A long noncoding RNA cluster-based genomic locus maintains proper development and visual function

Fei Wang[†]**, Dalong Ren**[†]**, Xiaolin Liang**[†]**, Shengwei Ke, Bowen Zhang, Bing Hu**[*]**, Xiaoyuan Song**[*] **and Xiangting Wang** [ID][*]

Hefei National Laboratory for Physical Sciences at the Microscale, CAS Key Laboratory of Brain Function and Disease, School of Life Sciences, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, Anhui 230026, China

ABSTRACT

Long noncoding RNAs (lncRNAs) represent a group of regulatory RNAs that play critical roles in numerous cellular events, but their functional importance in development remains largely unexplored. Here, we discovered a series of previously unidentified gene clusters harboring conserved lncRNAs at the nonimprinting regions in brain (*CNIBs*). Among the seven identified *CNIBs*, human *CNIB1* locus is located at Chr 9q33.3 and conserved from *Danio rerio* to *Homo sapiens*. Chr 9q33.3-9q34.11 microdeletion has previously been linked to human nail-patella syndrome (NPS) which is frequently accompanied by developmental and visual deficiencies. By generating *CNIB1* deletion alleles in zebrafish, we demonstrated the requirement of *CNIB1* for proper growth and development, and visual activities. Furthermore, we found that the role of *CNIB1* on visual activity is mediated through a regulator of ocular development-*lmx1bb*. Collectively, our study shows that *CNIB1* lncRNAs are important for zebrafish development and provides an lncRNA cluster-mediated pathophysiological mechanism for human Chr 9q33.3-9q34.11 microdeletion syndrome.

INTRODUCTION

A gene cluster consists of a group of genes closely spaced on a chromosome that are coordinately expressed or have related functions (1,2). This fascinating system has been a central model for improving our understanding of genomic function and evolution (2–4). However, most of our knowledge about gene clusters comes from protein coding genes. Recently, several types of small noncoding RNA (ncRNA) clusters have been explored. For example, mi-

croRNAs (miRNAs) are often found organized in clusters (5–11). The *Sfmbt2* miRNA cluster, one of the largest miRNA clusters in mouse, is located in intron 10 of the *Sfmbt2* gene and contains 72 miRNA precursors. Both *Sfmbt2* and its miRNAs are expressed predominantly from the paternal allele in the placenta. Loss of the entire *Sfmbt2* miRNA cluster in *Mus. musculus* leads to severely impaired placenta development, frequent lethality, and fetal defects (8). Prader-Willi Syndrome (PWS) is a neurodevelopmental disorder associated with the paternal Chr 15q11-q13 deletion (12,13). The imprinted small nuclear ribonuclear protein N (*SNRPN*) locus inside this deletion is likely the minimal critical region of PWS and has been found to host at least two small nucleolar RNA (snoRNA) clusters (14–16). Several studies have suggested that a deficiency of the snoRNA clusters confers most, if not all, of the phenotype of PWS (14,17,18), establishing a strong link between misexpression of the snoRNA clusters and central nervous system (CNS) diseases.

Long noncoding RNAs (lncRNAs) represent a large (e.g. ∼68% of transcripts in human) and heterogeneous family of RNAs that are over 200 nucleotides in length and regulate various cellular events at the RNA level (19,20). A few independent studies have suggested that lncRNAs may form 'clusters' as well (19–21). For instance, we have reported previously that several adjacent lncRNAs derived from the 5′ regulatory region of the *CCND1* gene (ncRNA$_{CCND1}$) repressed the transcription of the *CCND1* gene upon genotoxic stress (20). However, this work did not thoroughly analyze whether these lncRNAs were unique transcripts or processing products from a large single parent lncRNA. Similarly, Tomita S. *et al.* identified a series of intragenic lncRNAs called Eleanors (ESR1 locus enhancing and activating noncoding RNAs) that play important roles in *cis* activation of the estrogen receptor during breast cancer adaptation (21).

The human brain, one of the most enriched organs for lncRNAs, has been reported to contain ∼20 000 specifically expressed lncRNAs (22), making the number of human brain-expressed lncRNAs comparable to the total number of coding genes in the entire human genome (22). Mercer *et. al.* screened 1328 of Allen Brain Atlas (ABA) probes in *Mus. musculus* and discovered 849 brain-specific lncRNAs with precise spatiotemporal expression patterns by *in situ* hybridization (23). Functional studies have suggested that lncRNAs contribute to neuronal cell type identity, plasticity, synaptic transmission, advanced cognitive activity, and brain development (23–25). Although there is a growing body of information and an increasing number of novel identified lncRNAs, the full scope of CNS expressed lncRNAs is only beginning to be understood.

Danio rerio (zebrafish) is a well-established vertebrate model system for studying human diseases and has yielded advances in multiple fields including developmental biology, neurobiology, and evolutionary theory of protein encoding genes (26). Recently, several landmark studies have explored the characteristics of zebrafish lncRNAs (27–29). Similar to their mammalian counterparts (22–25,30), zebrafish lncRNAs have been reported to have low-level conservation across species (27,29,31,32). However, it is worth mentioning that although the conserved lncRNAs accounted for <9% of the total lncRNAs in zebrafish, 467 zebrafish lncRNA transcripts had orthologues in both primates and rodents (29). It has been speculated that although the majority of lncRNAs serve as regulators for sophisticated physiological events (33), a small portion of conserved lncRNAs may be indispensable for embryogenesis, viability, and development (31,34). Therefore, by taking advantage of well-established model organisms such as zebrafish, an investigation of the conserved lncRNAs may offer significant advances in understanding the biological roles and molecular consequences of human disease-related lncRNAs.

Here, we report the discovery of a type of novel genomic structure that is composed of lncRNA-based gene clusters and named as *CNIBs* (conserved, nonimprinting lncRNA clusters in the brain). Among the seven identified *CNIBs*, the *CNIB1* locus exhibits the highest conservation score from zebrafish to primate and possesses strong association with human Chr 9q33.3-9q34.11 microdeletion syndrome. The *CNIB1* locus is shown to be required for proper development, locomotor activity, and visual functions in zebrafish. These defects are observed in a portion of Chr 9q33.3-9q34.11 microdeletion patients.

MATERIALS AND METHODS

Experimental animals

Ten-week-old male BALB/c mice (Vital River Laboratories, Beijing, China) were used in this study and were housed in a temperature-controlled room under a 12-h light/12-h dark cycle, with free access to food and water *ad libitum*. The adult mice were euthanized by cervical dislocation, and their brains were carefully dissected using a stereomicroscope to isolate the four experimental brain regions of interest for this experiment (olfactory bulb, hypothalamus, hippocampus, and cerebellum) according to the mouse brain atlas (35).

AB zebrafish (*D. rerio*) were used as wild-type (WT) zebrafish in this study. Zebrafish embryos were maintained on a 14-h light/10-h dark cycle at 28.5°C in E3 embryo medium. *N*-Phenylthiourea (PTU, Sigma, USA) was used to prevent larval pigment formation. All zebrafish surgeries were performed after anesthetization with tricaine methane-sulfonate (Sigma, Cat. No. MS-222) treatment. To determine the temporal and spatial expression levels of *CNIB1* lncRNAs, fish tissue from brain, eye, heart, and intestine were dissected from 15, 30, 60, 120 dpf zebrafish according to a previous study (36).

All animal manipulations were conducted in strict accordance with the guidelines and regulations set forth by the University of Science and Technology of China (USTC) Animal Resources Center and University Animal Care and Use Committee. The protocol was approved by the Committee on the Ethics of Animal Experiments of the USTC (Permit Number: PXHG-SXY201510183 for mouse experiments and USTCACUC1103013 for zebrafish experiments).

Reverse transcription and quantitative PCR (qPCR)

Total RNAs from a set of three mouse brain tissues, three adult zebrafish, or a set of 20 zebrafish larvae were extracted using TRIzol Reagent (Ambion, Cat. No. 15596) following the manufacturers' instructions. RNase-free DNase (Promega, Cat. No. M6101) was used to remove residual DNA. Reverse transcription was performed using the commercially available reverse transcription system (Promega, Cat. No. A5001). The real-time qPCR experiment was performed on Bio-Rad CFX96 qPCR system according to the manufacturers' instructions (Vazyme, Cat. No. Q111-03). Gene expression levels were calculated relative to the reference gene β-actin. Three technical replicates were performed for each RT-qPCR reaction and at least three biological replicates were performed for each condition. The RT-qPCR primer sequences are listed in Supplementary Table S1.

LncRNA microarray

Arraystar Mouse LncRNA Microarray V3.0 (Arraystar Inc.) was designed for the global profiling of mouse lncRNAs and protein-coding transcripts. Approximately 32 368 lncRNAs and 23 395 coding transcripts can be detected using this LncRNA Microarray. Sample labeling and array hybridization were performed according to the Agilent One-Color Microarray-Based Gene Expression Analysis protocol (Agilent Technology) with minor modifications. Agilent Feature Extraction software (version 11.0.1.1) was used to analyze acquired array images. Quantile normalization and subsequent data processing were performed using the GeneSpring GX v12.1 software package (Agilent Technologies). After quantile normalization of the raw data, lncRNAs and mRNAs were chosen for further data analysis. Three biological replicates were used per condition for Microarray analysis, and three to ten mice were used for RT-qPCR assay to validate the Microarray results.

Empirical bayes (eBayes) method

The eBayes method is a practical approach for general Microarray experiments with arbitrary numbers of treatments and RNA samples (37). This approach is equivalent to shrinkage of the estimated sample variances towards a pooled estimate, resulting in greatly stabilized inference when the number of arrays is small. The eBayes method was used for screening region-specific lncRNAs/mRNAs in four brain regions in our research. The six key steps of the eBayes approach include constructing gene expression matrix and an experimental design matrix, building a comparison model and simulating log linear models, eBayes testing, and producing testing results. The eBayes approach can be implemented using the limma software package based on the R and Bioconductor platforms. We screened the region-specific lncRNAs according to three requirements: a normalized expression value larger than 8; a fold change >6 times; a *P*-value lower than 0.01. The normalized value >8 threshold was decided to ensure that all possible bias of false positive expressing transcripts were removed. 'Fold change >6' was decided based on comparing different parameter settings from >2, >4, >6 to >8. The normalized value >8 was found to show the strongest correlation between the brain function and the lncRNAs identified.

RNA-Seq data analysis

The published RNA-Seq data of wild type mouse olfactory bulb (OB) tissue (Gene Expression Omnibus (GEO) accession number, GSE93061) was downloaded and analyzed for the genome wide distribution of lncRNAs using the TopHat software. The 'FPKM (fragments per kilobase of transcript per million mapped reads) value >40' was then used to select the lncRNAs that were highly expressed in the OB in GSE93061. The 'FPKM > 40' threshold was used because that was roughly equivalent to the 'normalized value > 8' used in Microarray analysis. It is calculated by the expression level of common lncRNAs detected both in our Microarray data (GSE106832) and published RNA-Seq data (GSE93061).

LncRNA capping assay using anti-2,2,7-trimethylguanosine antibody

The lncRNA capping assay has been previously described by Li *et al.* (38). Briefly, total RNAs from mouse tissues were separated by TRIzol Reagent after being grounded and homogenized. The obtained RNA samples were then added to beads (15 µl per sample) coupled with anti-2,2,7-trimethylguanosine (anti-m^{2,2,7}G) (Calbiochem, Cat. No. NA02A) or negative control anti-HA (Proteintech, Cat. No. 51064-2-AP) and equilibrated in cap binding buffer (25 mM Tris 7.5, 10 mM MgCl$_2$, 0.5 mM DTT and 1 U/µl SU-PERase.IN (Invitrogen, Cat. No. AM2696)). After incubation for 1 h at 4°C, the beads were washed three times with cap binding buffer. The cap-binding RNA was finally extracted and detected by RT-qPCR using the same primers as listed in Supplementary Table S1.

In vitro transcription

The cDNAs of zebrafish *lmx1bb* and zebrafish and mouse *CNIB1* lncRNAs were cloned into pGEM-T easy vector (General Biosystems, China). The plasmids were linearized to produce DNA templates and RNAs were *in vitro* transcribed using T7 RNA polymerase (Invitrogen, Cat. No. 18033019) or MAXIscript SP6 Transcription Kit (Invitrogen, Cat. No. AM1330) in combination with normal NTPs (Invitrogen, Cat. No. 18109017). *In vitro* transcribed RNAs were further used in zebrafish for overexpression, northern blotting, or ISH assays directly or after biotin/digoxigenin modifications, as described below.

