered. PCR primers designed to easily distinguish between the two alleles allow counting the number of wild type and mutant alleles in ES cells. Four positive clones were obtained and one clone was injected into blastocysts of the C57Bl/6 mouse strain to make chimeras of SV129 and C57Bl/6. One of the chimeras transmitted the mutation through the germline, giving heterozygous individuals. I bred these individuals to homozygosity and therefore obtained a mouse line with the SNP version of the I56i sequence.

Figure 4.1: Strategy for the targeted replacement of *Dlx5/Dlx6* forebrain enhancer I56i with the variant version vI56i using homologous recombination in embryonic stem cells.

mI56i replacement (Knockin)



4.2 Results

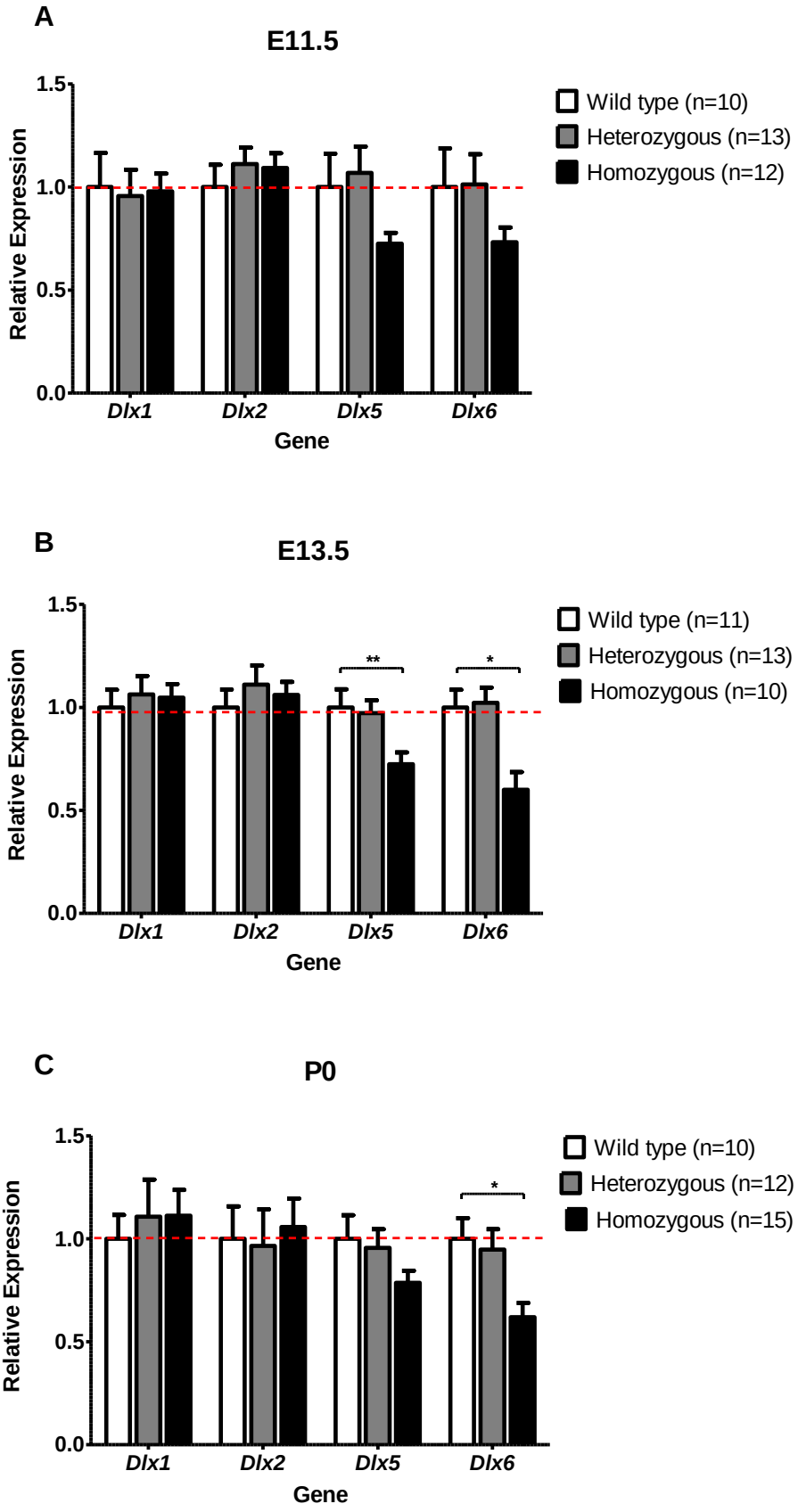
4.2.1 The I56i SNP and *Dlx* gene expression

First, to characterize the effects of the I56i SNP on gene expression levels, I performed real-time quantitative PCR (qPCR) on cDNA that was reverse-transcribed from RNA extracted from mouse forebrain tissue. These extracts were obtained at three different developmental stages i.e Embryonic day 11.5 (E11.5) and 13.5 (E13.5), and at post-natal day 0 (P0). Primers were designed to test for the expression levels of *Dlx* genes surrounding the I56i enhancer (*Dlx5* and *Dlx6*), as well as for *Dlx1* and *Dlx2* genes which are potentially regulated by Dlx5 (Poitras, Ghanem et al. 2007). I used wild-type siblings as my controls and compared the ratio of *Dlx*-to-*GAPDH* expression levels for vI56i heterozygous and homozygous mice using Livak method's (Livak and Schmittgen 2001).

At embryonic day 11.5 (E11.5), decreases of 27% and 28%, respectively, were seen in *Dlx5* and *Dlx6* expression levels of homozygous vI56i (Fig.4.2A). However, no significant changes were observed in heterozygous mice, suggesting that one mutant allele is probably not sufficient to alter *Dlx5* and *Dlx6* expression. Then, at E13.5, similar decreases in *Dlx5* and *Dlx6* expression were seen in homozygous vI56i mice compared to wild-type. Specifically, I observed a significant reduction of 28% in *Dlx5* ($T_{(20)}$=2.5184, P<0.02) and of 40% in *Dlx6* expression ($T_{(20)}$=3.2965, P<0.005) (Fig.4.2B). Finally at birth, a 21% decrease in *Dlx5* that did not reach statistical significance and a 38% decrease in *Dlx6* expression ($T_{(25)}$=3.2093, P<0.005) were observed in homozygous vI56i mice (Fig.4.2C). As observed at E11.5, mice heterozygous for the mutation did not appear to have impaired expression at E13.5 nor at P0. There was no significant change in *Dlx1* or *Dlx2* expression, in heterozygous or homozygous mutants, at each developmental stage and in postnatal mice,

suggesting that the expression of these genes is not dependent on the I56i regulatory element (Fig.4.2).

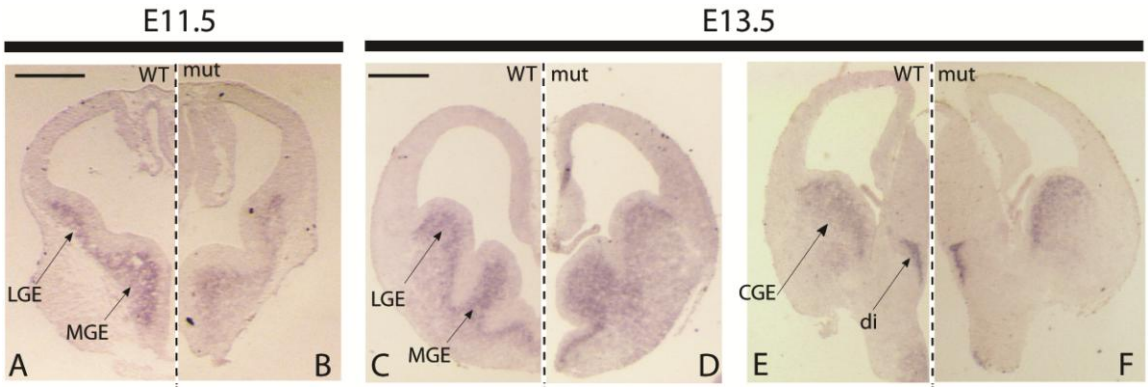
Figure 4.2: qPCR results corresponding to forebrain relative expression level of *Dlx* genes in wild-type, heterozygous and homozygous vI561 mice at E11.5 (A) E13.5 (B) and P0 (C). Relative expression level of wild-type mice was defined as 1.0. Numbers of individuals are indicated in parentheses and each genotype was tested in duplicate. Data are expressed as mean ± SEM. (*P <0.005) (**P < 0.02)


A
E11.5
Relative Expression
1.5
1.0
0.5
0.0
Dlx1
Dlx2
Dlx5
Dlx6
Gene
Wild type (n=10)
Heterozygous (n=13)
Homozygous (n=12)
B
E13.5
Relative Expression
1.5
1.0
0.5
0.0
**
*
Dlx1
Dlx2
Dlx5
Dlx6
Gene
Wild type (n=11)
Heterozygous (n=13)
Homozygous (n=10)
C
P0
Relative Expression
1.5
1.0
0.5
0.0
*
Dlx1
Dlx2
Dlx5
Dlx6
Gene
Wild type (n=10)
Heterozygous (n=12)
Homozygous (n=15)

To assess if the mutation leads to defects in *Dlx5* gene expression patterns in the forebrain, I used *in situ* hybridization with antisense RNA probes, on brain sections from two different developmental stages i.e. E11.5 and E13.5.

