A NOVEL ENHANCER OF DBX1 IS ABLE TO DRIVE NEURAL TISSUE SPECIFIC REPORTER GENE EXPRESSION

By

JENNIFER KIM

A thesis submitted to the

Graduate School – New Brunswick

Rutgers, The State University of New Jersey

And

The Graduate School of Biomedical Sciences

University of Medicine and Dentistry of New Jersey

in partial fulfillment of the requirements

for the degree of

Master of Science

Graduate Program in Biomedical Engineering

Written under the direction of

Dr. Li Cai

and approved by

_____

_____

_____

New Brunswick, New Jersey

October, 2010

ABSTRACT OF THE THESIS

A Novel Enhancer of Dbx1 Is Able To Drive Neural Tissue Specific Reporter Gene Expression

By JENNIFER KIM

Thesis Director:

Dr. Li Cai

Dbx1 is a homeobox gene that plays an essential role in the patterning of the central nervous system (CNS) during development. Its expression pattern is regionally restricted within the brain and spinal cord and has been linked to defining neuronal-specific domains. Enhancers, a major type of transcriptional regulatory element, are important in determining tissue specific gene expression. Here we have identified a novel enhancer of Dbx1 in order to understand mechanisms of spatial and temporal regulation of the Dbx1 gene. Using a multi-genome sequence alignment tool, a conserved non-coding region, Dbx1CR5 was identified as a putative enhancer of Dbx1. A green fluorescent protein (GFP) reporter assay system was used to verify the ability of Dbx1CR5 to drive neural tissue specific reporter gene expression in the developing chick CNS. GFP was detected between embryonic days 3 and 18 and was restricted to areas of the mesencephalon. Morphology and molecular staining of GFP+ tissues indicate that GFP is expressed in a heterogeneous population of cells. Within Dbx1CR5, multiple regions were identified to individually activate GFP expression in varying spatial patterns. Our findings support the notion that multiple binding elements reside within the 1135bp region of Dbx1CR5. Through this research we hope to identify factors that may one day be used to develop potential and novel therapeutics.

TABLE OF CONTENTS:

LIST OF TABLES

LIST OF ILLUSTRATIONS

BACKGROUND:

Central Nervous System Development:

Development of the central nervous system (CNS) begins from the neural tube. At early developmental stages, the neural tube separates into four primary regions: the forebrain (prosenecephalon), midbrain (mesencephalon), hindbrain (rhombencephalon), and spinal cord. As development continues, the forebrain divides into the telencephalon and diencephalon, while the hindbrain divides into the metencephalon and myelencephalon. These divisions then become the basic structures of the brain as illustrated below (fig.a).

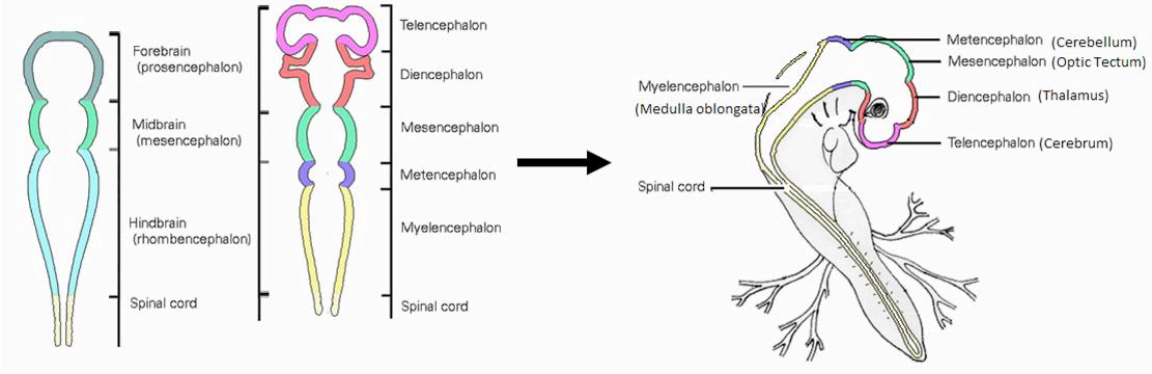

Figure a: Developmental Structures of the Neural Tube (modified from Developmental Biology Online[1])

On a cellular level, the neural tube begins with a layer of neuroepithelial cells which are arranged in a columnar pattern. As the neuroepithelium thickens with development, neuroepithelial cells become longer, thinner and demonstrate bipolar morphology[2]. The neuroepithelial cells give rise to radial glial cells and neural progenitors which then continue to either divide or differentiate into mature glia and neurons that eventually populate the rest of the CNS[3-4 5]. Dividing cells cluster at the

ventricular zone and travel up radial glia that function to guide migrating cells to their final positions in the developing brain[6-7]. Radial glia have also been found to give rise to precursor cells[5-6,8].

During neurogenesis, various molecules have been found to be fundamental in defining the pattern of tissue development which govern the basic structure and organization of the CNS[9]. These molecules are expressed in a spatio-temporal pattern and provide specific environmental cues that control migration and differentiation during development. Molecules such as Shh, BMP, and Wnts are involved in dorsoventral patterning[10], while others such as RA and homeobox genes have been identified to control rostrocaudal domains[11-12].

Dbx1:

Homeobox genes encode transcriptional regulators that are commonly known to specify spatial domains during development**.** Dbx1 is a homeobox gene that encodes for the Dbx1 transcription factor. Although very little has yet to be elucidated about the specific roles and functions that Dbx1 plays during development, researchers have established that Dbx1 plays an essential role in the patterning of the central nervous system (CNS) during embryogenesis[13]. Early studies have identified expression of Dbx1 to be extremely restricted in the CNS[14-15]. The majority of studies have been dedicated towards examining the role of Dbx1 in the brain and spinal cord of the mouse, while chick embryo studies have mainly focused on characterizing expression patterns only within the spinal cord at Hamburger Hamilton (HH) stages 15-20 (50-72hrs)[2].

Through mouse studies, researchers have defined Dbx1 to be expressed in restricted regions of the CNS within the telencephalon, diencephalon, dorsal mesencephalon, and spinal cord during early gestation[14]. At later gestational stages, Dbx1 was found to be continually expressed, however in a more restricted manner in the dorsal mesencephalon, diencephalon, as well as the primitive cerebellum[15]. Expression of Dbx1 has been observed to be limited to regions of active mitosis suggesting that Dbx1 is involved with early specification of Dbx1-expressing cells and their progenies[15]. Further evidence correlates strong Dbx1 expression to progenitor cells in the ventricular zone at the boundaries between dorsal and ventral parts of the neural tube[14,16-18]. Dbx1-derived cells have also been shown to have a high capacity to migrate from their sites of origin to populate different cortical regions[19-20].

In the spinal cord, distinct classes of neurons are generated from progenitor cells located at different regions which are regulated by the patterning of various proteins within the spinal cord[21-22]. Studies have shown that Dbx1 defines the distinct progenitor domain of V0 interneurons along the dorsoventral axis of the neural tube[16,21]. Cells derived from this Dbx1-expressing domain have also been linked to a small population of oligodendrocytes and astrocytes[23].

Dbx1's role in neuronal development has also been observed using mutant studies. For example, in one study where Dbx1 was knocked out, neural progenitors failed to generate V0 neurons and instead gave rise to characteristically different V1 neurons as determined by transcription factor profile, neurotransmitter phenotype, as well as the migratory pattern and axonal trajectory[24]. Other mutational studies have also

identified Dbx1-dependent interneurons to be key components in coordination of left-right motor activities necessary for the control of stepping movements in mammals[16,19].

As of today, only one enhancer of Dbx1 has been studied. It is described as a 5.7kbp region located directly 5' of the Dbx1 gene[13,25]. Through transgenic mouse studies using the 5.7kb region, the expression patterns of the reporter gene were found to be similar to the expression patterns of the Dbx1 gene but with mild differences[25]. These differences have therefore left the possibility of additional regulatory elements which are responsible for the fine-tuning of Dbx1 expression[13].

Gene Regulation by Enhancers:

With the completion of the human genome project, has come an even greater task of identifying and locating all cis-acting gene regulatory elements within non-coding regions[5-6]. The significance in identifying and characterizing gene regulatory elements not only lies in understanding, but also in identifying the important factors involved in gene regulation. By using this information, it is possible to not only understand global gene regulation, but more importantly offers a venue through which therapeutics may someday evolve and arise[26].

Approximately 98% of the human genome consists of non-protein coding regions, most of which have functions that are yet to be discovered[27]. This has led to the emergence of a new field of comparative genomics through which genome-based analysis methods have become a powerful tool in identifying functionally important genomic elements[28-31]. Through comparative genomics, researchers have established the

importance of sequence conservation between evolutionarily diverse species through the discovery of homology within non-coding regions[32].

Enhancers are one of the major types of gene regulatory elements. They are cis-acting, and have been linked to determining cell/tissue- and time-specific gene expression patterns[33]. Gene transcription is activated by the binding of activator proteins to the enhancer region. This binding is believed to cause the recruitment of transcriptional machinery to the promoter which then initiates gene expression[34]. An illustration of the transcriptional mechanisms is shown in figure b. Early research has demonstrated that enhancers act independent of position and orientation with respect to the promoter[35-36]. Thus enhancers can operate from distant chromosomes and are not limited to areas near the gene they control and can be found upstream, downstream, and even within the gene of control.

