An SNP in an ultraconserved regulatory element affects Dlx5/Dlx6 regulation in the forebrain

Luc Poitras1, Man Yu1, Cindy Lesage-Pelletier1, Ryan B. MacDonald1, Jean-Philippe Gagné2, Gary Hatch1, Isabelle Kelly3, Steven P. Hamilton4,5, John L. R. Rubenstein5,6, Guy G. Poirier2,3 and Marc Ekker1,*

SUMMARY
Dlx homeobox genes play a crucial role in the migration and differentiation of the subpallial precursor cells that give rise to various subtypes of γ-aminobutyric 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 in mice 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 reduced 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 led to the identification of a novel regulator of Dlx gene expression, general transcription factor 2 I (Gtf2i), which is among the genes most often deleted in Williams-Beuren syndrome, 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.

KEY WORDS: Autism, Enhancer, Polymorphism, Telencephalon, Transgenics, Mouse

INTRODUCTION
During embryonic development, two major types of neurons migrate and populate the cortex: the excitatory glutamate-releasing neurons and the inhibitory γ-aminobutyric acid (GABA)-releasing interneurons. It has been proposed that some forms of autism are the result 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 are controlled to a large extent by homeodomain-containing transcription factors of the Dlx family (Anderson et al., 1997a; Anderson et al., 1997b). Four Dlx genes, Dlx1, Dlx2, Dlx3 and Dlx6, are expressed in two distinct forebrain domains: the ventral telencephalon and parts of the diencephalon (Bulfone et al., 1993; Robinson et al., 1991). 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 et al., 1997a; Stuhmer et al., 2002a; Stuhmer et al., 2002b).

The Dlx genes are organized in convergently transcribed bigene clusters and have highly similar patterns of expression, possibly owing to shared cis-acting regulatory elements (CREs) (Nakamura et al., 1996; Porteus et al., 1991; Price et al., 1991; Robinson and Malon, 1994; Robinson et al., 1991; Scherer et al., 1995; Simeone et al., 1994; Stock et al., 1996). Through phylogenetic footprinting and transgenic analysis, two CREs, called I56i and I56ii, have been discovered in the Dlx5/Dlx6 intergenic region (Zerucha et al., 2000). In reporter assays, both elements are active in the telencephalon and diencephalon. In-depth analysis of their enhancer activities has shown that these individual enhancers may participate in Dlx gene regulation during the development of specific populations of GABAergic interneurons (Ghanem et al., 2007; Ghanem et al., 2008).

The Dlx genes have been associated with human autism spectrum disorders (ASDs) (Hamilton 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 that these two loci can be linked to ASDs (Liu et al., 2009; Nakashima et al., 2009). Liu and collaborators have shown that two polymorphisms found in the Dlx1/Dlx2 loci could potentially increase susceptibility or cause autism in cohorts of multiplex families in which there is a greater genetic component for these conditions (Liu et al., 2009). In a search for autism genomic variants in the Dlx1/2 and Dlx5/6 loci, we identified 31 single-nucleotide polymorphisms (SNPs) and two insertion/deletion polymorphisms. One adenine-to-guanine SNP was found at position 182 of I56i (Hamilton et al., 2005). The location of the SNP coincides with one of the two putative Dlx
binding sites (Feledy et al., 1999) within I56i (Zerucha et al., 2000). This region of I56i also encodes the second exon of the long polyadenylated non-coding RNA (lincRNA) Evf2 (Dlx5os1 – Mouse Genome Informatics) (Feng et al., 2006), which was recently shown to contribute to the mechanisms by which the Dlx5/Dlx6 bimere cluster is regulated (Bond et al., 2009). Furthermore, the I56i CRE was identified as one of 481 ultraconserved elements of the mouse genome (exhibiting 100% identity, with no insertion or deletions, on DNA fragments longer than 200 bp between orthologous regions of the human, rat and mouse genomes) (Bejerano et al., 2004). The 182-SNP was found to be part of a completely conserved 8 bp motif based on the sequence conservation between 40 vertebrate genomes (Fig. 1). Therefore, the unexpected identification of this SNP in the I56i enhancer suggested that it could have functional consequences for Dlx5/Dlx6 regulation during development. 

Here, we present evidence suggesting that the I56i SNP affects transcriptional regulation of the Dlx5/Dlx6 bimere 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 confirmed that Dlx proteins are bound at or near the SNP sequence and also led to the identification of Gtf2i as a new regulator of the Dlx5/Dlx6 bimere cluster. We propose that the impaired I56i enhancer activity resulting from the 182-SNP could impact on Dlx function and therefore cortical interneuron development and could constitute a contributing factor to developmental abnormalities underlying ASDs.