Northern blotting

Antisense or sense *in vitro* transcribed probes for *CNIB1* lncRNAs were labeled with biotin RNA labeling mix (Sigma, Cat. No. 11685597910), followed by Northern blotting as described previously (39). Briefly, 10 µg of RNAs from the eyes of 2-month-old mice or zebrafish were separated in a 1% denaturing gel. RNAs were then transferred to Hybond-N membrane and hybridized with the individual biotin-labeled probes in hybridization buffer (Invitrogen, Cat. No. AM8670). After hybridization, blots were washed and detected by chemiluminescent nucleic acid detection module (Thermo Scientific, Cat. No. 89880). The Northern probes sequences are listed in Supplementary Table S2.

In situ hybridization

Whole mount *in situ* hybridization was performed as previously described (40). Briefly, fixed larvae were incubated in 50% formamide hybridization buffer with digoxigenin-labeled *CNIB1*, *pbx3b* or *lmx1bb* RNA probes, which were synthesized by *in vitro* transcription as described above. Following probe hybridization to processed samples, the digoxin-labeled probes were recognized using anti-digoxin (Roche, Cat. No. 11093274910, 1:5000 dilution) antibody at 68°C for 16–20 h. This digoxin antibody was reacted with the substrate of alkaline phosphatase (Promega, Cat. No. 0000185009). Approximately 8–15 larvae were used in each ISH experiment. Three independent ISH experiments were conducted with the specific antisense probe, and the stained embryos were imaged with BX-60 and BX-16 microscope (Olympus, Japan), respectively. The ISH probe sequences are listed in Supplementary Table S3.

CRISPR-Cas9 mediated gene modification and identification

The sgRNAs were designed using CRISPR-DO software (41) to target the upstream site of *uc.273+* and the downstream site of *uc.277+* (one sgRNA for each site) in order to excise the whole genomic region of the five intronic lncRNAs (*uc.273+*, *uc.275+*, *uc.275−*, *uc.276+*, and *uc.277+*) in zebrafish. *In vitro* synthesized Cas9 mRNA obtained with the T7 mMESSAGE mMACHINE Kit (Ambion, Cat. No. AMB13345) and the sgRNAs synthesized with T7 Riboprobe Systems (Promega, Cat. No. P1460) were co-microinjected into one-cell zebrafish embryos. Genomic DNA was extracted from 10 embryos at 24 hours postfertilization (hpf) and used as templates for genotyping to test

the engineering efficiency. Three pairs of specific primers for polymerase chain reaction (PCR) were designed to distinguish wild type, heterozygous, or homozygous knockout (KO) alleles. Each resulting PCR product was confirmed by gel electrophoresis. Separated electrophoresis bands of the PCR products were then purified and sequenced to ensure the genotype of tested zebrafish. The sgRNA and genotype primer sequences are listed in Supplementary Table S4.

The rest of injected embryos (F0) were raised to adulthood and then crossed with wild-type zebrafish to produce F1 embryos. From each cross, fin-clipped DNA of 10 F1 embryos were collected for genotyping. The F1 embryos that carried heterozygous knockout alleles or WT alleles were raised to adulthood and further crossed with their same genetic background siblings to generate F2 fish for downstream experiments.

Knockdown of *lmx1bb* using morpholino oligonucleotides (MO)

Lmx1bb MOs were designed as previously described (42) to target the splicing site junction between exon 1 and intron 1. Mismatch *lmx1bb* control morpholino was synthesized as a negative control. The MOs were resuspended in embryo medium. For efficient knockdown, *lmx1bb* MO (4 ng) or mismatch MO (4 ng) was injected per WT embryo. About 300 embryos were injected in each experiment. The data were analyzed from three independent experiments. The morpholino sequences are listed in Supplementary Table S4.

RNAs microinjection in zebrafish embryos

Five *CNIB1* lncRNAs (*uc.273+*, *uc.275+*, *uc.275−*, *uc.276+*, and *uc.277+*), their antisense controls (except for *uc.275+/−*), and *lmx1bb* mRNA were obtained by *in vitro* transcription as described above in the 'In vitro transcription' section. About 2 ng of *lmx1bb* mRNA, 1 ng of *CNIB1* lncRNAs mix, or 1ng of antisense lncRNAs mix were injected into WT and/or mutant embryos. The expression levels of *CNIB1* lncRNAs, *lmx1bb* mRNA, and Lmx1bb protein were detected at 48 hpf. Each data point contained expression data for 30 embryos. Three biological repeats were used per condition.

Zebrafish locomotor activity analysis

Zebrafish locomotor activity analysis was performed as described previously with minor modifications (43). At 4 dpf, a single larva was placed in each of the 96 wells of a 96-well plate (32 WT larvae, 32 offspring of $CNIB1^{+/−}$ incross larval type 1, and 32 offspring of $CNIB1^{+/−}$ incross larval type 2), in a Viewpoint apparatus (France). The locomotor behavior of larvae was monitored for three consecutive days under the previously stated LD cycle using an automated video-tracking system (VideoTrack, ViewPoint), and the movement of each larva was recorded and analyzed using Zebralab 3.10 software (ViewPoint). Inside the ViewPoint chamber, constant infrared light was provided, and white light was provided from 08:30 A.M. to 10:30 P.M. Temperature controlling devices were placed in the chamber

to maintain a constant temperature of 28.5°C. The Video-Track quantization parameters were selected as described previously (44). Data were further analyzed using Graph-Pad software and Visual Basic macros for Microsoft Excel.

Measurement of the zebrafish larval optokinetic response (OKR)

WT or offspring of $CNIB1^{+/−}$ incross larval zebrafish at 5 dpf were conducted for OKR as described previously (45). A customized program was developed for sine-wave grating generation and real-time recording of eye movements based on LabVIEW 7.1 (National Instruments, USA) and NI-IAQ 3.7 (National Instruments, USA). The 5 dpf larvae were placed dorsal side up in 5% methylcellulose in the center of a transparent petri dish located under a cone-shaped white paper screen to immobilize them during the experiments (46). As rotating grating patterns were presented around eyes of the larvae, the elicited eye movement was recorded in real time by an infrared-sensitive charge-coupled device (CCD camera) (TCA-1.3BW, China). All tests were conducted at 05:00 P.M. The gain value, i.e. the ratio of eye velocity to the angular velocity of the grating, was calculated using the software as previously described (46).

Histology and immunohistochemistry

WT or offspring of $CNIB1^{+/−}$ incross larval zebrafish were raised to 5 dpf and used for tyrosine hydroxylase (Th) immunofluorescence staining. The larvae were fully dark adapted for ∼30 min in order to separate the pigment epithelium from the retina, and they were then fixed with 4% PFA in PBS. The fixed larvae were rinsed with PBS. The eyes were dissected out, and both the crystalline lens and the pigment epithelia were removed. The fixed retinas were incubated in blocking solution at 4°C overnight and then incubated in Th primary antibody (Millipore, Cat. No. MAB318, 1:1000) at 4°C overnight. The following day, the samples were rinsed with PBS, incubated with secondary antibody (Alexa 488-conjugated anti-mouse, Invitrogen, Cat. No. A-21202, 1:1000) for 3 h, and then rinsed with PBS. All the specimens were mounted on slides for fluorescence microscopy (BX60, Olympus).

Statistical analysis

All experiments were independently repeated at least three times. The data were analyzed with an unpaired, two-tailed *t*-test or one-way ANOVA using GraphPad Prism. The results are shown as the mean $\pm$ SEM. The level of significance was set to $P < 0.05$. *, **, and *** represented $P < 0.05$, $P < 0.01$, and $P < 0.001$, respectively.

RESULTS

Analysis of lncRNAs that are highly expressed in the brain

To decipher the spatial expression landscape of well-annotated lncRNAs in various brain regions, we dissected adult male mouse brain tissues from the olfactory bulb (OB), the hypothalamus (Hypo), the hippocampus (Hippo),

and the cerebellum (Cereb), and subjected them to an Arraystar Mouse LncRNA Microarray V3.0. In this study, to remove possible bias of false positive expression, we set up a normalized value >8 as 'expressed transcripts' for both mRNAs and lncRNAs. With this criterion, we found a total of 17 619 expressed mRNAs and 20 202 expressed lncRNAs in the 4 tested brain regions (GEO accession number: GSE106832).

To identify RNA transcripts specifically expressed in each brain region, we applied a stringency set of standards, with fold changes >6 between brain regions and *P*-value <0.01. The resulting lncRNAs were defined as regional specific lncRNAs expressed in brain. Using eBayes analysis, we identified 352 transcripts as regional specific expressed, including 183 lncRNAs (Supplementary Figure S1A) and 169 mRNAs (Supplementary Figure S1B). Except for 49 lncRNAs that could not be tested due to the difficulty of designing specific primers, we validated the rest of the 134 regional specific expressed lncRNAs (Supplementary Table S5).

Identification of lncRNA clusters that are conserved, located in nonimprinting regions and expressed in the brain (*CNIBs*)

In analyzing the brain highly expressed lncRNAs, we noticed that 34 lncRNA loci were localized close together, forming nine clusters on seven chromosomes, with a minimum number of three lncRNAs for each cluster (Table 1 and data not shown). The expression of these 34 lncRNAs was further confirmed by data mining of published RNA-Seq data generated from wild type (WT) mouse tissues by Li *et al.* (47) (data not shown). Due to the high conservation of various types of gene clusters that have been previously reported, we performed conservation analysis of the identified 9 lncRNA clusters using the *phastCons* method (Supplementary Figure S2). Among these clusters, seven were considered to be conserved clusters because the involved lncRNAs were derived from either ultraconserved elements (UCRs) discovered by Bejerano *et al.* (48) or from shared conserved patches (SCR; ~71–1066 bp, 65–100% identity; Table 1, Figure 1A, Supplementary Figure S3). All seven conserved clusters were located in nonimprinting regions according to the Geneimprint database (http://www.geneimprint.com/site/home; Supplementary Figure S2). We thus named these seven lncRNA clusters as *CNIB* lncRNA clusters, referring to conserved, nonimprinting lncRNA clusters in the brain. A flowchart of the analytical process is shown in Supplementary Figure S2.

Characterization of *CNIB1* lncRNAs

Among the seven *CNIBs*, *CNIB1* was particularly interesting because it consisted of the largest number (8) of lncRNAs. Moreover, the 8 lncRNAs in *CNIB1* included seven UCRs (*uc.280+*, *uc.277+*, *uc.276+*, *uc.275+*, *uc.275-*, *uc.274+*, and *uc.273+*) and 1 SCR (*ENSMUST00000155423*). Except for *uc.274+* (originating from *G. gallus*), the remaining six UCR and one SCR loci appeared since zebrafish (*Danio rerio*) (Figure 1A). At the level of DNA sequences, *uc.280+* showed 76% identity throughout evolution, while the remaining 7 *CNIB1* lncRNAs contained hundreds of nucleotide sequences that were

87%-100% identical during evolution. In addition, these *CNIB1* lncRNAs were conserved at their genomic location (derived from the transcriptional factor pbx3 locus) (Figure 1B) and overall organization (same synthetic order and strand) once they appeared.

To further investigate the characterization of *CNIB1* lncRNAs, we collected Encyclopedia of DNA Elements (ENCODE) chromatin immunoprecipitation sequencing (ChIP-Seq) data for the distribution of peaks of promoter marker (H3K4me3), enhancer markers (H3K27ac and H3K4me1), insulator marker CTCF, and RNA polymerase II binding sites in murine *CNIB1* genomic region. The ChIP-Seq data showed that each lncRNA had a unique transcriptional start site and was transcribed independently (Figure 1B). In addition, using the anti-m^{2,2,7}G antibody, we showed that murine *CNIB1* lncRNAs were capped at their 5′ ends (Figure 1C). According to our northern blotting results (Figure 1D and E), most of the *CNIB1* lncRNAs seemed to have multiple variants both in mouse and zebrafish. RT-qPCR (Figure 1F and G) experiments showed that all the five detectable *CNIB1* lncRNAs (*uc.273+*, *uc.275+*, *uc.275−*, *uc.276+*, *uc.277+*) exhibited similar temporal and spatial expression patterns in both mouse and zebrafish. It is worth noting that many *CNIB1* lncRNAs showed elevated expression levels in brain and eye of both mice and zebrafish (Figure 1F and G), which was further confirmed by whole-mount ISH performed on zebrafish larvae (Figure 1H and I). Taken together, the similar characterization and expression pattern of *CNIB1* lncRNAs in mouse and zebrafish indicated a conserved function for *CNIB1* during evolution.

