Induction and dorsal restriction of *Paired-box 3* (*Pax3)* gene expression in the caudal neuroectoderm is mediated by integration of multiple pathways on a short neural crest enhancer

Oraly Sanchez-Ferras, Guillaume Bernas, Emilie Laberge-Perrault and Nicolas Pilon.

Molecular Genetics of Development Laboratory, Department of Biological Sciences and BioMed Research Center, Faculty of Sciences, University of Quebec at Montreal (UQAM)

Running title: Regulation of *Pax3* neural expression by a Cdx-Zic2 complex

To whom correspondence should be addressed: Nicolas Pilon, Department of Biological Sciences and BioMed Research Center, Faculty of Sciences, University of Quebec at Montreal (UQAM), 141 President-Kennedy Ave, Montreal, PQ, Canada, H2X 3Y7; Tel.: 514-987-3000 x3342; Fax: 514-987-4647; Email: pilon.nicolas@uqam.ca

ABSTRACT:

Pax3 encodes a paired-box transcription factor with key roles in neural crest and neural tube ontogenesis. Robust control of *Pax3* neural expression is ensured by two redundant sets of cis-regulatory modules (CRMs) that integrate anterior-posterior (such as Wnt-βCatenin signaling) as well as dorsal-ventral (such as Shh-Gli signaling) instructive cues. In previous work, we sought to characterize the Wnt-mediated regulation of *Pax3* expression and identified the Cdx transcription factors (Cdx1/2/4) as critical intermediates in this process. We identified the neural crest enhancer-2 (NCE2) from the 5'-flanking region of *Pax3* as a Cdx-dependent CRM that recapitulates the restricted expression of *Pax3* in the mouse caudal neuroectoderm. While this is consistent with a key role in relaying the inductive signal from posteriorizing Wnt ligands, the broad expression of Cdx proteins in the tailbud region is not consistent with the restricted activity of NCE2. This implies that other positive and/or negative inputs are required and, here, we report a novel role for the transcription factor Zic2 in this regulation. Our data strongly suggests that Zic2 is involved in the induction (as a direct *Pax3*NCE2 activator and Cdx neural cofactor) as well as the maintenance of *Pax3* dorsal restriction (as a target of the ventral Shh repressive input). We also provide evidence that the inductive Cdx-Zic2 interaction is integrated on NCE2 with a positive input from the neural-specific transcription factor Sox2. Altogether, our data provide important mechanistic insights into the coordinated integration of different signaling pathways on a short *Pax3* CRM.

KEY WORDS: Cdx, Zic, Sox2, Pax3, neural crest cells, neural tube.

1-INTRODUCTION

Building an embryo from an initial population of equivalent cells requires precise spatiotemporal control of gene expression. Information to do this comes from just a few numbers of conserved signaling pathways, is transmitted by DNA binding proteins and interpreted at the cis-regulatory level on evolutionarily conserved genomic sequences. Redundant and different operating enhancers may exist to refine and protect expression of developmental genes from fluctuations in these signals or mutations in the genome. Subsequently, the output of gene expression and gene regulatory interactions provide the memory to maintain established expression patterns in the absence of signaling inputs. In this regard, a lot of work has been done to understand the molecular mechanisms of neural gene expression during establishment of the anterior-posterior (AP) as well as the dorsal-ventral (DV) axes. Nevertheless, how both AP and DV signaling inputs are coordinately integrated at the cis-regulatory level is still poorly understood.

Pax3/7 (Paired box 3 and 7) and *Zic* (Zinc finger protein of the cerebellum) family members (*Zic1-5*) encode transcription factors that exhibit overlapping expression domains in the neuroectoderm along both the AP and the DV axis. During neurulation, expression of these genes is similarly restricted to pre-migratory neural crest cells (NCC) and dorsal neural tube (NT) [1-9]. However, such overlap is less extensive along the AP axis and most especially in the caudal embryo where only *Pax3* and the *Zic2/Zic5* gene pair are expressed in the posterior neural plate (PNP) [1, 10]. Consistent with their wide expression pattern and key developmental role, loss-of-function mutations of *Pax3* and *Zic2* – as seen for example in the Splotch (*Pax3*) and Kumba (*Zic2*) mouse mutants – causes severe and similar NT and NCC defects affecting the entire AP axis such as spina bifida, craniofacial malformations, absence of dorsal root ganglia and pigmentary anomalies [6, 7, 10-12].

Work performed in several vertebrate species has revealed that AP instructive cues from Wnt (Wingless and Int-1 related) and FGF (fibroblast growth factor) pathways as well as DV instructive cues from BMP (bone morphogenetic protein) and Shh (Sonic Hedgehog) pathways are all involved in the induction and dorsal restriction of *Pax3/7* and *Zic* members [8, 13-22]. These studies notably point to a critical role for posteriorizing canonical Wnt signaling and intermediate levels of BMP molecules during induction in the neural plate [14, 16-18, 23, 24] whereas opposing gradients of dorsal BMP and ventral Shh signaling are subsequently implicated in the maintenance and dorsal restriction in the closed NT [19-21, 25]. Although the general role of these pathways is well accepted, some species-specific variations are also expected regarding their relative importance. In the case of *Pax3*, this is well exemplified by the comparison of its posterior expression domain between chick and mouse embryos. Indeed, in chick embryos, *Pax3* expression is initially induced in the whole PNP before becoming restricted to the dorsal NT whereas in mouse embryos, *Pax3* expression is already restricted to the lateral borders of the PNP during the induction phase [15, 26].

Multiple evolutionary conserved cis-regulatory modules (CRMs) have been identified for *Pax3*. These CRMs are clustered in two areas of the *Pax3* locus: one in the 5'-flanking region and the other in intron-4 [27-29]. The 5'-flanking region, named Neural Crest Enhancer (NCE), is located within the

1.6 Kb proximal promoter and is subdivided in two short CRMs of approximately 250 bp named NCE1 and NCE2 [27, 28]. The entire NCE is not only able to direct *Pax3* reporter expression in mouse NCC and dorsal NT along the hindbrain and trunk region, but also drive enough functional expression levels of *Pax3* to rescue the NT and NCC defects observed in *Pax3* Splotch mice [7]. Interestingly, targeted deletion studies in the mouse have suggested that NCE acts redundantly with a second evolutionary conserved region (ECR2) located in the 4th intron [29]. In fact, more recent work using the zebrafish as a model has demonstrated that the *Pax3* intron-4 contains at least two CRMs that appears to exhibit complementary activities in order to recapitulate the induction, dorsal restriction and maintenance of *Pax3* neural expression [18, 30]. Given that the effect of their deletion has not been documented so far, the requirement of any of these intron-4 CRMs for *Pax3* expression as well as their relative importance over the 5'-flanking NCE is currently unknown.

We have previously demonstrated that the NCE2 region alone is able to recapitulate both the induction and dorsal restriction of *Pax3* expression in the caudal NCC and NT [26], suggesting that this CRM is well suited for analyzing the coordinated integration of both AP and DV instructive cues. In this regard, we have already demonstrated that activity of this enhancer depends on a positive input from caudal-related homeobox (Cdx) transcription factors downstream of Wnt/βCatenin signaling [26]. Here we further show that, in addition to Cdx, Zic2 also directly regulates murine *Pax3* expression and acts as a Cdx neural cofactor. Importantly; we show that the NCE2 region integrates positive inputs from caudal Cdx, dorsal Zic2 as well as neural Sox2 transcription factors. Furthermore, we provide evidence for a putative role of Zic2/5 as mediators of the Shh-induced repressive input involved in the dorsal restriction of *Pax3* expression. Taken together with previous descriptions of other functional binding sites (e.g. Brn1 and Tead2) within NCE2 [28, 31], our data strongly suggest that this short CRM behaves as a "super-enhancer" [32] that mediates the spatiotemporal induction and dorsal restriction of *Pax3* expression in the mouse caudal neuroectoderm.

2-MATERIALS and METHODS

2.1-Ethics Statement

Experiments involving mice were performed following Canadian Council of Animal Care (CCAC) guidelines for the care and manipulation of animals used in medical research. Protocols involving the manipulation of animals were approved by the institutional ethics committee of the University of Quebec at Montreal (comité institutionnel de protection des animaux (CIPA)); Reference number 0513-C1-648-0514).

2.2-Plasmid constructs and site-directed mutagenesis

The *Pax3* cDNA vector pBH3.2 [1] was kindly provided by J. Epstein. Expression vectors for FLAG-tagged Cdx1 and GST-Cdx1 fusion proteins have been described previously [33]. Expression vectors for FLAG-tagged full-length and deletion mutant ZIC2 proteins were a generous gift from S. Tejpar [34]. HA-tagged Zic2 expression vector (pcDNA3-$_{HA}$Zic2) was kindly provided by J. Aruga [35]. Myc-tagged Sox2 expression vector (pcDNA3.1-$_{Myc}$Sox2) was a gift from M. Bani-Yaghoub [36]. $_{FLAG}$Cdx1 and $_{HA}$Zic2 expression vectors co-expressing GFP were generated by subcloning the respective cDNA

into the pIRES2-EGFP vector (Clontech). Wild type (wt) or CdxBS mutant versions of a *Pax3*NCE2-luciferase reporter constructs were as previously described [26]. For generating Zic binding site (BS) mutant versions, point mutations were introduced into the ZicBS of NCE2 by using the Quick Change Multisite-directed mutagenesis kit (Stratagene) in accordance with manufacturer's instructions. Sequence of the oligonucleotide used for site-directed mutagenesis of the ZicBS was: 5'-CGTCATATCCTGCTaaGGACACTTCAGCTCCTAGCCAAGA-3' (with ZicBS underlined and point mutations indicated by lower case letters).

2.3-In situ hybridization and immunofluorescence analyses

FVB mouse embryos were obtained from timed pregnancies, noon of the day on which a vaginal plug was detected being considered as embryonic day (e)0.5. Embryos to be compared were stage-matched in accordance with established criteria [37] and processed in parallel.

Whole-mount *in situ* hybridization of *Pax3* and *Zic2* mRNAs was performed using standard approach [38]. The *Pax3* probe was generated from the pBH3.2 plasmid while the *Zic2* probe was generated from the pcDNA3-$_{HA}$Zic2 plasmid. Transverse sections (100μm) were prepared using a vibrating blade microtome Microm HM 650V (Thermo Scientific) as previously described [39]. Images were acquired with a Leica DFC 495 camera mounted on a Leica M205 FA microscope (Leica Microsystems).

