

## Research Article

# Computational and Experimental Approach for the Identification of Mesencephalon-Specific *Cis*-Element of *Dbx1* Gene

Jennifer Kim<sup>#</sup>, Mohammed M. Islam<sup>#</sup>, Evangeline Tzatzalos, Hailing Hao, Ying Li, Haim Y. Tawil, Sungtae Doh and Li Cai\*

Department of Biomedical Engineering, Rutgers University, USA

<sup>#</sup>Both the authors contribute equally to this work

## Corresponding author

Li Cai, Department of Biomedical Engineering, Rutgers University, 599 Taylor Road, Piscataway, NJ 08854, USA, Tel: 848 445-6559; Fax: 732 445-3753; Email: lcai@rutgers.edu

Submitted: 05 September 2013

Accepted: 01 October 2013

Published: 03 October 2013

## Copyright

© 2013 Kim et al.

## OPEN ACCESS

## Keywords

- Gene regulation
- *cis*-Element
- Enhancer
- *Dbx1*
- Brain
- Mesencephalon
- Chick embryo

## Abstract

Gene regulation is central to the biological processes of differentiation, morphogenesis, and the creation of diversity of cell types. The identification of gene regulatory elements is thus essential for the understanding of the molecular mechanism underlying these processes. Here, we present an integrative approach for the identification and characterization of a *cis*-element of *Dbx1* gene. *Dbx1* (developing brain homeobox 1) plays an essential role during development of the central nervous system (CNS). *Dbx1* expression is regionally restricted within the brain and spinal cord to define neuronal specific domains. Although the functional roles of *Dbx1* have been established, the molecular mechanism that governs *Dbx1* regional specific expression is not well understood. Using comparative genomics method, we revealed a highly evolutionarily conserved noncoding DNA fragment of 624 bp located approximately 18 kb upstream of *Dbx1* locus (CR5) and is predicted as a *cis*-element. Using experimental approaches, we show that this *cis*-element CR5 directs mesencephalon-specific gene expression during chick embryonic CNS development. Cells with CR5 activity contain a heterogeneous neuronal population suggesting that CR5 is involved in multiple neuronal lineage development. Together, these findings provide additional insight into the complex regulation of *Dbx1* expression during CNS development. In addition, our study established an effective approach for the discovery and characterization of genomic regulatory elements.

## INTRODUCTION

Regulation of gene expression is essential for cell development and growth; it controls the processes of morphogenesis and the generation of different cell types [1-3]. The identification and characterization of the gene regulatory elements is essential for the understanding of the molecular mechanism underlying these biological processes and for the annotation of the genome.

Previously, the discovery of *cis*-regulatory elements typically relied on extensive experimentation. Tissue-specific *cis*-regulatory elements were generally identified by promoter/enhancer deletion studies in transgenic mice [4] or by DNase I hypersensitivity mapping in expressing tissues [5-7]. The DNase I hypersensitivity mapping technique is based on the observation that local disruptions of the regular nucleosomal array create

preferential targets for DNase I. Binding of transcription factors to the DNA at *cis*-regulatory elements is thought to cause the removal or dispersal of a nucleosome. Thus, DNase I hypersensitive sites are often found at active *cis*-regulatory elements. Genomic regions close to the promoter are sometimes further explored by footprinting and by electrophoretic mobility-shift assays, which seek to identify regulatory protein-binding DNA regions [5,8]. The ability to evaluate the expression of thousands of genes across various experimental conditions has allowed bioinformatics approaches to these problems. Clusters of genes that show similar patterns of expression were searched within their upstream sequences for over-represented or conserved sequence motifs. This method has worked well in bacteria and yeast [8]; however, due to the highly complex nature of the genome, finding *cis*-regulatory elements in vertebrates still remains as a challenge

[9]. Completion of the human genome sequencing along with many other species has allowed comparative genomic approach to identify evolutionarily conserved DNA sequences and with a great success [10-15]. However, detailed characterization of the spatiotemporal gene regulatory activities is still lacking. In this study, we use developing brain homeobox 1 (Dbx1) gene as an example to demonstrate a relatively rapid and effective approach for this problem.

During neurogenesis, various genes have been found to be fundamental in defining the basic structure and organization of the central nervous system (CNS) [16]. These genes are expressed in a spatiotemporal pattern and provide specific cues that control tissue development. Dbx1 has been identified to have specific regional expression, and it plays an essential role in the patterning of the CNS during embryogenesis [17,18]. Dbx1 is expressed in both the brain and spinal cord of developing mouse embryos [17-20]; the expression initially starts in the analogs of diencephalon and spinal cord at embryonic day 9.5 (E9.5) mouse brain. From E10.5 to E12.5, Dbx1 is expressed in the dorsal mesencephalon. The expression is also detected in the telencephalon and hindbrain. At later gestation, however, Dbx1 expression discontinues in the spinal cord and becomes even more restricted within the diencephalon, dorsal mesencephalon, as well as the primitive cerebellum [17-19]. On a cellular level, Dbx1 expression has been observed to be limited to regions of active mitosis, indicating its role in early specification of Dbx1-expressing cells and their progenies [17]. Further evidence correlates Dbx1 expression strongly to progenitor cells in the ventricular zone at the boundaries between dorsal and ventral parts of the neural tube [18-21]. Within the spinal cord, Dbx1 has long been established to define distinct progenitor domains of V0 interneurons along the dorsoventral axis [17-20,22]. In more recent years, studies have demonstrated that Dbx1-derived cells have a high capacity to migrate from their sites of origin to populate different cortical regions [23-25]. The Dbx1-expressing neural progenitors are responsible for the generation of a subset of oligodendrocytes [26], interneurons [27,28] in the spinal cord, Reelin-positive Cajal-Retzius cells [25,29], and a subtype of inhibitory neuron in the medial amygdala [30], indicating that these Dbx1-expressing cells are a source of neuronal diversity. Although the specific roles and functions of Dbx1 in the CNS development have been established, the transcriptional mechanism that governs the complex expression of Dbx1 is still not well understood. One transgenic mouse study shows that a 5.7 kbp region located directly 5' of Dbx1 gene can control Dbx1 expression similar to the expression pattern of the endogenous Dbx1 gene within both the brain and the spinal cord [31].