Microdeletion of Chr 9q33.3–9q34.11, encompassing the *CNIB1* locus, is associated with the clinical spectrum of developmental retardation and visual defects

As illustrated in Figure 2A, the human homologous sequence of *CNIB1* was located on Chr 9q33.3. In human, microdeletion of Chr 9q33.3-9q34.11 is associated with nail-patella syndrome (NPS) (49–53). NPS is an autosomal dominant disease, hallmarked by nail dysplasia and patellar aplasia or hypoplasia. Due to the complex genetic background, individuals with NPS display varying degrees of severity. Other areas of the body are often affected among NPS patients, including kidneys and the optical system. *LMX1B*, located on Chr 9q33.3, is a member of the LIM homeobox transcription factor and currently thought to be associated with NPS manifestations (54,55). However, it remains unclear whether the *CNIB1* region is also involved in the development of NPS (49,56).

To date, 10 NPS patients with relatively full clinical and genetic characterization have been reported (49–53) (Figure 2B). A summary of the clinical symptoms of each patient is listed in Supplementary Table S6. The common deletion region of *Lmx1b* is marked as Region b in Figure 2B. Among these 10 patients, 3 patients (patients 3, 5, and 6) had *CNIB1* deletion (Region a in Figure 2B). All three *CNIB1* deletion patients displayed severe developmental retardation and neurological manifestations. Moreover, two of these patients exhibited optical deficiencies (Supplementary Table S6). This information suggested that the visual
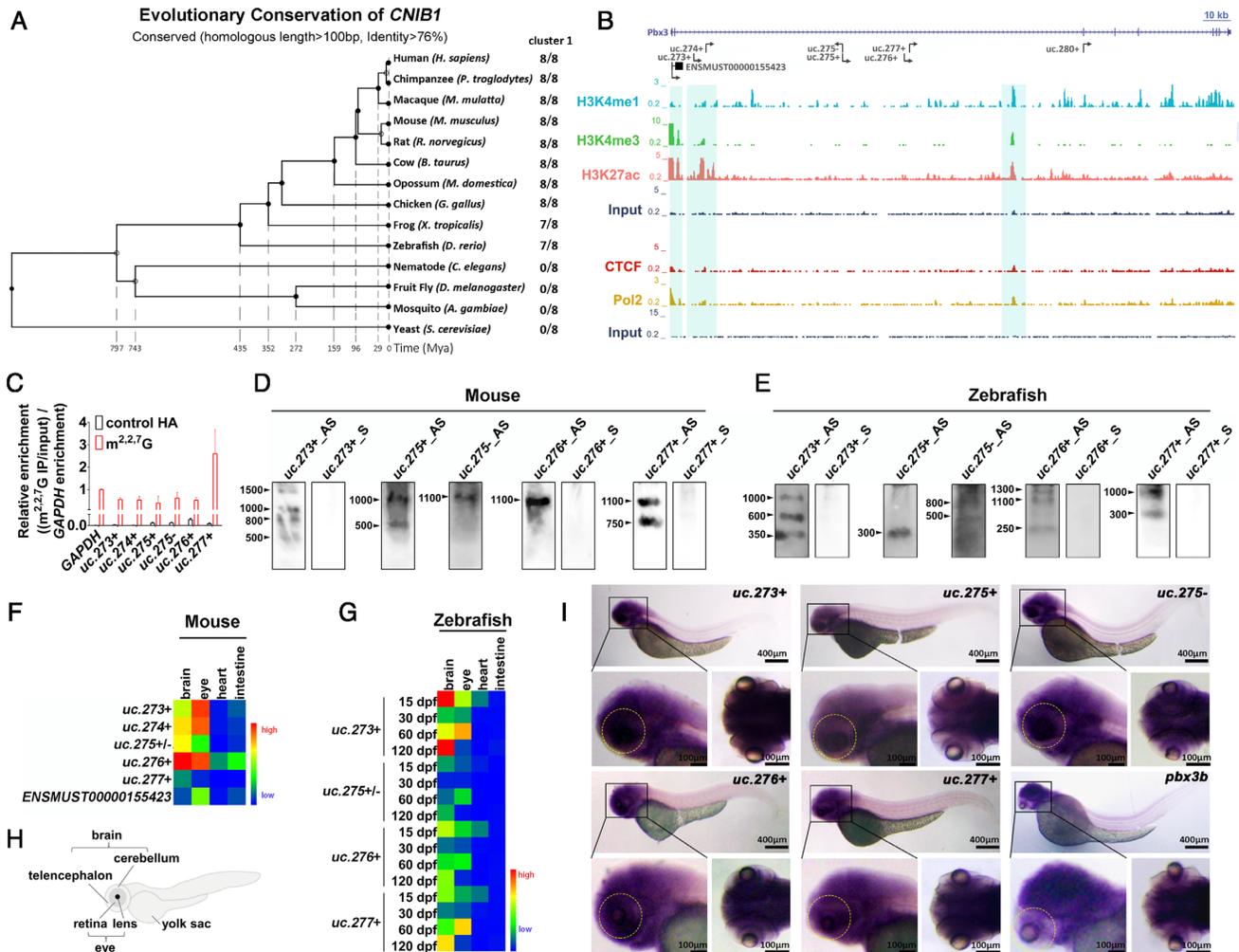
Figure 1. Characteristics of *CNIB1*. (**A**) Phylogenetic tree of *CNIB1* genomic locus in the 14 indicated species. (**B**) Comprehensive view of multiple peak distributions of promoter marker (H3K4me3), enhancer markers (H3K27ac and H3K4me1), insulator marker CTCF and Pol2 binding sites on *CNIB1* locus suggest these lncRNAs in *CNIB1* are transcribed as independent lncRNAs. The three regions highlighted in turquoise were predicted to be the transcription start sites for *CNIB1* lncRNAs. Scale bar: 10kb. (**C**) Cap assay of indicated *CNIB1* lncRNAs. Immunoprecipitation was performed using anti-m^{2,2,7}G antibody or control anti-HA, followed by RT-qPCR detection. Bar plots represent relative enrichment of RNAs immunoprecipitated by the antibody and normalized to *GAPDH*. (**D, E**) Northern blotting of *CNIB1* lncRNAs in the eyes of mouse (**D**) and of zebrafish (**E**) both at 2 months. (**F, G**) Heatmap of the expression level of *CNIB1* lncRNAs. qPCR assays were performed using the brain, eye, heart, and intestine tissues from three 10-week old male mice (**F**) and six zebrafish (**G**) at the indicated developmental time points (15, 30, 60 and 120 dpf, respectively), normalized to *β-actin*. (**H**) Schematic representation of zebrafish. (**I**) Whole-mount ISH of five indicated *CNIB1* lncRNAs (*uc.273+, uc.275+, uc.275-, uc.276+, uc.277+*) and *pbx3b* using corresponding antisense digoxigenin-labeled probes in larval zebrafish (*n* = 8). The yellow dotted oval marks the area where the eye of zebrafish is located.

disorders and other symptoms of NPS might be attributed to the *CNIB1* region.

Interestingly, besides the association between *CNIB1* locus and NPS syndrome, we found that three other *CNIB* (*CNIB2, CNIB3*, and *CNIB4*) loci were also linked to human CNS defects (Supplementary Figures S4 and S5 and Supplementary Tables S7–S9). *CNIB2* was transcribed from the *Meis2* gene. Among the 17 patients with CNS developmental defects and carrying a Chr 15q14 microdeletion, the region encompassing *CNIB2* and its nearby Meis2 was the minimal common key fragment (Supplementary Figure S4A–C, and Supplementary Table S7) (57–62). The *CNIB3* region was also the smallest shared part in over half of patients with Chr 13q32.2-33.2 deletion,

which manifested as severe neural tube defects, abnormalities of eye/hand/foot, and facial dysmorphisms (Supplementary Figure S5A–C) (63). *Dlx-5* and *Dlx-6*, the host genes of *CNIB4*, have been reported as critical regulators of neuronal-glial fate specification and oligodendrocyte lineage maturation (64). NR₋015388 and ENSMUST00000159827 (also known as *Evf2*) are within the *CNIB4* locus. These two lncRNAs are able to interact with *BRG1* to mediate chromatin remodeling events (65). Although genome mutations were not found for *CNIB4* lncRNAs, we discovered that the mutation of the RNA binding domain of BRG1 protein was associated with Coffin-Siris syndrome, a neurodevelopmental disorder (Supplementary Tables S8 and S9) (65).

Table 1. Summary of murine *CNIB* lncRNA clusters

CNIBs	chrom	seqname	strand	txStart	txEnd	Conservation
1	chr2	*uc.280+*	−	34 079 905	34 080 125	UCR
		uc.277+	−	34 142 371	34 142 647	UCR
		uc.276+	−	34 144 120	34 144 552	UCR
		uc.275-	+	34 166 204	34 166 459	UCR
		uc.275+	−	34 166 204	34 166 459	UCR
		uc.274+	−	34 215 213	34 215 540	UCR
		uc.273+	−	34 219 380	34 219 701	UCR
		ENSMUST00000155423	−	34 223 631	34 227 310	SCR
2	chr2	*uc.385-*	−	115 691 289	115 691 498	UCR
		ENSMUST00000140461	−	115 821 791	115 888 096	SCR
		AK043601	−	115 891 094	115 891 565	SCR
3	chr14	*AK049103*	−	122 856 718	122 857 397	SCR
		ENSMUST00000135485	−	122 869 505	122 872 801	SCR
		ENSMUST00000162218	+	123 312 296	123 336 691	SCR
4	chr6	*AK044144*	−	6 758 359	6 762 560	SCR
		AK089129	−	6 765 823	6 768 053	SCR
		ENSMUST00000159827	−	6 770 188	6 811 648	SCR
		NR_015388	−	6 770 545	6 819 533	SCR
		uc.220-	−	6 811 988	6 812 245	UCR
5	chr11	*uc007lit.1*	+	99 278 556	99 284 412	SCR
		ENSMUST00000132858	+	99 278 561	99 279 609	SCR
		ENSMUST00000128904	+	99 281 285	99 286 811	SCR
6	chr14	*ENSMUST00000145700*	+	13 453 983	13 566 712	SCR
		AK035448	−	13 507 280	13 510 560	SCR
		AK087257	−	13 519 086	13 597 291	SCR
7	chr8	*AK048779*	−	48 735 226	48 739 095	SCR
		CF174733	−	48 780 843	48 781 487	SCR
		AK013771	−	48 841 996	48 843 142	SCR

Abbreviations: UCR, ultraconserved elements; SCR, shared conserved patches.

$CNIB1^{+/-}$ zebrafish display developmental delay, decreased locomotor activity, and visual abnormalities

To investigate the *in vivo* function of *CNIB1*, we generated *CNIB1* mutant alleles in zebrafish using the CRISPR-Cas9 system. Zebrafish *CNIB1* contains 7 lncRNAs: *ENSMUST00000155423*, *uc.273+*, *uc.275+*, *uc.275-*, *uc.276+*, *uc.277+*, and *uc.280+*, which are listed from 5′ to 3′ of the sense strand of Chromosome 3. Of these, *uc.280+* and *ENSMUST00000155423* are overlapped with the two neighboring exons of *pbx3b*. The remaining 5 are consecutively distributed in the flanking intron of *pbx3b*, away from the splicing sites of *pbx3b*, and have no other known or annotated RNAs within this region. Thus, to avoid an effect caused by interference with *pbx3b*, we designed sgRNAs to delete the intronic region flanking the entire 5 *CNIB1* intronic lncRNAs (*uc.273+*, *uc.275+*, *uc.275-*, *uc.276+*, and *uc.277+*, Figure 3A), and referred to the resulting mutant fish as *CNIB1* knockout fish. Two independent deletion alleles were obtained in F1 generation by fin-DNA genotyping; $CNIB1^{+/-}$ 1 was the result of a 48 247 bp deletion and $CNIB1^{+/-}$ 2 had a 48 221 bp deletion (Figure 3B). These F1 heterozygous fish were incrossed to generate F2 offspring. Interestingly, we could not obtain any homozygous offspring and the ratio between wild-type (WT) and heterozygous offspring was about 1:2 (Figure 3C). Further investigation showed that the number of surviving eggs from

heterozygous incross was reduced to three quarters of that of WT incross at 24 hpf, although both groups had almost identical numbers of laid and living eggs at 4 and 12 hpf (Figure 3D). These results suggested that the $CNIB1^{-/-}$ zebrafish was embryonic lethal and the mortality might begin during segmentation.