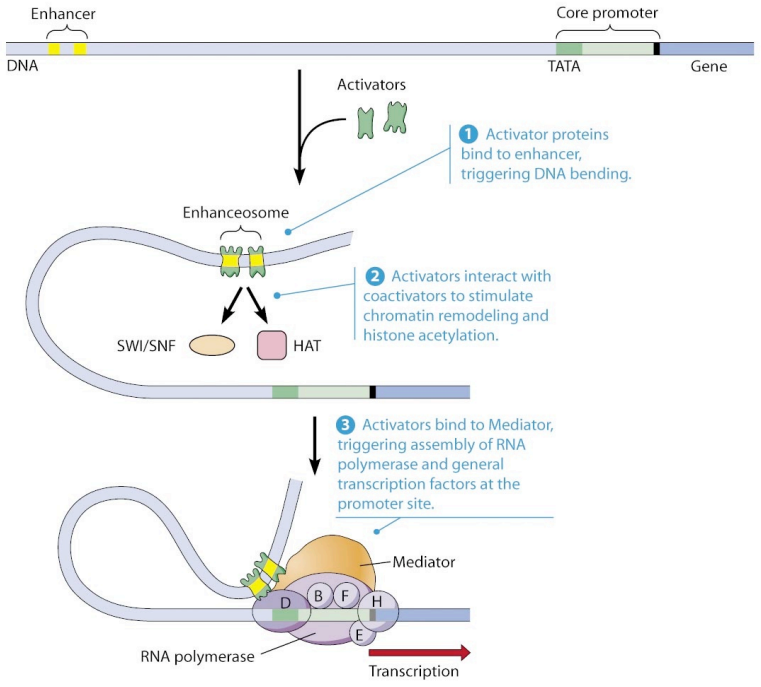

Figure b Transcriptional Regulation by Enhancers [37]

Enhancers have been defined by their unique ability to activate transcription within a reporter gene construct from a linked promoter[33,38]. The expression pattern of the reporter gene follows that of the enhancer despite having different gene origins from the promoter [38]. This characteristic has therefore allowed for the easy identification of enhancers when used in conjunction with a minimal promoter and has been implemented in numerous studies.

EXPERIMENTAL DESIGN/METHODS:

Selection of Conserved Region:

The Dbx1 conserved region 5 (Dbx1CR5) was selected using the Non-Coding Sequence Retrieval System (NCSRS)[29] which was developed in the Cai lab. The NCSRS is a program that performs a multi-genome alignment of non-coding sequences for homologous genes in order to retrieve and score putative gene regulatory elements based on length and level of conservation. The alignment for Dbx1CR5 is shown in figure 1a. Dbx1CR5 is located approximately 16kb upstream of the gene Dbx1. It is one of the sixteen predicted conserved regions linked to the gene Dbx1, and it is also the top ranked non-coding region based on the sequence similarity and length. The predicted conserved region has a sequence length of 624bp long and is denoted by a red bar in fig.1b.

Experimental Construct Design:

The Dbx1CR5 region was polymerase chain reaction (PCR) amplified and sub-cloned into a GFP reporter construct containing the minimal basal promoter, β-globin promoter (BGP) (fig.1c). Using the premise that the minimal promoter alone can not activate reporter gene expression, functionality of the putative enhancer was determined by the presence of GFP expression. GFP therefore correlates spatio-temporal enhancer activity and is only expressed in tissues maintaining the appropriate transcription factor profile.

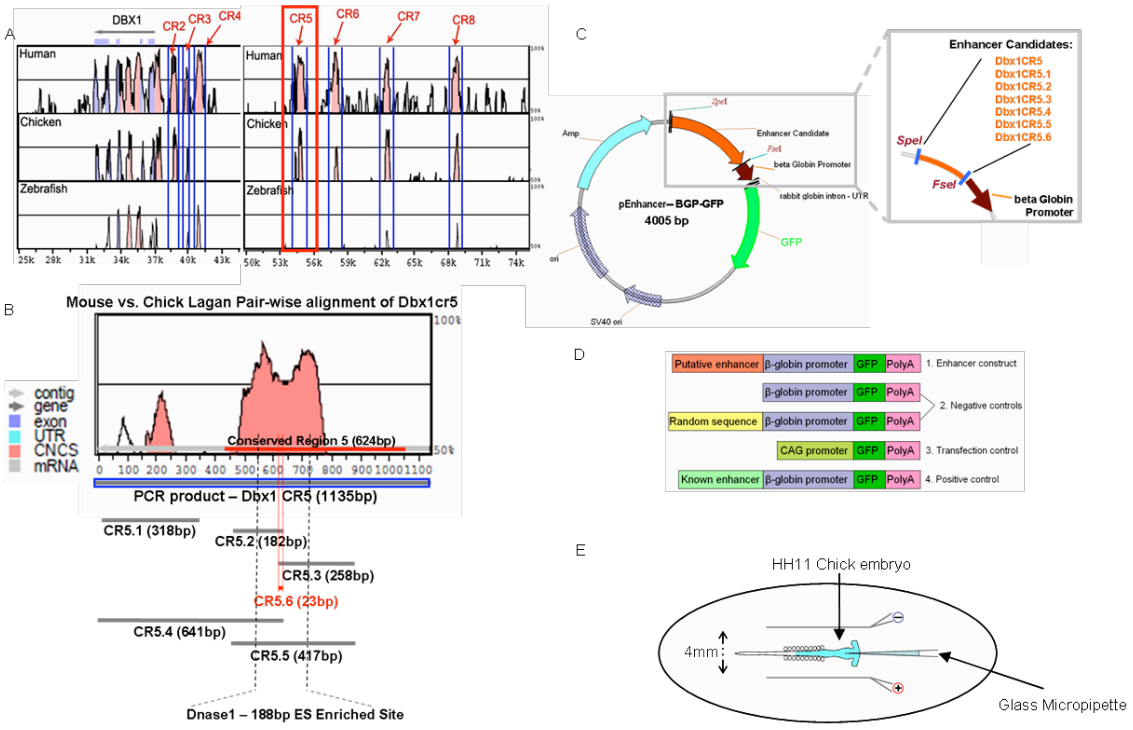

Figure 1: Dbx1 Sequence Alignment and Experimental Design.
A. Multiple-species alignment showing Dbx1 with correlating conserved regions (CR) of the gene; conservation of cow, dog, opossum, rat, and tetradon not depicted. CNCS, Conserved non-coding sequence (in pink peaks).
B. Lagan pair-wise alignment of the Dbx1CR5 region. The red bar indicates the actual predicted region, and the blue bar indicates the PCR product cloned into the reporter construct. Smaller fragments of the PCR product were separately cloned into the BGP-GFP and are indicated in the figure above. A Dnase1 protected site is also depicted in the figure above.
C. Map of BGP-GFP reporter construct with assembled enhancer constructs
D. Schematic of experimental and control constructs
E. I*n ovo* electroporation of a HH11 chick embryo is depicted with a glass micropipette. Electrodes are positioned 4mm apart. DNA, in green, is injected into the neural tube.

Control experiments for the β-globin promoter reporter construct were previously performed and verified *in ovo*. Positive control experiments utilized the β-globin promoter reporter construct linked to a known enhancer and resulted in appropriate GFP expression. Negative control experiments involved the use of the β-globin promoter reporter construct alone. A second negative control was also performed using the reporter construct linked to a random DNA sequence – both experiments resulted in negative GFP expression. Transfection controls were carried out using the pCAG-DsRed

fluorescent reporter construct. The control construct pCAG-DsRed contains a CAG promoter (chicken β-actin promoter with a CMV enhancer element), and is able to drive high levels of DsRed expression in all transfected cells. A pictorial description of all control experiments is shown in figure1d.

The Dbx1CR5 region was PCR amplified from genomic mouse tail template DNA. Primers were chosen to encompass regions greater than the actual conserved region to ensure that the whole predicted region was included. The PCR product resulted in a DNA fragment 1135bp long. The smaller sub-regions CR5.1-5.6 were amplified from the full-length Dbx1CR5 plasmid. Primers were selected at flanking regions to the sequence of interest. Forward primers included a SpeI restriction endonuclease sequence and reverse primers included an FseI restriction endonuclease cut site to guarantee directionality during ligation. All primers were ordered through UMDNJ-RWJMS IDT. The primer sequences used were as follows respective to their location on the original Dbx1CR5 sequence: **F:** (1bp) 5'-CGA TAT ATA CTA GTA AGC ATG AAG CAT GAA GGG CTG TG-3', (459bp) 5'-CGA TAT ATA CTA GTA GCT TTG AAT GTG CAA CCC GAT CC-3', (618bp) 5'CGA TAT ATA CTA GTA AGC AGA GGT TGT CCC TTG TTC CT-3'. **R:** (318bp) 5'-CGA TAT ATG GCC GGC CAC ACA CAC ACA CAC ACA CAC ACA C-3', (641bp) 5'-CGA TAT ATG GCC GGC CAG GAA CAA GGG ACA ACC TCT GCT T-3', (876bp) 5'-CGA TAT ATG GCC GGC CTA ATC AAG CCA CGG TGC TGT GAG A-3', (1135bp) 5'-CGA TAT ATG GCC GGC CAG GCC ATG AAC AAC CTG AGA GGA A-3'. DNA was then digested using SpeI/FseI enzymes, and subsequently gel purified (Quiagen). Purified DNA was then ligated together and transformed into competent DH5α cells (Invitrogen). Using

ampicillin selection, bacterial colonies were grown overnight on LB agar plates and single colonies were picked. The resultant plasmid DNA was verified using both PCR and digestion methods. Constructs were later further verified with sequencing provided by UMDNJ-RWJMS DNA core facility.

Glycerol stocks were made from verified colonies and stored at -80ºC. Bacterial colonies were grown in LB + ampicillin and plasmid was purified using Quiagen purification kits. DNA concentrations resulted in approximately 3-5ug/ul as determined by UV spectrophotometry.

In Ovo Electroporation:

Fertilized white leghorn chicken eggs (Sunshine Farms, Catskill NY) were incubated at 37.5ºC at 58-60% humidity and were protected from light. Eggs were incubated for approximately 40-45 hrs in order to achieve Hamburger Hamilton staged 10-12 embryos. Plasmid DNA was then injected into embryos through the dorsal surface of the neural tube using a glass-pulled micropipette needle. Fast green dye was mixed with plasmid DNA for visualization purposes at a 1:10 ratio. Co-transfections were performed at a ratio of 1:5 of control to enhancer plasmid. Injections were performed under a stereoscope at 1-3X magnification.

Gold-plated electrodes were set approximately 4mm apart to electroporate embryos using a BTX ECM 830 system (Harvard Apparatus) with generator settings of 18 volts, 5 pulses, 50 ms duration, and 950ms intervals. Eggs were then sealed with clear scotch tape and placed back into the incubator until appropriate harvest time points. An illustration of the *in ovo* electroporation procedure can be found in figure 1e.