Here, we report the development of an integrative approach for the identification and characterization of an evolutionarily conserved DNA element of 624 bp (CR5) located approximately 18 kbp upstream of the Dbx1 transcription start site. Using our newly developed reporter assay system in chick embryos, we found that CR5 acts as a novel regulatory element for brain-specific gene expression. In particular, CR5 is capable of directing reporter GFP expression specifically in the developing mesencephalon, distinct from the previously reported 5.7 kbp

sequence element of Dbx1 gene. Furthermore, several subregions of CR5 were also able to individually direct mesencephalon-specific GFP expression. Our findings indicate that CR5 is a novel *cis*-element for Dbx1 responsible for mesencephalon-specific gene expression during the chick brain development. In addition, CR5 activity was present in various neuronal cell types. Thus, we have established an integrative approach for the annotation and characterization of genomic regulatory elements.

## MATERIALS AND METHODS

### Computational prediction of Dbx1 cis-elements

Multiple genome sequence alignment methods were used to identify evolutionarily conserved non-coding DNA regions as putative gene regulatory elements. The genomic Dbx1 sequences and their orthologs from the various species including human, mouse, rat, and other vertebrate species were retrieved using non-coding sequence retrieval system, NCSRS [32]. These orthologous sequences were then aligned using multi-LAGAN [33] to identify regions with  $\geq 70\%$  identity and  $\geq 100$  bp. The percent identity and length of the conserved regions (CR) were used to calculate a ranking score for each CR (ranking score = percent identity + (length/60)). The minimum required length of 100 bp was to ensure significance in sequence conservation. Based on this scoring system, the percent identity was more heavily weighted to ensure that shorter (i.e., 200 bp) and very highly conserved sequences are not ranked below longer (i.e., 1000 bp) sequences with lower levels of conservation (Figure 1A).

### Design of plasmid constructs for reporter assay

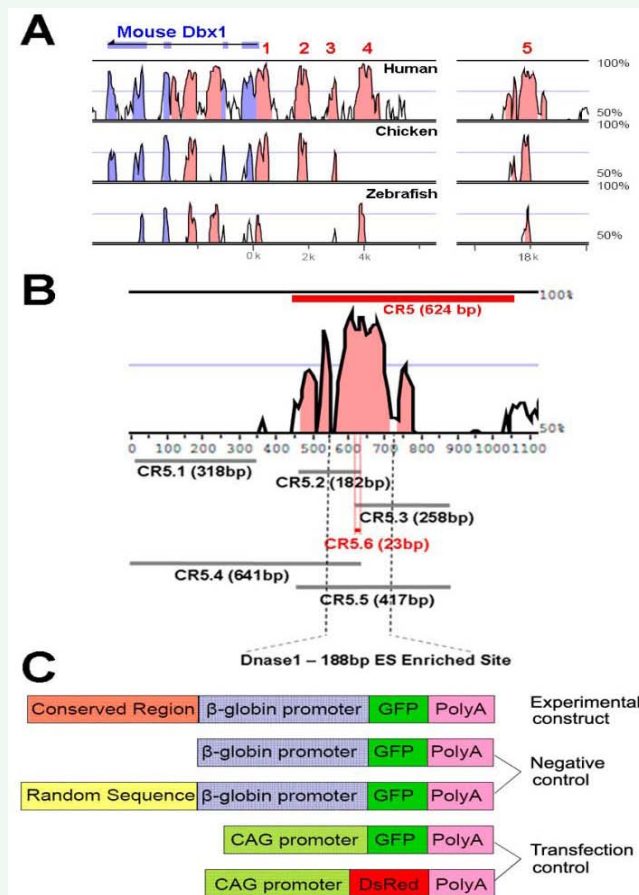
CR5 and its subregions (Figure 1B-C) were PCR amplified (Table 1) from mouse tail genomic DNA and individually subcloned into a GFP reporter construct containing a human basal promoter,  $\beta$ -globin promoter (BGP) [34]. Resulting constructs were verified by DNA sequencing.

For control experiments, we designed two negative control constructs (one contains a random DNA sequence coupled with BGP-GFP and the other is BGP-GFP alone without a conserved sequence); and two transfect control constructs (CAG-DsRed and CAG-GFP) (Figure 1C); CAG promoter is chicken  $\beta$ -actin promoter with a CMV enhancer element that is able to drive high levels of reporter expression in all transfected cells [35].

### Chicken embryos and *in ovo* electroporation

Fertilized pathogen-free (SPF) white leghorn chicken eggs (Sunshine Farms, Catskill, NY) were incubated (GQF manufacturing, Savannah, GA) at 38°C with 58-60% humidity and protected from light for approximately 40-45 hours to obtain embryos that were at the Hamburger Hamilton stage 10-12 (HH10-12; ~E2) [36]. All of the animal experiments were approved by the Animal Care and Facilities Committee at Rutgers University (approval ID 05-060).

Plasmid DNA (3-5 $\mu$ g/ $\mu$ l) was injected into the neural tube of chick embryos using a PCR micropipette glass-pulled needle. Fast green dye (0.0125%) was mixed with plasmid DNA for visualization purposes. Gold-plated electrodes were set



**Figure 1** Sequence analysis to predict Dbx1 *cis*-elements and experimental construct design for functional verification. (A). Multiple-species alignment of Dbx1 sequences from human, chicken, and zebrafish using mouse sequence as baseline. The sequence alignment reveals highly conserved regions (CR, pink peaks) of the noncoding Dbx1 locus as putative *cis*-elements. CR5 is the highest ranked region. Blue peaks represent exons of Dbx1. (B). Pair-wise alignment of the CR5 region from mouse and human. CR5.1-CR5.6 are six smaller fragments. They were individually cloned into the reporter construct. A DnaseI protected site is also depicted between the dotted lines, which overlaps with CR5.5. (C). Schematic of experimental and control reporter constructs for testing gene regulatory activity.