In *CNIB1* heterozygous fish, all 5 involved lncRNAs were significantly downregulated when compared with their WT siblings (Figure 3E). In contrast, *gapvd1* and *mvb12bb*, the two genes neighboring *CNIB1* and *pbx3b*, showed no significant changes (Figure 3E). In order to eliminate the possibility that $CNIB1^{+/-}$ affected splicing of the *pbx3b* gene, we next sequenced the exon 2–exon 3 junction of the *pbx3b* gene, and designed multiple PCR primer sets for alternative splicing events of the *pbx3b* gene. Our data showed that exon 2–exon 3 junction (Figure 3F) and the tested alternative splicing events of the *pbx3b* gene (Figure 3G and Supplementary Figure S6) remained the same under both WT and $CNIB1^{+/-}$ backgrounds. These data together indicated that $CNIB1^{+/-}$ had no effect on the splicing of host gene *pbx3b*.

Furthermore, we found that F2 $CNIB1^{+/-}$ zebrafish displayed developmental delay, decreased locomotor activity and visual abnormalities when compared with offspring generated from WT incross. First, at 48 hpf, more than 75% of the eggs laid by the WT incross group hatched, while the hatching rate of $CNIB1^{+/-}$ 1 or $CNIB1^{+/-}$ 2 incross
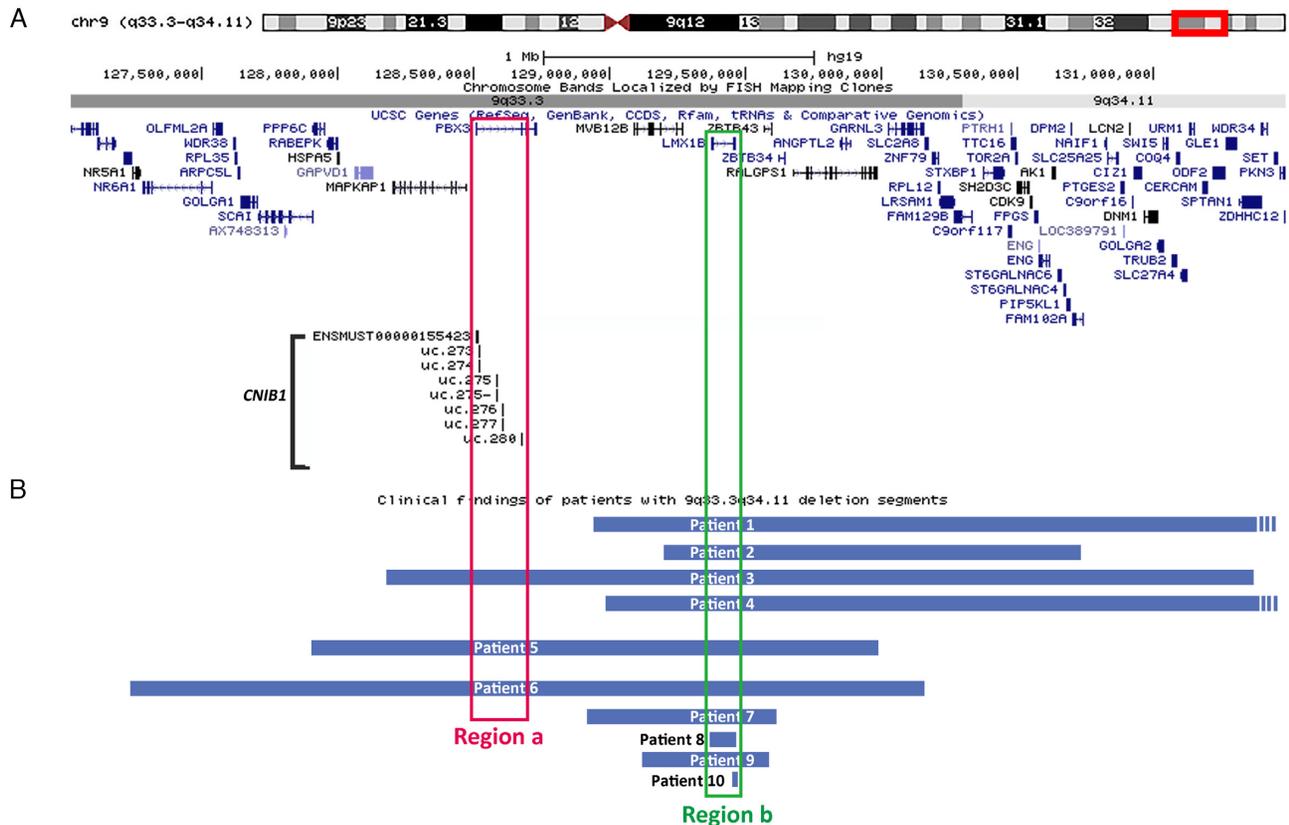
Figure 2. Schematic view of the genetic background of Chr 9q33.3–9q34.11 microdeletion patients. (**A**) Schematic representation of the Chr 9q33.3–9q34.11 region encompassing the *CNIB1* locus (*ENSMUST00000155423, uc.273+, uc.275+, uc.275-, uc.276+, uc.277+,* and *uc.280+*), *pbx3,* and *lmx1b.* (**B**) Schematic view of the genetic background of ten Chr 9q33.3–9q34.11 microdeletion patients as reported in published articles (see Supplementary Table S6). Information of patients 1-4 was adopted from the published article (85).

was less than 20% (Figure 4A and B). Second, the body lengths of larval at 72 hpf (Figure 4C and D) and adult fish at 4 months of age (Figure 4E and F) were significantly shorter in offspring from $CNIB1^{+/-}$ 1 or $CNIB1^{+/-}$ 2 incross. Similar results were observed in the F1 generation (Supplementary Figure S7A and B). Third, the locomotor activity assay showed that the average daytime and nighttime speeds of offspring from $CNIB1^{+/-}$ 1 or $CNIB1^{+/-}$ 2 incross were significantly decreased (Figure 4G–J). Finally, offspring from $CNIB1^{+/-}$ 1 or $CNIB1^{+/-}$ 2 incross exhibited smaller eyes (Figure 5A and B) and impaired eye velocity based on the decreased gain value in the optokinetic response (OKR) assay (66) (Figure 5C and D). To further study the effect of *CNIB1* on eye development and function, we tested whether the structure of the visual system was affected in the *CNIB1* knockout fish. By performing DAPI staining on the eyes, our results showed that there was no gross morphological difference in the retinas between $CNIB1^{+/-}$ 2 and WT zebrafish (Supplementary Figure S8). These results suggested the defect in $CNIB1^{+/-}$ 2 is not structural. Retinal dopamine, synthesized through the action of tyrosine hydroxylase (*TH*) in dopamine (DA) neurons, plays multiple trophic roles in light adaptation, cell survival, and eye growth (67–69). Th is a widely used marker of DA neurons (70). Therefore, we tested the number of retinal DA neuron and the expression level of *th*. Our results

showed a decreased number of Th-labeled DA neurons in isolated $CNIB1^{+/-}$ fish retinas (5 dpf) (Figure 5E and F) and a significant reduction of *th* mRNA level in $CNIB1^{+/-}$ 2 zebrafish (Figure 5G) when compared to their WT siblings. Taken together, our data indicated that the *CNIB1* locus was essential for the proper development, locomotor activity, and visual function in zebrafish.