Preparation of Embryonic Sections:

Transfected tissue samples were harvested at various developmental stages after injection and electroporation, and were placed in cold 1X PBS. Tissues were then fixed in 4% paraformaldehyde (in PBS) at 4ºC for 90 min. Fixed tissues were then washed in cold 1X PBS after which they were cryoprotected with 30% sucrose (in PBS) overnight at 4ºC. Cryoprotected tissue samples were embedded in OCT and stored at -80ºC until they were serially sectioned onto pre-cleaned slides in 15μm sections.

Immunostaining:

After cryosectioning at 10-15 μm/section, slides were washed with 1X PBS. Blocking solution with 10% goat serum in PBS was applied for 30 min, and primary antibodies diluted in blocking solution were added and left overnight at 4ºC. Antibodies used: 3CB2 (1:10, DSHB), vimentin (40E-C, 1:10, DSHB; H5, 1:10, DSHB), Pax6 (1:10, DSHB), NeuN (1:1000, Millipore), Lim1/2 (4F2, 1:400, DSHB), Evx1 (99.1-3A2, 1:10, DSHB), GFAP (1:250, Accurate), GS (1:250, Santa Cruz), Acetylcholine receptor (mab270, 1:10, DSHB), PKCα (1:400, Santa Cruz), Chx10 (1:50, Santa Cruz), Hb9 (81.5C10, 1:10, DSHB). Appropriate secondary antibodies were then used (1:300) and tissue sections were subsequently DAPI stained and mounted.

Microscopy:

Before embedding and sectioning, tissues were inspected using the Leica MZ16FA fluorescent dissection microscope. Tissue sections were inspected using the Zeiss Axio Imager A1. Images using the Leica and Zeiss microscopes were taken using

the Axiocam MRM digital camera (Zeiss). Confocal images were taken using the Olympus FluoView FV10i confocal microscope.

Electrophoretic Mobility Shift Assay (EMSA):

EMSA was used to demonstrate specific protein:DNA binding of nuclear extract harvested from chick CNS to the 23bp DNA fragment of interest (Dbx1cr5.6). Specific protein binding was demonstrated through a competition reaction of protein with unlabelled DNA. Using the 23bp sequence of Dbx1CR5.6 as DNA probe, specific binding of protein to the DNA probe was determined.

Nuclear extracts from embryonic day 4 and 6 (E4 and E6) chick brain and spinal cord tissues were harvested and protein concentrations were determined by NanoDrop. Forward and reverse strand DNA oligos ordered from UMDNJ-RWJMS IDT were biotinylated at the 3' end using a biotin labeling kit (Pierce). Probes were annealed at a 1:1 molar ratio to make double stranded probes. Protein-DNA binding reactions were prepared on ice and were performed at room temperature. EMSA gels were run at 100V on 8% non-denaturing gels and transferred onto biodyne plus 0.45um membranes using a semi-dry transfer system. Biotin labeled DNA was detected by chemiluminescence (Pierce).

Transcription Factor Binding Analysis:

Transcription factor binding analysis was performed using MatInspector (Genomtatix). By inputting the DNA sequence of interest, a list of predicted binding proteins is output along with their respective binding locations on either the forward or

reverse strand sequence. This analysis was performed for the entire 1135bp sequence for Dbx1CR5.

Statistical Analysis:

GFP expression for each sample was categorized based on anatomical location within the CNS: telencephalon, diencephalon, mesencephalon, metencephalon, myelencephalon, and spinal cord. Fluorescent whole mount images were used to distinguish the presence of GFP within each anatomical location. Brain regions were distinguished by folds in the brain and geometry of the CNS shown in fig.a. Frequency of GFP occurrence within each brain region at each embryonic stage for all constructs was calculated by summing the number of samples expressing GFP in a region and dividing by the total number of samples for that stage. Frequency of GFP expression was calculated for all constructs.

A two-sample analysis of variance (ANOVA) without replicates was performed on the frequency of GFP expression for Dbx1CR5, to examine variation in temporal and spatial expression patterns. In addition, a paired two-tailed t-test ($P<0.05$) comparing average temporal or spatial GFP expression was used to examine whether expression patterns for different embryonic stages or location in the CNS are significantly similar for all constructs. Using a paired two-tailed t-test ($P<0.05$) the different constructs (Dbx1CR5.1-6) were also compared against Dbx1CR5 for E6 to examine if average spatial and temporal GFP expression patterns are significantly similar. An F-test ($P<0.05$) determined whether variance of GFP expression between developmental stages

or CNS location is equal (paired). All analyses were performed in Excel (part of Microsoft Office, 2003).

RESULTS:

Dbx1CR5 drives tissue specific expression of GFP in the Developing CNS

We have isolated a 1135bp DNA fragment, Dbx1CR5, located ~16kb upstream of the Dbx1 gene that is able to drive brain-tissue specific reporter GFP expression in the developing chick CNS. Tissue specificity was determined by co-transfection of experimental construct with the control construct pCAG-DsRed. Co-transfections were performed at a high ratio of experimental to control DNA. This was done in order to statistically ensure equal transfection of experimental construct into each cell that was transfected with control. A high ratio of experimental:control DNA also enabled better visualization of the experimental gene expression; the signal strength of the control otherwise would have overpowered that of the experimental due to the higher capacity of CAG.

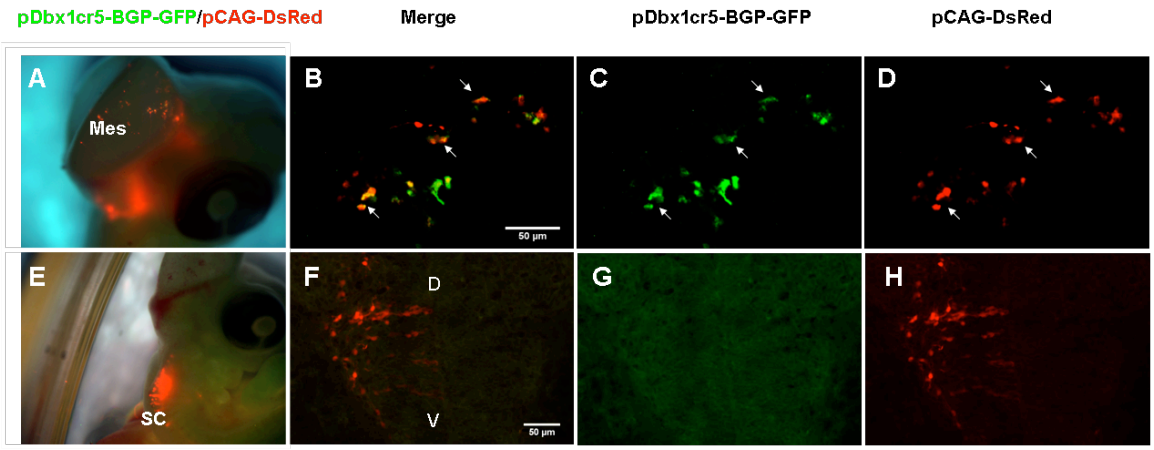

Figure 2 Tissue specific expression of GFP in Developing Chick CNS. Expression of GFP is found in the developing brain of E6 embryos co-transfected with experimental construct pDbx1CR5-BGP-GFP and control construct pCAG-DsRed (A) but not in the spinal cord (E). Transverse sections demonstrate the presence of GFP+ cells found in the dorsal mesencephalon (B-D), and absence of GFP+ cells in spinal cord tissue (F-H). White arrows point to cells with co-localized DsRed and GFP. Mes, mesencephalon, SC, spinal cord, D, dorsal, V, ventral. Scale bar = 50µm in B and F.

Brain and spinal cord tissues were separately targeted during *in ovo* electroporation. Transfected tissues harvested at E6 revealed tissue specific expression of GFP in the brain but not the spinal cord (fig.2). All brain tissues expressing DsRed also showed co-expression of GFP (n=4) (fig.2a-d); however this was not the case for all spinal cord tissues. Rather, positively transfected spinal cord tissues expressing DsRed did not show any expression of GFP (n=4) (fig.2e-g). Tissues were inspected both whole-mount and on cryo-sections in order to verify presence or absence of GFP.

GFP is Expressed Preferentially in the Mesencephalon

Time-specific activity of Dbx1CR5 was characterized by harvesting transfected chick embryos at various developmental stages. GFP expression was observed in all harvested brain samples from E3, E4, E8, E13, and E18 (fig.3a-d). Overall, GFP expression in harvested samples between time points was relatively consistent. Some samples however demonstrated higher levels of expression than what was normally observed. This may be attributed to loss of transfection over time as well as transfection variability.

In samples from E3 (n=3) and E4 (n=6), GFP expression was observed in a large number of cells and appeared in a large portion of the mesencephalon (fig.3a). GFP expression was found predominantly within regions of the mesencephalon and diencephalon but was also found to extend to areas of the metencephalon as well as the myencephalon. At later developmental stages E6 and on, GFP+ cells appeared less abundantly and seemed to be more localized and aggregated (fig3b,c) in comparison to earlier E3/4 samples (fig3a). GFP expression was consistently observed in the

mesencephalon, which at later stages correlated into the optic tectum, and less frequently appeared in other regions of the brain. GFP expression was also observed in lower instances within the forebrain, hindbrain, and cerebellum. Inspection of tissue sections revealed that no GFP+ cell bodies were within the spinal cord however GFP+ axon fibers were occasionally observed. These fibers extended down the full length of the spinal cord from the brain on the dorsal side in two distinct lateral tracts.