Whole mount immunostaining of Nkx6.1 in e9.5 mouse embryos was performed using the protocol described by [39] and consecutive vibratome transverse sections of the tail were prepared as described above. Immunofluorescence analyses on NT cryosections were performed as previously described [40]. Briefly, 20 μm frozen sections were prepared using a Leica CM1950 cryostat (Leica microsystems) and adjacent sections were used to compare the distribution of Cdx2 and Pax3 proteins. Sections were blocked with 10% fetal bovine serum, 0.1% TritonX-100 in PBS for 1 hour and then incubated at 4°C overnight with primary antibodies diluted in blocking solution. Nuclei were stained using DAPI (Molecular Probes) and slides were then incubated 1 hour at room temperature with secondary antibodies diluted in blocking solution. Antibodies and dilutions used were: mouse anti-Cdx2 1: 100 (Biogenex), mouse anti-Pax3 1:400 (R&D Systems), mouse IgG1 anti-NKX6.1 1:20 (Developmental Studies Hybridoma Bank) and Alexa-647 donkey anti-mouse 1:500 (Jackson Immunoresearch). Images were taken with a Nikon A1 laser-scanning confocal microscope.

2.4-Chromatin immunoprecipitation (ChIP) assays

ChIP assays in Neuro2a (N2a) cells were performed using the M-Fast Chromatin immunoprecipitation kit (ZmTech Scientific) in accordance with manufacturer's instructions. For immunoprecipitation of chromatin from $_{HA}$Zic2- or $_{FLAG}$Cdx1-transfected N2a cells (10^6 cells), 1μg of rabbit anti-HA (Abcam) and mouse anti-FLAG M2 (Sigma) antibodies were respectively used. A normal rabbit IgG serum was used as a negative control of immunoprecipitation. PCR amplifications were performed using the Platinum Taq DNA polymerase (Life Technologies) and consisted of 35 cycles of 30s at 96°C, 30s at 60°C and 30s at 72°C. The primers used for this study were: *Pax3* NCE2 Forward (5'-GGCACAATGGTACCTTCTCTAAGG-3') and Reverse (5'-CCCTTCTGAGAAGCGGGGACTTTAAA-3'). PCR products were resolved on a 2 % agarose gel and sequence-confirmed.

2.5-Electrophoretic mobility shift assays (EMSA)

Zic2 and Cdx1 binding to *Pax3*NCE2 sequences was assessed essentially as previously described [26]. Briefly, binding reactions were carried out in a previously described Zic binding buffer [41] with 8µg of nuclear extracts from mock-, $_{HA}$Zic2 or $_{FLAG}$Cdx1 transfected Cos7 cells. Supershift assay was performed using 1µg of rabbit anti-HA antibody (Abcam). Specificity of binding was assessed by competition with a 100-fold molar excess of unlabeled wild type or mutated cold probe. To test the interaction of Cdx1 and Zic2 in the presence of the target DNA, $_{HA}$Zic2 and $_{FLAG}$Cdx1 nuclear extracts were incubated for 30 min in Zic binding buffer at 4 ºC before the addition of the radiolabelled probe (either *Pax3*NCE2 CdxBS1, CdxBS3 or ZicBS). The sense strand of each wild type double-stranded probe was previously described [26]. The sense strand of a double-stranded probe harbouring the mutated *Pax3*NCE2 ZicBS was as described above for site-directed mutagenesis. The sense strand of a double stranded probe harbouring a previously described ZicBS [41] (used as positive control) was: 5'-GATCCTGTGATTTTCGTCTTGGGTGGTCTCCCTCG-3' (with ZicBS underlined).

2.6-Western blot, co-immunoprecipitation and GST pull-down assays

Primary antibodies and dilutions used for western blotting were: mouse anti-FLAG 1:1000 (M2, Sigma), rabbit anti-HA 1:2000 (ab9110, Abcam), mouse anti-Myc 1:100 (in house 9E10 hybridoma) and rabbit anti-GAPDH 1:2500 (sc25778, Santa Cruz Biotech).

For co-immunoprecipitation, Cos7 cells were transfected with expression vectors for $_{FLAG}$Cdx1 and $_{HA}$Zic2 using Genejuice transfection reagent (Novagen) and harvested in immunoprecipitation buffer (25 mm Hepes, pH 7.2, 0.5% Nonidet P-40, 150 mm NaCl, supplemented with protease inhibitors). Immunoprecipitations and western blots were carried out as previously described [33].

For pull-down assays, Glutathione *S*-transferase (GST)-Cdx1, GST-Cdx1 homeodomain (GST-Cdx1Homeo) and GST-Cdx1 N-terminus (GST-Cdx1Nterm) fusion proteins were produced and purified with glutathione-agarose beads (Sigma) as previously described [33]. FLAG-tagged full-length and deletion mutant ZIC2 proteins were *in vitro* synthesized using the T_NT T7 quick coupled transcription/translation system (Promega) and previously described plasmids [34]. *In vitro* pull-down assays were performed as previously described [33] and interactions were revealed by western blotting using mouse anti-FLAG M2 antibody (Sigma). To test the interaction between ZIC2 and Cdx1 in the presence of NCE2 sequences, *in vitro* synthesized $_{FLAG}$ZIC2 proteins were pre-incubated with increasing amounts of NCE2 DNA (50 or 200 ng) for 30 min in the same conditions as described for EMSA before adding GST or GST-Cdx1 homeodomain fusion proteins.

2.7-Cell culture, transfections and RT-PCR analyses

P19, Neuro2a (N2a) and Cos7 cell lines were propagated and transfected as previously described using Genejuice reagents (Novagen) in accordance with manufacturer's protocol [26].

Luciferase reporter assays were performed essentially as previously described [26]. For Zic2 or Sox2 dose response assays, N2a ($8x10^4$ cells per well) and P19 ($3x10^4$ cells per well) were transfected with 100 ng of *Pax3*NCE2-Luciferase reporter construct alone or with increasing amounts of $_{HA}$Zic2 (10 ng to 40 ng) or $_{Myc}$Sox2 (10 ng to 75 ng) expression vectors. For analysis of the synergistic effect of Cdx1 and Zic2, N2a cells were transfected with 100 ng of wild type or mutated *Pax3*NCE2-Luciferase reporter construct alone, or with fixed amount of $_{FLAG}$Cdx1 (5ng) and/or $_{HA}$Zic2 (10 ng) expression

vectors. When required, an empty expression vector was also included to complete the final amount of DNA per well to 150 ng. For analysis of the synergistic transactivation of *Pax3*NCE2 by Cdx1, Zic2 and Sox2 co-expression, fixed amounts of $_{FLAG}$Cdx1 (10 ng), $_{HA}$Zic2 (40 ng) and $_{Myc}$Sox2 (50 ng) were used and an empty expression vector was also included to ensure a total of 200 ng of DNA per well. All transfections were performed at least five times in triplicate. In all experiments, specificity of NCE2 activation was assessed by using a previously described luciferase reporter construct consisting of the *Pax3* 150bp minimal promoter alone [26], as a negative control. Forty-eight hours after transfection, cells were disrupted in 100 µl of lysis buffer (0.1 M Tris [pH 8.0], 1% Igepal, 1 mM dithiothreitol) and assessed for luciferase activity with a Berthold LB9507 luminometer (Berthold Technologies). Expression of exogenous $_{FLAG}$Cdx1, $_{HA}$Zic2 and $_{Myc}$Sox2 in cell lysates was verified by western blot.

To modulate Zic2 activity, N2a cells were seeded in 100-mm tissue culture plates ($2x10^6$ cells per plate) and transiently transfected with increasing doses (3, 6 and 9 µg) of $_{HA}$Zic2-IRES-GFP expression vector. Forty-eight hours after transfection, GFP-positive cells were recovered by FACS and analyzed by RT-PCR (for *Pax3* and *Gapdh* expression) as well as via western blot (for $_{HA}$Zic2 and Gapdh expression).

For dose-response analyses of Shh-induced repression of *Pax3*, *Zic2* and *Zic5* expression, N2a cells were seeded in 6 well plates ($4x10^5$ cells per well) and treated the following day with varying doses (50-200 ng/ml) of recombinant mouse Shh-N proteins (R&D systems) in absence or presence of 1ug/ml cyclopamine (R&D systems). Cells were harvested 24h post-treatment and analyzed by RT-PCR.

For cycloheximide (CHX) treatments, N2a cells were pre-treated for 30 min with 30 µg/mL of CHX or with the vehicle (DMSO) alone as previously described [26] and then treated for 24h with or without Shh-N recombinant protein (200 ng/mL) in the presence of 1 µg/mL of CHX.

For rescue experiments, N2a cells were seeded in 100-mm tissue culture plates ($2x10^6$ cells per plate) and transiently transfected with $_{HA}$Zic2-IRES-GFP or empty IRES-GFP expression vectors. The following day, cells were incubated in the presence of Ctl or Shh-enriched (200 ng/ml) medium and cultured for another 24h prior to recovery of GFP-positive cells by FACS.

RNA isolation and RT-PCR analyses were performed as previously described [26]. The primers used in this study were: *Pax3* (forward: 5'-CCTGCCAACATACCAGCTGTCG-3', reverse: 5'-CTGAGGTGAAAGGCCATTGCCG-3'); *Zic2* (forward:5'-AAGGTCTTCGCACGCTCCGAG-3', reverse: 5'-CGCAACGAGCTGGGATGCGTGT-3'); *Zic5* (forward: 5'-CAAGATCCACAAGCGCACTCATACA-3', reverse: 5'-TTGGGTCCAGCACAGGGGACAAAG-3'); *Gapdh* (forward: 5'-TCCTGCACCACCAACTGCTTAGC-3', reverse: 5'-AGGTCCACCACCCTGTTGCTGTA-3').

2.8- Statistical analyses

Statistical analyses were performed using GraphPad Prism software version 5.0. Differences between means were evaluated by one way ANOVA followed by a Tukey's post test and classed as not significant (n.s., $p>0.05$) or significant (*, $p<0.05$; **p<0.01; ***p<0.001).