CNIB1 intronic lncRNAs mediate visual function through *lmx1bb* in zebrafish

In order to determine if the *CNIB1* lncRNAs contributed to the visual defects observed in offspring generated from heterozygous incross, we microinjected a pool of *in vitro* transcripts of the 5 intronic *CNIB1* lncRNAs (*uc.273+, uc.275+, uc.275-, uc.276+,* and *uc.277+*) into the embryos of either WT or $CNIB1^{+/-}$ 2 incross. Beside the sense *CNIB1* lncRNAs (*CNIB1*_S) injections, several parallel injections were performed to the same batch of fish siblings serving as negative controls, including sham operation without RNA (Control-Injected), and a pool of antisense transcripts from single stranded *uc.273+, uc.276+,* and *uc.277+* (*CNIB1*_AS). Our RT-qPCR assay showed that all RNAs were overexpressed in 48 hpf embryos (Supplementary Figure S9A, B, C, and D). Further investigation showed that, compared to the 'Control-Injected' or '*CNIB1*_AS' controls, those F2 fish growing from $CNIB1^{+/-}$ 2 incrossed em-
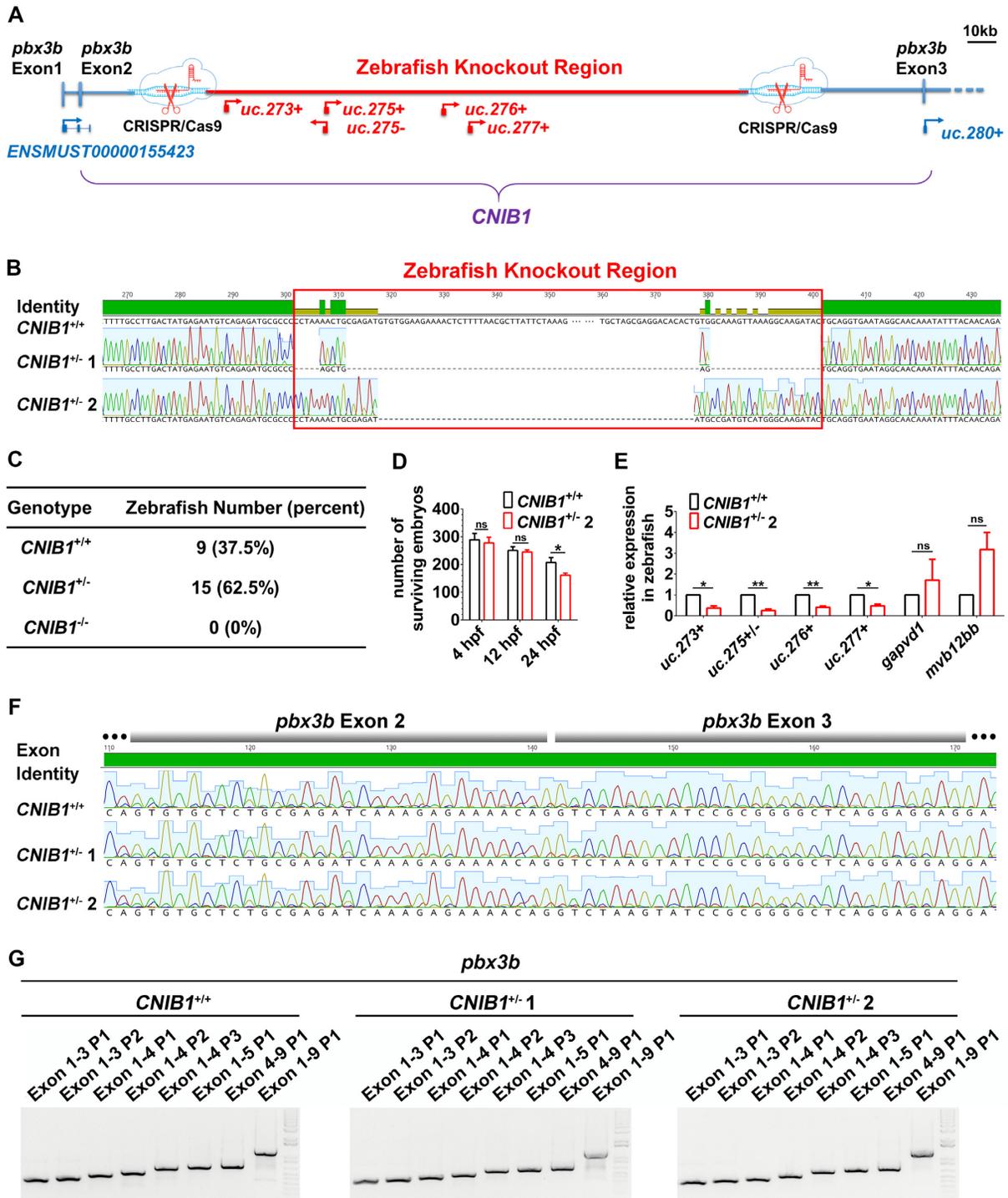
Figure 3. Generation of *CNIB1* knockout by CRISPR-Cas9 in zebrafish. (**A**) Schematic representation of the knockout strategy using CRISPR-Cas9. Scale bar: 10 kb. (**B**) The genotype sequencing results of *CNIB1*$^{+/+}$ and *CNIB1*$^{+/−}$ zebrafish confirmed by fin-DNA genotyping (*n* = 6). (**C**) The number and the percent of *CNIB1* zebrafish genotype (*n* = 24) from offspring of *CNIB1*$^{+/−}$ incross confirmed by embryo-DNA genotyping. (**D**) The average number of surviving embryos in offspring of both *CNIB1*$^{+/+}$ incross (*n* = 289 at 4 hpf, *n* = 250 at 12 hpf, *n* = 207 at 24 hpf) and *CNIB1*$^{+/−}$ 2 zebrafish incross (*n* = 277 at 4 hpf, *n* = 244 at 12 hpf, *n* = 161 at 24 hpf) at 4 hpf, 12 hpf, and 24 hpf, respectively (*n* = 4). (**E**) The expression levels of *CNIB1* lncRNAs and two nearby genes (*gapvd1* and *mvb12bb*) in 30 *CNIB1*$^{+/+}$ and 30 offspring of *CNIB1*$^{+/−}$ incross, normalized to *β-actin*. (**F**) Detection of the exon 2–exon 3 junction of *pbx3b* mRNA by cDNA sequencing in *CNIB1*$^{+/+}$ and *CNIB1*$^{+/−}$ zebrafish confirmed by fin-DNA genotyping (*n* = 6). (**G**) Alternative splicing events of *pbx3b* in *CNIB1*$^{+/+}$ and *CNIB1*$^{+/−}$ zebrafish confirmed by fin-DNA genotyping (*n* = 6).
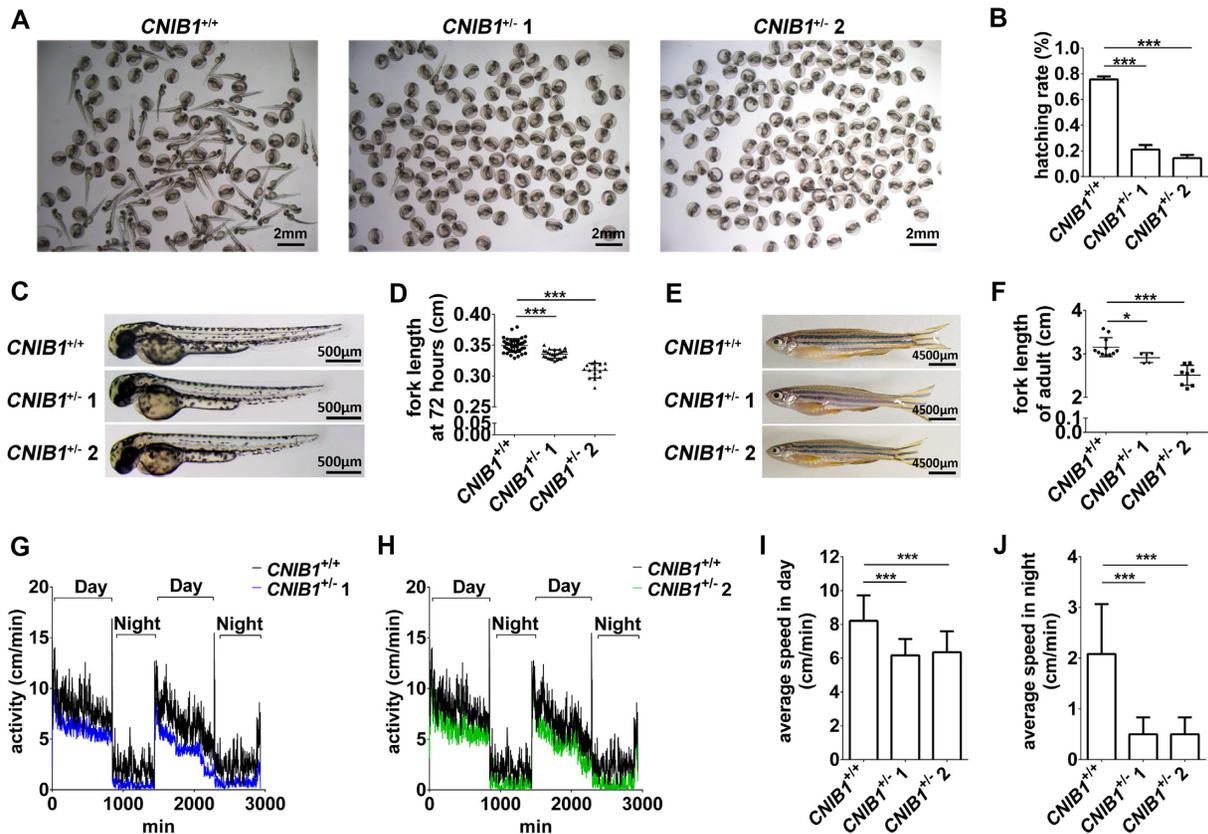
Figure 4. $CNIB1^{+/-}$ zebrafish exhibit developmental retardation and impaired locomotor activity. (**A**) Representative hatching status of the zebrafish offspring at 48 hpf from $CNIB1^{+/+}$ ($n = 130$), $CNIB1^{+/-}$ 1 ($n = 130$) and $CNIB1^{+/-}$ 2 ($n = 130$) incross. Scale bars: 2 mm. (**B**) Schematic representation of A. (**C, D**) Representative images and the schematic representation of the fork lengths of the zebrafish offspring at 5 dpf from $CNIB1^{+/+}$ ($n = 42$), $CNIB1^{+/-}$ 1 ($n = 27$) and $CNIB1^{+/-}$ 2 ($n = 15$) incross. Scale bars: 500 μm. (**E, F**) Representative images and the schematic representation of the fork lengths of the $CNIB1^{+/+}$ ($n = 11$), $CNIB1^{+/-}$ 1 ($n = 8$) and $CNIB1^{+/-}$ 2 ($n = 5$) adult zebrafish at 4 months. Scale bars: 4500 μm. (**G, H**) Activity records of the zebrafish offspring at 5 dpf from $CNIB1^{+/+}$ ($n = 16$), $CNIB1^{+/-}$ 1 ($n = 16$) and $CNIB1^{+/-}$ 2 ($n = 16$) incross at day (**G**) or night (**H**). (**I, J**) The schematic representation of the average locomotion speed gathered from figures G and H, respectively.

bryos that were injected with *CNIB1_S* exhibited reduced mortality rates (Supplementary Figure S10) and exhibited significant improvement in ameliorated gain value assayed by OKR (Figure 5H). Our results indicate that *CNIB1* intronic RNAs play critical roles in ocular function. However, a possible involvement of *CNIB1* locus is not ruled out.

Lmx1b, known as *lmx1bb* or *lmx1b.1* in zebrafish, the proposed causative gene of Chr 9q33.3-9q34.11 microdeletion syndrome (54,55), has also been reported to influence *fgf*-mediated retinal patterning during zebrafish eye development (42). To investigate the potential involvement of *lmx1bb* in *CNIB1*-mediated visual regulation, we first examined whether *CNIB1* regulated the distribution and expression of *lmx1bb* by ISH. Our results showed that the expression levels of *lmx1bb* in the eye and brain at 48 hpf were significantly reduced in $CNIB1^{+/-}$ 2 zebrafish when compared with WT zebrafish (Figure 6A, B). Moreover, this reduction could be completely restored to the normal level after overexpression of *CNIB1* lncRNAs (*CNIB1_S*) (Figure 6B).

It has been reported that both *lmx1bb*-mutant zebrafish (71) and *lmx1bb* morphants (42,72) exhibited eye developmental delay and optical defects. Similar to published results (42,71–76), our results showed that the *lmx1bb* mor-

phants exhibited visual impairment, which could be rescued by co-injection of *lmx1bb* mRNA (Figure 6C–E). Although overexpression of many transcription factors can cause abnormalities in WT embryos, it has been reported that overexpression of *lmx1bb* did not cause obvious developmental or visual abnormalities in frog (77), chicken, and mouse (78). Consistent with reported observations, when we overexpressed *lmx1bb* in WT zebrafish embryos, the fish embryos showed similar developmental processes and normal visual function when compared to the control injected fish (Supplementary Figure S11 and Figure 6H). Furthermore, overexpression of *lmx1bb* in offspring generated from $CNIB1^{+/-}$ 2 incross (Supplementary Figure S12 and Figure 6F and G) significantly promoted visual function compared with offspring generated from $CNIB1^{+/+}$ incross (Figure 6H). *foxc1a*, *eya2* and *pax8*, three known downstream targets of *lmx1b* in mouse (42,72), were also downregulated in $CNIB1^{+/-}$ 2 zebrafish (Supplementary Figure S13). However, only the expression level of *foxc1a* was shown to be reverted in *lmx1bb* overexpressed zebrafish. Taken together, our data suggests that the activity of *CNIB1* is achieved, at least partially, through *CNIB1* intronic lncRNAs-*lmx1bb*-*foxc1a*.
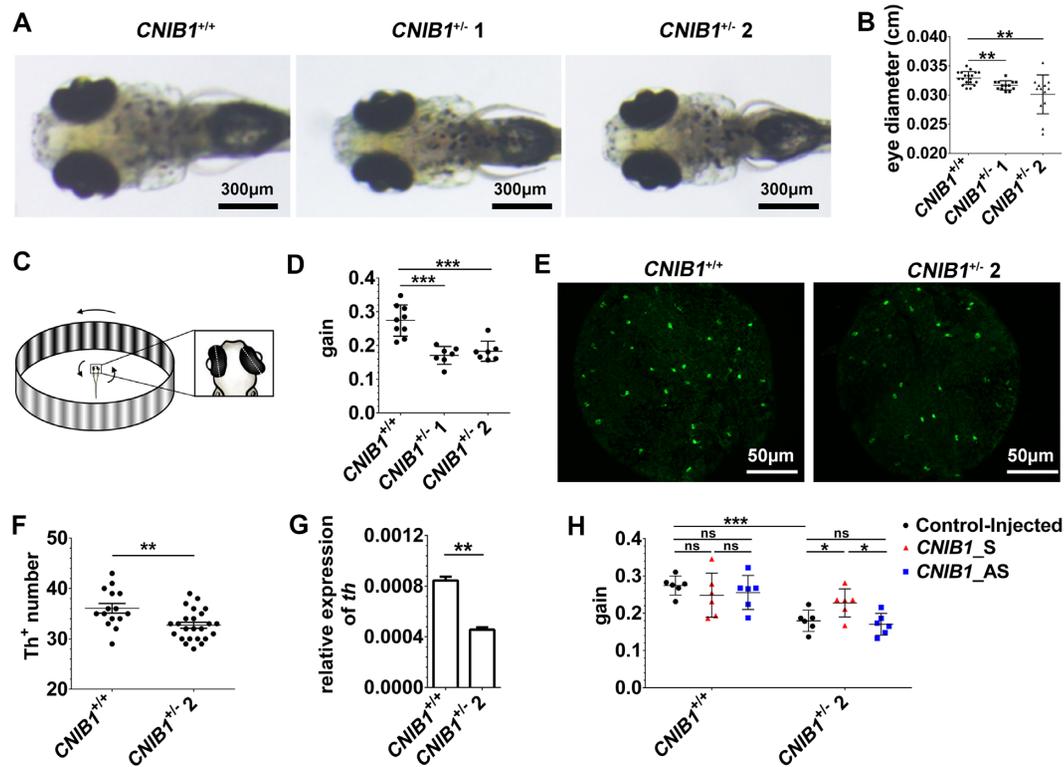
Figure 5. *CNIB1* influences visual function in zebrafish. (**A**) Representative eye images of the zebrafish offspring at 5 dpf from *CNIB1*$^{+/+}$, *CNIB1*$^{+/-}$ 1, and *CNIB1*$^{+/-}$ 2 incross. Scale bars: 300 μm. (**B**) The schematic representation of the eye diameters gathered from figure 5A. *CNIB1*$^{+/+}$, $n = 19$, *CNIB1*$^{+/-}$ 1, $n = 13$, *CNIB1*$^{+/-}$ 2, $n = 13$. (**C**) Schematic representation of the zebrafish larval optokinetic response (OKR) experiment. (**D**) The ratio of eye velocity to the angular velocity of the grating (gain value) assayed by OKR in zebrafish offspring at 5 dpf from *CNIB1*$^{+/+}$ ($n = 9$), *CNIB1*$^{+/-}$ 1 ($n = 7$), and *CNIB1*$^{+/-}$ 2 ($n = 7$) incross. (**E**) Representative immunofluorescence staining images of the Th positive (Th$^+$) DA neurons shown in the retinas of zebrafish offspring at 5 dpf from *CNIB1*$^{+/+}$ and *CNIB1*$^{+/-}$ 2 incross. Scale bars: 50 μm. (**F**) The schematic representation of E. *CNIB1*$^{+/+}$, $n = 15$; *CNIB1*$^{+/-}$ 2, $n = 24$. (**G**) The expression levels of *th* detected by RT-qPCR in zebrafish offspring from *CNIB1*$^{+/+}$ ($n = 30$) and *CNIB1*$^{+/-}$ 2 ($n = 30$) incross, normalized to *β-actin*. (**H**) The effects of *CNIB1* lncRNAs on the OKR assay in the indicated zebrafish offspring. Control-injected *CNIB1*$^{+/+}$, $n = 6$; *CNIB1*_S-injected *CNIB1*$^{+/+}$, $n = 6$; *CNIB1*_AS-injected *CNIB1*$^{+/+}$, $n = 6$; Control-injected *CNIB1*$^{+/-}$ 2, $n = 6$; *CNIB1*_S-injected *CNIB1*$^{+/-}$ 2, $n = 6$; *CNIB1*_AS-injected *CNIB1*$^{+/-}$ 2, $n = 6$.