At E11.5, in both wild type and mutant I56i SNP, a ventral domain of expression extending from the MGE towards the LGE was observed (Fig.4.3A-B). At E13.5, *Dlx5* was expressed in the subventricular zone (SVZ) and in the mantle zone (MZ) of the MGE and LGE. In more posterior sections of the forebrain, expression was detected in the CGE and the diencephalon, which is consistent with previous reports as described in (Liu, Ghattas et al. 1997) (Fig.4.3C-F). No apparent changes in expression were observed between wild type and vI56i mutant brains. Thus, the decrease in *Dlx5* expression detected by qPCR in homozygous vI56i mutant embryos at E11.5 and E13.5 could not be confirmed by in situ hybridization.

Figure 4.3: *In situ* hybrydization on transverse sections of embryonic forebrain tissue from both wild type (A,C,E) and homozygous mutant vI56i mice (B,D,F) at E.11.5 (A-B) and E13.5 (C-F) showing the expression of *Dlx5*. Scale bars: 1 mm in A-B; 2mm in C-F.


E11.5
E13.5
WT
mut
LGE
MGE
CGE
di
A
B
C
D
E
F

4.2.2 The I56i SNP and mouse behaviour

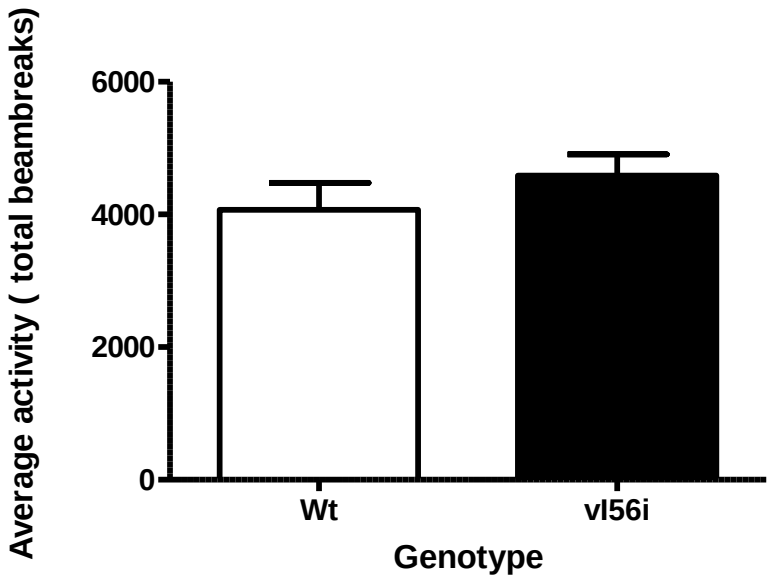
In order determine if the I56i SNP resulted in a behavioural phenotype, a series of four behavioural tests were used to evaluate potential autism-like behaviours. Mice homozygous for the variant enhancer (v156 n=21) were tested and compared to wild type (wt, n=7) littermates controls. It is necessary to use the littermate controls since the two strains that were originally used to create the 156i SNP mice, as SV129 stem cells were injected into a C57BL/6 background, have different baseline behavioural patterns

The first test presented is the beambreak test, which is a general measure for locomotor activity. The test measures spontaneous activity in a home cage and detects any gross abnormalities in motor behaviour. This is necessary to perform since the other autistic-like behaviours that I wanted to access all rely on measurements that are dependent on locomotor activity. Therefore it is necessary to first test if there is any gross locomotor difference between the vI56i and wt mice. Since mice are most active during the night, the test was performed overnight, for a period of 15 hours (Benstaali, Mailloux et al. 2001; Kopp 2001).

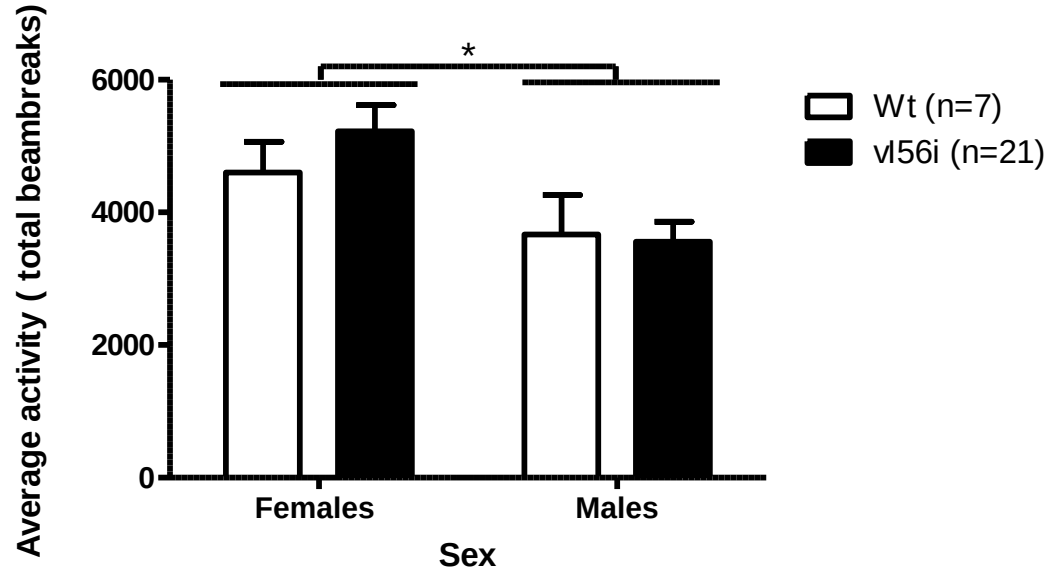
No significant differences were observed in locomotor activity between vI56i mutant and wt mice when they were assessed each hour (data not shown) or examined for total cumulative activity for the 15 hours (Fig.4.4A) Surprisingly, when we compared males and females, there was a significant differences in amount of activity, with the females being more active than males ($F_{(1.24)}$=5.915 P<0.02) (Fig4.4B).This is surprising because it has been shown in many studies that C57Bl/6 males and females usually have similar basal levels of activity in this type of test (Brookshire and Jones 2009; An, Zou et al. 2011)

Figure 4.4: Locomotor activity over a 15 hour period is shown for wild type and homozygous mutant vI56i mice (A) and for female and male individuals of both genotypes (B). Numbers of individuals are indicated in parentheses. Data are expressed as means ± SEM (*$P < 0.02$)

A


Average activity (total beambreaks)
6000
4000
2000
0
Wt
vl56i
Genotype

B


Average activity (total beambreaks)
*
6000
4000
2000
0
Females
Males
Sex
Wt (n=7)
vl56i (n=21)

Since one of the hallmark symptoms associated with autism is a lack of social interaction, I assessed social behaviour in vI56i and wt mice. Specifically, both the adult social interaction test and juvenile social interaction test were used since they are two independents tests of social behaviour that have been previously validated as measures of autism-like behaviour (Crawley 2007; Moy, Nadler et al. 2007; Moy and Nadler 2008; Moy, Nadler et al. 2009).

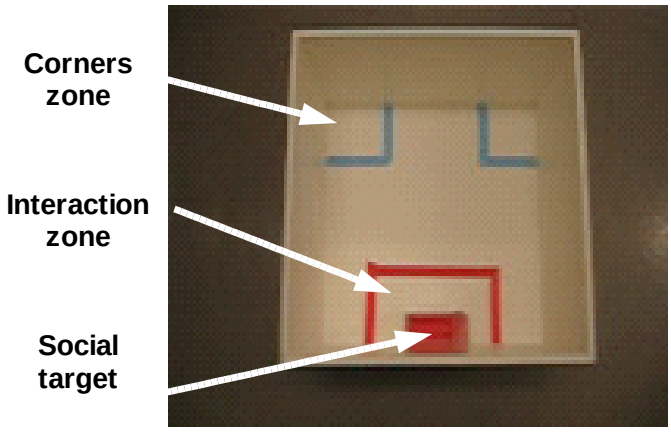
The adult social interaction test allows the evaluation of social behaviour, such as social affiliation and motivation, based on the ability of mice to recognize members of their own group (Moy, Nadler et al. 2004; Silverman, Yang et al. 2010). During this test, the test mouse is placed in the corner of an open field box containing a rectangular wire mesh cage (Fig.4.5A). The test mouse is left in the box to explore and get familiar with the environment and then removed from the box. Once the test mouse is removed another mouse, named the social target is placed inside the small rectangular wire mesh cage and the previous mouse is brought back into the wire mesh cage and allowed to interact with the test mouse. The time the test mouse interacts with the social target mouse in the interaction zone (around the wire mesh cage) and the time spent in the 2 corners (across from the wire mesh cage) is recorded.