MATERIALS AND METHODS

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 oligonucleotides I56i-1 for and I56i-316 rev (for oligonucleotide sequences, see Table S1 in the supplementary material). 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 reactions. Finally, the mutant enhancer was purified on a gel, TA-cloned and sequenced.

Transgenic experiments

The I56i and vI56i enhancers were subcloned into the XhoI 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 described (Zerucha et al., 2000). Two transgenic lines and three primary transgenic embryos were obtained for the I56i-lacZ construct. These transgenic embryos were compared with two mouse transgenic lines obtained with the I56i-lacZ construct. These transgenic embryos were compared with two mouse transgenic lines obtained with the I56i-lacZ construct (Ghanem et al., 2003; Ghanem et al., 2007; Zerucha et al., 2000).

Histology and double immunohistochemistry

Fixation of E13.5 telencephalon was carried out overnight in 4% paraformaldehyde (PFA) in 1×PBS. For adult brains, mice (P30) were deeply anesthetized and subjected to 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 2 hours 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 a cryostat (Leica CM3050 S). Double immunostaining was performed as previously described (Ghanem et al., 2007). An antibody from Abnova was used for Gtf2i immunohistochemistry (IHC-00273).

Nuclear extract preparation

Nuclear extracts were prepared from the brains of E13.5 mouse embryos (1 ml of tissue) as described (Sambrook et al., 1989). In addition, the protein extract was dialyzed against the resuspension buffer using a Spectra/Por dialysis cuvette (VWR) overnight at 4°C. Aliquots of 50 μl were frozen in liquid nitrogen and stored at −80°C.

DNase I footprinting analysis

A portion of the I56i enhancer corresponding to nucleotides 187-316 was amplified by PCR using primers zI56i-187 and zI56i-316. PCR fragments were TA-cloned in the pDrive vector (Qiagen). This fragment was recovered by digestion with EcoRI and KpnI followed by gel purification (QIAquick, Qiagen). Directional labeling of each fragment was performed by 5’ end fill using the large fragment of DNA polymerase I (Invitrogen). Footprinting reactions were carried out as described (Poitras et al., 2007).

Protein-DNA complex 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 the supplementary material) in 1× EMSA binding buffer [7 mM Tris pH 7.5, 81 mM NaCl, 2.75 mM dithiothreitol, 5 mM MgCl2, 0.05% NP40, 1 mg/ml bovine serum albumin, 25 μg/ml poly(dI-dC), 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 1× 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 2% SDS, 10% glycerol, 5% 2-mercaptoethanol, 0.002% Bromophenol Blue and 0.0625 M Tris-HCl). Removal of the sepharose beads was achieved by filtering the boiled reaction through a SPIN-X centrifuge tube filter (Corning).

Sample preparation for mass spectrometry analysis

Eluted proteins were resolved using a 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).

Liquid chromatography (LC)-tandem mass spectrometry (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 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 an LTQ linear ion-trap mass spectrometer equipped with a nanoelectrospray ion source (Thermo Electron, San Jose, CA, USA). Peptide separation took place within a PicoFrit BioBasic C18 column of 10 cm x 0.075 mm internal diameter (New Objective, Woburn, MA, USA) with a linear gradient from 2 to 50% solvent B (acetonitrile, 0.1% formic acid) in 30 minutes at 200 nl/minute. Mass spectra were acquired using the data-dependent acquisition mode (Xcalibur 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 second exclusion) and the relative collisional fragmentation energy was set to 35%.
Interpretation of tandem mass spectrometry spectra
All MS/MS samples were analyzed using Mascot (version 2.2.0; Matrix Science, London, UK). Mascot was set up to search against the rodent Uniref100 protein database (TaxID: 9989) assuming digestion with trypsin. Fragment and parent ion mass tolerance were 0.5 Da and 2.0 Da, respectively. Iodoacetamide derivative of cysteine was specified as a fixed modification and oxidation of methionine was specified as variable modification. Two missed cleavages were allowed.

Criteria for protein identification
Scaffold (version 02_02_03; Proteome Software, Portland, OR, USA) was used to validate MS/MS-based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 99.5% probability, as specified by the Peptide Prophet algorithm (Keller et al., 2002). Protein probabilities were assigned by the Protein Prophet algorithm (Nevinszhiikii 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 was less than 1%.

Co-transfection experiments
A zebrafish dlx2a cDNA (845 bp) 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 described previously (Poitras et al., 2007). The zebrafish dlx5a cDNA (1.2 kb) was amplified by PCR and cloned into the BglII/XhoI sites of the pT72 expression vector. The rat Gtf2i cDNA (accession BC085815) subcloned into the expression vector pExpress-1 was purchased from Open Biosystems.