**Table 1:** List of PCR primers for Dbx1CR5 and its subregions.

Conserved Region	PCR product length (bp)	Primer	Primer Sequence
CR5	1185	forward	CGA TAT ATA CTA GTA AGC ATG AAG CAT GAA GGG CTG TG
		reverse	CGA TAT ATG GCC GGC CAG GCC ATG AAC AAC CTG AGA GGA A
CR5.1	318	forward	CGA TAT ATA CTA GTA AGC ATG AAG CAT GAA GGG CTG TG
		reverse	CGA TAT ATG GCC GGC CAC ACA CAC ACA CAC ACA C
CR5.2	182	forward	CGA TAT ATA CTA GTA GCT TTG AAT GTG CAA CCC GAT CC
		reverse	CGA TAT ATG GCC GGC CAG GAA CAA GGG ACA ACC TCT GCT T
CR5.3	258	forward	CGA TAT ATA CTA GTA AGC AGA GGT TGT CCC TTG TTC CT
		reverse	CGA TAT ATG GCC GGC CTA ATC AAG CCA CGG TGC TGT GAG A
CR5.4	641	forward	CGA TAT ATA CTA GTA AGC ATG AAG CAT GAA GGG CTG TG
		reverse	CGA TAT ATG GCC GGC CAG GAA CAA GGG ACA ACC TCT GCT T
CR5.5	417	forward	CGA TAT ATA CTA GTA GCT TTG AAT GTG CAA CCC GAT CC
		reverse	CGA TAT ATG GCC GGC CTA ATC AAG CCA CGG TGC TGT GAG A
CR5.6	23	forward	CGA TAT ATA CTA GTA AGC AGA GGT TGT CCC TTG TTC CT
		reverse	CGA TAT ATG GCC GGC CAG GAA CAA GGG ACA ACC TCT GCT T

approximately 4mm apart to electroporate embryos using a BTX ECM 830 system (Harvard Apparatus, Holliston, MA) to deliver 5 pulses at 18 volts, 50 ms duration, and with 950ms intervals. Electroporated eggs were incubated until appropriate harvest time points.

### Tissue processing and immunostaining

Transfected tissues were harvested at various developmental stages. Tissue processing and immunostaining were performed as described previously [37] with the following antibodies: NeuN (1:1000, Millipore), GFAP (1:250, Accurate), GS (1:250, Santa Cruz), Pk $\alpha$  (1:400, Santa Cruz), Chx10 (1:50, Santa Cruz); and the following antibodies were obtained from the Developmental Studies Hybridoma Bank developed under the auspices of the NICHD and maintained by The University of Iowa, Department of Biology (Iowa City, IA 52242): Acetylcholine receptor (mab270), Hb9 (81.5C10), 3CB2, Vimentin (40E-C, and H5), Pax6, Lim1+2 (4F2), Evx1 (99.1-3A2).

### Image analysis

Whole-mount samples were examined using a Leica MZ16FA fluorescent dissection microscope. Tissue sections were analyzed using a Zeiss Axio Imager A1 and an Olympus Fluoview FV10i confocal microscope.

### Statistical analysis

GFP expression in each sample was categorized based on its anatomical location within the CNS: telencephalon, diencephalon, mesencephalon, metencephalon, myelencephalon, and spinal cord. Frequency of GFP occurrence within each CNS regions at each embryonic stage was determined by counting the number of GFP+ samples. Statistical analysis was performed using a paired two-tailed t-test, P<0.05 indicates a statistical significance.

## RESULTS AND DISCUSSION

### Prediction of evolutionarily conserved noncoding DNA of Dbx1 as cis-elements

To predict evolutionarily conserved noncoding sequence elements that may serve as transcription regulators for Dbx1 expression, we performed DNA sequence analysis. The intergenic sequences flanking the 5' and 3' regions of *Dbx1* from various species, including human, mouse, chicken, and other vertebrate species, were retrieved using a sequence retrieval system NCSRS [32], and aligned using multi-LAGAN [33,38]. Resulting alignment revealed that many non-coding fragments highly conserved among different species (Figure 1A). The highest conserved region CR5 is located approximately 18 kbp upstream of *Dbx1* transcription start site containing 624bp nucleotides (pink peak in Figure 1B), and it also contains a known DNaseI hypersensitivity site [39], indicating that CR5 is a strong candidate as a functional *cis*-element for *Dbx1* gene.

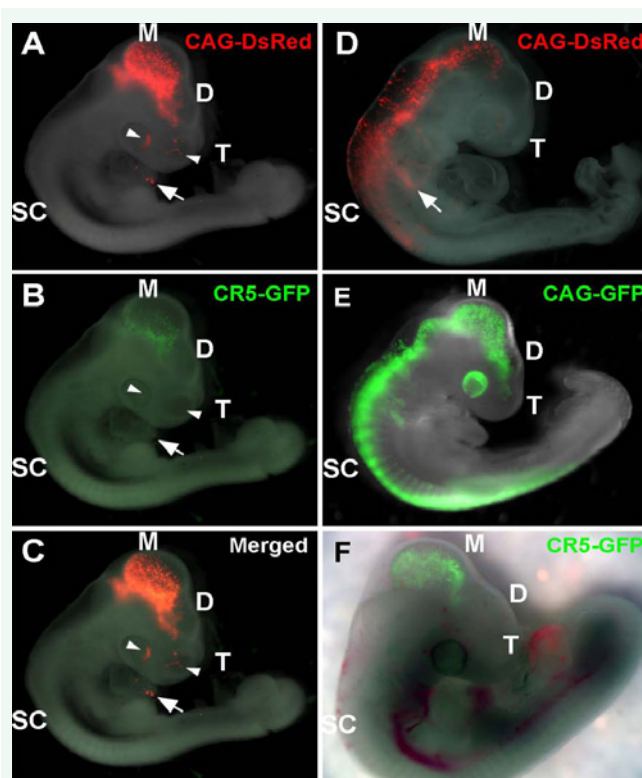
### CR5 directs tissue-specific GFP expression in the developing chick CNS