As expected, repeated measure ANOVA revealed that the mice spent significantly more time in the interacting zone when the mesh cage had the social target versus when the maze was empty ($F_{(1.24)}$=11.524, P<0.002). There was however no significant difference between the vI56i and wt mice in the time in the interaction zone when the social target was present or absent (Fig.4.5B). Repeated measure ANOVA also revealed that when the social target was present, the time spent in the interaction zone was significantly lower than the time spent in the corners zone ($F_{(1.24)}$=16.281, P<0.0005), and there was no significant difference between vI56i and wt mice (Fig.4.5C). On all of these measures there was no significant difference between male and female mice (data not shown). Together these results confirm that mice

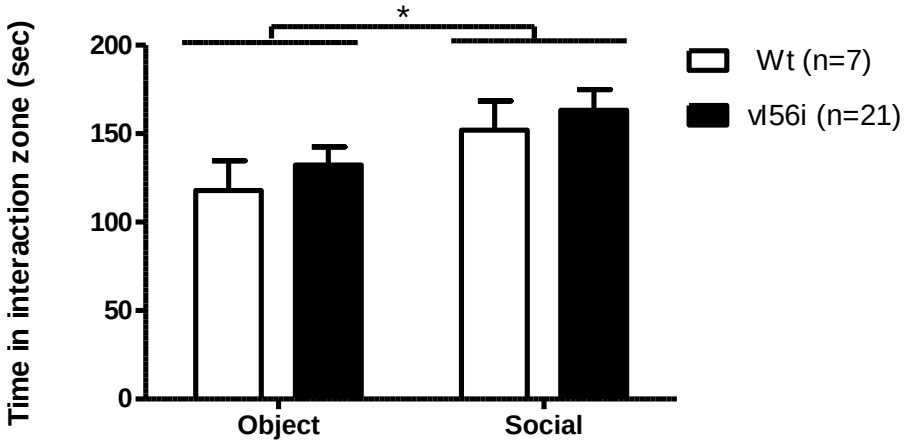
preferred social interaction (Gheusi G. 1994). Furthermore, since there was no effect of genotype on any outcome, the presence of the SNP in the I56i enhancer seems to have no impact in adult social interaction behaviour.

Figure 4.5: Adult social interaction of wild type and homozygous mutant vI56i mice. A) Schematic representation of the different zones of the test. Total duration of social interaction, scored as cumulative seconds spent in the interaction zone in presence or absence of the social target (B) and in the interaction zone or in the 2 corner zone in presence of the social target (C). The number of individuals is indicated in parentheses. Data are expressed as means ± SEM (*$P < 0.002$) (**$P < 0.0005$)

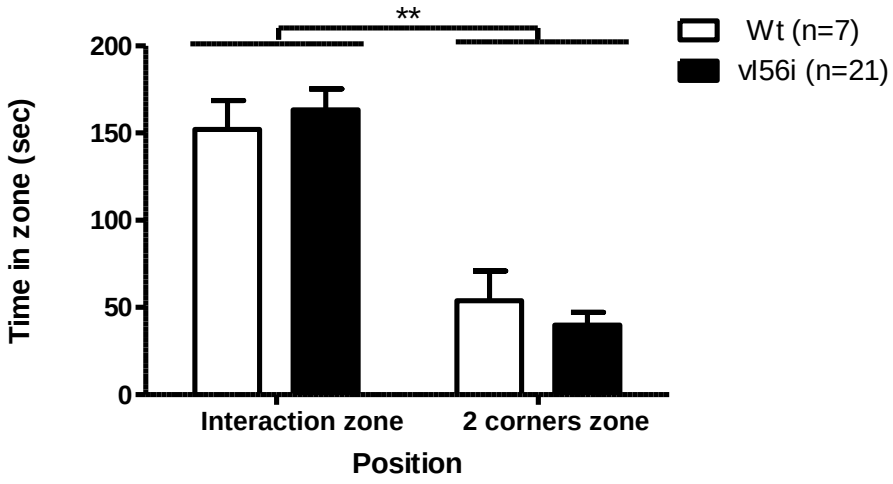
A


Corners zone
Interaction zone
Social target

B


*
Time in interaction zone (sec)
200
150
100
50
0
Object
Social
Wt (n=7)
vl56i (n=21)

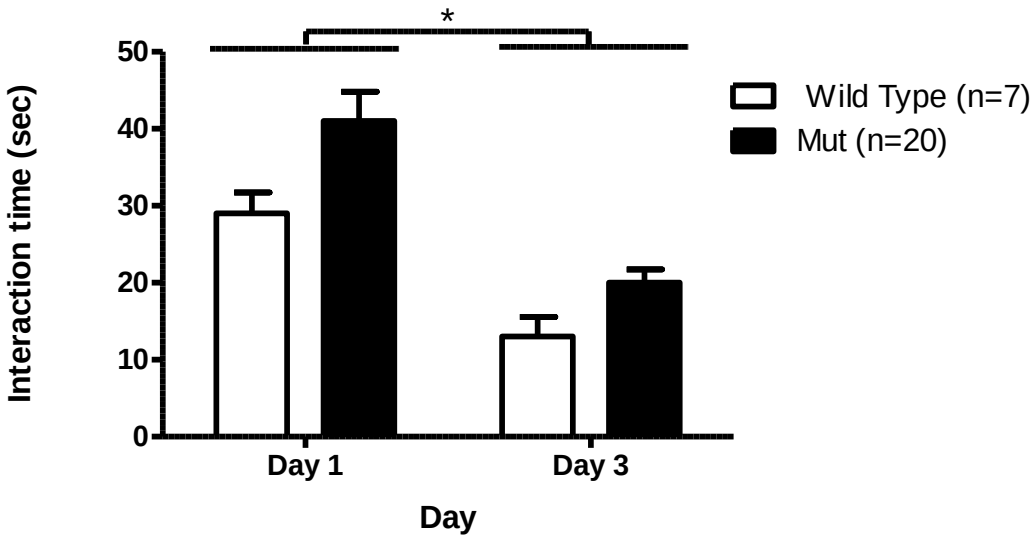
C


**
Time in zone (sec)
200
150
100
50
0
Interaction zone
2 corners zone
Position
Wt (n=7)
vl56i (n=21)

In addition to the adult social interaction test, I performed the juvenile interaction test since it is a standard test of sociability and social memory (McFarlane, Kusek et al. 2008; Silverman, Yang et al. 2010). Briefly, in the first trial of this test the adult test mouse (10-14 weeks) and a juvenile mouse (3 weeks) of the same strain were placed in a novel cage together at the same time, and allowed to explore. The time the test mouse interacted with the juvenile mouse was recorded. Three days later, a second trial is performed, in which the same experiment was done with the same two mice. The interaction time for the first exposure to the juvenile and for the second exposure to the same juvenile were compared.

As expected, repeated measure ANOVA has shown that mice spent a significantly greater amount of time interacting on day 1 compared to day 3 (Fig 4.6, $F_{(1.24)}$=58.056, P<0.000). This result is congruent with previous results showing that mice are able to remember a mouse they have interacted with in the past, and usually interact less time with it in the second trial (Crawley 2007). There was no significant difference in the interaction time between the wt and vI56i mice (P=0.052 NS) and no significant interaction effect between the different days and genotype (Fig.4.6). There was also no significant difference between male and female (data not shown). These results suggest that vI56i mice do not show impairment in sociability with a juvenile or in social memory compare to wild type mice.