DISCUSSION

A notable characteristic of lncRNAs is their low-level of conservation across evolution. Because of this outstanding feature for the majority of lncRNAs, lower organism such as *Danio rerio* is not considered a good model system for lncRNA-related human disease studies. Here, we discovered a previously unidentified genomic structure, designated as *CNIBs*, which are formed by highly conserved lncRNAs and appear as gene clusters. We further used *CNIB1* as an example to demonstrate that *Danio rerio* is an applicable model system to investigate the biological consequences of human disease-related lncRNAs which have orthologues in zebrafish.

Many lncRNAs regulate their neighboring or embedded protein encoding genes *in cis* (79). However, the intronic *CNIB1* lncRNAs we focused on, seem to be functionally independent of their host gene *pbx3 (b)*. First, the expression patterns of *pbx3b* and *CNIB1* lncRNAs were different. While *CNIB1* lncRNAs showed strong signals in mouse and zebrafish eyes, *pbx3b* was not found to be localized in zebrafish eyes in this study or the study by Armant *et al.* (80). Second, the phenotypes of *pbx3b* and *CNIB1* KO zebrafish were not similar. Zewdu *et al.* generated *pbx3b*^{scm8} mutant

zebrafish by creating an early stop codon at exon 3 of *pbx3b* (81). They found that *pbx3b*$^{scm8/scm8}$ larvae were viable and normal during development. In contrast, *CNIB1*$^{-/-}$ larvae showed to be embryonic lethal. The *pbx3b*$^{scm8/scm8}$ zebrafish (5 dpf) did not exhibit any obvious eye defects, while the *CNIB1*$^{+/-}$ zebrafish exhibited smaller eyes and visual abnormalities at the same developmental stage. It is worth mentioning that *pbx3b*$^{scm8/scm8}$ fish have intact expression levels of *CNIB1* lncRNAs in theory, but a loss of DNA binding activity which is located at its exon 6. Therefore, the discrepancy in phenotypes between *pbx3b*$^{scm8/scm8}$ and *CNIB1* KO fish strongly indicated that the biological function of *pbx3b*, at least as a transcriptional factor through DNA binding activity, is not associated with *CNIB1*. During the review process for this manuscript, Parenteau *et al.* and Morgan *et al.* showed that intron RNAs can help cultured yeast cells to adapt to environmental changes associated with nutrient depletion. Similarly to that of *CNIB1* lncRNAs, these intron RNAs play different roles from their host protein encoded genes (82,83). The functional RNAs investigated by Parenteau *et al.* and Morgan *et al.* were unspliced pre-mRNAs or excised/debranched introns. In contrast, *CNIB1* intronic RNAs identified in the current study
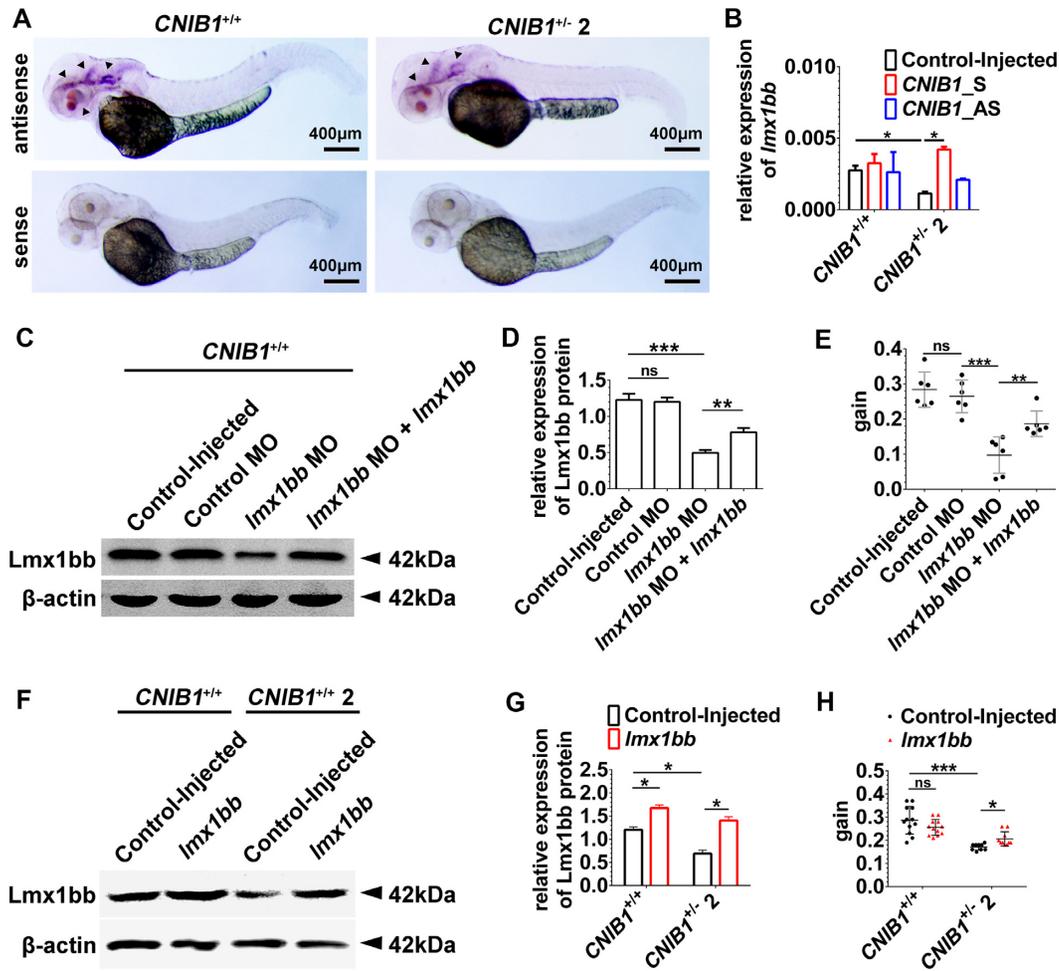
Figure 6. *CNIB1* mediates visual function via *lmx1bb* in zebrafish. (**A**) *In situ hybridization* of *lmx1bb* (indicated by arrowheads) in zebrafish offspring of *CNIB1*$^{+/+}$ and *CNIB1*$^{+/-}$ 2 incross at 48 hpf. Scale bars: 400 μm. (**B**) Relative expression level of *lmx1bb* mRNA detected by RT-qPCR in the indicated zebrafish offspring (*CNIB1*$^{+/+}$ or *CNIB1*$^{+/-}$ 2) that were either Control-injected or injected with *in vitro* transcribed *CNIB1* sense (*CNIB1*_S) or antisense RNAs (*CNIB1*_AS). The data were normalized to *β-actin*. (**C**) Representative images of Western blot for Lmx1bb and β-actin in *CNIB1*$^{+/+}$ zebrafish with the indicated modifications. Control MO, mismatch *lmx1bb* MO; *lmx1bb* MO, *lmx1bb* morphants; *lmx1bb* MO+*lmx1bb*, co-injection of *lmx1bb* morphants and *in vitro* transcribed *lmx1bb* mRNA. (**D**) Quantified expression level of Lmx1bb protein (normalized to β-actin) in the zebrafish shown in C, *n* is 30 in each group. (**E**) Gain value of the OKR assay in the zebrafish shown in C. Control-injected, *n* = 6; Control MO, *n* = 6; *lmx1bb* MO, *n* = 6; *lmx1bb* MO + *lmx1bb*, *n* = 6. (**F**) Representative images of Western blot for Lmx1bb and β-actin in zebrafish offspring from either *CNIB1*$^{+/+}$ or *CNIB1*$^{+/-}$ 2 incross with the indicated modifications. *lmx1bb*, injection with *in vitro* transcribed *lmx1bb* mRNA. (**G**) Quantified expression levels of Lmx1bb protein (normalized to β-actin) in the zebrafish shown in F, *n* = 20. (**H**) Gain value of the OKR assay in the zebrafish shown in F. Control-injected *CNIB1*$^{+/+}$, *n* = 12; *lmx1bb* *CNIB1*$^{+/+}$, *n* = 12; Control-injected *CNIB1*$^{+/-}$ 2, *n* = 9; *lmx1bb* *CNIB1*$^{+/-}$ 2, *n* = 10.

belonged to the lncRNA family. The evolutionary importance of these intron-derived RNAs being independent of their protein gene host requires further investigation.

From zebrafish to human, the *CNIB1* cluster is without exception located on the same chromosome as the *lmx1b* gene. In human, a large genetic deletion encompassing *CNIB1* and *LMX1B* on Chr 9q33.3-9q34.11 has been linked to osteo-onychodysplasia, growth retardation, glomerulopathy, and optical anomalies, which produce NPS (49–53). Genetic studies in both mouse and zebrafish have demonstrated that *Lmx1b* is an essential regulator of eye development and optical function (42,73–76,84), suggesting a conserved role for *Lmx1b* in the optical system. Similar to *CNIB1*$^{-/-}$ zebrafish, *Lmx1b*$^{-/-}$ mice are embryonic lethal. In zebrafish, *lmx1bb* has been reported to pro-

mote the survival of periocular mesenchymal cells by targeting *foxc1a* and *eya2* and to influence *fgf*-mediated retinal patterning (42). We found that the *CNIB1* locus acts as an upstream regulatory signal for *lmx1bb*, DA neuron formation/maintenance, and *foxc1a*. Both heterozygous offspring of *CNIB1*$^{+/-}$ zebrafish and *Lmx1b* mutant animals exhibited developmental delay and abnormal eyes/ocular function (42,72–76,84). In addition, the defective visual activity in *CNIB1*$^{+/-}$ zebrafish could be rescued by providing exogenous *Lmx1bb*. Taken together, our study suggests a conserved role of *CNIB1-Lmx1b* axis during development and an involvement of *CNIB1* in Chr 9q33.3-9q34.11 microdeletion syndrome.