To examine the possibility that CR5, conserved among human, mouse, chick and other vertebrate species, might represent *cis*-regulatory activity, we explored its ability to direct tissue-specific

gene expression during chick CNS development with an *in ovo* reporter assay. Mouse CR5 was cloned upstream of a human beta-globin minimal promoter ( $\beta$ GP) [34] coupled to a reporter gene, green fluorescent protein (GFP)(Figure 1C). Plasmid constructs CAG-GFP or CAG-DsRed were used as transfection controls. CAG (chicken  $\beta$ -actin promoter with CMV enhancer), a strong and ubiquitous promoter, can drive gene expression in all transfected cell types [35]. The experimental and transfection control constructs were injected into the developing chick neural tube at Hamilton-Hamburger stage 11-12 (~ embryonic day 2, E2), and followed by electroporation to facilitate the uptake of DNA constructs into the neural stem/progenitor cells lining the neural tube. The patterns of reporter expression from CR5-GFP and CAG-DsRed were compared at various stages from E3 to E18. The control CAG-DsRed+ cells (Figures 2A, D) and CAG-GFP+ cells (Figure 2E) were observed in both CNS tissues (e.g., telencephalon, mesencephalon, spinal cord, eye) and non-CNS tissues (e.g., skin, limb, and heart), while CR5-GFP+ cells were restricted to the developing CNS (Figures 2B, F; n=22).

### CR5 is preferentially active in the developing mesencephalon

To determine temporal specificity of CR5 in directing gene



**Figure 2** CR5 directs mesencephalon-specific GFP expression in developing chick CNS. Expression of GFP was examined in the developing brain of E3-E4 embryos after *in ovo* electroporation at E2 (i.e., HH stage10-12). The control CAG-DsRed+ cells (A and D) and CAG-GFP+ cells (E) were found in the eye, telencephalon, diencephalon, mesencephalon, spinal cord of the CNS and also in the non-CNS tissues, e.g., skin (arrowheads) and heart (arrow). Expression of CR5-GFP was restricted in the mesencephalon (B, C and F). T, telencephalon; D, diencephalon; M, mesencephalon; SC, spinal cord.

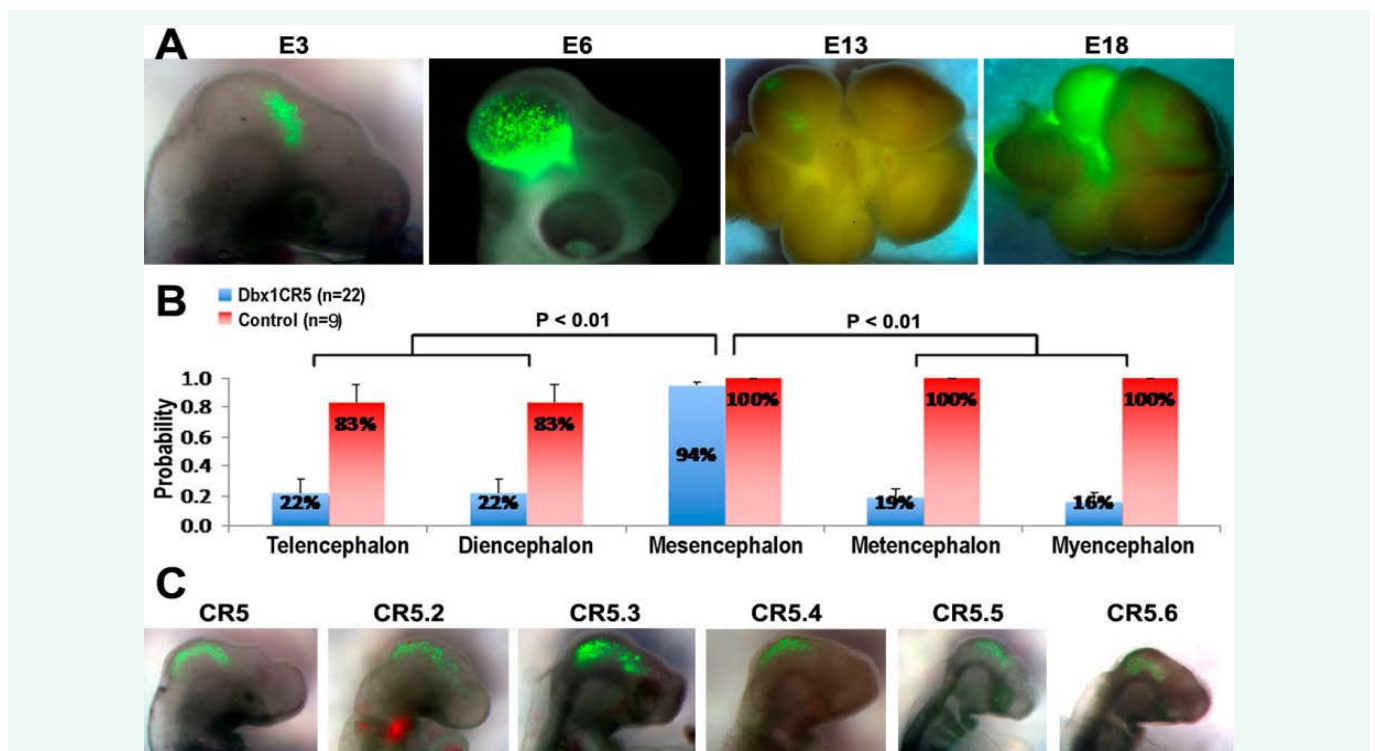
expression during chick CNS development. GFP expression was examined in transfected chick embryos at various developmental stages from E3 to E18. The differences in reporter expression between the control (CAG-GFP or CAG-DsRed) and experimental group (CR5-GFP) were determined by categorizing reporter expression pattern into different CNS regions with respect to time (Figure 3). CR5-GFP expression was found predominantly within the mesencephalon and occasionally the diencephalon or other regions (Figure 3B). No CR5-GFP<sup>+</sup> cells were observed within the spinal cord (data not shown). The pattern of CR5-GFP expression was largely consistent within the developing mesencephalon (Figure 3A-B). This result indicates that CR5 activity exists primarily in the mesencephalon of the developing chick brain and not in the spinal cord.