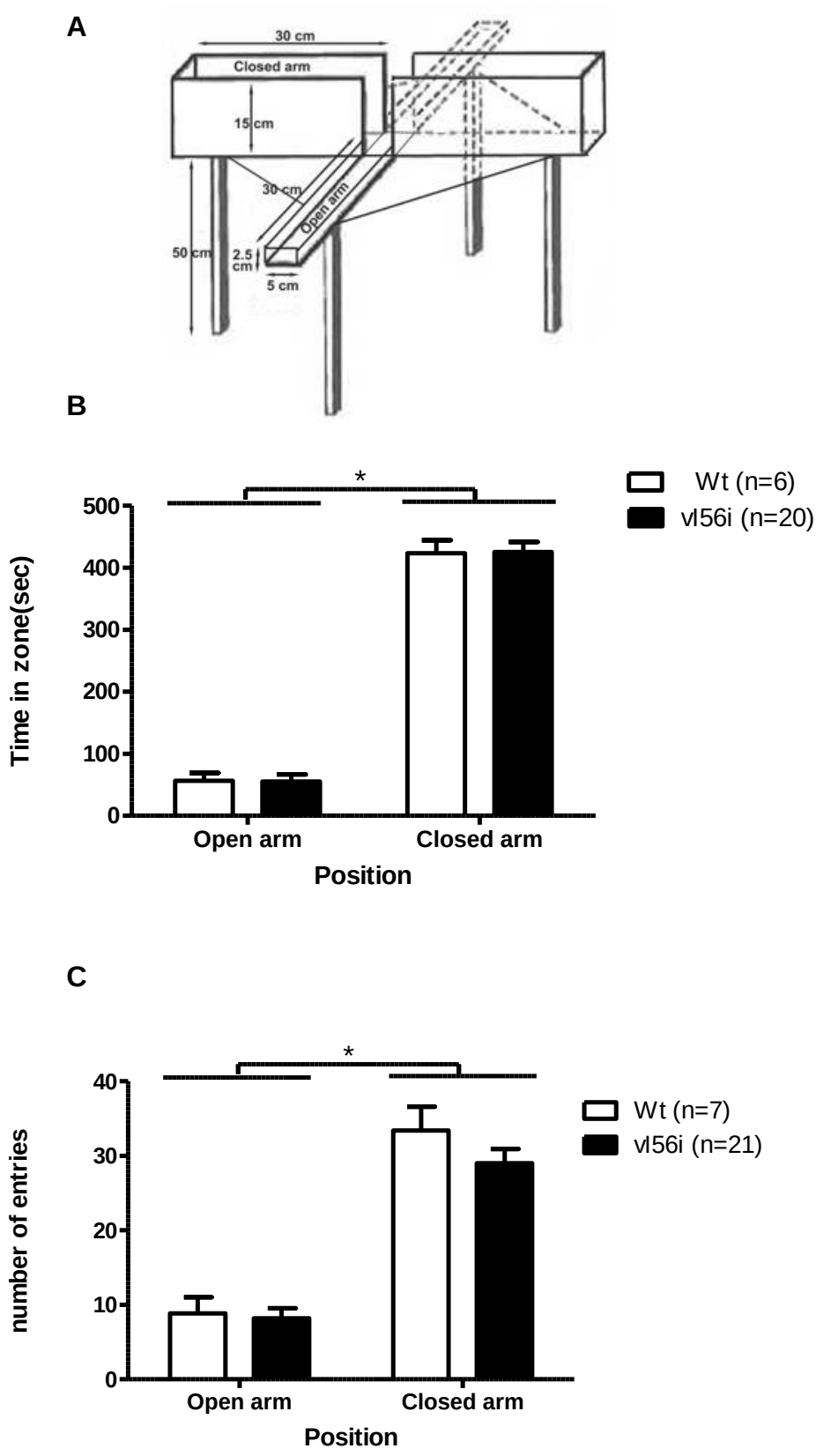
Figure 4.6: Juvenile social interaction of wild type and homozygous mutant vI56i mice. Total duration of social interactions is scored as cumulative seconds spent by the test mouse in interacting with a juvenile. Numbers of studied mice are indicated in parentheses. Data are expressed as means ± SEM. (*$P < 0.000$)


*
50
40
30
20
10
0
Interaction time (sec)
Day 1
Day 3
Day
Wild Type (n=7)
Mut (n=20)

The elevated plus maze test was also used in our behavioural assessment since it is a standard test for fear and anxiety (Griebel, Belzung et al. 2000; Paterson, Iwunze et al. 2010). Anxiety is an associated symptom of autism which occur in subsets of autistic individuals (Rogers and Ozonoff 2005; Silverman, Yang et al. 2010). The test is based on an animal's preference for dark and enclosed places, compared to bright and exposed places (Griebel, Belzung et al. 2000; Paterson, Iwunze et al. 2010). The mouse is placed in the center of a plus-shaped maze and allowed to freely explore the maze (Fig.4.7A). The number of entries and the time spent in each arm were recorded.

As expected for this test, repeated measure ANOVA revealed that the time spent and the number of entries in closed arms is higher than in opened arms (Fig.4.7B-C, respectively $F_{(1.22)}$=205.106, P<0.000 and $F_{(1.24)}$=140.964, P<0.000). Between the wt and vI56i mice there was no significant difference in total entries and in time in each arm. There was also no significant difference between male and female mice (data not shown). These results suggest that the vI56i mice have similar levels of anxiety compared to the wild type mice.

Figure 4.7: Elevated plus maze (EPM) test of wild type and homozygous mutant vI56i mice. A) Schematic representation of the apparatus used for the test; B) time spent in the open or closed arms of the elevated plus maze; C) number of entries in the open or closed arms of the elevated plus maze. Number of individuals is indicated in parentheses. Data are expressed as means ± SEM. (*$P < 0.000$)


A
30 cm
Closed arm
15 cm
30 cm
Open arm
50 cm
2.5 cm
5 cm
B
*
Wt (n=6)
vl56i (n=20)
Time in zone(sec)
500
400
300
200
100
0
Open arm
Closed arm
Position
C
*
Wt (n=7)
vl56i (n=21)
number of entries
40
30
20
10
0
Open arm
Closed arm
Position

5 Discussion

Dlx homeobox genes encode transcription factors involved in the differentiation and migration of GABAergic inhibitory interneurons in the developing forebrain (Stuhmer, Puelles et al. 2002). It was suggested that autism could be due to an increase in the ratio of excitation/inhibition of cortical neurons, which may be due, for example, to a defect in inhibitory interneurons activity (Rubenstein and Merzenich 2003). Furthermore, chromosomal positions of human DLX genes coincide with autism susceptibility loci, as determined by linkage studies (Risch, Spiker et al. 1999; Anney, Klei et al. 2010). A SNP was found in a child with autism in the *cis*-regulatory element I56i, located in the intergenic region between the *Dlx5* and *Dlx6* genes (Hamilton, Woo et al. 2005). Finding a SNP in the I56i enhancer was very surprising considering that it is an ultraconserved *cis*-regulatory element of the mouse genome (Bejerano, Pheasant et al. 2004). This suggests that the SNP could be causally related to the behavioural defects seen in the autistic individuals that carry it. During this study, I described some effects of the SNP on the enhancer activity of I56i at the molecular and developmental level, as well as at the behavioural level.

5.1 The SNP has functional consequences on I56i forebrain enhancer activity

Interestingly, the SNP in I56i is located in a position that lies directly within a previously characterized Dlx binding site in the I56i enhancer, as we determined by DNase I footprint analysis. Therefore, to test how the SNP may affect enhancer activity, transgenic mice were produced using either the wild type or variant I56i enhancers in front of a *lacZ* reporter cassette. We have shown that the expression of *lacZ*, driven by the I56i enhancer, was reduced in the SVZ and MZ of the LGE, MGE and CGE regions of the forebrain in mice with the vI56i transgene, compared to mice with the normal version of the enhancer.

Reduction of I56i activity was also observed for the variant enhancer compared to wild type in a number of cells expressing the transgene that migrate tangentially towards the cortex at E13.5 and in GABAergic interneurons of the adult cortex. This result showed that the SNP is likely to have important functional consequences in brain development and neuronal function. In order to determine the impact of the SNP on the genetic interactions taking place at the I56i enhancer *in vitro*, we have shown that the binding of Dlx transcription factors to the I56i enhancer is partially impaired by the presence of the SNP. These results suggest that the Dlx proteins have more affinity for the wild type binding site than for the variant sequence, as determined with EMSA. We have also demonstrated that the SNP has an impact on the transcriptional activity of the I56i enhancer in response to the Dlx2 or Dlx5 proteins through co-transfection assays which showed that transcriptional activation of I56i by Dlx proteins is impaired by the SNP. Affinity purification using part of the I56i sequence allowed us to identify Gtf2i as a putative new regulator of *Dl*x gene expression. Interestingly Gtf2i is involved in Williams-Beuren syndrome (WBS), a rare neurodevelopmental disorder (Francke 1999). Since a putative Gtft2i binding site is located between the two Dlx binding sites on the I56i enhancer, we have determined the impact of Gtf2i in *Dlx5* and *Dlx6* gene regulation via the I56i enhancer using additional co-transfections assays. We found that Dlx and Gtf2i can synergize in activating transcription using the I56i enhancer suggesting a role for interactions between Dlx and Gtf2i proteins in *Dlx* regulation. The results of the above experiments showed that introduction of the A → G SNP in I56i has drastic effects on I56i enhancer activity, when tested in vitro with protein-DNA binding assays and when tested in transgenic animals with reporter constructs.

5.2 The vI56i results in decreased *Dlx5/Dlx6* expression

I showed that the SNP located in an ultraconserved Dlx binding site of the I56i enhancer leads to a 20 to 40% decrease of *Dlx5* and *Dlx6* expression in the brain of mice homozygous for the mutation at E11.5, E13.5, as well as P0 (see Figure 4.2). As we showed previously *in vitro* with EMSA, different homeodomain proteins, including but not limited to Dlx , bind the protected site on I56i enhancer. Then, the binding of one of them may be differentially affected by the SNP mutation. Since enhancers act to increase the gene transcription level, the activation of *Dlx*5 and *Dlx*6 transcription seems to be impaired by this single nucleotide mutation. This decreased expression may result in impaired migration and differentiation of GABAergic interneurons in the telencephalon, which could be further linked to the autism phenotype. To test this hypothesis, it will be relevant to perform immunohistochemistry experiments with markers for various types of GABAergic interneurons. These markers include calretinin, calbindin somatostatin and parvalbumin (Ghanem, Yu et al. 2007). Decreased inhibitory tone provided by these GABAergic inhibitory interneurons can possibly affect the behaviour of individuals that carry the SNP.

The expression of *Dlx1* and *Dlx2* doesn't seem to be affected by the presence of the SNP, with no difference compared to the wild-type. This is not surprising since it corroborates previous results showing that *Dlx1* and *Dlx2* are upstream regulators of *Dlx5* and *Dlx6* by their binding to the I56i enhancer (Anderson, Qiu et al. 1997; Zerucha, Stuhmer et al. 2000; Stuhmer, Anderson et al. 2002; Zhou, Le et al. 2004).