In the present work, we focused on the biological role of *CNIB1*. Quite promisingly, we found evidence linking other

CNIBs to human CNS defects as well. The *CNIBs* identified in this study are unique because of their special genomic structure, location, conservation, and association with human CNS diseases. LncRNA-based gene clusters are likely to be existed in other systems. These clusters may serve as valuable resources to deepen our understanding of lncRNAs during evolution and human diseases.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

FUNDING

REFERENCES

1. Fabre,P.J., Benke,A., Manley,S. and Duboule,D. (2015) Visualizing the HoxD gene cluster at the nanoscale level. *Cold Spring Harb. Symp. Quant. Biol.*, **80**, 9–16.
2. Duboule,D. (2007) The rise and fall of Hox gene clusters. *Development*, **134**, 2549–2560.
3. Ferrier,D.E. and Holland,P.W. (2001) Ancient origin of the Hox gene cluster. *Nat. Rev. Genet.*, **2**, 33–38.
4. Hart,C.P., Awgulewitsch,A., Fainsod,A., McGinnis,W. and Ruddle,F.H. (1985) Homeo box gene complex on mouse chromosome 11: molecular cloning, expression in embryogenesis, and homology to a human homeo box locus. *Cell*, **43**, 9–18.
5. Ullmann,P., Rodriguez,F., Schmitz,M., Meurer,S.K., Qureshi-Baig,K., Felten,P., Ginolhac,A., Antunes,L., Frasquilho,S., Zugel,N. *et al.* (2018) The miR-371 approximately 373 cluster represses colon cancer initiation and metastatic colonization by inhibiting the TGFBR2/ID1 signaling axis. *Cancer Res.*, **78**, 3793–3808.
6. De Lorenzi,M., Rohrer,U. and Birnstiel,M.L. (1986) Analysis of a sea urchin gene cluster coding for the small nuclear U7 RNA, a rare RNA species implicated in the 3′ editing of histone precursor mRNAs. *Proc. Natl. Acad. Sci. U.S.A.*, **83**, 3243–3247.
7. Pane,A., Jiang,P., Zhao,D.Y., Singh,M. and Schupbach,T. (2011) The cutoff protein regulates piRNA cluster expression and piRNA production in the Drosophila germline. *EMBO J.*, **30**, 4601–4615.
8. Inoue,K., Hirose,M., Inoue,H., Hatanaka,Y., Honda,A., Hasegawa,A., Mochida,K. and Ogura,A. (2017) The rodent-specific MicroRNA cluster within the Sfmbt2 gene is imprinted and essential for placental development. *Cell Rep.*, **19**, 949–956.
9. Jha,P., Agrawal,R., Pathak,P., Kumar,A., Purkait,S., Mallik,S., Suri,V., Chand Sharma,M., Gupta,D., Suri,A. *et al.* (2015) Genome-wide small noncoding RNA profiling of pediatric high-grade gliomas reveals deregulation of several miRNAs, identifies downregulation of snoRNA cluster HBII-52 and delineates H3F3A and TP53 mutant-specific miRNAs and snoRNAs. *Int. J. Cancer*, **137**, 2343–2353.
10. Ullmann,P., Rodriguez,F., Schmitz,M., Meurer,S.K., Qureshi-Baig,K., Felten,P., Ginolhac,A., Antunes,L., Frasquilho,S., Zugel,N. *et al.* (2018) The microRNA-371∼373 cluster represses colon cancer initiation and metastatic colonization by inhibiting the TGFBR2/ID1 signaling axis. *Cancer Res.*, **78**, 3793–3808.
11. Sapetschnig,A. and Miska,E.A. (2014) Getting a grip on piRNA cluster transcription. *Cell*, **157**, 1253–1254.
12. Manning,K.E., Tait,R., Suckling,J. and Holland,A.J. (2018) Grey matter volume and cortical structure in Prader-Willi syndrome compared to typically developing young adults. *NeuroImage Clin.*, **17**, 899–909.
13. Angulo,M.A., Butler,M.G. and Cataletto,M.E. (2015) Prader-Willi syndrome: a review of clinical, genetic, and endocrine findings. *J. Endocrinol. Invest.*, **38**, 1249–1263.
14. Gallagher,R.C., Pils,B., Albalwi,M. and Francke,U. (2002) Evidence for the role of PWCR1/HBII-85 C/D box small nucleolar RNAs in Prader-Willi syndrome. *Am. J. Hum. Genet.*, **71**, 669–678.
15. Duker,A.L., Ballif,B.C., Bawle,E.V., Person,R.E., Mahadevan,S., Alliman,S., Thompson,R., Traylor,R., Bejjani,B.A., Shaffer,L.G. *et al.* (2010) Paternally inherited microdeletion at 15q11.2 confirms a significant role for the SNORD116 C/D box snoRNA cluster in Prader-Willi syndrome. *Eur. J. Hum. Genet.: EJHG*, **18**, 1196–1201.
16. Runte,M., Varon,R., Horn,D., Horsthemke,B. and Buiting,K. (2005) Exclusion of the C/D box snoRNA gene cluster HBII-52 from a major role in Prader-Willi syndrome. *Hum. Genet.*, **116**, 228–230.
17. Coulson,R.L., Powell,W.T., Yasui,D.H., Dileep,G., Resnick,J. and LaSalle,J.M. (2018) Prader-Willi locus Snord116 RNA processing requires an active endogenous allele and neuron-specific splicing by Rbfox3/NeuN. *Hum. Mol. Genet.*, **27**, 4051–4060.
18. Bortolin-Cavaille,M.L. and Cavaille,J. (2012) The SNORD115 (H/MBII-52) and SNORD116 (H/MBII-85) gene clusters at the imprinted Prader-Willi locus generate canonical box C/D snoRNAs. *Nucleic Acids Res.*, **40**, 6800–6807.
19. Moretto,F., Wood,N.E., Kelly,G., Doncic,A. and van Werven,F.J. (2018) A regulatory circuit of two lncRNAs and a master regulator directs cell fate in yeast. *Nat. Commun.*, **9**, 780.
20. Wang,X., Arai,S., Song,X., Reichart,D., Du,K., Pascual,G., Tempst,P., Rosenfeld,M.G., Glass,C.K. and Kurokawa,R. (2008) Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. *Nature*, **454**, 126–130.
21. Tomita,S., Abdalla,M.O., Fujiwara,S., Matsumori,H., Maehara,K., Ohkawa,Y., Iwase,H., Saitoh,N. and Nakao,M. (2015) A cluster of noncoding RNAs activates the ESR1 locus during breast cancer adaptation. *Nat. Commun.*, **6**, 6966.
22. Derrien,T., Johnson,R., Bussotti,G., Tanzer,A., Djebali,S., Tilgner,H., Guernec,G., Martin,D., Merkel,A., Knowles,D.G. *et al.* (2012) The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.*, **22**, 1775–1789.
23. Mercer,T.R., Dinger,M.E., Sunkin,S.M., Mehler,M.F. and Mattick,J.S. (2008) Specific expression of long noncoding RNAs in the mouse brain. *Proc. Natl. Acad. Sci. U.S.A.*, **105**, 716–721.
24. Yao,P., Lin,P., Gokoolparsadh,A., Assareh,A., Thang,M.W. and Voineagu,I. (2015) Coexpression networks identify brain region-specific enhancer RNAs in the human brain. *Nat. Neurosci.*, **18**, 1168–1174.
25. Lin,N., Chang,K.Y., Li,Z., Gates,K., Rana,Z.A., Dang,J., Zhang,D., Han,T., Yang,C.S., Cunningham,T.J. *et al.* (2014) An evolutionarily conserved long noncoding RNA TUNA controls pluripotency and neural lineage commitment. *Mol. Cell*, **53**, 1005–1019.
26. Dooley,K. and Zon,L.I. (2000) Zebrafish: a model system for the study of human disease. *Curr. Opin. Genet. Dev.*, **10**, 252–256.

27. Pauli,A., Valen,E., Lin,M.F., Garber,M., Vastenhouw,N.L., Levin,J.Z., Fan,L., Sandelin,A., Rinn,J.L., Regev,A. *et al.* (2012) Systematic identification of long noncoding RNAs expressed during zebrafish embryogenesis. *Genome Res.*, **22**, 577–591.

28. Kaushik,K., Leonard,V.E., Kv,S., Lalwani,M.K., Jalali,S., Patowary,A., Joshi,A., Scaria,V. and Sivasubbu,S. (2013) Dynamic expression of long non-coding RNAs (lncRNAs) in adult zebrafish. *PLoS One*, **8**, e83616.

29. Chen,W., Zhang,X., Li,J., Huang,S., Xiang,S., Hu,X. and Liu,C. (2018) Comprehensive analysis of coding-lncRNA gene co-expression network uncovers conserved functional lncRNAs in zebrafish. *BMC Genomics*, **19**, 112.

30. Kung,J.T., Colognori,D. and Lee,J.T. (2013) Long noncoding RNAs: past, present, and future. *Genetics*, **193**, 651–669.

31. Ulitsky,I., Shkumatava,A., Jan,C.H., Sive,H. and Bartel,D.P. (2011) Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell*, **147**, 1537–1550.

32. Sarangdhar,M.A., Chaubey,D., Bhatt,A., Km,M., Kumar,M., Ranjan,S. and Pillai,B. (2017) A novel long non-coding RNA, durga modulates dendrite density and expression of kalirin in zebrafish. *Front. Mol. Neurosci.*, **10**, 95.

33. Wei,S., Chen,H., Dzakah,E.E., Yu,B., Wang,X., Fu,T., Li,J., Liu,L., Fang,S., Liu,W. *et al.* (2019) Systematic evaluation of C. elegans lincRNAs with CRISPR knockout mutants. *Genome Biol.*, **20**, 7.

34. Goff,L.A., Groff,A.F., Sauvageau,M., Trayes-Gibson,Z., Sanchez-Gomez,D.B., Morse,M., Martin,R.D., Elcavage,L.E., Liapis,S.C., Gonzalez-Celeiro,M. *et al.* (2015) Spatiotemporal expression and transcriptional perturbations by long noncoding RNAs in the mouse brain. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, 6855–6862.

35. Paxinos,G. and Franklin,K.B. (2004) *The Mouse Brain in Stereotaxic Coordinates*. Gulf Professional Publishing.

36. Gupta,T. and Mullins,M.C.(2010) Dissection of organs from the adult zebrafish. *J. Vis. Exp.*, **37**, e1717.

37. Smyth,G.K. (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.*, **3**, doi:10.2202/1544-6115.1027.

38. Li,D., Zhang,J., Wang,M., Li,X., Gong,H., Tang,H., Chen,L., Wan,L. and Liu,Q. (2018) Activity dependent LoNA regulates translation by coordinating rRNA transcription and methylation. *Nat. Commun.*, **9**, 1726.

39. Wilusz,J.E., Freier,S.M. and Spector,D.L. (2008) 3′ end processing of a long nuclear-retained noncoding RNA yields a tRNA-like cytoplasmic RNA. *Cell*, **135**, 919–932.

40. Thisse,C. and Thisse,B. (2008) High-resolution in situ hybridization to whole-mount zebrafish embryos. *Nat. Protoc.*, **3**, 59–69.

41. Ma,J., Koster,J., Qin,Q., Hu,S., Li,W., Chen,C., Cao,Q., Wang,J., Mei,S., Liu,Q. *et al.* (2016) CRISPR-DO for genome-wide CRISPR design and optimization. *Bioinformatics*, **32**, 3336–3338.

42. McMahon,C., Gestri,G., Wilson,S.W. and Link,B.A. (2009) Lmx1b is essential for survival of periocular mesenchymal cells and influences Fgf-mediated retinal patterning in zebrafish. *Dev. Biol.*, **332**, 287–298.

43. Prober,D.A., Rihel,J., Onah,A.A., Sung,R.J. and Schier,A.F. (2006) Hypocretin/orexin overexpression induces an insomnia-like phenotype in zebrafish. *J. Neurosci.*, **26**, 13400–13410.

44. Appelbaum,L., Wang,G., Yokogawa,T., Skariah,G.M., Smith,S.J., Mourrain,P. and Mignot,E. (2010) Circadian and homeostatic regulation of structural synaptic plasticity in hypocretin neurons. *Neuron.*, **68**, 87–98.

45. Mueller,K.P. and Neuhauss,S.C. (2010) Quantitative measurements of the optokinetic response in adult fish. *J. Neurosci. Methods*, **186**, 29–34.

46. Nie,K., Wang,K., Huang,D.F., Huang,Y.B., Yin,W., Ren,D.L., Wang,H. and Hu,B. (2018) Effects of circadian clock protein Per1b on zebrafish visual functions. *Chronobiol. Int.*, **35**, 160–168.