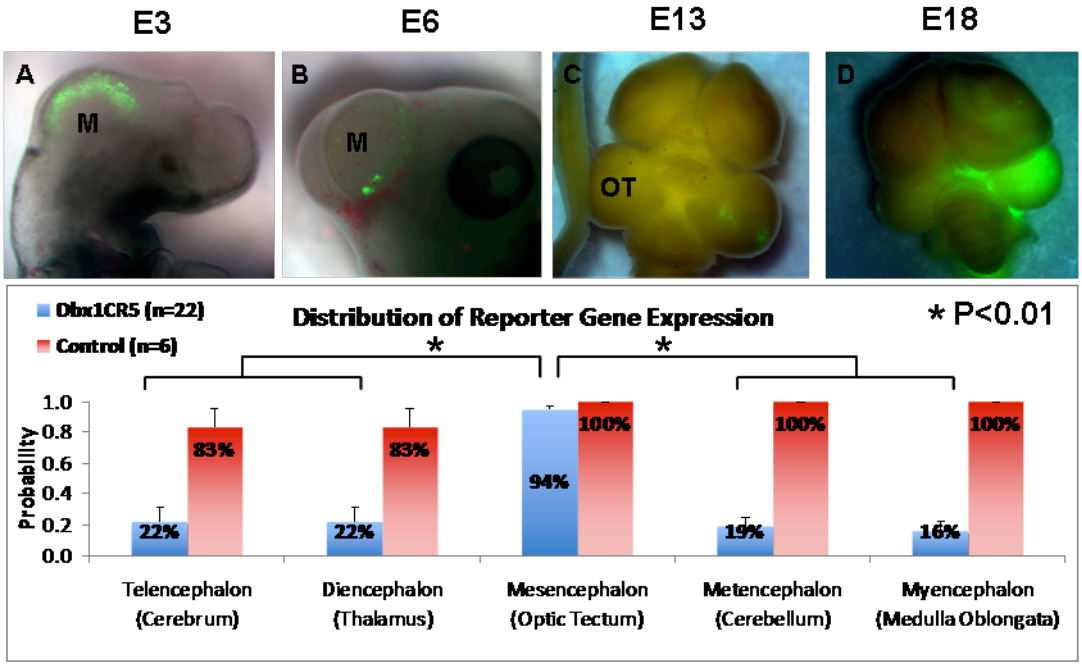

Figure 3 GFP Expression Derived From Transfection of Dbx1CR5 Construct Throughout Embryonic Development. Dbx1CR5 is able to drive expression of GFP from E3 through E18 (A-D). A distribution analysis shows that GFP is expressed predominantly in the mesencephalon (E). GFP expression in the mesencephalon was found to be statistically significant in comparison to all other regions ($P<0.01$). DsRed expression derived from control construct is shown in red.

In order to analyze spatial and temporal differences of GFP expression, a two-factor ANOVA analysis was first performed on the bulk of the data; GFP expression was categorized into different brain regions with respect to time. ANOVA analysis revealed significant differences of GFP expression between brain regions ($P<0.01$), but found no significant differences between the various time points. This allowed us to then lump

together data from all time points in order to analyze spatial expression alone. We found that GFP was expressed within the mesencephalon in 94% of the harvested samples and less frequently in other brain regions. A full distribution of GFP expression representative of all collected samples is shown in fig.3e.

Using a two tailed t-test, further analysis was performed comparing expression within brain regions in order to find which regions were showing significant differences in GFP expression. The results of this analysis revealed significant differences ($P<0.01$) within the mesencephalon in comparison to all other brain regions while comparisons between non-mesencephalic regions showed no statistical differences ($P>0.6$).

GFP+ Cells Make a Heterogeneous Population

Morphology:

At early stages E3 and E6, GFP+ cells spanned the developing neuroepithelium where the majority of cells observed displayed bipolar morphology (fig.4a). GFP+ cells were commonly found in clusters within neuroepithelium of transfected brain tissues. These cells displayed elongated processes and had small rounded cell bodies. The majority of cell bodies were found within the intermediate zone of the cortex and their processes reached from the inner ventricular zone towards the outer pial layer in a parallel arrangement (fig.4a-b). This distinctive morphology was found throughout each observed stage of development from E3, E6, E13 and E18. At thicker cortical regions, GFP+ cells in E13 and E18 samples no longer spanned the full cortex, but rather

appeared to be migratory within the cortex. Cells also appeared to be more commonly found near the ventricular zones of the brain in tissues harvested at E13 and E18.

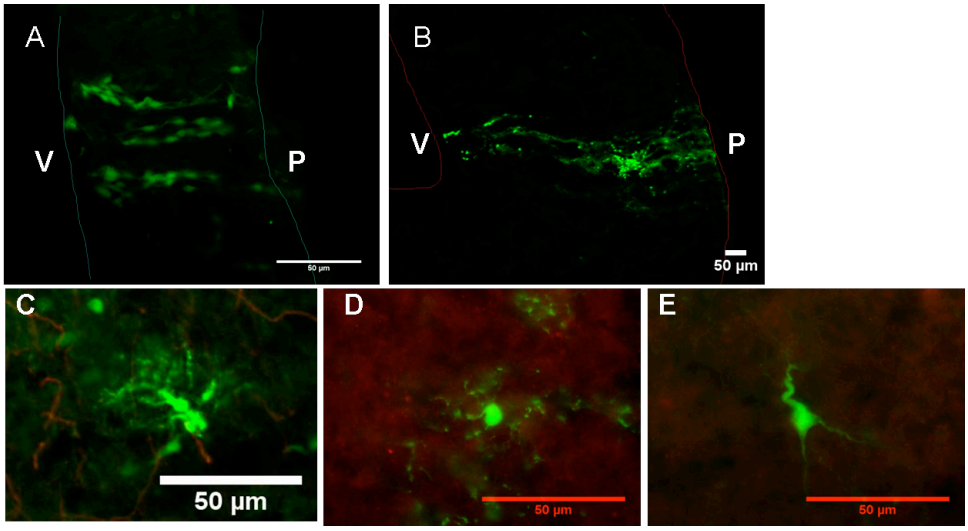

Figure 4 Morphologies of GFP expressing cells. Transverse sections of E6 (A) and E13 (B) mesencephalic tissues show bipolar GFP+ cells that span the width of the developing tectum. In E13 and E18 optic tectum tissue samples, GFP+ cells resemble astrocytes (C), oligodendrocytes (D), and neurons (E), as well as other cell morphologies. V, ventricle, P, pial layer

In E13 tissues the majority of cells maintained bipolar morphologies. Also, cell processes did not always span the full length of the tectum as they did in earlier stages giving them a migratory appearance. Cells with different morphologies were also observed to emerge in small numbers within the tectum. These cells displayed characteristics of mature cell types having larger cell bodies with multiple axonal processes. Some of the morphological phenotypes observed were similar to oligodendrocytes, astrocytes, and neurons (fig.4c-d).

In E18 tissues, many more cells were observed to have taken on mature cell phenotypes in comparison to E13 samples. While a large number continued to maintain radial glial-like phenotype, many GFP+ cells displayed an array of various phenotypes with multiple processes which were present within multiple layers of the cortex. These

cells were commonly found separate from each other in a widely dispersed manner unlike the more common clustered and highly organized phenotype seen with the bipolar cells. To further confirm the cellular identity of Dbx1cr5-GFP+ cells, we performed immunohistochemistry using neural cell-type specific antibodies.

Immunostaining:

Transfected tissues were stained using various antibodies marking distinct cell types in order to characterize the GFP+ cells. During neural tube development, the neuroepithelium consists of stem cells, neural progenitor cells and radial glial cells. As development advances, cells continue to divide and differentiate into mature neurons and glial cells. In our experiments, we have used antibodies marking these progenitor cell types in order to first globally classify GFP+ cells and then to indicate cell type specific lineages. Of all antibodies used, staining resulted in colocalization with NeuN, GFAP, Lim1/2, and Evx1. A summary of antibody staining can be found in table1.

Transfected tissues were initially stained using Pax6, a marker for neural stem cells. Samples collected from E4, E6, and E13 all yielded negative co-localization with antibody staining indicating that GFP+ cells are not neural stem cells. Tissues were then stained using multiple antibodies for vimentin, a radial glial marker. Vimentin staining resulted in negative colocalization with samples collected at E3, E4, and E18. For samples harvested at E6, neither positive nor negative co-localization with GFP+ processes could be confirmed.