3-RESULTS

3.1-A positive Cdx input is necessary but not sufficient to refine the Pax3 expression domain in the caudal neuroectoderm

Cdx genes (*Cdx1/2/4*) encode homeodomain transcription factors exhibiting functional redundancy and overlapping expression patterns in the caudal embryo [42-48]. Using mouse transgenic reporters and murine cell line perturbation analyses, we have previously demonstrated that Cdx proteins are required, downstream of canonical Wnt signaling, to induce *Pax3* expression in the mouse caudal neuroectoderm [26]. However, as shown in Fig.1A by immunofluorescence analysis of Cdx2 and Pax3 distribution in e9.5 caudal NT sections, *Cdx* and *Pax3* expression patterns only partially overlap [1, 42, 44, 45]. Indeed, Cdx proteins are broadly distributed in the entire tailbud and, more anteriorly, they are detected along the whole DV axis of the recently closed NT. In contrast, *Pax3* expression is restricted to the lateral borders of the open PNP and, more anteriorly, to the dorsal NT. This implies that non-Cdx inputs are required to refine the *Pax3* expression domain in the caudal neuroectoderm and leaves us with three possibilities that may involve either one or two mechanisms (Fig.1B). The first possibility (Mechanism-1 only) would involve one or several Cdx cofactors that exhibit more-or-less restricted neural expression in the lateral PNP and dorsal NT. The second possibility (Mechanism-2 only) would involve a negative input from a repressor expressed in the medial PNP and ventral NT. The third possibility would involve a combination of both mechanisms.

3.2-The neural and dorsally restricted zinc finger transcription factor Zic2 regulates Pax3 expression

Based on their overlapping expression pattern with Cdx and Pax3, we first hypothesized that Zic2/5 transcription factors might contribute to Mechanism-1. Indeed, Zic proteins can transactivate target genes and of the five murine *Zic* genes, only the *Zic2-Zic5* gene pair (sharing the same 5'-flanking region) is expressed in the caudal most region of the neuroectoderm [10]. Moreover, *Pax3* expression has been previously reported to be significantly reduced in NCC homozygous for a knockdown allele of *Zic2* while *Zic2*-null mutants exhibit posterior NT and NCC defects that phenocopy those observed in *Pax3*-null mutants [12]. To validate the candidacy of Zic2 as a *Pax3* regulator, we began by carrying out a detailed analysis of their expression pattern using whole-mount *in situ* hybridization in stage-matched e9.0 and e9.5 mouse embryos. In accordance with previous studies, this analysis revealed a striking and extensive overlap in the neural expression domain of both genes along the AP axis (Fig.2A-J). Interestingly, as evidenced by its more posterior limit of expression in the open PNP, this analysis notably demonstrated that *Zic2* is in fact induced earlier than *Pax3* in the caudal embryo (compare Figs.2C,E,G with 2D,F,H). Following NT closure, *Pax3* and *Zic2* are then found to be equally expressed in a dorsal domain that covers ~40% of the DV neural axis (bracket in Fig.2I-J). It is noteworthy that such overlapping expression pattern is neural specific since, outside the NT, *Pax3* and *Zic2* are expressed in a mutually exclusive manner in the somites (asterisk in Fig.2I,J). Thus, these data support the candidacy of Zic2 and further suggest that it might be involved in both the induction and dorsal restriction of neural *Pax3* expression. To more formally test our hypothesis that Zic2 regulates *Pax3* expression, we then overexpressed Zic2 in the murine NCC-derived N2a cell line and evaluated the effect on *Pax3* expression by RT-PCR. As shown in Fig.2K, overexpression of Zic2 resulted in a

specific and dose-dependent increase of *Pax3* expression. In summary, our data strongly suggest that Zic2 is a regulator of *Pax3* expression.

3.3-Zic2 directly binds and transactivates Pax3NCE2

We previously showed that the *Pax3*NCE2 CRM is sufficient for recapitulating induction and dorsal restriction of *Pax3* expression in the caudal neuroectoderm [26]. As we hypothesised that Zic2 is a Cdx neural cofactor, we reasoned that Zic2 should be able to activate *Pax3* expression via this Cdx-dependent CRM. We tested this possibility via luciferase assays in N2a cells and found that Zic2 transactivated a *Pax3*NCE2 reporter in a dose-dependent manner, similarly to endogenous *Pax3* (Fig.3A and 2K). EMSAs were then used to determine whether Zic2 directly binds to *Pax3*NCE2. Zic proteins are known to bind GC-rich elements but very few ZicBS have been previously identified in CRMs of Zic-regulated genes [10]. Thus, to avoid excluding any unpredicted binding site, we scanned the whole NCE2 sequence with eight previously described overlapping probes [26]. This approach revealed that one of these probes was bound by Zic2 as efficiently as a probe containing a previously described ZicBS [41] (Fig.3B). We confirmed the presence of $_{HA}$Zic2 in the complex formed with the NCE2 probe by supershift with an anti-HA antibody (Fig. 4E'). In accordance with previous reports, sequence analysis of the NCE2 probe revealed the presence of a GC-rich 9-nucleotide core sequence (5'-CTGCTGGGG-3'). Mutation of guanosine residues in position 6 and 7 prevented Zic2 binding to a cold probe bearing these point mutations, thus demonstrating specificity of Zic2 binding to this NCE2 target sequence (Fig.4E'). We next sought to determine whether Zic2, like Cdx proteins, occupies the endogenous *Pax3*NCE2 using ChIP-PCR. To this end – and because we failed to obtain a ChIP-grade anti-Zic2 antibody – we overexpressed $_{HA}$Zic2 in N2a cells and evaluated its presence on NCE2 using anti-HA. $_{FLAG}$Cdx1-overexpressing cells were used for comparison purposes. As shown in Fig.3C, *Pax3*NCE2 was specifically amplified in sonicated chromatin samples immunoprecipitated with either anti-HA or anti-FLAG antibodies, while no amplification was obtained when a mouse pre-immune serum was used for IP. Altogether, these results demonstrate that Zic2 directly binds and transactivates *Pax3*NCE2.

3.4-Zic2 directly interacts with Cdx1

Analysis of NCE2 sequences revealed that the ZicBS is flanked by CdxBS in close proximity (Fig.5B), further supporting a putative Cdx cofactor function for Zic2. To directly test this possibility, we carried out protein-protein interaction assays beginning with co-IP in $_{HA}$Zic2- and/or $_{FLAG}$Cdx1-transfected Cos7 cells. As shown in Fig.4A, mouse Zic2 was found to specifically interact with Cdx1 when both are co-expressed. Then, to assess whether this interaction was direct, we performed GST-based pull-down assays using *in vitro* translated FLAG-tagged human ZIC2 and bacterially produced GST fusion proteins containing either the full-length Cdx1, the Cdx1 homeodomain or the Cdx1 N-terminal domain (GST-Cdx1, GST-Cdx1Homeo and GST-Cdx1Nterm). Western blotting using an anti-FLAG antibody showed that $_{FLAG}$ZIC2 directly interacts with both the full-length Cdx1 and the Cdx1 homeodomain whereas no interaction was observed with the Cdx1 N-terminus (Fig.4B). Mapping of Cdx1 and ZIC2 interacting domains was next assessed using a series of *in vitro* translated $_{FLAG}$ZIC2 deletion-containing proteins and the GST-Cdx1Homeo fusion protein (Fig.S1A). This analysis revealed that ZIC2 and Cdx1 interact via their respective DNA binding domain since no interaction with the Cdx1

335 homeodomain was observed when the Zinc finger domain of ZIC2 was deleted (construct 1-255)
336 (Fig.4C-D and Fig.S1B). Regions located N-terminus (constructs 140-532 and 255-532) or C-terminus
337 (1-415) to the Zinc finger domain have apparently no major impact on this interaction (Fig.4C-D).
338 Intrigued by this discovery, we then tested whether Zic2 and Cdx1 could still physically interact with
339 each other in the presence of their respective target sequences using EMSA as well as GST pull-down
340 assays. In EMSA, we found that pre-incubation of $_{HA}$Zic2- with $_{FLAG}$Cdx1-expressing Cos7 nuclear
341 extracts did not preclude target DNA binding and even resulted in the appearance of novel bands of
342 higher molecular weight either on the ZicBS (E'), the CdxBS1 (E'') or the CdxBS3 (E'''), suggesting
343 the formation of a Zic2-Cdx1-DNA complex on these elements. Consistent with these outcomes, pre-
344 incubation of $_{FLAG}$ZIC2 with increasing amounts of the whole NCE2 did not impair its interaction with
345 the Cdx1 homeodomain in GST pull-down assays (Fig.S1C). These data demonstrate that, although the
346 Zic2-Cdx1 physical interaction is mediated by their respective DNA binding domain, this does not
347 prevent the ability of both Zic2 and Cdx1 to bind their respective target sequences.
348
349 ***3.5-Robust Zic2 and Cdx1 functional interaction on Pax3NCE2 requires a positive input from the***
350 ***neural specific Sox2 transcription factor***
351 Last validation of Zic2 as a Cdx1 cofactor required the demonstration of a functional interaction with
352 Cdx1 and we assessed this via luciferase assays in both N2a and undifferentiated P19 cell lines. In N2a
353 cells, Cdx1-Zic2 co-expression resulted in a weak synergistic activation of the *Pax3*NCE2 luciferase
354 reporter (Fig.5A). Interestingly, mutation of the three functional CdxBS alone or in combination with
355 mutation of the ZicBS significantly abrogated this Cdx1-Zic2 cooperative effect. This is consistent with
356 the central role of Cdx1 in activation of *Pax3*NCE2 [26] and suggests that binding of Cdx1 to this
357 CRM is important for the Cdx1-Zic2 functional interaction. Intriguingly, although mutation of the
358 ZicBS also abrogated the Cdx1-Zic2 cooperative effect, these analyses further revealed that Zic2 ability
359 to transactivate *Pax3*NCE2 can be independent of the ZicBS as well as the three CdxBS (Fig.5A).
360 Moreover, in the undifferentiated P19 cell line, Zic2 failed to transactivate *Pax3*NCE2 and no
361 cooperative effect between Cdx1 and Zic2 was observed (data not shown). Therefore, this suggests that
362 interaction with another neural factor is necessary for the Zic2-Cdx1 complex to robustly transactivate
363 *Pax3*NCE2. Based on several lines of evidence, Sox2 quickly emerged as a primary candidate in this
364 regard. Indeed, at least one potential SoxBS is present in *Pax3*NCE2 (Fig.5B), SoxB1 family members
365 have recently been implicated as positive regulators of *Pax3* expression [30] and the SoxB1 family
366 member Sox2 has been shown to directly interact with Cdx1 [33]. To assess the putative contribution
367 of Sox2, we first evaluated whether this transcription factor alone was able to transactivate the
368 *Pax3*NCE2 luciferase reporter in N2a cells. Consistent with our hypothesis, Sox2 strongly
369 transactivated *Pax3*NCE2 in a dose-dependent manner (Fig.5C). Of note, such effect was not observed
370 in P19 cells, suggesting that the Sox2 capacity to transactivate this CRM relies on the presence of
371 other, as yet undefined, neural factors (Fig.5C). To demonstrate the importance of the Sox2 input in the
372 functional interaction between Cdx1 and Zic2, we co-transfected N2a cells with Cdx1, Zic2 and Sox2
373 and evaluated the activation of the *Pax3*NCE2 luciferase reporter. Interestingly, a significant
374 synergistic effect was observed when the three factors were co-expressed (Fig.5D; see Fig.S2 for
375 western blot validations). Together, these results suggest that *Pax3*NCE2 is robustly activated by a
376 concerted interaction between Cdx1, Zic2, and Sox2 trans-acting factors.