47. Li,J., Wang,C., Zhang,Z., Wen,Y., An,L., Liang,Q., Xu,Z., Wei,S., Li,W., Guo,T. *et al.* (2018) Transcription factors Sp8 and Sp9 coordinately regulate olfactory bulb interneuron development. *Cereb. Cortex*, **28**, 3278–3294.

48. Bejerano,G., Pheasant,M., Makunin,I., Stephen,S., Kent,W.J., Mattick,J.S. and Haussler,D. (2004) Ultraconserved elements in the human genome. *Science*, **304**, 1321–1325.

49. Nambot,S., Masurel,A., El Chehadeh,S., Mosca-Boidron,A.L., Thauvin-Robinet,C., Lefebvre,M., Marle,N., Thevenon,J., Perez-Martin,S., Dulieu,V. *et al.* (2016) 9q33.3q34.11 microdeletion: new contiguous gene syndrome encompassing STXBP1, LMX1B and ENG genes assessed using reverse phenotyping. *Eur. J. Hum. Genet.*, **24**, 830–837.

50. Marini,M., Bocciardi,R., Gimelli,S., Di Duca,M., Divizia,M.T., Baban,A., Gaspar,H., Mammi,I., Garavelli,L., Cerone,R. *et al.* (2010) A spectrum of LMX1B mutations in Nail-Patella syndrome: new point mutations, deletion, and evidence of mosaicism in unaffected parents. *Genet. Med.*, **12**, 431–439.

51. Schlaubitz,S., Yatsenko,S.A., Smith,L.D., Keller,K.L., Vissers,L.E., Scott,D.A., Cai,W.W., Reardon,W., Abdul-Rahman,O.A., Lammer,E.J. *et al.* (2007) Ovotestes and XY sex reversal in a female with an interstitial 9q33.3-q34.1 deletion encompassing NR5A1 and LMX1B causing features of Genitopatellar syndrome. *Am. J. Med. Genet. A*, **143a**, 1071–1081.

52. Bongers,E.M., de Wijs,I.J., Marcelis,C., Hoefsloot,L.H. and Knoers,N.V. (2008) Identification of entire LMX1B gene deletions in nail patella syndrome: evidence for haploinsufficiency as the main pathogenic mechanism underlying dominant inheritance in man. *Eur. J. Hum. Genet.*, **16**, 1240–1244.

53. Jiang,S., Zhang,J., Huang,D., Zhang,Y., Liu,X., Wang,Y., He,R. and Zhao,Y. (2014) A microdeletion of chromosome 9q33.3 encompasses the entire LMX1B gene in a Chinese family with nail patella syndrome. *Int. J. Mol. Sci.*, **15**, 20158–20168.

54. Negrisolo,S., Carraro,A., Fregonese,G., Benetti,E., Schaefer,F., Alberti,M., Melchionda,S., Fischetto,R., Giordano,M. and Murer,L. (2018) Could the interaction between LMX1B and PAX2 influence the severity of renal symptoms? *Eur. J. Hum. Genet.: EJHG*, **26**, 1708–1712

55. Knoers,N.V., Bongers,E.M., van Beersum,S.E., Lommen,E.J., van Bokhoven,H. and Hol,F.A. (2000) Nail-patella syndrome: identification of mutations in the LMX1B gene in Dutch families. *J. Am. Soc. Nephrol.: JASN*, **11**, 1762–1766.

56. Ehret,J.K., Engels,H., Cremer,K., Becker,J., Zimmermann,J.P., Wohlleber,E., Grasshoff,U., Rossier,E., Bonin,M., Mangold,E. *et al.* (2015) Microdeletions in 9q33.3-q34.11 in five patients with intellectual disability, microcephaly, and seizures of incomplete penetrance: is STXBP1 not the only causative gene? *Mol. Cytogenet.*, **8**, 72.

57. Shimojima,K., Ondo,Y., Okamoto,N. and Yamamoto,T. (2017) A 15q14 microdeletion involving MEIS2 identified in a patient with autism spectrum disorder. *Hum. Genome Var.*, **4**, 17029.

58. Brunetti-Pierri,N., Sahoo,T., Frioux,S., Chinault,C., Zascavage,R., Cheung,S.W., Peters,S. and Shinawi,M. (2008) 15q13q14 deletions: phenotypic characterization and molecular delineation by comparative genomic hybridization. *Am. J. Med. Genet. A*, **146a**, 1933–1941.

59. Crowley,M.A., Conlin,L.K., Zackai,E.H., Deardorff,M.A., Thiel,B.D. and Spinner,N.B. (2010) Further evidence for the possible role of MEIS2 in the development of cleft palate and cardiac septum. *Am. J. Med. Genet. A*, **152a**, 1326–1327.

60. Roberti,M.C., Surace,C., Digilio,M.C., D'Elia,G., Sirleto,P., Capolino,R., Lombardo,A., Tomaiuolo,A.C., Petrocchi,S. and Angioni,A. (2011) Complex chromosome rearrangements related 15q14 microdeletion plays a relevant role in phenotype expression and delineates a novel recurrent syndrome. *Orphanet. J. Rare Dis.*, **6**, 17.

61. Johansson,S., Berland,S., Gradek,G.A., Bongers,E., de Leeuw,N., Pfundt,R., Fannemel,M., Rodningen,O., Brendehaug,A., Haukanes,B.I. *et al.* (2014) Haploinsufficiency of MEIS2 is associated with orofacial clefting and learning disability. *Am. J. Med. Genet. A*, **164a**, 1622–1626.

62. Chen,C.P., Chen,C.Y., Chern,S.R., Wu,P.S., Chen,Y.N., Chen,S.W., Chen,L.F., Yang,C.W. and Wang,W. (2016) Prenatal diagnosis and molecular cytogenetic characterization of a de novo 4.858-Mb microdeletion in 15q14 associated with ACTC1 and MEIS2 haploinsufficiency and tetralogy of Fallot. *Taiwan J. Obstet. Gynecol.*, **55**, 270–274.

63. Paulussen,A.D., Schrander-Stumpel,C.T., Tserpelis,D.C., Spee,M.K., Stegmann,A.P., Mancini,G.M., Brooks,A.S., Collee,M., Maat-Kievit,A., Simon,M.E. *et al.* (2010) The unfolding clinical

spectrum of holoprosencephaly due to mutations in SHH, ZIC2, SIX3 and TGIF genes. *Eur. J. Hum. Genet.*, **18**, 999–1005.

64. Jeong,J., Li,X., McEvilly,R.J., Rosenfeld,M.G., Lufkin,T. and Rubenstein,J.L. (2008) Dlx genes pattern mammalian jaw primordium by regulating both lower jaw-specific and upper jaw-specific genetic programs. *Development*, **135**, 2905–2916.
65. Cajigas,I., Leib,D.E., Cochrane,J., Luo,H., Swyter,K.R., Chen,S., Clark,B.S., Thompson,J., Yates,J.R. 3rd, Kingston,R.E. *et al.* (2015) Evf2 lncRNA/BRG1/DLX1 interactions reveal RNA-dependent inhibition of chromatin remodeling. *Development*, **142**, 2641–2652.
66. Zou,S., Yin,W., Zhang,M., Hu,C., Huang,Y. and Hu,B. (2010) Using the optokinetic response to study visual function of zebrafish. *J. Vis. Exp.*, **36**, e1742.
67. Witkovsky,P., Veisenberger,E., Haycock,J.W., Akopian,A., Garcia-Espana,A. and Meller,E. (2004) Activity-dependent phosphorylation of tyrosine hydroxylase in dopaminergic neurons of the rat retina. *J. Neurosci.*, **24**, 4242–4249.
68. Harris,J.P., Calvert,J.E., Leendertz,J.A. and Phillipson,O.T. (1990) The influence of dopamine on spatial vision. *Eye*, **4**, 806–812.
69. Li,L. and Dowling,J.E. (2000) Effects of dopamine depletion on visual sensitivity of zebrafish. *J. Neurosci.*, **20**, 1893–1903.
70. Matsushita,N., Okada,H., Yasoshima,Y., Takahashi,K., Kiuchi,K. and Kobayashi,K. (2002) Dynamics of tyrosine hydroxylase promoter activity during midbrain dopaminergic neuron development. *J. Neurochem.*, **82**, 295–304.
71. Schibler,A. and Malicki,J. (2007) A screen for genetic defects of the zebrafish ear. *Mech. Dev.*, **124**, 592–604.
72. O'Hara,F.P., Beck,E., Barr,L.K., Wong,L.L., Kessler,D.S. and Riddle,R.D. (2005) Zebrafish Lmx1b.1 and Lmx1b.2 are required for maintenance of the isthmic organizer. *Development*, **132**, 3163–3173.
73. Pressman,C.L., Chen,H. and Johnson,R.L. (2000) LMX1B, a LIM homeodomain class transcription factor, is necessary for normal development of multiple tissues in the anterior segment of the murine eye. *Genesis*, **26**, 15–25.
74. Thaung,C., West,K., Clark,B.J., McKie,L., Morgan,J.E., Arnold,K., Nolan,P.M., Peters,J., Hunter,A.J., Brown,S.D. *et al.* (2002) Novel ENU-induced eye mutations in the mouse: models for human eye disease. *Hum. Mol. Genet.*, **11**, 755–767.
75. Cross,S.H., Macalinao,D.G., McKie,L., Rose,L., Kearney,A.L., Rainger,J., Thaung,C., Keighren,M., Jadeja,S., West,K. *et al.* (2014) A dominant-negative mutation of mouse Lmx1b causes glaucoma and is semi-lethal via LDB1-mediated dimerization [corrected]. *PLos Genet.*, **10**, e1004359.
76. Liu,P. and Johnson,R.L. (2010) Lmx1b is required for murine trabecular meshwork formation and for maintenance of corneal transparency. *Dev. Dyn.*, **239**, 2161–2171.
77. Haldin,C.E., Masse,K.L., Bhamra,S., Simrick,S., Kyuno,J. and Jones,E.A. (2008) The lmx1b gene is pivotal in glomus development in Xenopus laevis. *Dev. Biol.*, **322**, 74–85.
78. Chizhikov,V.V. and Millen,K.J. (2004) Control of roof plate development and signaling by Lmx1b in the caudal vertebrate CNS. *J. Neurosci.*, **24**, 5694–5703.
79. Luo,S., Lu,J.Y., Liu,L., Yin,Y., Chen,C., Han,X., Wu,B., Xu,R., Liu,W., Yan,P. *et al.* (2016) Divergent lncRNAs Regulate Gene Expression and Lineage Differentiation in Pluripotent Cells. *Cell Stem. Cell*, **18**, 637–652.
80. Armant,O., Marz,M., Schmidt,R., Ferg,M., Diotel,N., Ertzer,R., Bryne,J.C., Yang,L., Baader,I., Reischl,M. *et al.* (2013) Genome-wide, whole mount in situ analysis of transcriptional regulators in zebrafish embryos. *Dev. Biol.*, **380**, 351–362.
81. Zewdu,R., Risolino,M., Barbulescu,A., Ramalingam,P., Butler,J.M. and Selleri,L. (2016) Spleen hypoplasia leads to abnormal stress hematopoiesis in mice with loss of Pbx homeoproteins in splenic mesenchyme. *J. Anat.*, **229**, 153–169.
82. Morgan,J.T., Fink,G.R. and Bartel,D.P. (2019) Excised linear introns regulate growth in yeast. *Nature*, **565**, 606–611.
83. Parenteau,J., Maignon,L., Berthoumieux,M., Catala,M., Gagnon,V. and Abou Elela,S. (2019) Introns are mediators of cell response to starvation. *Nature*, **565**, 612–617.
84. Gu,W.X. and Kania,A. (2010) Identification of genes controlled by LMX1B in E13.5 mouse limbs. *Dev. Dyn.*, **239**, 2246–2255.
85. Nambot,S., Masurel,A., Chehadeh,S.E., Mosca-Boidron,A.-L., Thauvin-Robinet,C., Lefebvre,M., Marle,N., Thevenon,J., Perez-Martin,S., Dulieu,V. *et al.* (2016) 9q33.3q34.11 microdeletion: new contiguous gene syndrome encompassing *STXBP1*, *LMX1B* and *ENG* genes assessed using reverse phenotyping. *Eur. J. Hum. Genet.*, **24**, 830–837.