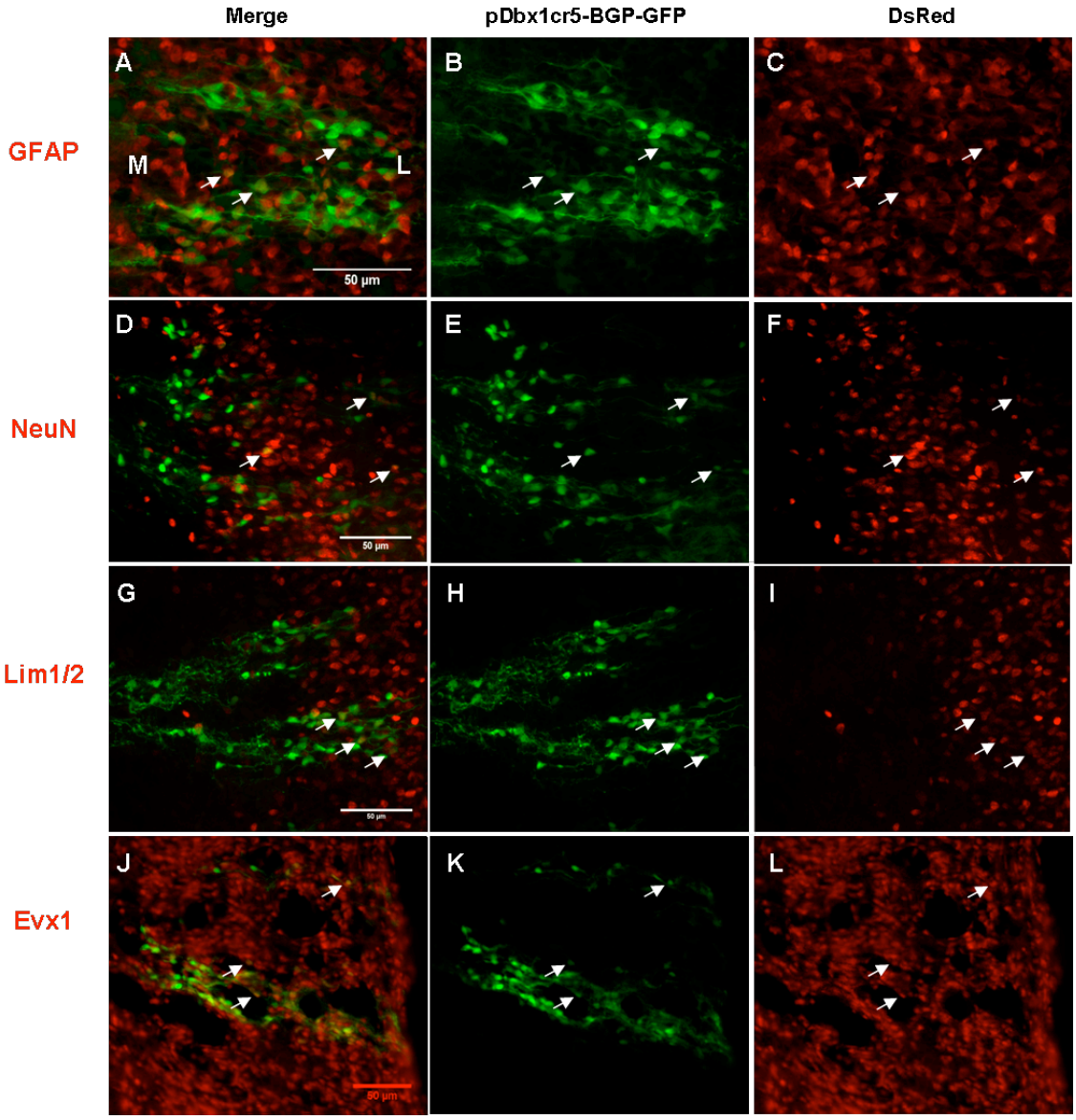

Figure 5 Co-localization of Antibody Staining and Dbx1CR5 derived GFP+ Cells in E13 embryos. Stained transverse sections of optic tectum from E13 embryos show colocalization of GFP with a small population of GFAP+(a-c), NeuN+(d-f), Lim1/2+(g-i), and Evx1+(j-l) cells. White arrows point at cells with colocalization of GFP and antibody staining. M, medial, L, lateral. Scale bar = 50µm in A, D, G, J.

NeuN, a post-mitotic neuronal progenitor marker and GFAP, an astrocytic cell marker were also used to classify either a neural or glial population. NeuN staining revealed no colocalization in tissues from E4 and E6, however staining of samples taken at E13 (fig.5d-f) and E18 did show colocalization with a small number of GFP+ cells.

GFAP staining also revealed colocalization with a small number of GFP+ cells at E13 and is shown in fig.5a-c.

Further staining using more mature cell markers also resulted in positive colocalization with antibodies for Lim1/2 and Evx1. Lim1/2 staining resulted in colocalization with E6 and E13 tissues (fig.5g-i), but not in E4 tissues. Evx1 staining resulted in colocalization with GFP+ cells at all stages tested (E13 staining is shown in fig.5j-l). Confocal images were taken to further confirm colocalization of GFP+ cells in E18 samples with nuclear Evx1 (fig.6).

Staining with all other antibodies revealed no colocalization at any time points tested. Engrailed1 antibody was used on E6, E13, and E18 samples. Islet1 antibody was used on E6 and E13 tissues. Hb9, Lim3, and Chx10 were tested on E6 samples. Glutamine synthetase antibody was tested on E13 tissue. Antibodies for Acetylcholine receptor and PKCα were used on E18 tissues.

	Vimentin	Pax6	En1	Isl1	NeuN	Lim1/2	Evx1	GFAP	Hb9, Lim3, Chx10	GS	Ach-receptor, PKCα
E3	X										
E4	X	X			X	X					
E6		X	X	X	X	✓	✓		X		
E13		X	X	X	✓	✓	✓	✓		X	
E18	X		X		✓		✓				X

X = No colocalization
✓ = Colocalization

Table 1: Antibody Staining Summary. Red boxes represent negative colocalization with GFP+ cells, and yellow boxes represent positive colocalization with GFP+ cells. NC, not confirmed

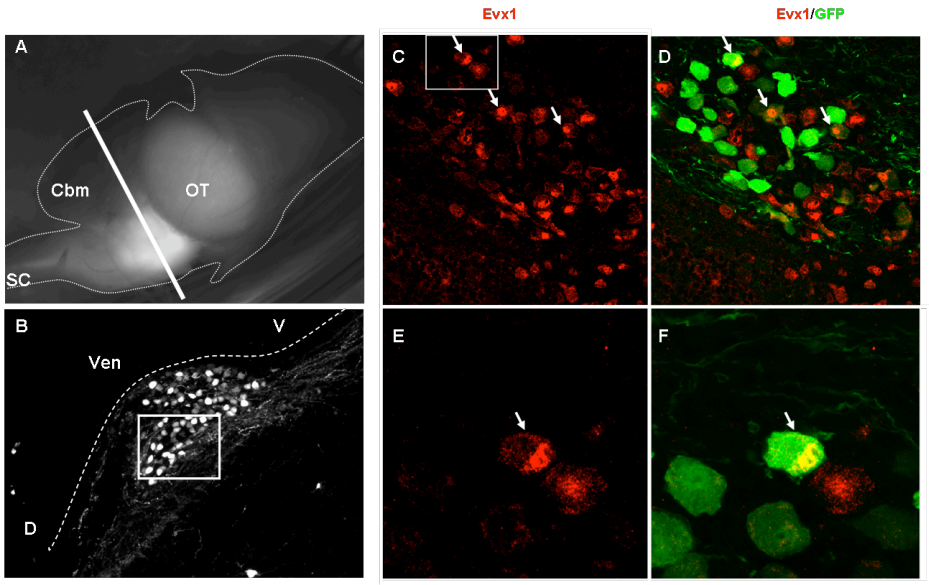

Figure 6 Dbx1CR5-GFP+ cells express neuronal marker Evx1. Side view of transfected E18 brain sample in grey scale (A). Transverse section as indicated by the white bar in grey scale (B). Confocal images of stained tissue sections are shown in (C-F). Brain region of white box inset of (B) is shown in (C-D). Brain region in white box inset of (C) is shown in (E-F). Evx1 antibody staining colocalization with GFP+ cells is indicated by white arrows. SC, spinal cord, Cbm, cerebellum, OT, optic tectum, D, dorsal, V, ventral, Ven, ventricle.

Dbx1CR5 Contains Multiple Regulatory Elements

To identify functional/active regions of Dbx1CR5, the 1.13kb region of Dbx1cr5 was broken into smaller sub-regions and the smaller sub-regions were individually cloned into the BGP-GFP reporter construct. Six separate constructs, CR5.1-5.6, were prepared and tested by *in ovo* electroporation into HH11-12 chick neural tube. Of the six constructs, five (except Dbx1CR5.1) were able to drive GFP expression in E3 embryos (fig.7a-f). Expression patterns of GFP however, were found to vary from construct to construct as well as stage to stage (fig.7g-h).

The design of each construct is illustrated in fig.1b. DNA regions were selected based off of alignment results of the mouse and chicken genome sequences of Dbx1CR5. Peaks in conservation guided the division of regions and were separately cloned into the reporter construct. The two regions represented by CR5.2 and CR5.3 demonstrated >75% homology and were included within the computationally predicted region. CR5.1 encompassed the less conserved peak that was included into the PCR product of Dbx1CR5, but was not included within the predicted region. CR5.4 was designed to overlap regions CR5.1 and CR5.2, while CR5.5 was designed to overlap regions CR5.2 and CR5.3. CR5.6, which represents the 23bp overlapping region of CR5.2 and CR5.3, was not cloned until after CR5.1-CR5.5 were tested *in ovo*.

In ovo electroporation of the six DNA constructs resulted in consistent GFP expression of constructs CR5.2-CR5.6 in E3 embryos. CR5.1 was not able to drive GFP expression in the CNS of harvested embryos (n=7) and was ruled out of further analyses. Consequently, CR5.4 was also ruled out with the reasoning of being equivalent to CR5.2.

From these results, we were able to deduce that the necessary regulatory elements resided within CR5.5, which includes regions CR5.2, CR5.3, and CR5.6.

In comparing GFP expression patterns between the original Dbx1CR5 construct and CR5.2, CR5.3, CR5.5, and CR5.6, obvious differences in time and tissue specific expression were observed. While GFP expression for Dbx1CR5 was never categorized anywhere but the mesencephalon in E3 samples, each of the other constructs had varying amounts of GFP expression in different regions of the brain (fig.7g). GFP expression also seemed to become more restricted when comparing E3 and E6 expression patterns, where GFP expression is found in multiple regions of the neural tube in E3 samples, but is only found in the mesencephalon, metencephalon, and myencephalon in E6 (fig.7g,h). The most evident change observed is CR5.6's loss of ability to drive GFP expression in E6 embryos. Although CR5.6 was able to drive GFP expression in E3 embryos (n=9), E6 embryos (n=6) yielded no GFP (fig.7g,h).