The wild-type and 182-SNP variant of the mouse 156i 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 et al., 2000). Luciferase assays were performed according to the manufacturer’s protocol (Promega).

Electrophoretic mobility shift assay (EMSA)
EMSA were performed using double-stranded oligonucleotides corresponding to the identified putative binding site. Complementary 21 nucleotide oligonucleotides (m156i FP3.5or and m156i FP3.rev) were annealed and end-labeled using [32P]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. EMSAs were carried out as described (Poitras 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 wild-type oligonucleotide competitor or a version of equal length containing the 182-SNP (see Fig. 1A) was used.

The competition assay with recombinant Dlx proteins was carried out using double-stranded 32P-labeled oligonucleotides corresponding to a triplicate of the core FP3 site (ATAATT), termed m156i FP3 X3, with fixed amounts of GST-Dlx5a protein. These binding reactions were competed with increasing amounts (1, 3 and 10 ng) of the unlabeled double-stranded wild-type oligonucleotide competitor (156i FP3 X3) or the variant version, v156i FP3 X3. Quantification of the percentage of the remaining complexes was achieved using ImageJ (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).

Western blots
Proteins from the previously described E13.5 nuclear extract were separated on an 8% SDS-PAGE gel and electrophoretically transferred to Hybond-C membrane (GE Healthcare) in 25 mM Tris, 192 mM glycine, 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% dried milk. The membranes were then incubated with an anti-Gtf2i antibody (Abnova, IHC-00273; 1/1000) in PBS-T containing 1% dried milk overnight at 4°C. After washing with PBS-T containing 1% dry milk, blots were incubated with horseradish peroxidase-conjugated goat anti-rabbit IgG (1/5000) in PBS-T containing 1% dry milk. The blots were developed using an HRP chemiluminescence substrate kit (Immun-Star).

RESULTS
An SNP affects the activity of the 156i enhancer
We previously identified seven variants in CREs of Dlx genes from a collection of DNA samples from autistic probands (Hamilton et al., 2005). Using sequence conservation analysis in orthologous sequences (see Fig. S1 in the supplementary material) and DNase1 footprinting to localize binding of trans-acting (Poitras et al., 2007) (data not shown; see also Fig. 4A), we found that position 182 in I56i (base 96,641,429 on chromosome 7, see Fig. 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 156i enhancer from 40 vertebrate genomes, we found that this adenine (underlined) was part of a completely conserved 8 bp motif (ATAATTAG) (Fig. 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 156i or its corresponding 182-SNP-containing variant (hereafter referred to as v156i) in combination with a β-globin minimal promoter-lacZ cassette (Yee and Rigby, 1993). As previously reported, the 156i enhancer targets the expression of the reporter cassette to the developing forebrain and branchial arches (Ghanem et al., 2003; Ghanem et al., 2007; Zerucha et al., 2000) (Fig. 2B). In the telencephalon of E11.5 mouse embryos, I56i targets lacZ expression to the subventricular zone (SVZ) and mantle zone (MZ) of the medial and lateral ganglionic eminences (MGE and LGE) (Fig. 2C). By contrast, lacZ expression driven by v156i-lacZ was reduced in all the aforementioned regions of the forebrain. A more pronounced effect was observed in the MGE (Fig. 2D-F). The reductions in transgene expression varied when comparing embryos from two independent transgenic lines and three primary transgenic animals, ranging from slight reductions to almost total absence of transgene expression (Fig. 2C-F; data not shown). As we never observed such variations in activity in at least nine primary transgenic embryos and four established lines expressing I56i-lacZ (Zerucha et al., 2000) (data not shown), it is most unlikely that the decreases we observed with v156i-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. 2G,H). Moreover, almost no v156i-lacZ-expressing cells could be detected in the SVZ and MZ of the caudal ganglionic eminence (CGE) (Fig. 2LJ). 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 156i enhancer can be observed in GABAergic interneurons that are migrating tangentially from the ventral telencephalon to the cortex (Ghanem et al., 2007). Hence, we evaluated the activity of the v156i-lacZ transgene in migrating interneurons at E13.5. In the brains of 156i-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. 3A-D). By contrast, only a few scattered migrating cells expressed the vI56i-lacZ transgene at both levels of the embryonic telencephalon (Fig. 3E-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 major subtypes of interneurons: the calcium-binding proteins calbindin, calretinin (calbindin 2) and parvalbumin (CB, CR and PV) and the neuropeptide somatostatin (SOM). These four markers were specifically chosen on the basis of previous findings showing that most CB+, PV+ and SOM+ interneurons originate from the MGE, whereas the CGE is one major origin of CR+ interneurons (Fogarty et al., 2007; Wonders and Anderson, 2006; Wonders et al., 2008).