3.6-Zic2 is a potential mediator of the Shh-induced repression of Pax3

Having confirmed that Zic2 is a *bona fide* Cdx cofactor, and thus contribute to Mechanism-1, we next investigated the possibility that it could also contribute to the repressor-dependent Mechanism-2 (Fig.1B). Indeed, a ventrally Shh-induced repressive pathway is a prominent candidate to consider and both *Zic* members and *Pax3* are known to be affected by such regulatory cascade. Studies in multiple species, including mice, indicate that Shh signals secreted by the notochord and floorplate are required to prevent ectopic expression of *Pax3* and *Zic* members in the ventral NT [20, 21, 25, 30, 49-51]. While the molecular mechanism of this regulation is poorly understood for *Zic* members, recent work in zebrafish and chick embryos suggest that the Shh target Nkx6.1 represses *Pax3* expression via direct binding to a *Pax3* CRM located in the 4th intron [29, 30]. However, whether such Shh-Nkx6.1 circuit is at work in the mouse and active at all AP levels has not been investigated. To verify this, we compared the expression profile of *Pax3* and *Zic2* with the distribution of Nkx6.1 proteins in the caudal neuroectoderm of stage matched e9.5 mouse embryos (Fig.6A-O). Surprisingly, we found that Nkx6.1 is not detected in the open PNP and recently closed NT whereas *Pax3* is already induced and properly restricted in these structures (compare Fig.6A-B with K-L). Therefore, the Shh-Nkx6.1 pathway cannot be responsible for the restriction of *Pax3* expression in the caudal neuroectoderm of mouse embryos. On the other hand, we found that *Zic2* is dynamically expressed along the DV axis of the caudal neuroectoderm. *Zic2* transcripts are indeed first detected in the whole PNP and along the entire DV axis of the recently closed NT before becoming dorsally restricted more anteriorly (Fig.6F-J). In more anterior regions, a perfect overlap is finally observed between *Zic2* and *Pax3* neural expression domains (Fig.2I-J). Very interestingly, we also noticed that the onset of *Zic2* repression in the ventral NT is correlated with the onset of Nkx6.1 expression (see arrow in Fig.6H and M). On one hand, all these observations suggest that, while the Cdx-Zic2 neural input is necessary for *Pax3* induction, a Shh-Nkx6.1-independent input must also be required to restrict *Pax3* expression in the open PNP and most caudal NT. On the other hand, these expression data also strongly suggest that Zic2 might be an intermediary in the mechanism of *Pax3* repression by the Shh-Nkx6.1 pathway. We sought to validate this hypothesis using the murine N2a cell line and first validated this line as a good model to study the mechanism of *Pax3* repression by Shh (Fig.S3). We then performed a dose-response assay to evaluate whether *Zic2* and *Pax3* were both repressed by increasing doses of Shh (Fig.6P). As expected from previous work [20, 21, 25, 30, 49-51] , Shh repressed the expression of *Pax3* as well as expression of the *Zic2-Zic5* gene pair. To finally confirm our hypothesis, we tried to rescue *Pax3* expression by overexpressing Zic2 in Shh-treated cells. In agreement with a Zic-dependent mechanism, transfection with a Zic2 expression vector specifically rescued *Pax3* expression from the repressive input of Shh (Fig.6Q). Taken together, these evidences strongly support an intermediary role for Zic2 in the mechanism of *Pax3* repression by Shh in the mouse.

4-DISCUSSION

In this work, we report a novel direct role for Zic2 in the regulation of *Pax3* neural expression. Our data indicate that Zic2 can transactivate *Pax3* at least via the NCE2 CRM, which we previously showed to be sufficient for recapitulating caudal *Pax3* expression in pre-migratory NCC and dorsal NT [26].

Interestingly, we also provide evidence that Zic2 is involved in both the Wnt-mediated induction (by interacting with Cdx proteins) and the Shh-induced dorsal restriction (by being itself repressed by Shh signals) of *Pax3* expression. Therefore, our studies provide important mechanistic insights into how AP and DV instructive cues can be coordinately integrated on a short CRM.

4.1-Role of the Zic2-Cdx complex in Wnt-mediated induction of Pax3 neural expression

In previous studies, we presented evidence that a Wnt-Cdx circuit is critically required for activation of a *Pax3*NCE2 reporter in transgenic mice [26]. These data support a model in which Cdx proteins convey the posteriorizing Wnt-βCatenin signals to *Pax3* in order to mediate induction of trunk NCC [26, 52, 53]. Interestingly, indirect activation of *Pax3* expression by canonical Wnt signaling – involving either AP2a or Gbx2 as intermediates – has also been described in *Xenopus* embryos during induction of cranial NCC [17, 54]. On the other hand, other groups have in addition reported that *Pax3* could be directly regulated by Wnt-βCatenin signaling through CRMs distinct from NCE2 [29, 30, 55]. This suggests that regulation of *Pax3* expression by the canonical Wnt pathway could be controlled via both direct and indirect means in order to buffer against perturbations in the Wnt input. In the caudal neuroectoderm, a third level of regulation would even be possible given that Cdx proteins can directly and positively interact with the βcatenin-Lef/Tcf complex on Lef/Tcf binding sites [33]. However, the very broad distribution of Cdx proteins in both neural and non-neural cells of the caudal embryo [42, 56] implies that other players must cooperate with the Wnt-Cdx inductive input to spatially refine the *Pax3* expression domain.

The present study demonstrates that Zic2 not only acts as a *Pax3* regulator, but also as a neural-specific Cdx cofactor. Indeed, we found that Zic2 and Cdx1 cooperate in the activation of *Pax3*NCE2 and physically interact via their respective DNA binding domain (Fig.4 and 5). Such involvement of DNA binding domains in protein-protein interactions is not unusual and examples implicating either the Zic2 C2H2 zinc finger region or the Cdx1 homeodomain have been previously described [33, 34]. On the other hand, to the best of our knowledge, this is the first report showing that the Zic zinc finger region or the Cdx homeodomain can simultaneously mediate protein-protein as well as protein-DNA interactions (Fig.4, 5 and Fig.S1). This bifunctionality is most likely due to the fact that, for each of these proteins, establishment of direct contacts with DNA bases is not carried out by the whole DNA binding domain *per se* but by specific subregions only. Indeed, previous studies have suggested that only zinc fingers 3-5 of Zic2 might be implicated in DNA recognition [10, 57, 58] whereas, for homeodomain proteins, DNA binding activity is known to be specifically conferred by the 3rd α-helix [59]. Thus, this suggests that the Zic2-Cdx1 interaction might be mediated by the zinc fingers 1-2 of Zic2 and the 1st and/or 2nd α-helix of the Cdx1 homeodomain. Regardless of the exact minimal interacting domain on each protein, it should be noted that functional interactions between different members of each family are expected given the very high conservation of both Zic and Cdx DNA binding domains [58, 60].

Interestingly, our data also indicate that, although both Zic2 and Cdx1 have the ability to bind *Pax3*NCE2 sequences in a direct as well as in an indirect manner (Fig.4), robust activation by the Zic2-Cdx1 complex requires direct binding of each protein to this CRM as well as the participation of Sox2 (Fig.5) – a previously described Cdx interacting partner [33]. Such contribution of Sox2 is in line with recent work suggesting the implication of SoxB members in the regulation of neural *Pax3* expression

via another CRM located in the 4th intron [30]. On the other hand, it is interesting to note that in contrast to Cdx proteins [26], neither Zic2 nor Sox2 can activate *Pax3* expression in non-neural cells (Fig.5C and data not shown). Taken together with the indispensable role of CdxBS1-3 for *Pax3*NCE2 activity [26], this strongly suggests that Cdx proteins have a pioneer-like activity required to attract and tether neural-specific transcriptional activators on this CRM [61]. Interestingly, a similar pioneer-like activity has been also previously suggested for Cdx2 in the regulation of the transcriptional program of intestinal epithelial cells [62] as well as for Pdx1 – another member of the Cdx-containing paraHox family – in the regulation of several genes important for pancreatic β cell development and function [63, 64].

4.2-Restriction of Pax3 expression to the lateral borders of the PNP and the dorsal NT

A role for notochord- and floor plate-secreted Shh in the regulation of *Pax3* restriction in the dorsal NT has been described in diverse model organisms [20, 30, 65]. In the mouse, Shh gain-of-function in *Ptch1*-null mutants results in a nearly complete loss of *Pax3* expression whereas Shh loss-of-function in Shh-null embryos results in expansion of *Pax3* expression along the entire DV neural axis [21, 66]. Recent studies suggest that this Shh-induced repressive input could be mediated by the Nkx6.1 homeodomain transcriptional repressor – a known direct target of the Shh-Gli pathway [67]. Indeed, Moore et al. [30] reported that Nkx6.1 can directly bind a *Pax3* CRM located in intron-4 (called CNE3) and that mutation of a putative binding site for homeodomain proteins within this CRM is required to turn on ubiquitous but mosaic expression of a reporter gene driven by this CRM in zebrafish embryos. Combination of this work with Tcf3 ChIP-seq data from ES cells [68] led these authors to propose a model whereby Shh-induced Nkx6.1 could counteract canonical Wnt signaling in order to restrict induction of *Pax3* expression to the lateral borders of the PNP [30]. However, in addition to the fact that Tcf3 is not a direct mediator of Wnt-βCatenin positive inputs [69], this model is not consistent with the Nkx6.1 expression profile in the mouse. Indeed, we and others have shown that Nkx6.1 is not expressed in the open PNP and recently closed NT, whereas *Pax3* is already induced and properly restricted in these regions (Fig.6) [67]. Moreover, *in situ* hybridization analyses have failed to detect *Shh* transcripts in the caudal neuroectoderm [70], implying that no other Shh targets could be involved in this mechanism. In agreement with this, *Zic2*, a known Shh-regulated gene in the mouse [49], is broadly expressed in the entire PNP and recently closed NT (Fig.6). Taking into account all these observations, the proposed Shh-Nkx6.1 repressive circuit cannot contribute to the induction phase of *Pax3* expression and can only be implicated in the maintenance of *Pax3* dorsal restriction.