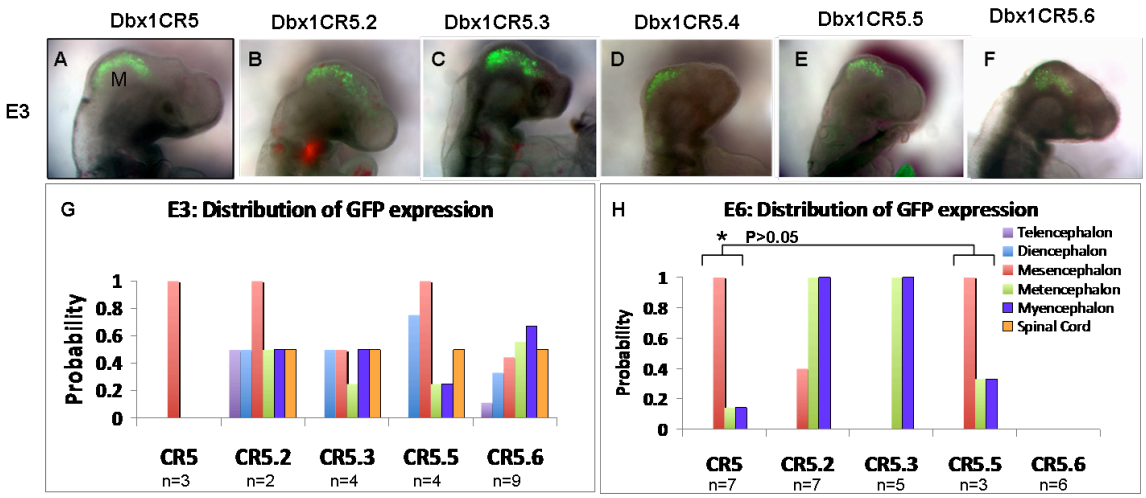

Figure 7 Smaller DNA fragments of Dbx1CR5 are able to individually drive reporter gene expression. Regions CR5.2-CR5.6 (B-F) each demonstrate the ability to drive GFP expression at E3 in the developing brain. Distribution of GFP expression for all constructs is shown for E3(g) and E6(h) harvested embryos. GFP expression patterns of CR5 and CR5.5 were most similar (P>0.05) in comparison to all other regions. M, mesencephalon.

In order to narrow down the major regulatory element responsible for regulating the tissue and time specific expression of the larger Dbx1CR5, E6 expression patterns were used. Differences between the constructs were more pronounced and less variable at this stage. A comparison of expression profiles revealed that CR5.5 shared the most similar pattern with the larger CR5.

CR5.2 and CR5.3 were found to have dramatically different profiles from the original Dbx1CR5. GFP was detected in the metencephalon and myencephalon in all CR5.2 and CR5.3 samples, whereas the full-length construct CR5 had low to non-existing amounts of GFP in those reasons. Similarly, CR5 was expressed in high amounts in the mesencephalon, but CR5.2 and CR5.3 were not (fig.7h). Interestingly, CR5.2 and CR5.3 share similar profiles with the exception that CR5.3 does not have GFP expression in the mesencephalon. CR5.6 was immediately ruled out as the major regulatory element due to the absence of expression at this time point. A comparison of CR5 and CR5.5 revealed identical expression profiles with the exception that CR5.5 had a higher instance of GFP in the metencephalon and myencephalon.

Statistical analysis of GFP expression patterns within each construct revealed similar results and verified that CR5.5 was most similar to CR5. For each construct, GFP expression patterns were categorized by brain region from all collected samples. A t-test was then performed to look for differences in expression for each region. Analysis of CR5.2, CR5.3 and CR5.6 revealed that means between regions of expression were not significantly different indicating non-tissue specific expression ($P>0.05$). Analysis of CR5.5 however, revealed significant differences ($P<0.05$) in GFP expression only between the mesencephalon in comparison to the telencephalon, metencephalon and

myencephalon. Although these results coincide with those of CR5, significant differences between expression between diencephalon and mesencephalon are not present.

Specific Protein Binds Dbx1CR5.6

Since Dbx1CR5.6 demonstrated its ability to drive reporter GFP expression, we began to identify the specific trans-acting factors that might be involved in binding the Dbx1CR5.6 region. EMSA binding studies were performed using the 23bp DNA sequence as a DNA probe. Both double stranded (DS) and single stranded probes were used for these studies in order to narrow down binding to either the forward strand (FS) or reverse strand (RS), which would essentially help to narrow down potential binding factors during transcription factor binding sequence (TFBS) analysis. Computational analysis of known binding factors resulted in the identification of three potential factors (Sox9.02, LEF1.01, and HPF1.01) which were predicted to bind only on the reverse strand (fig.8a). These three factors all maintained binding anchors within the CR5.6. HNF1 and GATA were found to overlap the CR5.6 region by 1-2bp, but do not maintain anchors within CR5.6, therefore are not candidates for binding. TFBS analysis for the full Dbx1CR5.5 with all possible binding factors is shown in figure 8a.

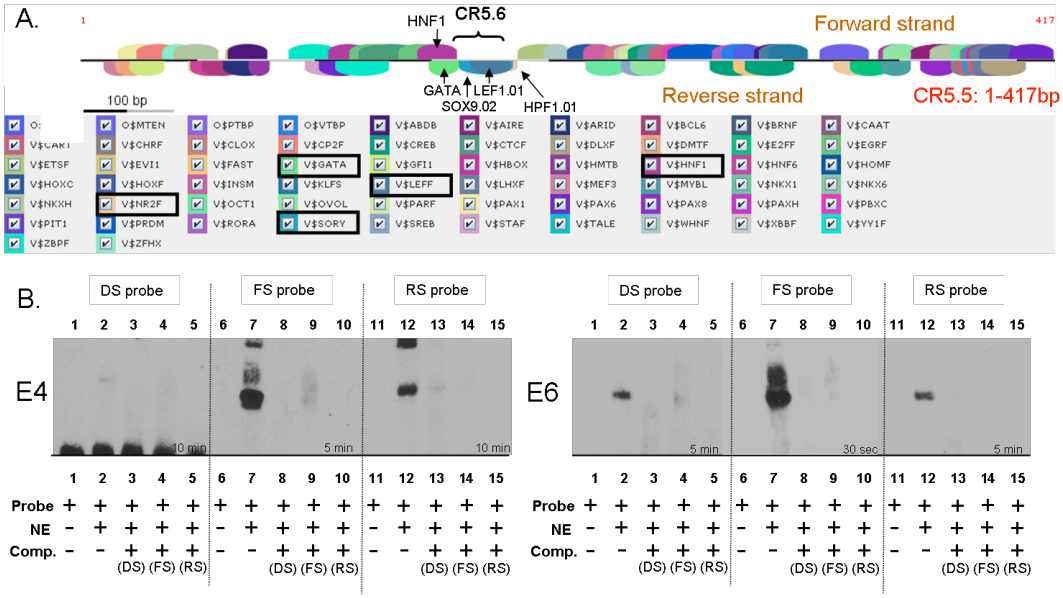

Figure 8 TFBS analysis and EMSA results for CR5.6. TFBS analysis shows all possible binding factors for the sub-region CR5.5 which includes the 23bp of CR5.6. Specific factors for CR5.6 are labeled and correlate with black boxes (A). EMSA gel results using E4 and E6 nuclear extracts show specific protein binding (lanes 2,7,12) and competition (lanes 3-5, 8-10, 13-15) using CR5.6 probes (B). Exposure times for each condition are labeled. DS, double strand probe, FS, forward strand probe, RS, reverse strand probe.

Using E4 and E6 nuclear extracts from chick CNS, EMSA studies showed that the sequence for CR5.6 was able to bind specific protein as annealed double stranded (DS) probe as well as separate forward (FS) and reverse strand (RS) probes. Specific binding of DS probe in E4 conditions was weaker than for E6 conditions (Fig.8b-lane2). Unexpectedly, binding with the FS probe was seen at both E4 and E6 (fig.8b-lane7) and was much stronger than binding with DS (lane2) and RS probe (lane12) for both conditions. EMSA studies were carried out in duplicate and indicated equivalent results.

DISCUSSION:

Importance of Conservation

NCSRS takes advantage of homologies of non-coding regions in order to predict putative enhancers in non-coding DNA. In this study we used the NCSRS to predict an evolutionarily conserved enhancer of Dbx1, and experimentally validated its ability to drive reporter gene expression. Through this, we were able to demonstrate evolutionary conservation of enhancer function by using mammalian DNA in the less sophisticated chick vertebrate system. This information not only helps to validate our system of predicting enhancers, but also propels the idea of conservation and its significance in biological systems.

Discovery of a putative enhancer of Dbx1

In this study, we were able to take a computationally predicted enhancer of Dbx1 and verify its functionality as a transcriptional regulatory element. Tissue and time specific expression have long been thought to be a major characteristic of enhancers. In reporter gene constructs, enhancers are dominant in conferring tissue specificity to a linked gene and therefore determine the pattern of expression for a reporter construct[38]. By implementing this common method of identifying enhancers, we were able to identify Dbx1CR5 to be a functional regulatory element defined by its ability to drive tissue specific reporter gene expression[38-39]. Further support of enhancer function comes in the form overlapping regions of Dbx1CR5 with a known DNaseI hypersentivity site (ENSR00000058286), which are specific to enhancer activity[33,39-40]. The implications of

this study are two fold; one in which we have been able to further validate the importance of conservation in identifying gene regulatory elements, and more importantly being the discovery of a putative enhancer of Dbx1.

Expression patterns of Dbx1 have been thoroughly characterized in murine studies which have shown Dbx1 to display spatial restriction in both the brain and spinal cord of the developing mouse CNS[14]. We were able to characterize the tissue specific expression of Dbx1CR5 by separately targeting brain and spinal cord tissues during co-transfection of experimental and control DNA. Brain tissues that showed positive transfection of the control construct also showed expression of the experimental construct while spinal cord tissues showing transfection of the control did not exhibit expression of the experimental construct – thereby showing tissue specificity. An analysis of expression patterns within the brain further indicate that Dbx1CR5 plays a specific role in mesencephalic development. These results correlate with previous mouse studies which have defined Dbx1 to be expressed in a highly restricted manner especially during later stages of development; Dbx1 was most strongly found in the dorsal mesencephalon, and was not expressed in the spinal cord at late stages[15]. Our findings lead us to believe that we have identified a regulatory element involved in the specific regulation of the Dbx1 gene in brain tissues and shows consistent expression patterns with Dbx1.

The question of whether Dbx1CR5 drives expression in the spinal cord during early stages however remains unanswered. Previous studies have recorded Dbx1 expression in spinal cord tissues during early development which have been defined with HH15-20 (50-72hrs) in developing chick. In this study we were unable to correlate expression of GFP to these times. A possible reason for this may be due to the

limitations of our experimental model that did not allow for the visualization of GFP earlier than E3. A more likely explanation for these differences however may be found in the fact that Dbx1CR5 is involved with the specific spatial expression of Dbx1 in the mesencephalon only.