We observed a significant decrease in co-labeling in the somatosensory cortex of the vI56i-lacZ mice as compared with the I56i-lacZ mice (Fig. 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. By 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 (Ghanem et al., 2007). Taken together, these results suggest that the 182-SNP has important functional consequences for I56i enhancer activity at both
may present reduced levels of Dlx5/Dlx6 expression, and that this in turn may affect the development of cortical interneurons

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 DNasel footprinting analysis using mouse E13.5 telencephalon nuclear protein extract. Seven DNA-protein complexes were observed, four of which are shown in Fig. 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. 2A) was incubated with the E13.5 telencephalon nuclear extract, four DNA-protein complexes were observed (Fig. 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. 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 weights of the E3 and E4 complexes were comparable to those of complexes produced by a recombinant Dlx2 (Fig. 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 of similar size.

To identify factors that bind to the I56i at and near the FP3 binding site, we carried out an affinity purification assay using a DNA fragment that encompasses 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 (see Table S2 in the supplementary material).

Consistent with previous observations that Dlx2 is able to bind and activate transcription through the I56i enhancer (Zerucha et al., 2000; Zhou et al., 2004), Dlx proteins were identified among the purified proteins. However, the most abundant protein isolated from the purification was general transcription factor 2I (Gtf2i), with 49 distinct peptides identified by mass spectrometry (see Table S2 in the supplementary material). Interestingly, a potential Gtf2i binding site (Chimge et al., 2008) is located halfway between the two Dlx binding sites of the sequence used for affinity purification (see Fig. S1 in the supplementary material).

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

We evaluated the affinity of the Dlx proteins for the FP3 binding site in its wild-type or 182-SNP-containing versions by 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 wild-type FP3 binding site than for the variant sequence (Fig. 6B). Next, we examined the effects of the 182-SNP on the transcriptional activity of the I56i enhancer in response to Dlx2 or Dlx5 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. 6C).

To test a possible involvement of Gtf2i in Dlx5/Dlx6 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. 6D). The transfection of a Gtf2i-expressing plasmid per se did not stimulate 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 with 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 gene regulation via the I56i enhancer.

DISCUSSION

In this report, we demonstrated that an SNP, originally isolated in a screen of autistic probands and located in an ultraconserved Dlx5/Dlx6 CRE (I56i), has functional consequences for the activity of the enhancer throughout forebrain development, continuing into adulthood. The identification of this functionally important SNP in the I56i enhancer provides 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 among orthologous regions of the human, rat and mouse genomes, the I56i enhancer is considered as one region of the genome that is not susceptible to variation (Bejerano et al., 2004). Furthermore, the SNP was found in a region of I56i that is identical in 40 vertebrates species surveyed (Fig. 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 is located within the FP3 protein binding site. This site was previously identified as one of two Dlx binding sites in the I56i enhancer. These Dlx binding sites play crucial roles in the auto- or cross-regulation mechanisms of the Dlx5/6 bigene (Zerucha et al., 2000). Thus, the presence of the 182-SNP within one of these important binding sites might, by itself, have functional consequences for the regulatory cascades controlling Dlx expression.

The hypothesis that the 182-SNP would affect auto- or cross-regulatory mechanisms at I56i is supported by our EMSA and co-transfection experiments. Dlx affinity for the FP3 sequence was impaired by the presence of the 182-SNP as was the transcriptional activity of I56i in response to Dlx2 or Dlx5. Co-transfection assays provided further 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 a similar manner to the E3 and E4 nuclear complexes. A potential explanation for the presence of two complexes is that Dlx2 binds to the FP3 site as a monomer and/or as a homodimer. It has been demonstrated that Dlx proteins can bind DNA as hetero- and homodimers (Zhang et al., 1997). Supershift experiments on these complexes with a number of anti-Dlx antibodies were inconclusive (data not shown). However, our affinity purification experiment enabled us to identify a new regulator of the Dlx genes: 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 plus Gtf2i combination was affected by the SNP. The mechanism by which Gtf2i interacts with Dlx proteins on the I56i enhancer is still unknown, but possibilities include: (1) Gtf2i interacts directly, or through an intermediate factor, with Dlx proteins independently of the ability of Gtf2i to bind to DNA [attempts at co-immunoprecipitating Gtf2i and Dlx2 in vitro have not been successful (data not shown)]; and (2) Gtf2i interacts with 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 gene 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) has been 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 by Francke, 1999). Although WBS children show high sociability, almost half of them display socio-communicative deficits characteristic of ASD (Klein-Tasman et al., 2007; Lincoln 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 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 the possibility that the mental retardation and autistic traits observed in this patient were due to a dosage effect of \textit{GTF2I} gene expression. Finally, a case of an autistic individual with a 7q11.23 duplication was recently reported (Van der Aa et al., 2009).