To determine minimum active region, CR5 subregions (CR5.1-5.6 in Figure 1B) were tested for their gene regulatory activity. Except CR5.1, all other five subregions, CR5.2-CR5.6, were shown to drive GFP expression in E3 chick embryos (Figure 3C). GFP expression pattern resulted from each of CR5 subregions was analyzed at E3 (Table 3) and at E6 (Table 4). The observation revealed a predominant GFP expression in the mesencephalon (all 4 samples at E3 and all 3 samples at E6) resulted from CR5.5 construct. While CR5-GFP expression was never observed anywhere but the mesencephalon in E3 samples, each of the subregional constructs (CR5.2-CR5.6) had varying amounts of GFP expression in different regions of the brain

(e.g., telencephalon, diencephalon, spinal cord, etc.) in addition to mesencephalon. These results indicate each or combination of CR5.2 – 5.6 subregions may be important in restricting GFP expression specifically in the mesencephalon.

### CR5 is active in various neuronal subtypes

To determine the cellular identity of CR5-GFP<sup>+</sup> cells, cryosections of E13 and E18 brain tissues were immunostained using various antibodies marking a specific neuronal cell type (Figure 4). E13 samples were analyzed, because the chick neurogenesis is near completion at this stage [40], and many differentiated neuron cell types can be determined by immunostaining [37,40-42]. Immunostaining showed that CR5-GFP<sup>+</sup> cells were labeled with neuronal markers, e.g., Evx1 (21±0.07%; Figure 4A) and Lim1+2 (15±0.06%; Figure 4B). Confocal microscopic analysis further confirmed co-labeling of CR5-GFP<sup>+</sup> cells in E18 samples with Evx1 labeling (Figure 5). Lim1+2 [43] and Evx1 [44] are markers for interneurons. However, no CR5-GFP<sup>+</sup> cells were co-labeled with the glial marker GFAP (*data not shown*). To our surprise, markers for neural stem/progenitor cells and radial glia (e.g., Vimentin and Pax6) did not co-label with GFP<sup>+</sup> cells, indicating CR5 is not active in neural stem/progenitor cells. Staining with other neuronal markers (e.g., GABA, Tyrosine Hydroxylase, Brn3a, Engrailed-1; see summary in Table 2) failed to co-label with CR5-GFP<sup>+</sup> cells at any time points examined. Other regionally expressed molecules were also examined on CR5-GFP<sup>+</sup> cells, including Islet1 on E6



**Figure 3** CR5 and its subregions direct GFP expression at various stages during chick embryonic development. (A) GFP expression mediated by CR5 in chick developing brain at E3, E6, E13, and E18 after *in ovo* electroporation at E2 (HH stage 10-12).

(B) Quantification shows that CR5 directs GFP expression predominantly in the mesencephalon (For each stage, n ≥ 4. p < 0.01).

(C) GFP expression derived by CR5 subregions (CR5.2-CR5.6) in chick developing brain at E3 after *in ovo* electroporation at E2 (HH stage 10-12). The number of samples with reporter GFP expression in various developing subregions of the CNS is listed in Tables 3 and 4.

**Table 2:** Summary of antibody staining with CR5-GFP transfected tissue samples.

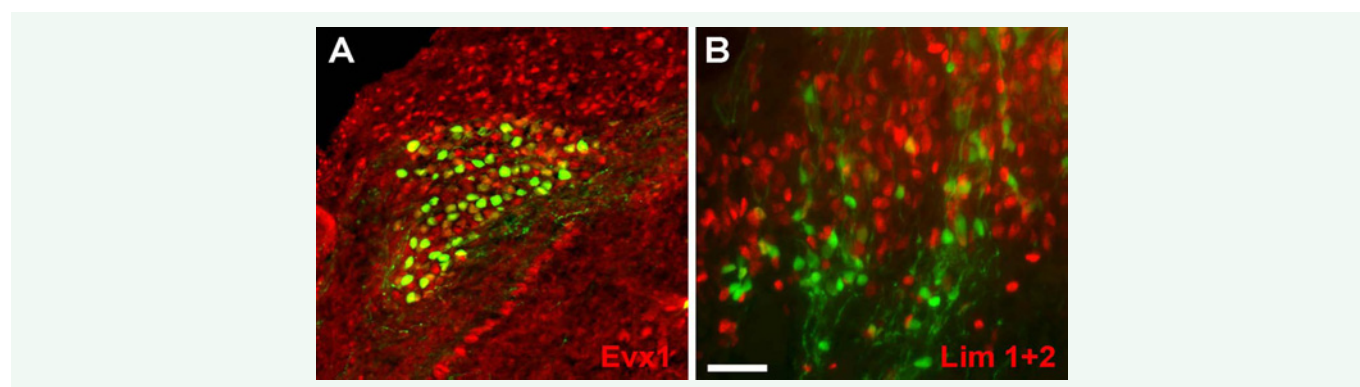
	NeuN	Lim1/2	Evx1	GFAP	Vimentin	Pax6	En1	Isl1	Hb9	Lim3	Chx10	GS	Ach-receptor	PKC $\alpha$
E3					-									
E4	-	-			-	-								
E6	-	+	+		-	-	-	-	-	-	-			
E13	+	+	+	-	-	-	-	-				-		
E18	+	+	+	-	-	-	-	-					-	-