Multiple Binding Elements Reside within Dbx1CR5

In order to condense the sequence size of Dbx1CR5 to encompass only its major regulatory region, smaller DNA fragments encompassing different parts of the larger region were individually cloned and tested. What was expected was a singular functional construct, however GFP expression was generated by multiple DNA fragments that each produced distinctively different expression patterns both spatially and temporally. These findings indicate that multiple binding regions reside throughout the Dbx1CR5 region and are each able to generate individual tissue and time specific expression patterns.

GFP+ tissues varied from consistent expression in the mesencephalon with Dbx1CR5 and CR5.5 to limited expression only in the metencephalon and myencephalon with constructs CR5.2 and CR5.3. These differences in expression pattern can be explained with the possibility that the larger regulatory region may contain multiple binding sites. These binding sites could then recruit factors that result in inhibition or activation of GFP in different neural tissues. Therefore, when inhibitory binding is disrupted by the removal of either CR5.2 or CR5.3, expression is no longer restricted to the mesencephalon. Conversely, when the inhibitory binding sites are left intact, represented by Dbx1CR5 and CR5.5, mesencephalon specific expression is observed and does not shift. Interestingly, CR5.6 demonstrated time-specific GFP expression in E3

tissues in all regions of the CNS, yet no GFP was detected in E6 tissues. These results signify a temporal preference and indicate that CR5.6 may play a role in early neuronal tissues.

In an attempt to identify specific binding factors responsible for activating GFP expression in CR5.6, TFBS analysis and EMSA were performed. EMSA results using E4 and E6 nuclear extracts revealed binding of protein with all probe conditions. Competition using unlabelled DNA probes also rendered positive results and fully competed protein away. Forward stranded probes were able to bind better to protein than reverse stranded probes, and double stranded probes binded least strongly. In searching for possible binding transcription factors, we found that no proteins were predicted to bind anywhere on the forward sequence of CR5.6. Protein-DNA binding however was observed to be the strongest for single strand forward sequence probes. This data suggest that there may be novel proteins binding to the DNA region.

Characterization of Dbx1CR5 and its role in determining cell fate

In order to characterize and elucidate the role of Dbx1CR5 during neural development, cell phenotypes of GFP+ cells in terms of morphology, molecular markers, as well as spatial arrangement were analyzed through transverse tissue sections. GFP expression was detected in tissues as early as E3 and as late as E18. Because of the early presence of GFP+ cells in transfected tissues, we believe that Dbx1CR5 is active in early progenitor cells and continues to play a role in determining cell fate of numerous cell types which have been identified both chemically and morphologically. Interestingly, staining with Pax6 antibody did not colocalize with GFP+ cells in any transfected tissues.

Pax6 is a common marker for neural stem cells, however it that is regionally expressed within the chick CNS. Within the developing chick brain, Pax6 is regionally expressed only in the telencephalon and diencephalon[41]. This regional difference makes clear that colocalization is not likely to occur and indicates that Pax6 is not an appropriate marker for all neural stem cells, especially those found within the mesencephalon. Therefore, negative colocalization with GFP+ cells does not dismiss the prospect that Db1CR5 is active in neural stem cells. Rather, lack of colocalization with Pax6 further supports the idea that Dbx1CR5 is regionally expressed in the mesencephalon.

The majority of GFP+ cells maintained long radial processes which extended from the ventricular zones to the outer pial layer in early stages and at later stages tended to look migratory. During early development, cells found within the neuroepithelium with bipolar morphology are characteristic of either neuroepithelial or radial glial cells. Neural studies however have indicated that radial glia are generated at E3 during chick neural development[42]. Therefore, using the rationale that radial glial cells have not yet had the opportunity to develop, we have categorized GFP+ cells in E3 tissues to neuroepithelial cells. GFP+ cells were also commonly found near the ventricular zones which are defined to be the origin of progenitor cells. This evidence correlates with previous findings that have implicated Dbx1 to be expressed in actively mitotic cells and further indicate the role of Dbx1CR5 role in regulating gene expression in early progenitor cells[14].

In later staged embryos non-bipolar cells were also found in small numbers and increased with developmental stage. These cells displayed morphologies of mature neurons and were identified, some of which were astrocytic, oligodendritic, and

pyramidal. The majority of GFP+ cells however continued to demonstrate bipolar morphology and were still found in clusters. Also, many GFP+ cells were no longer observed to span the entire of the neuroepithelium which is significantly thickened in later stages, and appeared to be migratory. These results are consistent with findings within the literature which have defined Dbx1-expressing cells to be highly migratory [20].

In an effort to properly identify GFP+ cell types, multiple cell marker antibodies were used. One of the major limitations encountered with using a chick system was the difficulty in finding working antibodies to chick tissues. While many of the conventional cell markers were not reactive with chick tissues, other cell markers were not expressed in chick tissues and significantly narrowed the span of our study. Cell markers that colocalized with a sparse number of GFP+ cells showed signs of neural and astrocytic cell lineages at E13 and E18.

Numerous radial glial filament markers were used in order to characterize early bipolar cells, however none of the staining yielded conclusive results even with the use of confocal imaging. In many tissues, it was difficult to distinguish whether GFP+ cells were actually radial glial cells or whether they were migratory cells following the path of radial glial cells. It is likely that GFP+ cells are migrating progenitors and neurons. Staining results further support this idea.

Cells from E13 and older tissues, positively labeled with NeuN, a neural marker indicating that Dbx1CR5 plays a role in determining cells of neural lineage. Interestingly, tissues from E4 and E6 did not positively label GFP+ cells with NeuN. Cells likely did not co-label due the yet immature state of GFP+ cells at these early times.

GFAP, an intermediate filament protein for astrocytes was also shown to co-label GFP+ cells at E13 showing that Dbx1CR5 is involved in glial lineage.

Further staining with Lim1/2 also showed colocalization indicating that GFP+ cells may be specific in determining interneuronal lineage; lim1/2 is essential in determining dorsal interneurons within the spinal cord[43] Staining with Evx1 also correlates with identifying cells with interneuronal lineage. These findings are particularly relevant in relation to Dbx1. Within the spinal cord, Dbx1 has been reported to control V0 interneuron from V1 interneuron development. V0 interneurons within the spinal cord have been defined to express the transcription factor Evx1 whereas V1 interneurons express En1[24]. In this study, GFP+ cells have been found to maintain similar expression Evx1 and non-expression of En1. Although specific transcription factor expression patterns have not been defined within the brain as they have in the spinal cord, the results from cell staining indicate that may also be V0 interneuronal.

Physiological Interpretation

In interpreting the physiological role that GFP+ cells may play during development, spatial locations of GFP+ cells were correlated to areas of known physiological function. GFP expression was commonly found in mesencephalic areas of the neural tube and later in the lobes of the optic tectum in E13 and E18 samples. During brain development of the chick, the mesencephalon, starts to divide into the optic tectum, which laterally rotate as development continues [44]. The mesencephalon is therefore the source of GFP+ cells found in the optic lobes at later stages. The optic tectum acts as the primary integrating center for eye movement and has been linked to have motor as well

as sensory related properties [45]. Within the chick mesencephalon are the main axonal tracts that connect the cerebral hemispheres with the posterior parts of the brain and spinal cord [46]. Transverse tissue sections of E13 and E18 tissues have revealed a large number of axonal processes which extend from the right hemisphere to the left hemisphere and confer the role in GFP+ cells in relaying information across hemispheres.

Interestingly, a very specific cluster of GFP+ cells were observed in a nucleus within the medulla oblongata of E18 tissue (fig.6). Some of these cells labeled positively with Evx1, but the majority of cells have yet to be defined. The specific clustering of GFP+ cells suggests that Dbx1CR5 is specific to defining that particular nucleus of cells. Using a stereotaxic atlas of chick brain, the cluster of GFP+ cells could possibly be identified to be part of the nucleus nervi trochlearis, nucleus vestibularis medialis or nucleus nervi glossopharynei et nucleus motorus dorsalis nervi vagi [47].

Axon fibers were also observed to extend down both dorsal lateral regions of the spinal cord. This indicates that the cells are involved in descending tracts which located within the lateral corticospinal or posterior spinocerebellar tracts. Other possible functions of Dbx1CR5 are regulation of complex muscular movements as well as involuntary reflexes as suggested by the presence of GFP+ cells in the cerebellum and myencephalon. Cell staining using multiple neurotransmitters was also implemented in trying to correlate Dbx1CR5 function. Unfortunately, many antibodies were unreactive with chick tissues. Cell staining with GS, however was successful and indicated that some GFP+ cells were able to metabolize glutamate.

Limitations of Study

Although *in ovo* electroporation methods offered a quick method of screening predicted regulatory elements, learning the techniques were difficult and time consuming. The procedure itself required precise micromanipulations if not performed properly would drastically affect embryo survival rate. One of the major problems encountered after ascertaining surviving embryos was the lack of samples that expressed stable transfection of plasmid. Also, in spite of specifically targeting specific tissues within the CNS, oftentimes embryos were not consistently transfected within the same regions. These were all factors which added to variability and made analysis difficult.

Another limitation of this study was the lack of established research in the chick system. Previous studies were largely based off of mammalian murine systems which have enormously different brain structures than vertebrates. The major differences in organization and lack of structured architecture make the vertebrate system less than optimal for brain studies.

Future studies and Significance

Following the characterization of Dbx1CR5 in the chick vertebrate system, we have invested our resources into creating a transgenic mouse. Transgenic studies will therefore allow us to overcome many of the limitations related to an *in ovo* system. The use of a transgenic mouse will ensure stable transfection and integration of DNA into every cell and will allow for the complete and thorough characterization of Dbx1CR5. Also available to us will be a wider array of antibodies that can be used to better

characterize GFP expressing cells. The largest benefit towards using the transgenic mouse will be the ability to make direct comparisons to current mouse studies which may provide relevant insight into the major functions of Dbx1CR5.

In studying Dbx1CR5 and the major factors that regulate its activity, we hope to trigger a cascade of knowledge that may some day result in global gene regulation of Dbx1. By understanding global gene regulation, researchers can then work towards identifying major factors involved in genetic or neurological diseases which in turn have the potential to allow for the development of novel therapeutics.