Interestingly, our data indicate that an alternative pathway involving repression of the *Pax3* activator Zic2 could also mediate the Shh-induced repression of *Pax3* expression (Fig.6). This hypothesis is strongly supported by data from Moore et al. [30] showing that Nkx6.1 overexpression has a much more dramatic impact on repression of *Pax3* expression in comparison to the CNE3 reporter bearing the mutated homeodomain binding site. Moreover, our expression data in Fig.6 – showing that the onset of *Zic2* repression is correlated with the onset of Nkx6.1 expression in the ventral NT – are also consistent with the idea that *Zic2* might be a direct target of Nkx6.1. Although it is currently highly speculative, we propose that this alternative Shh-Nkx6.1-Zic2-Pax3 pathway targeting NCE2 would be functionally redundant with the more straightforward Shh-Nkx6.1-Pax3 pathway targeting CNE3 [30]. Given the co-expression of the *Zic2-Zic5* gene pair as well as the known functional redundancy

503 between Zic members [4, 71-74], we also believe that Zic5 acts redundantly with Zic2 in this
504 mechanism. Furthermore, given that Zic and Gli proteins have similar DNA binding specificity and Zic
505 inhibit Gli transcriptional activity [41], it is tempting to speculate that cross-repression between Zic2/5
506 and the Shh pathway would be important for proper DV patterning of the NT.
507
508 ***4.3-Coordinated integration of AP and DV instructive cues and redundancy between Pax3 CRMs***
509 As evidenced by our work and the work of others, *Pax3* gene expression is regulated by two sets of
510 CRMs that are clustered in the 5'-flanking region or in intron-4, respectively (Fig.5B) [18, 26-30, 55].
511 Interestingly, although the requirement of intron-4 CRMs has not yet been formally addressed via gene
512 targeting approaches, deletion of 5'-located CRMs in the mouse is indicative of functional redundancy
513 between each regulatory cluster [29]. It is also interesting to note that, although each set of CRMs is
514 expected to respond to the same AP and DV instructive cues, different mechanisms appears to be used
515 (e.g. Shh-Nkx6.1-Pax3 vs Shh-Nkx6.1-Zic2-Pax3) and different complementarities are seen among
516 CRMs of each cluster. Within the 5' cluster, NCE1 and NCE2 appear to be complementary for the
517 spatial control of *Pax3* expression along the AP axis. Indeed, while we have shown that NCE2 drives
518 reporter gene expression in a caudal-specific and Cdx-dependent manner, others have shown that
519 NCE1 is involved in the regulation of cranial *Pax3* expression in a Pbx1-dependent manner [26, 75]. In
520 contrast, within the intron-4 cluster, CNE3 (also known as ECR2 or IR1) and CNE1 (also known as
521 IR2) appear to exhibit temporal complementarities in the control of *Pax3* expression, with CNE3
522 suggested to be involved in the induction phase in a Lef/Tcf-dependent manner and CNE1 suggested to
523 be involved in the maintenance phase in a Pax3/7-dependent manner [18, 29, 30]. In accordance with
524 the key developmental role of Pax3, such functional redundancy at both the *cis* and *trans* level most
525 likely allows to protect its expression from the otherwise deleterious impacts of genetic mutations in *cis*
526 or perturbations of signaling pathways in *trans*.
527 Our work on NCE2 indicates that it may serve as a good model to further understand how AP and DV
528 instructive cues can be coordinately integrated on a single CRM. Indeed, in addition to allow
529 recapitulating induction of *Pax3* expression in the PNP, NCE2-driven expression is properly restricted
530 [26]. It is also noteworthy that we and others have shown that NCE2 seems particularly enriched in
531 binding sites for a wide array of transcription factors including members of the Cdx, Pou class-III,
532 SoxB, Zic and Tead families [26, 28, 31]. In an effort to gather all this information into a dynamic
533 network (Fig.7), we propose a model in which the posteriorizing cue from the Wnt-Cdx circuit, in
534 cooperation with Zic and SoxB positive inputs, is first required to induce *Pax3* expression in the PNP.
535 Later on (i.e. more rostrally), the Shh-Nkx repressive input is then most likely involved in the
536 maintenance of the dorsally-restricted *Pax3* expression domain via repression of *Zic2/5* expression. In
537 absence of Cdx factors more rostrally, maintenance is also most likely ensured by the positive input of
538 other transcription factors such as members of the Pou class III and Tead families [28, 31]. While this
539 model clearly allows understanding how the *Pax3*NCE2 CRM is regulated by AP instructive cues
540 (Wnt-Cdx), this is much less clear regarding the integration of DV instructive cues. Indeed, our studies
541 clearly show that an unknown Shh-independent factor is required during the induction phase in order to
542 limit expression to the lateral borders of the PNP. More work will definitely be necessary for the
543 identification of this key player, once again in accordance with the mechanisms described in Fig.1B.
544

4.4-Conclusion

This study allowed confirming the central role of Cdx proteins in the induction of *Pax3* neural expression as well as identifying novel molecular functions for Zic2 as a Cdx co-factor and activator of *Pax3* expression. By allowing extension of the list of transcription factors able to bind and transactivate *Pax3*NCE2, our work also suggests that this CRM behaves as a "super-enhancer" region.

ACKNOWLEDGEMENTS: The authors thank Denis Flipo (UQAM) for FACS analyses and assistance with confocal imaging. Jonathan Epstein, Sabine Tejpar, Jun Aruga, Mahmud Bani-Yaghoub and David Lohnes are thanked for kindly agreeing to provide expression vectors. This work was supported by a grant from the Canadian Institute for Health Research (CIHR grant number MOP-111130) to NP. OSF holds an Alexander-Graham-Bell scholarship from the Natural Science and Engineering Research Council (NSERC) of Canada. NP is a Fonds de la Recherche du Québec – Santé (FRQS) Junior2 scholar.

REFERENCES

[1] M.D. Goulding, G. Chalepakis, U. Deutsch, J.R. Erselius, P. Gruss, Pax-3, a novel murine DNA binding protein expressed during early neurogenesis, Embo J, 10 (1991) 1135-1147.

[2] C. Gaston-Massuet, D.J. Henderson, N.D. Greene, A.J. Copp, Zic4, a zinc-finger transcription factor, is expressed in the developing mouse nervous system, Developmental dynamics : an official publication of the American Association of Anatomists, 233 (2005) 1110-1115.

[3] T. Nagai, J. Aruga, S. Takada, T. Gunther, R. Sporle, K. Schughart, K. Mikoshiba, The expression of the mouse Zic1, Zic2, and Zic3 gene suggests an essential role for Zic genes in body pattern formation, Developmental biology, 182 (1997) 299-313.

[4] T. Inoue, M. Hatayama, T. Tohmonda, S. Itohara, J. Aruga, K. Mikoshiba, Mouse Zic5 deficiency results in neural tube defects and hypoplasia of cephalic neural crest derivatives, Developmental biology, 270 (2004) 146-162.

[5] D.J. Epstein, M. Vekemans, P. Gros, Splotch (Sp2H), a mutation affecting development of the mouse neural tube, shows a deletion within the paired homeodomain of Pax-3, Cell, 67 (1991) 767-774.

[6] R. Auerbach, Analysis of the developmental effects of a lethal mutation in the house mouse, J Exp Zoology, 127 (1954) 305–329.

[7] J. Li, K.C. Liu, F. Jin, M.M. Lu, J.A. Epstein, Transgenic rescue of congenital heart disease and spina bifida in Splotch mice, Development, 126 (1999) 2495-2503.

[8] J. Aruga, The role of Zic genes in neural development, Mol Cell Neurosci, 26 (2004) 205-221.

[9] A. Mansouri, A. Stoykova, M. Torres, P. Gruss, Dysgenesis of cephalic neural crest derivatives in Pax7-/- mutant mice, Development, 122 (1996) 831-838.

[10] C.S. Merzdorf, Emerging roles for zic genes in early development, Dev Dyn, 236 (2007) 922-940.

[11] P. Elms, P. Siggers, D. Napper, A. Greenfield, R. Arkell, Zic2 is required for neural crest formation and hindbrain patterning during mouse development, Developmental biology, 264 (2003) 391-406.

[12] T. Nagai, J. Aruga, O. Minowa, T. Sugimoto, Y. Ohno, T. Noda, K. Mikoshiba, Zic2 regulates the kinetics of neurulation, Proceedings of the National Academy of Sciences of the United States of America, 97 (2000) 1618-1623.

[13] A.H. Monsoro-Burq, E. Wang, R. Harland, Msx1 and Pax3 cooperate to mediate FGF8 and WNT signals during Xenopus neural crest induction, Dev Cell, 8 (2005) 167-178.

[14] L.A. Taneyhill, M. Bronner-Fraser, Dynamic alterations in gene expression after Wnt-mediated induction of avian neural crest, Mol Biol Cell, 16 (2005) 5283-5293.