There are no significant differences observed in the expression of *Dlx* genes for the heterozygous vI56i at each developmental stage tested. This suggests that one allele is sufficient for proper *Dlx5* and *Dlx6* expression in heterozygous vI56i.

In order to determine if the decrease in *Dlx5* and *Dlx6* expression has an impact on development, it will be interesting to test some downstream targets of Dlx5 and Dlx6 such as *Gad1* and *Gad2*, that encode for the enzyme responsible for the synthesis of GABA, the neurotransmitter of GABAergic interneurons (Stuhmer, Puelles et al. 2002).

Recent studies have also shown that *Wnt5a* is a putative transcriptional target of the Dlx5 protein in neural progenitor cells (Paina, Garzotto et al. 2011). Wnt proteins are involved in proliferation and differentiation of neural progenitor cells during embryonic development (Ille and Sommer 2005). These proteins activate various pathways in the cell that can be categorized into the canonical and non-canonical Wnt pathways (Nelson and Nusse 2004). The canonical Wnt pathway is β-catenin-dependent and is essential to maintain proliferation of subpallial neural progenitor cells (Gulacsi and Anderson 2008; Paina, Garzotto et al. 2011). In contrast, the non-canonical pathway, initiated by Wnt5a, allows neuronal differentiation, such as GABAergic differentiation in olfactory bulb (Paina, Garzotto et al. 2011). The *Wnt5a* locus contains predicted Dlx binding sites and it is conserved within mammals (Paina, Garzotto et al. 2011). Therefore, the authors of this study have shown that Dlx transcription factors can transcriptionnally regulate *Wnt5a* expression (Paina, Garzotto et al. 2011). Wnt5a expression was found to be down-regulated in Dlx5 null mutants (Paina, Garzotto et al. 2011). This strongly suggests that the decrease in *Dlx5* expression observed in vI56i mutant mice can have an impact on the expression of *Wnt5a*, especially in the olfactory bulb.

In situ hybridization was done to confirm that the mutation in I56i leads to defects in *Dlx* expression. However, we did not see obvious differences in the *Dlx5* expression signal intensity in wild type mice compared to homozygous vI56i mice on comparable sections. Taken together, these results suggest that the decrease in *Dlx* expression found in qPCR

experiments is possibly uniform over the forebrain rather than in just one place, which may explain why it cannot be distinguished by *in situ* hybridization. This technique is more anatomically descriptive than qPCR that gives quantitative expression, but can have limitations with regard to sampling efficiency since only a few thin forebrain sections are assumed to be representative of all the tissue (Evans, Watson et al. 2003). The *in situ* hybridization was only done with a *Dlx5* riboprobe, therefore it will be interesting to do it also with a *Dlx6* riboprobe since a decrease in expression was also observed in the qPCR experiment.

5.3 The vI56i seems to have no effect on mouse behaviour

Motor activity dysfunction can have an impact on the result of behavioural tests that require locomotion, such as social interaction and elevated plus maze, because the mice cannot explore properly their test environment (Silverman, Yang et al. 2010). Therefore to circumvent this problem, each mouse had to be evaluated for general activity before doing other tests. Since no differences were observed between genotypes, this potential artefact has been eliminated. However, the female locomotor activity seems to be higher than male activity. We cannot rule out that our gender effect could be due to the fact that experiments for males and females were done on two consecutive days since there were not enough chambers available to complete all tests at once. In order to determine if the day of the experiment was a confounding variable, futures experiments will need to counterbalance the setup so that both males and females are tested on the same day to eliminate a potential effect of the day of testing on locomotor activity. However, since other tests have shown no significant gender effect on any of the parameters when incorporated into the analysis, gender was not taken into account and the results were pooled.

Mice are a social species and they have a lot of reciprocal social interactions (Silverman, Yang et al. 2010). As autism is a disorder in which social skills are affected, it was relevant to determine whether there was a social deficit in the vI56i mice. However, I did not observe any difference in social behaviour in the adult interaction test as well as in the juvenile interaction test between genotypes. A reduced social interest and interaction was expected as an autism phenotype (Ey, Leblond et al. 2011). Therefore my results indicate that there is no obvious phenotype of absence of social interaction reported for these mice. We cannot rule out the possibility that the phenotype is so subtle that it cannot be seen with those tests and additional social experiments are still needed.

For example, we could use a test for social transmission of food preference. This test involves an interaction between a test mice and a cagemate which ate a new flavoured food. The test mouse will eat more of the above food than of a completely new food because the cagemate will confer familiarity with the above flavour (Galef 2003; Silverman, Yang et al. 2010).

Also, a relevant test to be done to assess a possible deficit in communication, another characteristic symptom of autism, will be the ultrasonic vocalizations test. Complex ultrasonic vocalizations are emitted by mice in social situations such as weaning of the pups, juvenile play, intruder interactions, and male-female sexual interactions. The quantification of ultrasonic vocalizations is used as a measure of the communication deficits in mouse models of neurodevelopmental disorders such as autism that are characterized by social and communication deficits (Branchi, Santucci et al. 2001; Scattoni, Gandhy et al. 2008; Scattoni, Crawley et al. 2009; Silverman, Yang et al. 2010).

Autism behaviour is characterized by different symptoms including deficit in social interactions and communication and by repetitive behaviour. Therefore, to detect a subtle

phenotype, a test for repetitive behaviour would also be required. This test consists of recording and noting stereotypical behaviours such as circling, jumping, backflips and self-grooming that persist for unusually long periods (Ryan, Young et al. 2010; Silverman, Yang et al. 2010).

General anxiety is an associated symptom of autism which occurs only in a subset of individuals and can be measured in mice with the elevated plus maze test (Silverman, Yang et al. 2010). My results showed that vI56i mice are not more anxious than wild type mice. However, considering the behavioural diversity of ASD, it is possible that the SNP in the I56i enhancer had no impact in the anxiety of mice.

The absence of differences in behavioural phenotype observed between vI56i and wild type may be attributed, in part, to the small number of mice of each genotype in the cohort. I used only seven wild type mice compared to twenty-one vI56i. A bigger cohort of wild type mice sibling will be required for next testing.

The absence of differences can also be attributed to the fact that all the mice in the cohort were not exactly of the same age. There was a difference of 4 weeks between the youngest and the oldest adult mice. However, there were wild type mice and vI56i mice in both age groups and, therefore, age may not have a big influence. It will be interesting to repeat experiments with a cohort of mice that are all born in a shorter time span.

We also need to get rid of the mixed genetic SV129-C57Bl/6 background to introduce the vI56i mutation in a pure C57Bl/6 genetic background. A mixed background can have an impact on the results because each mouse background differs in neural and behavioural phenotypes and the percentage of each background can be slightly different in each mouse (Gerlai 1996; Wolfer, Crusio et al. 2002). Furthermore, the *Neo* cassette, used to screen stem cells that have incorporated the mutation, has to be removed in case it interferes with the

phenotypes. That can be done using a strategy based on the properties of the Cre/loxP system. Mice carrying the mutant allele have two loxP sites flanking the *Neo* gene (fig. 4.1) and can be crossed with a transgenic mouse expressing the CRE recombinase. The CRE recombinase causes the excision of the sequence between the loxP sites. Finally, to obtain a pure background, these mice will be crossed with a pure C57Bl/6 mouse. The average percentage of the genetic material of the offspring that is derived from that pure background increases after each crossing. The result, after sufficient reiterations, is an animal with the vI56i allele in the desired genetic background, with the percentage of genetic material from the original SV129 stem cells reduced to a minimum.

6 Conclusion

In conclusion, our results support that the SNP in the I56i cis-regulatory element of *Dlx5* and *Dlx6*, probably has functional consequences for the development of the forebrain. We showed that the presence of the SNP affects the binding and the transcription activity induced by the Dlx proteins. A reduced binding of Dlx and possibly of other transcription factors results in decreased expression of *Dlx5* and *Dlx6* in the developing forebrain. This can results in decreased GABAergic inhibitory interneuron migration to the cortex as seen in transgenic mice. Even if no obvious behavioural consequences were observed, this decrease in migration may possibly lead to a subtle phenotype in behaviour of mice homozygous for the mutation. *Dlx* are likely key members of genetic cascades responsible for the establishment of a diversity of neuronal types such as GABAergic interneurons in the central nervous system (Stuhmer, Puelles et al. 2002). Even subtle disruptions in these cascades such as the SNP in the I56i enhancer seem to have major effects on the developing brain. This SNP is rare with an incidence of 0,3% in individual with autism, or one out of 161 individuals sequenced (Hamilton, Woo et al. 2005). This is not surprising since individuals with autism often exhibit different phenotypes associated with the disease. Thus, various independent risk factors increase the probability of having a full autistic phenotype, hence the name "autism spectrum disorder". Since ASD is genetically heterogeneous, future studies would be to cross our vI56i line with other ASD candidate genes (Folstein and Rosen-Sheidley 2001). The neuroligin family (*Nlgn1,2,3 and 4*) and neurexin1α (*Nrxn1α*) knockout mice would be good candidates since they have an important role in excitatory and inhibitory synaptic contacts (Dean and Dresbach 2006; Moy and Nadler 2008). A synergistic phenotype could be observed in a double mutant. This highlights the behavioural diversity and

complexity of mouse models of ASD. However, research for candidate genes underlying defects in ASD, psychological disorder for which little is known is important because it is affecting a large number of individuals and their families.