REFERENCES:

1 *Developmental Biology Online*, <http://www.uoguelph.ca/zoology/devobio/dbindex.htm> (

2 Tamamaki, N. Radial glias and radial fibers: what is the function of radial fibers? *Anat Sci Int* **77**, 2-11, doi:10.1046/j.0022-7722.2002.00013.x (2002).

3 Alvarez-Buylla, A., Garcia-Verdugo, J. M. & Tramontin, A. D. A unified hypothesis on the lineage of neural stem cells. *Nat Rev Neurosci* **2**, 287-293, doi:10.1038/3506758235067582 [pii] (2001).

4 Miyata, T. Development of three-dimensional architecture of the neuroepithelium: role of pseudostratification and cellular 'community'. *Dev Growth Differ* **50 Suppl 1**, S105-112, doi:DGD980 [pii]10.1111/j.1440-169X.2007.00980.x (2008).

5 Campbell, K. & Gotz, M. Radial glia: multi-purpose cells for vertebrate brain development. *Trends Neurosci* **25**, 235-238, doi:S0166223602021562 [pii] (2002).

6 Noctor, S. C., Flint, A. C., Weissman, T. A., Dammerman, R. S. & Kriegstein, A. R. Neurons derived from radial glial cells establish radial units in neocortex. *Nature* **409**, 714-720, doi:10.1038/35055553 (2001).

7 Haines, D. E. (Churchhill Livingstone, Philadelphia, PA, 2002).

8 Merkle, F. T., Tramontin, A. D., Garcia-Verdugo, J. M. & Alvarez-Buylla, A. Radial glia give rise to adult neural stem cells in the subventricular zone. *P Natl Acad Sci USA* **101**, 17528-17532, doi:DOI 10.1073/pnas.0407893101 (2004).

9 Turing, A. M. The Chemical Basis of Morphogenesis (Reprinted from Philosophical Transactions of the Royal Society (Part B), Vol 237, Pg 37-72, 1953). *B Math Biol* **52**, 153-197 (1990).

10 Ulloa, F. & Marti, E. Wnt won the war: antagonistic role of Wnt over Shh controls dorso-ventral patterning of the vertebrate neural tube. *Dev Dyn* **239**, 69-76, doi:10.1002/dvdy.22058 (2010).

11 McGinnis, W. & Krumlauf, R. Homeobox genes and axial patterning. *Cell* **68**, 283-302, doi:0092-8674(92)90471-N [pii] (1992).

12 Ensini, M., Tsuchida, T. N., Belting, H. G. & Jessell, T. M. The control of rostrocaudal pattern in the developing spinal cord: specification of motor neuron subtype identity is initiated by signals from paraxial mesoderm. *Development* **125**, 969-982 (1998).

13 Lu, S., Shashikant, C. S. & Ruddle, F. H. Separate cis-acting elements determine the expression of mouse Dbx gene in multiple spatial domains of the central nervous system. *Mech Dev* **58**, 193-202 (1996).

14 Lu, S., Bogarad, L. D., Murtha, M. T. & Ruddle, F. H. Expression pattern of a murine homeobox gene, Dbx, displays extreme spatial restriction in embryonic forebrain and spinal cord. *Proc Natl Acad Sci U S A* **89**, 8053-8057 (1992).

15 Lu, S., Wise, T. L. & Ruddle, F. H. Mouse homeobox gene Dbx: sequence, gene structure and expression pattern during mid-gestation. *Mech Dev* **47**, 187-195 (1994).

16 Pierani, A., Brenner-Morton, S., Chiang, C. & Jessell, T. M. A sonic hedgehog-independent, retinoid-activated pathway of neurogenesis in the ventral spinal cord. *Cell* **97**, 903-915, doi:S0092-8674(00)80802-8 [pii] (1999).

17 Medina, L. *et al.* Expression of Dbx1, Neurogenin 2, Semaphorin 5A, Cadherin 8, and Emx1 distinguish ventral and lateral pallial histogenetic divisions in the developing mouse claustroamygdaloid complex. *J Comp Neurol* **474**, 504-523, doi:10.1002/cne.20141 (2004).

18 Shoji, H. *et al.* Regionalized expression of the Dbx family homeobox genes in the embryonic CNS of the mouse. *Mech Dev* **56**, 25-39 (1996).

19 Lanuza, G. M., Gosgnach, S., Pierani, A., Jessell, T. M. & Goulding, M. Genetic identification of spinal interneurons that coordinate left-right locomotor activity necessary for walking movements. *Neuron* **42**, 375-386, doi:S0896627304002491 [pii] (2004).

20 Bielle, F. *et al.* Multiple origins of Cajal-Retzius cells at the borders of the developing pallium. *Nat Neurosci* **8**, 1002-1012, doi:nn1511 [pii]10.1038/nn1511 (2005).

21 Jessell, T. M. Neuronal specification in the spinal cord: inductive signals and transcriptional codes. *Nat Rev Genet* **1**, 20-29, doi:10.1038/35049541 (2000).

22 Briscoe, J., Pierani, A., Jessell, T. M. & Ericson, J. A homeodomain protein code specifies progenitor cell identity and neuronal fate in the ventral neural tube. *Cell* **101**, 435-445, doi:S0092-8674(00)80853-3 [pii] (2000).

23 Fogarty, M., Richardson, W. D. & Kessaris, N. A subset of oligodendrocytes generated from radial glia in the dorsal spinal cord. *Development* **132**, 1951-1959, doi:Doi 10.1242/Dev.01777 (2005).

24 Pierani, A. *et al.* Control of interneuron fate in the developing spinal cord by the progenitor homeodomain protein Dbx1. *Neuron* **29**, 367-384, doi:S0896-6273(01)00212-4 [pii] (2001).

25 Matise, M. P., Lustig, M., Sakurai, T., Grumet, M. & Joyner, A. L. Ventral midline cells are required for the local control of commissural axon guidance in the mouse spinal cord. *Development* **126**, 3649-3659 (1999).

26 Crawford, G. E. *et al.* Identifying gene regulatory elements by genome-wide recovery of DNase hypersensitive sites. *Proc Natl Acad Sci U S A* **101**, 992-997, doi:10.1073/pnas.0307540100307540100 [pii] (2004).

27 Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921 (2001).

28 Brudno, M. *et al.* LAGAN and Multi-LAGAN: efficient tools for large-scale multiple alignment of genomic DNA. *Genome Res* **13**, 721-731, doi:10.1101/gr.926603GR-9266R [pii] (2003).

29 Doh, S. T., Zhang, Y., Temple, M. H. & Cai, L. Non-coding sequence retrieval system for comparative genomic analysis of gene regulatory elements. *BMC Bioinformatics* **8**, 94, doi:1471-2105-8-94 [pii]10.1186/1471-2105-8-94 (2007).

30 Berman, B. P. *et al.* Computational identification of developmental enhancers: conservation and function of transcription factor binding-site clusters in Drosophila melanogaster and Drosophila pseudoobscura. *Genome Biol* **5**, -, doi:Artn R61 (2004).

31 Ureta-Vidal, A., Ettwiller, L. & Birney, E. Comparative genomics: Genome-wide analysis in metazoan eukaryotes. *Nature Reviews Genetics* **4**, 251-262, doi:Doi 10.1038/Nrg1043 (2003).

32 Woolfe, A. *et al.* Highly conserved non-coding sequences are associated with vertebrate development. *Plos Biol* **3**, 116-130, doi:ARTN e7DOI 10.1371/journal.pbio.0030007 (2005).

33 Bulger, M., , Groudine, M. in *Encyclopedia of Life Sciences* (John Wiley and Sons, 2005).

34 Arnosti, D. N. & Kulkarni, M. M. Transcriptional enhancers: Intelligent enhanceosomes or flexible billboards? *J Cell Biochem* **94**, 890-898, doi:Doi 10.1002/Jcb.20352 (2005).

35 Bondarenko, V. A., Liu, Y. V., Jiang, Y. I. & Studitsky, V. M. Communication over a large distance: enhancers and insulators. *Biochem Cell Biol* **81**, 241-251, doi:Doi 10.1139/003-051 (2003).

36 Sipos, L. & Gyurkovics, H. Long-distance interactions between enhancers and promoters - The case of the Abd-B domain of the Drosophila bithorax complex. *Febs J* **272**, 3253-3259, doi:DOI 10.1111/j.1742-4658.2005.04757.x (2005).

37 Becker, W. M., Kleinsmith, Lewis, J., Hardin, Jeff., Bertoni, Gregory Paul. (Benjamin Cummings, 2008).

38 Banerji, J., Olson, L. & Schaffner, W. A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. *Cell* **33**, 729-740, doi:0092-8674(83)90015-6 [pii] (1983).

39 Bulger, M. & Groudine, M. Enhancers: the abundance and function of regulatory sequences beyond promoters. *Dev Biol* **339**, 250-257, doi:S0012-1606(09)01401-8 [pii]10.1016/j.ydbio.2009.11.035 (2010).

40 Flicek, P. *et al.* Ensembl 2008. *Nucleic Acids Res* **36**, D707-714, doi:gkm988 [pii]10.1093/nar/gkm988 (2008).

41 Matsunaga, E., Araki, I. & Nakamura, H. Pax6 defines the di-mesencephalic boundary by repressing En1 and Pax2. *Development* **127**, 2357-2365 (2000).

42 LaVail, J. H. & Cowan, W. M. The development of the chick optic tectum. I. Normal morphology and cytoarchitectonic development. *Brain Res* **28**, 391-419 (1971).

43 Muroyama, Y., Fujihara, M., Ikeya, M., Kondoh, H. & Takada, S. Wnt signaling plays an essential role in neuronal specification of the dorsal spinal cord. *Genes Dev* **16**, 548-553, doi:10.1101/gad.937102 (2002).

44 Pearson, R. (Academic Press Inc, London, 1972).

45 Kuhlenbeck, H. Vol. 4 (S.Karger, Basel, 1975).

46 Patten, B. M. (McGraw-Hill Book Company, Inc., York, PA, 1951).

47 Kuenzel, W. J. (The Johns Hopkins University Press, Baltimore and London, 1988).