[15] A.G. Bang, N. Papalopulu, C. Kintner, M.D. Goulding, Expression of Pax-3 is initiated in the early neural plate by posteriorizing signals produced by the organizer and by posterior non-axial mesoderm, Development, 124 (1997) 2075-2085.
[16] A.G. Bang, N. Papalopulu, M.D. Goulding, C. Kintner, Expression of Pax-3 in the lateral neural plate is dependent on a Wnt-mediated signal from posterior nonaxial mesoderm, Dev Biol, 212 (1999) 366-380.
[17] N. de Croze, F. Maczkowiak, A.H. Monsoro-Burq, Reiterative AP2a activity controls sequential steps in the neural crest gene regulatory network, Proceedings of the National Academy of Sciences of the United States of America, 108 (2011) 155-160.
[18] A.T. Garnett, T.A. Square, D.M. Medeiros, BMP, Wnt and FGF signals are integrated through evolutionarily conserved enhancers to achieve robust expression of Pax3 and Zic genes at the zebrafish neural plate border, Development, 139 (2012) 4220-4231.
[19] K.F. Liem, Jr., G. Tremml, H. Roelink, T.M. Jessell, Dorsal differentiation of neural plate cells induced by BMP-mediated signals from epidermal ectoderm, Cell, 82 (1995) 969-979.
[20] M.D. Goulding, A. Lumsden, P. Gruss, Signals from the notochord and floor plate regulate the region-specific expression of two Pax genes in the developing spinal cord, Development, 117 (1993) 1001-1016.
[21] L.V. Goodrich, L. Milenkovic, K.M. Higgins, M.P. Scott, Altered neural cell fates and medulloblastoma in mouse patched mutants, Science, 277 (1997) 1109-1113.
[22] J. Ericson, S. Morton, A. Kawakami, H. Roelink, T.M. Jessell, Two critical periods of Sonic Hedgehog signaling required for the specification of motor neuron identity, Cell, 87 (1996) 661-673.
[23] T. Sato, N. Sasai, Y. Sasai, Neural crest determination by co-activation of Pax3 and Zic1 genes in Xenopus ectoderm, Development, 132 (2005) 2355-2363.
[24] A.H. Monsoro-Burq, R.B. Fletcher, R.M. Harland, Neural crest induction by paraxial mesoderm in Xenopus embryos requires FGF signals, Development, 130 (2003) 3111-3124.
[25] J. Aruga, T. Tohmonda, S. Homma, K. Mikoshiba, Zic1 promotes the expansion of dorsal neural progenitors in spinal cord by inhibiting neuronal differentiation, Developmental biology, 244 (2002) 329-341.
[26] O. Sanchez-Ferras, B. Coutaud, T. Djavanbakht Samani, I. Tremblay, O. Souchkova, N. Pilon, Caudal-related homeobox (Cdx) protein-dependent integration of canonical Wnt signaling on paired-box 3 (Pax3) neural crest enhancer, J Biol Chem, 287 (2012) 16623-16635.
[27] T.A. Natoli, M.K. Ellsworth, C. Wu, K.W. Gross, S.C. Pruitt, Positive and negative DNA sequence elements are required to establish the pattern of Pax3 expression, Development, 124 (1997) 617-626.
[28] R.C. Milewski, N.C. Chi, J. Li, C. Brown, M.M. Lu, J.A. Epstein, Identification of minimal enhancer elements sufficient for Pax3 expression in neural crest and implication of Tead2 as a regulator of Pax3, Development, 131 (2004) 829-837.
[29] K.R. Degenhardt, R.C. Milewski, A. Padmanabhan, M. Miller, M.K. Singh, D. Lang, K.A. Engleka, M. Wu, J. Li, D. Zhou, N. Antonucci, L. Li, J.A. Epstein, Distinct enhancers at the Pax3 locus can function redundantly to regulate neural tube and neural crest expressions, Dev Biol, 339 (2010) 519-527.
[30] S. Moore, V. Ribes, J. Terriente, D. Wilkinson, F. Relaix, J. Briscoe, Distinct Regulatory Mechanisms Act to Establish and Maintain Pax3 Expression in the Developing Neural Tube, PLoS genetics, 9 (2013) e1003811.
[31] S.C. Pruitt, A. Bussman, A.Y. Maslov, T.A. Natoli, R. Heinaman, Hox/Pbx and Brn binding sites mediate Pax3 expression in vitro and in vivo, Gene Expr Patterns, 4 (2004) 671-685.
[32] W.A. Whyte, D.A. Orlando, D. Hnisz, B.J. Abraham, C.Y. Lin, M.H. Kagey, P.B. Rahl, T.I. Lee, R.A. Young, Master transcription factors and mediator establish super-enhancers at key cell identity genes, Cell, 153 (2013) 307-319.
[33] M. Beland, N. Pilon, M. Houle, K. Oh, J.R. Sylvestre, P. Prinos, D. Lohnes, Cdx1 autoregulation is governed by a novel Cdx1-LEF1 transcription complex, Mol Cell Biol, 24 (2004) 5028-5038.
[34] R. Pourebrahim, R. Houtmeyers, S. Ghogomu, S. Janssens, A. Thelie, H.T. Tran, T. Langenberg, K. Vleminckx, E. Bellefroid, J.J. Cassiman, S. Tejpar, Transcription factor Zic2 inhibits Wnt/beta-catenin protein signaling, The Journal of biological chemistry, 286 (2011) 37732-37740.
[35] Y. Koyabu, K. Nakata, K. Mizugishi, J. Aruga, K. Mikoshiba, Physical and functional interactions between Zic and Gli proteins, The Journal of biological chemistry, 276 (2001) 6889-6892.

[36] M. Bani-Yaghoub, R.G. Tremblay, J.X. Lei, D. Zhang, B. Zurakowski, J.K. Sandhu, B. Smith, M. Ribecco-Lutkiewicz, J. Kennedy, P.R. Walker, M. Sikorska, Role of Sox2 in the development of the mouse neocortex, Developmental biology, 295 (2006) 52-66.
[37] M.H. Kaufman, The Atlas of Mouse Development, Academic Press, London, 1992.
[38] D.G. Wilkinson, In " *In Situ* Hybridisation. A Practical Approach.", in: I. Press (Ed.), New York, 1992.
[39] B. Coutaud, N. Pilon, Characterization of a novel transgenic mouse line expressing Cre recombinase under the control of the Cdx2 neural specific enhancer, Genesis, (2013) 1-8.
[40] J. Jeong, A.P. McMahon, Growth and pattern of the mammalian neural tube are governed by partially overlapping feedback activities of the hedgehog antagonists patched 1 and Hhip1, Development, 132 (2005) 143-154.
[41] K. Mizugishi, J. Aruga, K. Nakata, K. Mikoshiba, Molecular properties of Zic proteins as transcriptional regulators and their relationship to GLI proteins, J Biol Chem, 276 (2001) 2180-2188.
[42] B.I. Meyer, P. Gruss, Mouse Cdx-1 expression during gastrulation, Development, 117 (1993) 191-203.
[43] J.G. Savory, N. Pilon, S. Grainger, J.R. Sylvestre, M. Beland, M. Houle, K. Oh, D. Lohnes, Cdx1 and Cdx2 are functionally equivalent in vertebral patterning, Dev Biol, 330 (2009) 114-122.
[44] F. Beck, T. Erler, A. Russell, R. James, Expression of Cdx-2 in the mouse embryo and placenta: possible role in patterning of the extra-embryonic membranes, Dev Dyn, 204 (1995) 219-227.
[45] L.W. Gamer, C.V. Wright, Murine Cdx-4 bears striking similarities to the Drosophila caudal gene in its homeodomain sequence and early expression pattern, Mech Dev, 43 (1993) 71-81.
[46] J.G. Savory, M. Mansfield, F.M. Rijli, D. Lohnes, Cdx mediates neural tube closure through transcriptional regulation of the planar cell polarity gene Ptk7, Development, 138 (2011) 1361-1370.
[47] E. van den Akker, S. Forlani, K. Chawengsaksophak, W. de Graaff, F. Beck, B.I. Meyer, J. Deschamps, Cdx1 and Cdx2 have overlapping functions in anteroposterior patterning and posterior axis elongation, Development, 129 (2002) 2181-2193.
[48] J. van Nes, W. de Graaff, F. Lebrin, M. Gerhard, F. Beck, J. Deschamps, The Cdx4 mutation affects axial development and reveals an essential role of Cdx genes in the ontogenesis of the placental labyrinth in mice, Development, 133 (2006) 419-428.
[49] L.Y. Brown, A.H. Kottmann, S. Brown, Immunolocalization of Zic2 expression in the developing mouse forebrain, Gene expression patterns : GEP, 3 (2003) 361-367.
[50] K.B. Rohr, S. Schulte-Merker, D. Tautz, Zebrafish zic1 expression in brain and somites is affected by BMP and hedgehog signalling, Mechanisms of development, 85 (1999) 147-159.
[51] A.M. Holtz, K.A. Peterson, Y. Nishi, S. Morin, J.Y. Song, F. Charron, A.P. McMahon, B.L. Allen, Essential role for ligand-dependent feedback antagonism of vertebrate hedgehog signaling by PTCH1, PTCH2 and HHIP1 during neural patterning, Development, 140 (2013) 3423-3434.
[52] N. Pilon, K. Oh, J.R. Sylvestre, N. Bouchard, J. Savory, D. Lohnes, Cdx4 is a direct target of the canonical Wnt pathway, Dev Biol, 289 (2006) 55-63.
[53] N. Pilon, K. Oh, J.R. Sylvestre, J.G. Savory, D. Lohnes, Wnt signaling is a key mediator of Cdx1 expression in vivo, Development, 134 (2007) 2315-2323.
[54] B. Li, S. Kuriyama, M. Moreno, R. Mayor, The posteriorizing gene Gbx2 is a direct target of Wnt signalling and the earliest factor in neural crest induction, Development, 136 (2009) 3267-3278.
[55] T. Zhao, Q. Gan, A. Stokes, R.N. Lassiter, Y. Wang, J. Chan, J.X. Han, D.E. Pleasure, J.A. Epstein, C.J. Zhou, beta-catenin regulates Pax3 and Cdx2 for caudal neural tube closure and elongation, Development, (2013).
[56] S. Takada, K.L. Stark, M.J. Shea, G. Vassileva, J.A. McMahon, A.P. McMahon, Wnt-3a regulates somite and tailbud formation in the mouse embryo, Genes Dev, 8 (1994) 174-189.
[57] K. Mizugishi, M. Hatayama, T. Tohmonda, M. Ogawa, T. Inoue, K. Mikoshiba, J. Aruga, Myogenic repressor I-mfa interferes with the function of Zic family proteins, Biochem Biophys Res Commun, 320 (2004) 233-240.
[58] R. Houtmeyers, J. Souopgui, S. Tejpar, R. Arkell, The ZIC gene family encodes multi-functional proteins essential for patterning and morphogenesis, Cell Mol Life Sci, 70 (2013) 3791-3811.
[59] W.J. Gehring, Y.Q. Qian, M. Billeter, K. Furukubo-Tokunaga, A.F. Schier, D. Resendez-Perez, M. Affolter, G. Otting, K. Wuthrich, Homeodomain-DNA recognition, Cell, 78 (1994) 211-223.