7 Bibliography

Acampora, D., G. R. Merlo, et al. (1999). "Craniofacial, vestibular and bone defects in mice lacking the *Distal-less*-related gene *Dlx5*." Development **126**: 3795-3809.

Ahituv, N., Y. Zhu, et al. (2007). "Deletion of Ultraconserved Elements Yields Viable Mice." PLoS Biol **5**(9): e234.

Alberts, B., Lewis, Raff, Roberts, D Watson (2007). Molecular Biology of the Cell, Garland Science.

An, X. L., J. X. Zou, et al. (2011). "Strain and sex differences in anxiety-like and social behaviors in C57BL/6J and BALB/cJ mice." Exp Anim **60**(2): 111-23.

Anderson, S. A., D. D. Eisenstat, et al. (1997). "Interneuron migration from basal forebrain to neocortex: dependence on Dlx genes." Science **278**: 474-476.

Anderson, S. A., D. D. Eisenstat, et al. (1997). "Interneuron migration from basal forebrain to neocortex: dependence on *Dlx* genes." Science **278**: 474-476.

Anderson, S. A., M. Qiu, et al. (1997). "Mutations of the homeobox genes Dlx-1 and Dlx-2 disrupt the striatal subventricular zone and differentiation of late-born striatal neurons." Neuron **19**: 27-37.

Anderson, S. A., M. Qiu, et al. (1997). "Mutations of the homeobox genes *Dlx-1* and *Dlx-2* disrupt the striatal subventricular zone and differentiation of late-born striatal neurons." Neuron **19**: 27-37.

Anney, R., L. Klei, et al. (2010). "A genome-wide scan for common alleles affecting risk for autism." Hum Mol Genet **19**(20): 4072-82.

Bejerano, G., M. Pheasant, et al. (2004). "Ultraconserved elements in the human genome." Science **304**(5675): 1321-5.

Ben-Ari, Y. (2002). "Excitatory actions of gaba during development: the nature of the nurture." Nat Rev Neurosci **3**(9): 728-39.

Benes, F. M. and S. Berretta (2001). "GABAergic interneurons: implications for understanding schizophrenia and bipolar disorder." Neuropsychopharmacology **25**(1): 1-27.

Benstaali, C., A. Mailloux, et al. (2001). "Circadian rhythms of body temperature and motor activity in rodents their relationships with the light-dark cycle." Life Sci **68**(24): 2645-56.

Bernardet, M. and W. E. Crusio (2006). "Fmr1 KO mice as a possible model of autistic features." ScientificWorldJournal **6**: 1164-76.

Bischoff, S., S. Tsai, et al. (2009). "Characterization of Conserved and Nonconserved Imprinted Genes in Swine." Biol Reprod.

Bishop, D. V., M. Maybery, et al. (2006). "Characteristics of the broader phenotype in autism: a study of siblings using the children's communication checklist-2." Am J Med Genet B Neuropsychiatr Genet **141**(2): 117-22.

Blundell, J., C. A. Blaiss, et al. (2010). "Neuroligin-1 deletion results in impaired spatial memory and increased repetitive behavior." J Neurosci **30**(6): 2115-29.

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

Branchi, I., D. Santucci, et al. (2001). "Ultrasonic vocalisation emitted by infant rodents: a tool for assessment of neurobehavioural development." Behav Brain Res **125**(1-2): 49-56.
Brookshire, B. R. and S. R. Jones (2009). "Direct and indirect 5-HT receptor agonists produce gender-specific effects on locomotor and vertical activities in C57 BL/6J mice." Pharmacol Biochem Behav **94**(1): 194-203.
Bulfone, A., L. Puelles, 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." J. Neurosci. **13**: 3155-3172.
Chadman, K. K., S. Gong, et al. (2008). "Minimal aberrant behavioral phenotypes of neuroligin-3 R451C knockin mice." Autism Res **1**(3): 147-58.
Chakrabarti, S. and E. Fombonne (2001). "Pervasive developmental disorders in preschool children." Jama **285**(24): 3093-9.
Cheh, M. A., J. H. Millonig, et al. (2006). "En2 knockout mice display neurobehavioral and neurochemical alterations relevant to autism spectrum disorder." Brain Res **1116**(1): 166-76.
Cheng, H., F. Zhang, et al. (2008). "Isolation and imprinting analysis of the porcine DLX5 gene and its association with carcass traits." Anim Genet **39**(4): 395-9.
Chimge, N., A. Makeyev, et al. (2008). "Identification of the TFII-I family target genes in the vertebrate genome." Proc Natl Acad Sci U S A **105**(26): 9006-10.
Chudley, A. E., E. Gutierrez, et al. (1998). "Outcomes of genetic evaluation in children with pervasive developmental disorder." J Dev Behav Pediatr **19**(5): 321-5.
Church, D. M., L. Goodstadt, et al. (2009). "Lineage-specific biology revealed by a finished genome assembly of the mouse." PLoS Biol **7**(5): e1000112.
Cobos, I., M. E. Calcagnotto, et al. (2005). "Mice lacking Dlx1 show subtype-specific loss of interneurons, reduced inhibition and epilepsy." Nat Neurosci **8**(8): 1059-68.
Constantino, J. N., C. Lajonchere, et al. (2006). "Autistic social impairment in the siblings of children with pervasive developmental disorders." Am J Psychiatry **163**(2): 294-6.
Cook, E. H., Jr., R. Y. Courchesne, et al. (1998). "Linkage-disequilibrium mapping of autistic disorder, with 15q11-13 markers." Am J Hum Genet **62**(5): 1077-83.
Crawley, J. (2007). What is wrong with my mouse ?: behavioral phenotyping of transgenic and knockout mice Wiley-Liss.
Dean, C. and T. Dresbach (2006). "Neuroligins and neurexins: linking cell adhesion, synapse formation and cognitive function." Trends Neurosci **29**(1): 21-9.
Depew, M. J., J. K. Liu, et al. (1999). "*Dlx5* regulates regional development of the branchial arches and sensory capsules." Development **126**: 3831-3846.
Depew, M. J., T. Lufkin, et al. (2002). "Specification of jaw subdivisions by *Dlx* genes." Science **298**: 381-385.
Dollé, P., M. Price, et al. (1992). "Expression of the murine *Dlx-1* homeobox gene during facial, ocular and limb development." Differentiation **49**: 93-99.
Edelmann, L., A. Prosnitz, et al. (2007). "An atypical deletion of the Williams-Beuren syndrome interval implicates genes associated with defective visuospatial processing and autism." J Med Genet **44**(2): 136-43.
Eisenstat, D. D., J.-K. Liu, et al. (1999). "DLX-1, DLX-2 and DLX-5 expression define distinct stages of basal forebrain differentiation." J. Comp. Neurol. **414**: 217-237.
Ekker, M., M.-A. Akimenko, et al. (1992). "Regional Expression of three homeobox transcripts in the inner ear of zebrafish embryos." Neuron **9**: 27-35.

Erlander, M. G. and A. J. Tobin (1991). "The structural and functional heterogeneity of glutamic acid decarboxylase: a review." Neurochem Res **16**(3): 215-26.
Etherton, M. R., C. A. Blaiss, et al. (2009). "Mouse neurexin-1alpha deletion causes correlated electrophysiological and behavioral changes consistent with cognitive impairments." Proc Natl Acad Sci U S A **106**(42): 17998-8003.
Evans, S. J., S. J. Watson, et al. (2003). "Evaluation of sensitivity, performance and reproducibility of microarray technology in neuronal tissue." Integr Comp Biol **43**(6): 780-5.
Ey, E., C. S. Leblond, et al. (2011). "Behavioral profiles of mouse models for autism spectrum disorders." Autism Res **4**(1): 5-16.
Feledy, J. A., M. I. Morasso, et al. (1999). "Transcriptional activation by the homeodomain protein Distal-less 3." Nucl. Acids Res. **27**: 764-770.
Feng, J., C. Bi, 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): 1470-84.
Fogarty, M., M. Grist, et al. (2007). "Spatial genetic patterning of the embryonic neuroepithelium generates GABAergic interneuron diversity in the adult cortex." J Neurosci **27**(41): 10935-46.
Folstein, S. E. and B. Rosen-Sheidley (2001). "Genetics of autism: complex aetiology for a heterogeneous disorder." Nat Rev Genet **2**(12): 943-55.
Fombonne, E., S. Quirke, et al. (2009). "Prevalence and interpretation of recent trends in rates of pervasive developmental disorders." Mcgill J Med **12**(2): 73.
Francke, U. (1999). "Williams-Beuren syndrome: genes and mechanisms." Hum Mol Genet **8**(10): 1947-54.
Freitag, C. M. (2007). "The genetics of autistic disorders and its clinical relevance: a review of the literature." Mol Psychiatry **12**(1): 2-22.
Galef, B. G., Jr. (2003). "Social learning of food preferences in rodents: rapid appetitive learning." Curr Protoc Neurosci **Chapter 8**: Unit 8 5D.
Gehring, W. J. (1993). "Exploring the homeobox." Gene **135**(1-2): 215-21.
Gehring, W. J. (1994). "Homeodomain proteins." Ann. Rev. Biochem. **63**: 487-526.
Gerlai, R. (1996). "Gene-targeting studies of mammalian behavior: is it the mutation or the background genotype?" Trends Neurosci **19**(5): 177-81.
Ghanem, N., O. Jarinova, et al. (2003). "Regulatory roles of conserved intergenic domains in vertebrate *Dlx* bigene clusters." Genome Res. **13**: 533-543.
Ghanem, N., M. Yu, 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): 5012-22.
Ghanem, N., M. Yu, et al. (2007). "Distinct *cis*-regulatory elements from the *Dlx1/Dlx2* locus mark distinct progenitor cell populations in the ganglionic eminences and different subtypes of adult cortical interneurons." J. Neurosci. **27**: 5012-5022.
Ghanem, N., M. Yu, 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): 415-24.
Gheusi G., B. R., Goodball G., Dantzer R. (1994). "Social and individual recognition in rodents: methodological aspects and neurobiological
bases." Behav Processes **33**: 59-88.