**Table 3:** Number of samples with reporter GFP expression in various developing subregions of the CNS at E3.

	Telencephalon	Diencephalon	Mesencephalon	Metencephalon	Myelencephalon	Spinal Cord
CR5 (n=3)	0/3	0/3	3/3	0/3	0/3	0/3
CR5.2 (n=2)	1/2	1/2	2/2	1/2	1/2	1/2
CR5.3 (n=4)	0/4	2/4	2/4	1/4	2/4	2/4
CR5.5 (n=4)	0/4	3/4	4/4	1/4	1/4	2/4
CR5.6 (n=9)	1/9	3/9	4/9	5/9	6/9	5/9

**Table 4:** Number of samples with reporter GFP expression in various developing subregions of the CNS at E6.

	Telencephalon	Diencephalon	Mesencephalon	Metencephalon	Myelencephalon	Spinal Cord
CR5 (n=7)	0/7	0/7	7/7	1/7	1/7	0/7
CR5.2 (n=7)	0/7	0/7	3/7	7/7	7/7	0/7
CR5.3 (n=5)	0/5	0/5	0/5	5/5	5/5	0/5
CR5.5 (n=3)	0/3	0/3	3/3	1/3	1/3	0/3
CR5.6 (n=6)	0/6	0/6	0/6	0/6	0/6	0/6



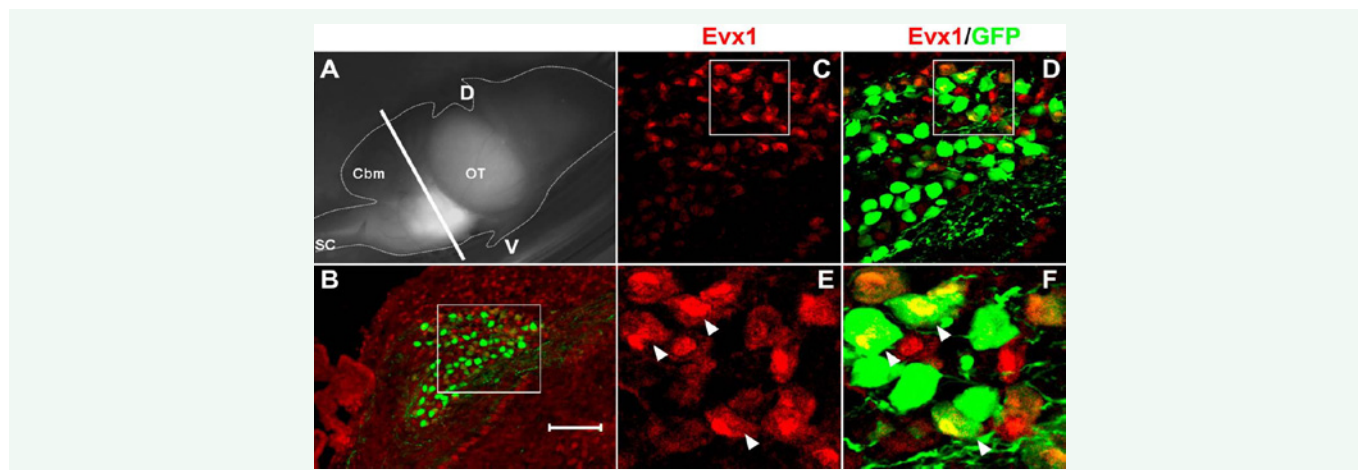
**Figure 4** CR5-GFP<sup>+</sup> cells in late chick embryos contain various neuronal types. Transverse sections of optic tectum (mesencephalon) from E18 (A) and E13 (B) embryos after *in ovo* electroporation at E2. CR5-GFP<sup>+</sup> cells were co-labeled with Evx1 (A) and Lim1+2 (B). Scale bar = 50 $\mu$ m.

and E13 samples, and Hb9, Lim3, and Chx10 on E6 samples, glutamine synthetase on E13 samples, acetylcholine receptor, and PKC $\alpha$  on E18 tissues, which all failed to co-label with CR5-GFP<sup>+</sup> cells (Table 2). These results indicate that CR5 activity exists in different neuronal cell types, but not in glial cells at the stages we examined.

## CONCLUSIONS

In this study, we demonstrate that CR5, a 624bp noncoding DNA fragment, acts as a *cis*-element that regulates gene

expression during chick brain development using an integrative approach combining computational prediction and experimental verification. CR5 is able to direct CNS-specific reporter GFP expression in the developing mesencephalon during embryonic days 3 (E3) through E18. CR5-GFP<sup>+</sup> cells contain a heterogeneous neuronal population indicating that CR5 may be involved in regulating gene expression for the development of multiple neuronal lineages. The capability of several subregions within CR5 (CR5.2-CR5.6) in directing GFP expression indicates that multiple *cis*-regulatory modules are involved in restricting mesencephalon-specific gene expression in the developing chick



**Figure 5** CR5-GFP<sup>+</sup> cells co-labeled with a neuronal marker Evx1. (A) Side view of transfected E18 brain sample in grey scale. (B) Transverse section as indicated by the white bar in grey scale. (C-F) Confocal images of stained tissue sections. Brain region of white box inset of (B) is shown in (C-D). Brain region in white boxes inset of (C-D) are shown in (E-F). Evx1 antibody staining co-localization with GFP<sup>+</sup> cells is indicated by white arrowheads. SC, spinal cord, Cbm, cerebellum, OT, optic tectum, D, dorsal, V, ventral. Scale bar = 50 μm.

CNS.

In previous studies, an enhancer of *Dbx1* of 5.7 kb DNA fragment that contains the transcriptional start site has been reported [31]. The 5.7kb enhancer showed reporter gene expression in the forebrain, midbrain, and spinal cord [17]. However, CR5-GFP expression is highly restricted in the mesencephalon, which indicates that 1) *cis*-element CR5 plays a specific role in the midbrain (particularly the mesencephalon) development, instead of having a global effect on overall *Dbx1* gene expression (e.g., 5.7 kb enhancer) [31]; 2) the regulation of *Dbx1* expression is a complex process involves multiple *cis*-regulatory elements for various patterns of *Dbx1* expression in the CNS. The fact that GFP expression is not detected in the spinal cord suggests that CR5 is not involved in the regulation of gene expression in the developing spinal cord.