[60] D. Lohnes, The Cdx1 homeodomain protein: an integrator of posterior signaling in the mouse, Bioessays, 25 (2003) 971-980.
[61] K.S. Zaret, J.S. Carroll, Pioneer transcription factors: establishing competence for gene expression, Genes Dev, 25 (2011) 2227-2241.
[62] M.P. Verzi, H. Shin, A.K. San Roman, X.S. Liu, R.A. Shivdasani, Intestinal master transcription factor CDX2 controls chromatin access for partner transcription factor binding, Mol Cell Biol, 33 (2013) 281-292.
[63] B.G. Hoffman, G. Robertson, B. Zavaglia, M. Beach, R. Cullum, S. Lee, G. Soukhatcheva, L. Li, E.D. Wederell, N. Thiessen, M. Bilenky, T. Cezard, A. Tam, B. Kamoh, I. Birol, D. Dai, Y. Zhao, M. Hirst, C.B. Verchere, C.D. Helgason, M.A. Marra, S.J. Jones, P.A. Hoodless, Locus co-occupancy, nucleosome positioning, and H3K4me1 regulate the functionality of FOXA2-, HNF4A-, and PDX1-bound loci in islets and liver, Genome Res, 20 (2010) 1037-1051.
[64] K. Ohneda, R.G. Mirmira, J. Wang, J.D. Johnson, M.S. German, The homeodomain of PDX-1 mediates multiple protein-protein interactions in the formation of a transcriptional activation complex on the insulin promoter, Mol Cell Biol, 20 (2000) 900-911.
[65] Y. Litingtung, C. Chiang, Control of Shh activity and signaling in the neural tube, Developmental dynamics : an official publication of the American Association of Anatomists, 219 (2000) 143-154.
[66] C. Chiang, Y. Litingtung, E. Lee, K.E. Young, J.L. Corden, H. Westphal, P.A. Beachy, Cyclopia and defective axial patterning in mice lacking Sonic hedgehog gene function, Nature, 383 (1996) 407-413.
[67] K.A. Peterson, Y. Nishi, W. Ma, A. Vedenko, L. Shokri, X. Zhang, M. McFarlane, J.M. Baizabal, J.P. Junker, A. van Oudenaarden, T. Mikkelsen, B.E. Bernstein, T.L. Bailey, M.L. Bulyk, W.H. Wong, A.P. McMahon, Neural-specific Sox2 input and differential Gli-binding affinity provide context and positional information in Shh-directed neural patterning, Genes & development, 26 (2012) 2802-2816.
[68] A. Marson, S.S. Levine, M.F. Cole, G.M. Frampton, T. Brambrink, S. Johnstone, M.G. Guenther, W.K. Johnston, M. Wernig, J. Newman, J.M. Calabrese, L.M. Dennis, T.L. Volkert, S. Gupta, J. Love, N. Hannett, P.A. Sharp, D.P. Bartel, R. Jaenisch, R.A. Young, Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells, Cell, 134 (2008) 521-533.
[69] C.I. Wu, J.A. Hoffman, B.R. Shy, E.M. Ford, E. Fuchs, H. Nguyen, B.J. Merrill, Function of Wnt/beta-catenin in counteracting Tcf3 repression through the Tcf3-beta-catenin interaction, Development, 139 (2012) 2118-2129.
[70] F. Gofflot, M. Hall, G.M. Morriss-Kay, Genetic patterning of the developing mouse tail at the time of posterior neuropore closure, Developmental dynamics : an official publication of the American Association of Anatomists, 210 (1997) 431-445.
[71] T. Inoue, M. Ota, K. Mikoshiba, J. Aruga, Zic2 and Zic3 synergistically control neurulation and segmentation of paraxial mesoderm in mouse embryo, Developmental biology, 306 (2007) 669-684.
[72] T. Inoue, M. Ota, M. Ogawa, K. Mikoshiba, J. Aruga, Zic1 and Zic3 regulate medial forebrain development through expansion of neuronal progenitors, The Journal of neuroscience : the official journal of the Society for Neuroscience, 27 (2007) 5461-5473.
[73] J. Aruga, T. Inoue, J. Hoshino, K. Mikoshiba, Zic2 controls cerebellar development in cooperation with Zic1, J Neurosci, 22 (2002) 218-225.
[74] T. Inoue, M. Ogawa, K. Mikoshiba, J. Aruga, Zic deficiency in the cortical marginal zone and meninges results in cortical lamination defects resembling those in type II lissencephaly, J Neurosci, 28 (2008) 4712-4725.
[75] C.P. Chang, K. Stankunas, C. Shang, S.C. Kao, K.Y. Twu, M.L. Cleary, Pbx1 functions in distinct regulatory networks to pattern the great arteries and cardiac outflow tract, Development, 135 (2008) 3577-3586.

FIGURE LEGENDS

FIGURE 1: Partially overlapping Cdx and Pax3 expression domains indicate that Cdx proteins are necessary but not sufficient to establish the *Pax3* expression domain in the caudal neuroectoderm.
(A) Immunofluorescence analysis of Cdx2 and Pax3 distribution in tailbud sections of a 25-somite stage mouse embryo (e9.5). The dashed red lines in the left panel indicate the AP level at which the sections shown on the right were made *(A', A'', A''')*. The dotted white lines in the right panel delimit the posterior neural plate (PNP) and the neural tube (NT) structures. Note that Cdx2 is broadly expressed in the tailbud including the whole PNP (A') and, more rostrally, in the entire NT (A'', A''') while Pax3 domain is limited to the lateral borders of the PNP (A') and to the dorsal part of the NT (A'', A'''). *(B)* Schematic drawing of two putative mechanisms to explain the induction and dorsal restriction of *Pax3* neural expression (pink) via Cdx proteins (green). *(B') Pax3* expression at the lateral borders of the PNP might be established by collaboration between Cdx proteins and a positive input from a neural-specific protein with more-or-less restricted expression: i.e. a Cdx cofactor expressed in the PNP (yellow) or at the lateral borders of the PNP (purple) [Mechanism 1]. Restriction of *Pax3* expression at the PNP border may also be achieved by integration of the positive Cdx input with the negative input from a repressor expressed in the medial PNP (red) [Mechanism 2]. *(B'')* At more rostral levels, where Cdx expression is restricted to the NT, the Cdx input might be integrated with the positive input from a dorsally restricted Cdx cofactor (blue) [Mechanism 1] and/or the negative input from a ventrally expressed repressor (red) [Mechanism 2].

FIGURE 2: The Zic2 transcription factor regulates *Pax3* expression.
(A-J) Overlapping spatiotemporal *Zic2* and *Pax3* expression patterns are consistent with a role for Zic2 in regulating *Pax3* expression. *(A-B)* Lateral view of 20-somite stage embryos (~e9.0) with anterior to the left showing *Zic2 (A)* and *Pax3 (B)* expression domains assessed by whole-mount *in situ* hybridisation. (*C-J)* Comparison of *Zic2* and *Pax3* expression domains in the caudal end of 24-somite stage embryos (~e9.5). The dotted lines in the lateral views shown in C-D indicate the level at which the transverse sections shown in G-J were cut. The red brackets in E-F designate the length of the caudal tip of the tailbud region (dorsal view) that is devoid of *Zic2* and *Pax3* gene expression. Note that, compared to *Pax3*, the *Zic2* expression domain in the open PNP extends more posteriorly (C-H). Also note the extensive overlap in *Pax3* and *Zic2* expression in the dorsal NT (C-D; red brackets in I-J). *Pax3* and *Zic2* expression patterns do not overlap in the developing somites. The presomitic mesoderm (C, D) and somites (I, J) are denoted by asterisks. *(K)* Analyses of endogenous *Pax3* expression in N2a cells transfected with increasing amounts of a $_{HA}$Zic2-IRES-GFP expression vector. GFP-positive N2a cells were recovered by FACS and exogenous $_{HA}$Zic2 as well as endogenous *Pax3* expression were analyzed by western blot and RT-PCR, respectively. Note that increasing levels of $_{HA}$Zic2 (left panel) result in a dose-dependent increase of endogenous *Pax3* expression but not *Gapdh* used as control (right panel).

FIGURE 3: Zic2 activates and directly binds the *Pax3* neural crest enhancer NCE2.
(A) Co-transfection assays in N2a cells using luciferase reporter constructs driven by the *Pax3*NCE2 region plus the *Pax3* 150bp minimal promoter (*Pax3NCE2*-min-Luc) or only driven by the *Pax3* 150bp

minimal promoter (*Pax3*min-Luc) as well as a $_{HA}$Zic2 expression vector. The luciferase quantification results are expressed as fold activation compared to each reporter vector alone. n=5 independent experiments performed in triplicate; one-way ANOVA: P<0.0001; Tukey's post test: n.s. P>0.05; * P<0.05; ***: P<0.001. Error bars indicate s.e.m. Note that increasing Zic2 expression significantly increases the activation of the *Pax3*NCE2 reporter but not the *Pax3*min reporter. *(B)* Identification of Zic binding sites in the *Pax3*NCE2 by electrophoretic mobility shift assay (EMSA). Eight overlapping double-stranded oligonucleotide probes (represented as numbers 1 to 8 in boxes at the top of each panel) were used to scan the whole NCE2 for Zic binding sites (ZicBS). A probe bearing a consensus Gli/ZicBS [41] was used as a positive control for Zic binding. All *in vitro* binding reactions were performed in parallel under identical conditions and each probe was either not incubated (-), or incubated with nuclear extracts from Mock- (transfected without DNA) or $_{HA}$Zic2-transfected Cos7 cells. Note that, similarly to the shifted band observed with the positive control (ZicBS), a specific shifted band is observed for radiolabelled probe #3 (red box) when incubated with $_{HA}$Zic2 nuclear extracts. *(C)* Chromatin immunoprecipitation assays in N2a cells showing the occupancy of *Pax3*NCE2 by both $_{FLAG}$Cdx1 and $_{HA}$Zic2. Anti-FLAG-, anti-HA- or normal rabbit IgG-immunoprecipitated chromatin extracts from $_{FLAG}$Cdx1- and $_{HA}$Zic2- transfected N2a cells were purified and amplified by PCR using primers flanking the *Pax3*NCE2 region. Amplification products, resolved on a 2% agarose gel, were of the expected size (246bp) and sequence-confirmed. Lanes 1, 2 and 3 represent PCR amplifications of serial dilutions of DNA. Note that *Pax3*NCE2 is amplified from chromatin samples immunoprecipitated with anti-FLAG (Cdx1) and anti-HA (Zic2), but not with the non-specific IgG.