Gillberg, C. (1998). "Chromosomal disorders and autism." J Autism Dev Disord **28**(5): 415-25.
Griebel, G., C. Belzung, et al. (2000). "Differences in anxiety-related behaviours and in sensitivity to diazepam in inbred and outbred strains of mice." Psychopharmacology (Berl) **148**(2): 164-70.
Gulacsi, A. A. and S. A. Anderson (2008). "Beta-catenin-mediated Wnt signaling regulates neurogenesis in the ventral telencephalon." Nat Neurosci **11**(12): 1383-91.
Hallmayer, J., J. M. Hebert, et al. (1996). "Autism and the X chromosome. Multipoint sib-pair analysis." Arch Gen Psychiatry **53**(11): 985-9.
Hamilton, S. P., J. M. Woo, et al. (2005). "Analysis of four DLX homeobox genes in autistic probands." BMC Genet **6**: 52.
Hogg, S. (1996). "A review of the validity and variability of the elevated plus-maze as an animal model of anxiety." Pharmacol Biochem Behav **54**(1): 21-30.
Horike, S., S. Cai, et al. (2005). "Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome." Nat Genet **37**(1): 31-40.
Ille, F. and L. Sommer (2005). "Wnt signaling: multiple functions in neural development." Cell Mol Life Sci **62**(10): 1100-8.
Jamain, S., H. Quach, et al. (2003). "Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism." Nat Genet **34**(1): 27-9.
Jamain, S., K. Radyushkin, et al. (2008). "Reduced social interaction and ultrasonic communication in a mouse model of monogenic heritable autism." Proc Natl Acad Sci U S A **105**(5): 1710-5.
Justice, M. J., L. D. Siracusa, et al. (2011). "Technical approaches for mouse models of human disease." Dis Model Mech **4**(3): 305-10.
Keller, A., A. Nesvizhskii, et al. (2002). "Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search." Anal Chem **74**(20): 5383-92.
Kim, Y. S., B. L. Leventhal, et al. (2011). "Prevalence of Autism Spectrum Disorders in a Total Population Sample." Am J Psychiatry.
Kimura, M., Y. Kazuki, et al. (2004). "Dlx5, the mouse homologue of the human-imprinted DLX5 gene, is biallelically expressed in the mouse brain." J Hum Genet **49**(5): 273-7.
Klein-Tasman, B., C. Mervis, et al. (2007). "Socio-communicative deficits in young children with Williams syndrome: performance on the Autism Diagnostic Observation Schedule." Child Neuropsychol **13**(5): 444-67.
Kopp, C. (2001). "Locomotor activity rhythm in inbred strains of mice: implications for behavioural studies." Behav Brain Res **125**(1-2): 93-6.
Kraus, P. and T. Lufkin (2006). "Dlx homeobox gene control of mammalian limb and craniofacial development." Am J Med Genet A **140**(13): 1366-74.
Kuemerle, B., F. Gulden, et al. (2007). "The mouse Engrailed genes: a window into autism." Behav Brain Res **176**(1): 121-32.
Latchman, D. S. (1997). "Transcription factors: an overview." Int J Biochem Cell Biol **29**(12): 1305-12.
Le, T., G. Du, 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): 19071-81.

Lee, E. C., D. Yu, et al. (2001). "A highly efficient Escherichia coli-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA." Genomics **73**(1): 56-65.
Lettice, L., A. Hill, et al. (2008). "Point mutations in a distant sonic hedgehog cis-regulator generate a variable regulatory output responsible for preaxial polydactyly." Hum Mol Genet **17**(7): 978-85.
Lincoln, A., Y. Searcy, et al. (2007). "Social interaction behaviors discriminate young children with autism and Williams syndrome." J Am Acad Child Adolesc Psychiatry **46**(3): 323-31.
Liu, J. K., I. Ghattas, 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.
Liu, J. K., I. Ghattas, 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**: 498-512.
Liu, X., N. Novosedlik, et al. (2009). "The DLX1and DLX2 genes and susceptibility to autism spectrum disorders." Eur J Hum Genet **17**(2): 228-35.
Livak, K. J. and T. D. Schmittgen (2001). "Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method." Methods **25**(4): 402-8.
Long, J., I. Cobos, et al. (2009). "Dlx1&2 and Mash1 transcription factors control MGE and CGE patterning and differentiation through parallel and overlapping pathways." Cereb Cortex **19 Suppl 1**: i96-106.
Long, J., C. Swan, et al. (2009). "Dlx1&2 and Mash1 transcription factors control striatal patterning and differentiation through parallel and overlapping pathways." J Comp Neurol **512**(4): 556-72.
Mao, R., D. T. Page, et al. (2009). "Reduced conditioned fear response in mice that lack Dlx1 and show subtype-specific loss of interneurons." J Neurodev Disord **1**(3): 224-36.
Marin, O. and J. L. R. Rubenstein (2001). "A long, remarkable journey: tangential migration in the telencephalon." Nat. Rev. Neurosci. **2**: 780-790.
McFarlane, H. G., G. K. Kusek, et al. (2008). "Autism-like behavioral phenotypes in BTBR T+tf/J mice." Genes Brain Behav **7**(2): 152-63.
Missler, M., W. Zhang, et al. (2003). "Alpha-neurexins couple Ca2+ channels to synaptic vesicle exocytosis." Nature **423**(6943): 939-48.
Morasso, M. I., K. A. Mahon, et al. (1995). "A *Xenopus* Distal-less gene in transgenic mice: conserved regulation in distal limb epidermis and other sites of epithelial-mesenchymal interactions." Proc. Natl. Acad. Sci. USA **92**: 3968-3972.
Moy, S. S. and J. J. Nadler (2008). "Advances in behavioral genetics: mouse models of autism." Mol Psychiatry **13**(1): 4-26.
Moy, S. S., J. J. Nadler, et al. (2004). "Sociability and preference for social novelty in five inbred strains: an approach to assess autistic-like behavior in mice." Genes Brain Behav **3**(5): 287-302.
Moy, S. S., J. J. Nadler, et al. (2009). "Social approach in genetically engineered mouse lines relevant to autism." Genes Brain Behav **8**(2): 129-42.
Moy, S. S., J. J. Nadler, et al. (2007). "Mouse behavioral tasks relevant to autism: phenotypes of 10 inbred strains." Behav Brain Res **176**(1): 4-20.

Nakamura, S., D. W. Stock, et al. (1996). "Genomic analysis of a new mammalian distal-less gene: Dlx7." Genomics **38**: 314-324.
Nakashima, N., T. Yamagata, et al. (2009). "Expression analysis and mutation detection of DLX5 and DLX6 in autism." Brain Dev.
Nelson, W. J. and R. Nusse (2004). "Convergence of Wnt, beta-catenin, and cadherin pathways." Science **303**(5663): 1483-7.
Nesvizhskii, A., A. Keller, et al. (2003). "A statistical model for identifying proteins by tandem mass spectrometry." Anal Chem **75**(17): 4646-58.
Newschaffer, C. J., D. Fallin, et al. (2002). "Heritable and nonheritable risk factors for autism spectrum disorders." Epidemiol Rev **24**(2): 137-53.
Nurmi, E. L., Y. Bradford, et al. (2001). "Linkage disequilibrium at the Angelman syndrome gene UBE3A in autism families." Genomics **77**(1-2): 105-13.
Okita, C., M. Meguro, et al. (2003). "A new imprinted cluster on the human chromosome 7q21-q31, identified by human-mouse monochromosomal hybrids." Genomics **81**(6): 556-9.
Paina, S., D. Garzotto, et al. (2011). "Wnt5a is a transcriptional target of Dlx homeogenes and promotes differentiation of interneuron progenitors in vitro and in vivo." J Neurosci **31**(7): 2675-87.
Panganiban, G. and J. L. R. Rubenstein (2002). "Developmental functions of the *Distal-less/Dlx* homeobox genes." Development **129**: 4371-4386.
Park, B. K., S. M. Sperber, et al. (2004). "Intergenic enhancers with distinct activities regulate Dlx gene expression in the mesenchyme of the branchial arches." Dev Biol **268**(2): 532-45.
Paterson, N. E., M. Iwunze, et al. (2010). "Comparison of the predictive validity of the mirror chamber and elevated plus maze tests in mice." J Neurosci Methods **188**(1): 62-70.
Petryniak, M. A., G. B. Potter, et al. (2007). "Dlx1 and Dlx2 control neuronal versus oligodendroglial cell fate acquisition in the developing forebrain." Neuron **55**(3): 417-33.
Poitras, L., N. Ghanem, et al. (2007). "The proneural determinant MASH1 regulates forebrain Dlx1/2 expression through the I12b intergenic enhancer." Development **134**(9): 1755-65.
Porteus, M. H., A. Bulfone, et al. (1991). "Isolation and characterization of a novel cDNA clone encoding a homeodomain that is developmentally regulated in the ventral forebrain." Neuron **7**: 221-229.
Price, M., M. Lemaistre, et al. (1991). "A mouse gene related to distal-less shows a restricted expression in the developing forebrain." Nature **351**: 748-751.
Purves, D., G. J. Augustine, et al. (2008). Neuroscience, Sinauer Associates and Sumanas, Inc.
Qiu, M., A. Bulfone, 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**: 165-184.
Radyushkin, K., K. Hammerschmidt, et al. (2009). "Neuroligin-3-deficient mice: model of a monogenic heritable form of autism with an olfactory deficit." Genes Brain Behav **8**(4): 416-25.