Altogether, there is accumulating evidence suggesting that members of the \textit{GTF2I} gene family play an important role in WBS, even though it remains unclear whether the \textit{GTF2I} genes can account for the 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 might affect the early process of GABAergic interneuron differentiation, as Dlx proteins have been shown to regulate \textit{Gad1} and \textit{Gad2} expression (Long et al., 2009a; Long et al., 2009b). As for the impact of the SNP on precursor migration, it has been shown that neuropilin 2 (Nrp2) is ectopically expressed when Dlx1 and Dlx2 are inactivated (Le et al., 2007). Nrp2 is a transmembrane protein that is implicated in repulsive axon guidance and, therefore,
it was proposed that downregulation of Nrp2 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 gene regulation in GABAergic interneurons of newborn or adult mice remains to be established. In mice, the brain of Dlx-I-null mutants displays no overt phenotype at birth, but later in life they develop epileptic seizures (Cobos et al., 2005). Dlx5/6+/- mice, the cortex of which 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 PV-expressing interneurons in Dlx5/6-/- and Dlx5+/+ mutants (Wang et al., 2010). Thus, Dlx proteins could adopt a specific role in the adult brain or subtle alterations that occur during development could have delayed consequences. In either case, a decrease in 15eti 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 et al., 2005). This SNP was paternally transmitted and was absent from an affected sibling (Richler et al., 2006). This situation is not exceptional considering that siblings, who share half the genetic risk, often display a different autist 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 reported that an increased proportion of unaffected siblings has subsyndromal concepts of autism, two independent studies have reported that an increased proportion of unaffected siblings has subsyndromal autistic impairments (Bishop et al., 2006; Constantino et al., 2006), supporting the notion of independent risk factors or genetic susceptibility factors.

Paternal imprinting of the Dlx5 locus was recently documented (Horike et al., 2005; Okita et al., 2003) and would have phenotypic implications for an SNP in 15eti, 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 the misregulation of several genes, including Dlx5, resulting from the loss of their imprinting, might contribute to the etiology of this syndrome. Imprinting of Dlx5 has also been observed in human and swine loci (Cheng et al., 2008; Okita et al., 2003). However, paternal imprinting of Dlx5 has been challenged by other studies reporting biallelic expression of the gene (Bischoff et al., 2009; Kimura et al., 2004; Schule et al., 2007). When one copy of the Dlx5/Dlx6 locus is silenced by imprinting, the presence of the SNP on the remaining allele will result in a decrease in Dlx5/Dlx6 expression that is much more substantial than in the absence of imprinting.

The reason for the exceptional sequence conservation of some non-coding regulatory elements in vertebrate genomes is still puzzling. Although it is unlikely to be due to chance that these regions have been maintained for hundreds of millions of years, it is also unclear what their function is, given the lack of a phenotype when some of the ultraconserved regions are deleted (Ahituv 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.

Acknowledgements
We thank Mélanie Debiais-Thiabaud and Fabien Avaron for helpful discussions and Adriana Gamarbottta 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). Work in the laboratory of J.L.R.R. was supported by research grants from: Nina Ireland, the Althea Foundation, Cure Autism Now and R37 MH049428. Work in the laboratory of G.G.P. was supported by grants from Genome Canada and Génome Québec. G.G.P. is a recipient of a Canadian Research Chair in Proteomics.

Competing interests statement
The authors declare no competing financial interests.

Author contributions
L.P. performed experiments in Figs 2–6 and contributed to the writing of the manuscript. M.Y. carried out the immunolocalization experiments presented in Figs 2–4. C.L.P. carried out transfections and EMSA experiments shown in Fig. 6. R.B.M. participated in the EMSA experiments shown in Fig. 6. J.-P.G., I.K. and G.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.P.H. and J.L.R.R. identified the SNP presented in this study and contributed to the writing of this manuscript. M.E. participated in the design of the experiments and contributed to the writing of the manuscript.

Supplementary material
Supplementary material for this article is available at http://dev.biologists.org/lookup/suppl/doi:10.1242/dev.051052/-/DC1

References

Supplementary material
Supplementary material for this article is available at http://dev.biologists.org/lookup/suppl/doi:10.1242/dev.051052/-/DC1