The previously characterized 5.7 kb *Dbx1* enhancer should be a control for this study, however, the corresponding regions (i.e., CR3 and CR4 in Figure 1A) are not highly conserved in chick, and their ability to regulate gene expression in chick was not demonstrated (data not shown). This indicates that 1) there is a difference in regulating *Dbx1* gene expression between the two species of mouse and chicken; 2) not all conserved noncoding DNA sequence can function as a gene expression regulator.

Our study identified that CR5 is a functional *cis*-regulatory element outside of 5.7 kb promoter/enhancer region located directly upstream of the transcription start site. It provides an additional piece of information in the complex transcription regulation of *Dbx1* expression. There are several other conserved regions, and their gene regulatory activity needs to be determined in order to gain a full understanding the complex *Dbx1* expression. In addition, the mechanism of CR5 in regulating gene expression is likely via its interactions with specific protein factors. Future research should include the determination of such specific CR5 binding factors and their role in regulating *Dbx1* expression. Finally, our integrative approach used in this study has proven to be effective. Using this approach, we have successfully identified

and characterized *cis*-elements in the *Notch1* [45], *CD44* [46], and *Foxn4* [47] genes.

## ACKNOWLEDGEMENTS

We thank Drs. Richard S. Nowakowski and Mladen-Roko Rasin for discussion and comments on the manuscript. Antibodies against *En1*, *Evx1*, *Lim1/2*, *Lim3*, *Vimentin*, and *Pax6* were obtained from the Developmental Studies Hybridoma Bank developed under the auspices of the NICHD and maintained by The University of Iowa, Department of Biology, Iowa City, IA 52242. This work was supported in part by grants from the NIH (EY018738), the New Jersey Commission on Spinal Cord Research (08-3074-SCR-E-0 and 10-3091-SCR-E-0), Busch Biomedical Research Award, and The Aresty Center for Undergraduate Research.

## REFERENCES

1. Naval-Sánchez M, Potier D, Haagen L, Sánchez M, Munck S, Van de Sande B, et al. Comparative motif discovery combined with comparative transcriptomics yields accurate targetome and enhancer predictions. *Genome Res.* 2013; 23: 74-88.
2. Vandebon A, Kumagai Y, Akira S, Standley DM. A novel unbiased measure for motif co-occurrence predicts combinatorial regulation of transcription. *BMC Genomics.* 2012; 13 Suppl 7: S11.
3. Peter IS, Faure E, Davidson EH. Predictive computation of genomic logic processing functions in embryonic development. *Proc Natl Acad Sci U S A.* 2012; 109: 16434-16442.
4. Pfeffer PL, Payer B, Reim G, di Magliano MP, Busslinger M. The activation and maintenance of *Pax2* expression at the mid-hindbrain boundary is controlled by separate enhancers. *Development.* 2002; 129: 307-318.
5. Thomas GH, Elgin SC. Protein/DNA architecture of the DNase I hypersensitive region of the *Drosophila hsp26* promoter. *EMBO J.* 1988; 7: 2191-2201.
6. Elgin SC. The formation and function of DNase I hypersensitive sites in the process of gene activation. *J Biol Chem.* 1988; 263: 19259-19262.
7. Kleinjan DA, van Heyningen V. Long-range control of gene expression:

- emerging mechanisms and disruption in disease. *Am J Hum Genet.* 2005; 76: 8-32.
8. Pilpel Y, Sudarsanam P, Church GM. Identifying regulatory networks by combinatorial analysis of promoter elements. *Nat Genet.* 2001; 29: 153-159.
9. Vavouri T, Walter K, Gilks WR, Lehner B, Elgar G. Parallel evolution of conserved non-coding elements that target a common set of developmental regulatory genes from worms to humans. *Genome Biol.* 2007; 8: R15.
10. Pennacchio LA, Ahituv N, Moses AM, Prabhakar S, Nobrega MA, Shoukry M, et al. In vivo enhancer analysis of human conserved non-coding sequences. *Nature.* 2006; 444: 499-502.
11. Poulin F, Nobrega MA, Plajzer-Frick I, Holt A, Afzal V, Rubin EM, et al. In vivo characterization of a vertebrate ultraconserved enhancer. *Genomics.* 2005; 85: 774-781.
12. Visel A, Bristow J, Pennacchio LA. Enhancer identification through comparative genomics. *Semin Cell Dev Biol.* 2007; 18: 140-152.
13. Visel A, Minovitsky S, Dubchak I, Pennacchio LA. VISTA Enhancer Browser--a database of tissue-specific human enhancers. *Nucleic Acids Res.* 2007; 35: D88-92.
14. Visel A, Taher L, Girgis H, May D, Golonzhka O, Hoch RV, et al. A high-resolution enhancer atlas of the developing telencephalon. *Cell.* 2013; 152: 895-908.
15. De Val S, Chi NC, Meadows SM, Minovitsky S, Anderson JP, Harris IS, et al. Combinatorial regulation of endothelial gene expression by ets and forkhead transcription factors. *Cell.* 2008; 135: 1053-1064.
16. Gilbert SF, ed. *Developmental Biology.* 9th ed. 2010, Sinauer Associates, Inc.: Sunderland, MA. 711.
17. Lu S, Wise TL, Ruddle FH. Mouse homeobox gene *Dbx*: sequence, gene structure and expression pattern during mid-gestation. *Mech Dev.* 1994; 47: 187-195.
18. Lu S, Bogarad LD, Murtha MT, Ruddle FH. 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.* 1992; 89: 8053-8057.
19. Shoji H, Ito T, Wakamatsu Y, Hayasaka N, Ohsaki K, Oyanagi M, et al. Regionalized expression of the *Dbx* family homeobox genes in the embryonic CNS of the mouse. *Mech Dev.* 1996; 56: 25-39.
20. Pierani A, Brenner-Morton S, Chiang C, Jessell TM. A sonic hedgehog-independent, retinoid-activated pathway of neurogenesis in the ventral spinal cord. *Cell.* 1999; 97: 903-915.
21. 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 claustroramygdaloid complex. *J Comp Neurol.* 2004; 474: 504-23.
22. Jessell TM. Neuronal specification in the spinal cord: inductive signals and transcriptional codes. *Nat Rev Genet.* 2000; 1: 20-29.
23. Picardo MC, Weragalaarachchi KT, Akins VT, Del Negro CA. Physiological and morphological properties of *Dbx1*-derived respiratory neurons in the pre-Botzinger complex of neonatal mice. *J Physiol.* 2013; 591: 2687-2703.
24. Causeret F, Ensini M, Teissier A, Kessar N, Richardson WD, Lucas de Couville T, et al. *Dbx1*-expressing cells are necessary for the survival of the mammalian anterior neural and craniofacial structures. *PLoS One.* 2011; 6: e19367.
25. Griveau A, Borello U, Causeret F, Tissir F, Boggetto N, Karaz S, et al. A novel role for *Dbx1*-derived Cajal-Retzius cells in early regionalization of the cerebral cortical neuroepithelium. *PLoS Biol.* 2010; 8: e1000440.
26. Fogarty M, Richardson WD, Kessar N. A subset of oligodendrocytes generated from radial glia in the dorsal spinal cord. *Development.* 2005; 132: 1951-1959.
27. Lanuza GM, Gosgnach S, Pierani A, Jessell TM, Goulding M. Genetic identification of spinal interneurons that coordinate left-right locomotor activity necessary for walking movements. *Neuron.* 2004; 42: 375-386.
28. Pierani A, Moran-Rivard L, Sunshine MJ, Littman DR, Goulding M, Jessell TM. Control of interneuron fate in the developing spinal cord by the progenitor homeodomain protein *Dbx1*. *Neuron.* 2001; 29: 367-384.
29. Bielle F, Griveau A, Narboux-Nême N, Vigneau S, Sigrist M, Arber S, et al. Multiple origins of Cajal-Retzius cells at the borders of the developing pallium. *Nat Neurosci.* 2005; 8: 1002-1012.
30. Hirata T, Li P, Lanuza GM, Cocas LA, Huntsman MM, Corbin JG. Identification of distinct telencephalic progenitor pools for neuronal diversity in the amygdala. *Nat Neurosci.* 2009; 12: 141-149.
31. Lu S, Shashikant CS, Ruddle FH. Separate cis-acting elements determine the expression of mouse *Dbx* gene in multiple spatial domains of the central nervous system. *Mech Dev.* 1996; 58: 193-202.
32. Doh ST, Zhang Y, Temple MH, Cai L. Non-coding sequence retrieval system for comparative genomic analysis of gene regulatory elements. *BMC Bioinformatics.* 2007; 8: 94.
33. Brudno M, Do CB, Cooper GM, Kim MF, Davydov E, NISC Comparative Sequencing Program, Green ED, et al. LAGAN and Multi-LAGAN: efficient tools for large-scale multiple alignment of genomic DNA. *Genome Res.* 2003; 13: 721-731.
34. Yee SP, Rigby PW. The regulation of myogenin gene expression during the embryonic development of the mouse. *Genes Dev.* 1993; 7: 1277-1289.
35. Matsuda T, Cepko CL. Electroporation and RNA interference in the rodent retina in vivo and in vitro. *Proc Natl Acad Sci U S A.* 2004; 101: 16-22.
36. Hamburger V, Hamilton HL. A series of normal stages in the development of the chick embryo. 1951. *Dev Dyn.* 1992; 195: 231-272.
37. Doh ST, Hao H, Loh SC, Patel T, Tawil HY, Chen DK, et al. Analysis of retinal cell development in chick embryo by immunohistochemistry and in ovo electroporation techniques. *BMC Dev Biol.* 2010; 10: 8.
38. Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I. VISTA: computational tools for comparative genomics. *Nucleic Acids Res.* 2004; 32: W273-279.
39. Hubbard TJ, et al. Ensembl 2009. *Nucleic Acids Res.* 2009. 37(Database issue): D690-7.
40. Hollyday M. Neurogenesis in the vertebrate neural tube. *Int J Dev Neurosci.* 2001; 19: 161-173.
41. Cheung AF, Pollen AA, Tavare A, DeProto J, Molnár Z. Comparative aspects of cortical neurogenesis in vertebrates. *J Anat.* 2007; 211: 164-176.
42. Abdel-Mannan O, Cheung AF, Molnár Z. Evolution of cortical neurogenesis. *Brain Res Bull.* 2008; 75: 398-404.
43. Tsuchida T, Ensini M, Morton SB, Baldassare M, Edlund T, Jessell TM, et al. Topographic organization of embryonic motor neurons defined by expression of LIM homeobox genes. *Cell.* 1994; 79: 957-970.
44. Moran-Rivard L, Kagawa T, Saueressig H, Gross MK, Burrill J, Goulding



- M. Evx1 is a postmitotic determinant of v0 interneuron identity in the spinal cord. *Neuron*. 2001; 29: 385-399.
45. Tzatzalos E, Smith SM, Doh ST, Hao H, Li Y, Wu A, et al. A cis-element in the Notch1 locus is involved in the regulation of gene expression in interneuron progenitors. *Dev Biol*. 2012; 372: 217-228.
46. Smith SM, Cai L. Cell specific CD44 expression in breast cancer requires the interaction of AP-1 and NF-κB with a novel cis-element. *PLoS One*. 2012; 7: e50867.
47. Islam MM, et al. Meis1 regulates Foxn4 expression during retinal progenitor cell differentiation. *Biology Open*. 2013. doi: 10.1242/bio.20132279.

#### Cite this article

Kim J, Islam MM, Tzatzalos E, Hao H, Li Y, Tawil HY, et al. (2013) Computational and Experimental Approach for the Identification of Mesencephalon-Specific Cis-Element of *Dbx1* Gene. *JSM Biotechnol Bioeng* 1(3): 1017.