Richler, E., J. G. Reichert, et al. (2006). "Autism and ultraconserved non-coding sequence on chromosome 7q." Psychiatr Genet **16**(1): 19-23.
Risch, N., D. Spiker, et al. (1999). "A genomic screen of autism: evidence for a multilocus etiology." Am J Hum Genet **65**(2): 493-507.
Robinson, G. W. and K. A. Mahon (1994). "Differential and overlapping expression domains of *Dlx-2* and *Dlx-3* suggest distinct roles for *Distal-less* homeobox genes in craniofacial development." Mech. Devel. **48**: 199-215.
Robinson, G. W., S. Wray, et al. (1991). "Spatially restricted expression of a member of a new family of murine distal-less homeobox genes in the developing forebrain." New Biol. **3**: 1183-1194.
Rogers, S. J. and S. Ozonoff (2005). "Annotation: what do we know about sensory dysfunction in autism? A critical review of the empirical evidence." J Child Psychol Psychiatry **46**(12): 1255-68.
Rubenstein, J. L. and M. M. Merzenich (2003). "Model of autism: increased ratio of excitation/inhibition in key neural systems." Genes Brain Behav **2**(5): 255-67.
Ryan, B. C., N. B. Young, et al. (2010). "Social deficits, stereotypy and early emergence of repetitive behavior in the C58/J inbred mouse strain." Behav Brain Res **208**(1): 178-88.
Sambrook, J., E. F. Fritsch, et al. (1989). Molecular cloning, a laboratory manual, 2nd Ed. New York, Cold Spring Harbor Laboratory Press.
Sanes D, Reh T.A, et al. (2011). Development of the Nervous System Academic Press.
Sanes, D., T. A. Reh, et al. (2011). Development of the Nervous System Academic Press.
Scattoni, M. L., J. Crawley, et al. (2009). "Ultrasonic vocalizations: a tool for behavioural phenotyping of mouse models of neurodevelopmental disorders." Neurosci Biobehav Rev **33**(4): 508-15.
Scattoni, M. L., S. U. Gandhy, et al. (2008). "Unusual repertoire of vocalizations in the BTBR T+tf/J mouse model of autism." PLoS One **3**(8): e3067.
Scherer, S. W., H. H. Q. Heng, et al. (1995). "Assignment of the human homolog of mouse Dlx3 to chromosome 17q21.3-q22 by analysis of somatic cell hybrids and fluorescence in situ hybridization." Mammal. Genome **6**: 310-311.
Schroer, R. J., M. C. Phelan, et al. (1998). "Autism and maternally derived aberrations of chromosome 15q." Am J Med Genet **76**(4): 327-36.
Schule, B., H. H. Li, et al. (2007). "DLX5 and DLX6 expression is biallelic and not modulated by MeCP2 deficiency." Am J Hum Genet **81**(3): 492-506.
Scott, M. P., J. W. Tamkun, et al. (1989). "The structure and function of the homeodomain." Biochim. Biophys. Acta **989**: 25-48.
Silverman, J. L., M. Yang, et al. (2010). "Behavioural phenotyping assays for mouse models of autism." Nat Rev Neurosci **11**(7): 490-502.
Simeone, A., D. Acampora, et al. (1994). "Cloning and characterization of two members of the vertebrate *Dlx* gene family." Proc. Natl. Acad. Sci. USA **91**: 2250-2254.
Skuse, D. H. (2007). "Rethinking the nature of genetic vulnerability to autistic spectrum disorders." Trends Genet **23**(8): 387-95.
Smith, A., J. Zhang, et al. (2008). "Gene expression analysis on sections of zebrafish regenerating fins reveals limitations in the whole-mount in situ hybridization method." Dev Dyn **237**(2): 417-25.
Stock, D. W., D. L. Ellies, et al. (1996). "The evolution of the vertebrate Dlx gene family." Proc. Natl. Acad. Sci. USA **93**: 10858-10863.

Stock, D. W., D. L. Ellies, et al. (1996). "The evolution of the vertebrate Dlx gene family." Proc. Natl. Acad. Sci. USA **93**: 10858-10863.
Stuhmer, T., S. A. Anderson, et al. (2002). "Ectopic expression of the Dlx genes induces glutamic acid decarboxylase and Dlx expression." Development **129**: 245-252.
Stuhmer, T., S. A. Anderson, et al. (2002). "Ectopic expression of the *Dlx* genes induces glutamic acid decarboxylase and *Dlx* expression." Development **129**: 245-252.
Stuhmer, T., L. Puelles, et al. (2002). "Expression from a Dlx gene enhancer marks adult mouse cortical GABAergic neurons." Cereb Cortex **12**(1): 75-85.
Szatmari, P., A. D. Paterson, et al. (2007). "Mapping autism risk loci using genetic linkage and chromosomal rearrangements." Nat Genet **39**(3): 319-28.
Umeda, T., N. Takashima, et al. (2010). "Evaluation of Pax6 mutant rat as a model for autism." PLoS One **5**(12): e15500.
Valenzuela, D. M., A. J. Murphy, et al. (2003). "High-throughput engineering of the mouse genome coupled with high-resolution expression analysis." Nat Biotechnol **21**(6): 652-9.
Van der Aa, N., L. Rooms, et al. "Fourteen new cases contribute to the characterization of the 7q11.23 microduplication syndrome." Eur J Med Genet **52**(2-3): 94-100.
Varoqueaux, F., G. Aramuni, et al. (2006). "Neuroligins determine synapse maturation and function." Neuron **51**(6): 741-54.
Wang, B., T. Lufkin, et al. (2011). "Dlx6 regulates molecular properties of the striatum and central nucleus of the amygdala." J Comp Neurol.
Wang, Y., C. Dye, et al. (2010). "Dlx5 and Dlx6 regulate the development of parvalbumin-expressing cortical interneurons." J Neurosci **30**(15): 5334-45.
Wang, Y., C. A. Dye, et al. (2010). "Dlx5 and Dlx6 regulate the development of parvalbumin-expressing cortical interneurons." J Neurosci **30**(15): 5334-45.
Wassink, T. H., J. Piven, et al. (2001). "Evidence supporting WNT2 as an autism susceptibility gene." Am J Med Genet **105**(5): 406-13.
Wolfer, D. P., W. E. Crusio, et al. (2002). "Knockout mice: simple solutions to the problems of genetic background and flanking genes." Trends Neurosci **25**(7): 336-40.
Wolpert, L., R. Beddington, et al. (1998). Principles of development, Current biology Ltd.
Wonders, C. and S. Anderson (2006). "The origin and specification of cortical interneurons." Nat Rev Neurosci **7**(9): 687-96.
Wonders, C., L. Taylor, et al. (2008). "A spatial bias for the origins of interneuron subgroups within the medial ganglionic eminence." Dev Biol **314**(1): 127-36.
Yee, S.-P. and P. W. J. Rigby (1993). "The regulation of myogenin gene expression during the embryonic development of the mouse." Genes Dev. **7**: 1277-1289.
Zerucha, T. and M. Ekker (2000). "*Distal-less*-related homeobox genes of vertebrates: evolution, function, and regulation." Biochem. Cell Biol. **78**: 593-601.
Zerucha, T., T. Stuhmer, et al. (2000). "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**: 709-721.
Zerucha, T., T. Stuhmer, et al. (2000). "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**: 709-721.
Zhang, H., G. Hu, et al. (1997). "Heterodimerization of Msx and Dlx homeoproteins results in functional antagonism." Mol. Cell. Biol. **17**: 2920-2932.

Zhang, Y., F. Buchholz, et al. (1998). "A new logic for DNA engineering using recombination in *Escherichia coli*." Nature Genet. **20**: 123-128.
Zhou, Q. P., T. N. Le, 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): 884-92.