FIGURE 4: Zic2 and Cdx1 directly interact via their respective DNA binding domain**.**
(A) Co-immunoprecipitation assay using whole cell lysates from Cos7 cells transfected with $_{FLAG}$Cdx1 and/or $_{HA}$Zic2 expression vectors. Total lysates were immunoprecipitated with anti-FLAG antibody and analyzed by immunoblotting using an anti-HA antibody (top panel). Expression of Cdx1 was assessed by reprobing the blot with an anti-FLAG antibody (middle panel). Inputs represent 5% of the total lysate used for immunoprecipitation. *(B-C)* Mapping of Cdx1 and Zic2 interacting domains via GST pull-down assays. Inputs represent 15% of the total *in vitro*-translated proteins used for pull-downs. *(B)* Western blot using an anti-FLAG antibody showing that human $_{FLAG}$ZIC2 specifically interacts with both full-length Cdx1 (GST-Cdx1) and Cdx1 homeodomain (GST-Cdx1Homeo) but not with the Cdx1 N-terminal region (GST-Cdx1Nterm). *(C)* The Zinc finger domain of ZIC2 is essential for the interaction with the Cdx1 homeodomain. Different $_{FLAG}$ZIC2 deletion constructs were used to identify ZIC2 regions that interact with the Cdx1 homeodomain. Note that deletion of the five Cys2His2-type zinc fingers (construct 1-255) specifically abolishes interaction with the Cdx1 homeodomain. *(D)* **S**chematic representation of ZIC2 and its deletion constructs (adapted from [34] and [58]) and their interaction with Cdx1. *(E)* Characterization of the Zic binding site (ZicBS) contained in NCE2 and evaluation of the interaction of Zic2 and Cdx1 in the presence of their target DNA by EMSA. Results with a probe containing the new ZicBS (NCE2 scanning probe 3) are shown in E' while results with probes containing the previously described CdxBS1 (NCE2 scanning probe 5) or CdxBS3 (NCE2 scanning probe 2) are shown in E'' and E''', respectively. All *in vitro* binding reactions were performed in parallel under identical conditions. Binding of $_{HA}$Zic2 and $_{FLAG}$Cdx1 as well as anti-HA supershift (SS) and non-specific bindings (NS) are indicated by arrows. Brackets and asterisks denote

the presence of higher molecular weight bands when $_{HA}$Zic2- and $_{FLAG}$Cdx1-overexpressing nuclear extracts are used in combination. Note that specificity of Zic2 binding to the new ZicBS was evaluated by pre-incubation of nuclear extracts from $_{HA}$Zic2- or mock-transfected Cos7 cells with a 100-fold molar excess of non-radiolabeled wild type (wt) or mutated (mt) probes. Sequences of the ZicBS as well as its mutated version used to assess Zic2 binding specificity by EMSA are at the bottom of E'; point mutations are denoted in green.

FIGURE 5: Cdx1 and Zic2 functionally interact and synergize with the SoxB1 family member Sox2 in the transactivation of *Pax3*NCE2.
(A) Cdx1 and Zic2 cooperatively transactivate *Pax3*NCE2 in N2a cells and Cdx binding sites (CdxBS) are essential for this functional interaction. A wild type (wt) *Pax3*NCE2-luciferase reporter (1) or mutant versions bearing mutations in the three CdxBS (2), the ZicBS (3) or a combination of these mutations (4) were evaluated for Cdx1 and Zic2 transactivation in N2a cells. Note that $_{FLAG}$Cdx1 and $_{HA}$Zic2 co-transfection results in a transactivation of the wt *Pax3*NCE2-luc reporter stronger than simple addition of each single transfections, and that mutation of CdxBS and/or ZicBS abrogates this cooperative effect. In contrast, the transactivating effect of Zic2 is not affected by mutation of CdxBS and/or ZicBS. *(B)* Schematic representation of the *Pax3* genomic locus showing the relative position of *Pax3* regulatory regions. The 5' regulatory region (represented as a blue box) is named NCE and is subdivided in two CRMs called NCE1 and NCE2 [28]. The intron-4 regulatory region (red box) contains the CNE3 and CNE1 modules [30]. Black boxes represent exons. (*B'*) Magnification view of *Pax3*NCE2 DNA sequence showing the location of the ZicBS as well as the three functional CdxBS (in boxes) and in relation to a putative SoxBS (MatInspector analyses) as well as the previously described Brn1/2 [31] and Tead binding sites [28] (underlined). *(C)* Sox2 strongly activates *Pax3*NCE2 in N2a cells, but cannot activate this CRM in undifferentiated P19 cells. N2a and P19 cells were equally transfected with the *Pax3*NCE2-luc reporter alone or with increasing amounts of the Sox2 expression vector. *(D)* Cdx1, Zic2 and Sox2 act synergistically in the transactivation of *Pax3*NCE2 in N2a cells. Cells were transiently transfected with the *Pax3*NCE2-luc reporter alone or co-transfected with $_{FLAG}$Cdx1, $_{HA}$Zic2 or $_{Myc}$Sox2 expression vector alone or in combination. In A, C and D, results are expressed as fold activation compared to the relevant reporter vector alone. n=6-10 independent experiments performed in triplicate; one-way ANOVA: $P<0.0001$; Tukey's post test: n.s. $P>0.05$; * $P<0.05$; ***: $P<0.001$. Error bars indicate s.e.m.

FIGURE 6: *Zic2* and *Zic5* are potential intermediates in the Shh-induced repression of *Pax3* expression.
(A-O) Comparison of *Pax3* and *Zic2* expression domains (assessed by whole-mount *in situ* hybridisation), with Nkx6.1 distribution (assessed by whole-mount immunofluorescence) in serial vibratome sections of the posterior end of 24-somite (e9.5) stage-matched embryos (caudal to rostral direction). Note that restricted *Pax3* expression in the open PNP (A) and recently closed NT (B) is established before the onset of Nkx6.1 expression in the ventral NT (red arrow in M). Also note that the onset of *Zic2* repression in the ventral NT (red arrow in H) correlates with the presence of Nkx6.1 in this region (red arrow in M). Going forward in the posterior to anterior direction, *Zic2* and Nkx6.1 expression domains are complementary along the dorsal-ventral axis [i.e.: ~90% (H) vs ~10% (M);

~60% (I) vs ~40% (N); ~40% (J) vs ~60% (O)]. In contrast, *Pax3* and Nkx6.1 expression domains only become complementary at the most rostral levels [~40% (E) vs ~60% (O)]. *(P-Q)* Analysis of *Pax3, Zic2* and *Zic5* gene expression in Shh-treated N2a cells via semi-quantitative RT-PCR. Results shown are representative of at least two independent experiments. *(P)* The *Zic2-Zic5* gene pair is, like *Pax3,* expressed in N2a cells and negatively regulated by Shh in a dose-dependent manner. Prior to RT-PCR analysis, cells were cultured in absence or presence of increasing doses of Shh over a fixed 24-hour period of time. *(Q)* Zic2 is an intermediary in the mechanism of *Pax3* repression by Shh. Prior to RT-PCR analysis, cells were transfected with $_{HA}$Zic2 or empty IRES-GFP expression vector and then cultured for 24h in Ctl or Shh-enriched medium. Note that overexpression of Zic2 prevents the Shh-induced repression of *Pax3* in FACS-recovered N2a cells.

FIGURE 7: Current model for the Cdx-dependent control of *Pax3* expression in the caudal neuroectoderm via the NCE2 CRM.

Neural-specific induction of *Pax3* expression by the previously described Wnt-Cdx circuit requires an interaction with Zic2/5 and Sox2. Supplemental unknown (?) positive and/or negative inputs are required to restrict the *Pax3* expression domain to the lateral borders of the PNP. Maintenance of restricted NCE2 activity in the dorsal NT is most likely mediated by the dorsally restricted Zic2/5 transcription factors, which act as Cdx cofactors and are the predicted immediate early targets of the Shh-Nkx6.1 repression pathway. As elongation proceeds and the Cdx input disappears, *Pax3* expression in the NT is also later maintained by Tead2 and Brn1/2 transcription factors.

FIGURE S1: ZIC2 and Cdx1 can still physically interact in presence of the NCE2.

(A) Anti-FLAG western blot showing the expected size of *in vitro*-translated FLAG-tagged full-length ($_{FLAG}$ZIC2(1-532)) and truncated ZIC2 proteins ($_{FLAG}$ZIC2(140-532); $_{FLAG}$ZIC2(1-255); $_{FLAG}$ZIC2(1-415); $_{FLAG}$ZIC2(255-532)). *(B)* ZIC2 and Cdx1 bind to each other via their DNA binding domain. GST and GST-Cdx1 homeodomain (GST-Cdx1Homeo) fusion proteins were bound to glutathione-agarose beads and then incubated with *in vitro*-translated full-length or truncated $_{FLAG}$ZIC2 proteins. Presence of $_{FLAG}$ZIC2 in the precipitated complexes was then analysed by Western blot using an anti-FLAG antibody. Inputs represent 15% of the total *in vitro*-translated proteins used for pull-downs. Note that specific bindings are observed with $_{FLAG}$ZIC2(1-532), $_{FLAG}$ZIC2(140-532), $_{FLAG}$ZIC2(1-415) and $_{FLAG}$ZIC2(255-532) proteins but not with the truncated version $_{FLAG}$ZIC2 1-255 lacking the zinc finger domain. *(C)* Anti-FLAG immunoblot analysis of the interaction between $_{FLAG}$ZIC2(1-532) and GST-Cdx1Homeo in presence of the target NCE2 DNA. Note that, in comparisons to standard pull-down conditions (lanes 1-3), using EMSA buffer conditions does not affect the interaction (lanes 4-5). Note also that variations in signal intensity detected in lanes 6-7 are independent of the amount of NCE2 sequences.

FIGURE S2: Western blot validation of the expression of Cdx1, Zic2 and Sox2 in N2a lysates used for luciferase reporter assays.

Expression vectors for $_{FLAG}$Cdx1, $_{HA}$Zic2 as well as $_{Myc}$Sox2 were transfected alone or in different combinations together with a luciferase reporter construct driven by the *Pax3NCE2* and the *Pax3*

907 minimal promoter (*Pax3NCE2*-min-Luc). Production of exogenous proteins was evaluated by western
908 blot using anti-FLAG, anti-HA as well as anti-Myc antibodies. Shown are the luciferase quantification
909 results (expressed as fold activation compared to the reporter vector alone) as well as western blot
910 validations for one representative experiment of the results displayed in Fig.5D.
911
912 **FIGURE S3:** The N2a cell line is a good model for studying the Shh-induced repression of *Pax3*
913 expression.
914 *(A,B)* RT-PCR analysis demonstrating that *Pax3* expression is indirectly repressed by Shh signaling in
915 N2a cells. Prior to RT-PCR, cells were cultured for 24h in absence or presence of Shh and/or absence
916 or presence of cyclopamine or cycloheximide (CHX). Results are representative of at least n=3
917 independent experiments. *(A) Pax3* expression is repressed by Shh in a dose-dependent manner (left
918 panel) and this repression is abrogated in presence of cyclopamine, an inhibitor of the Shh receptor
919 Smo (right panel). *(B) De novo* protein synthesis is required for the Shh-mediated repression of *Pax3*
920 expression, as shown here in presence of the translation inhibitor CHX.
921
922

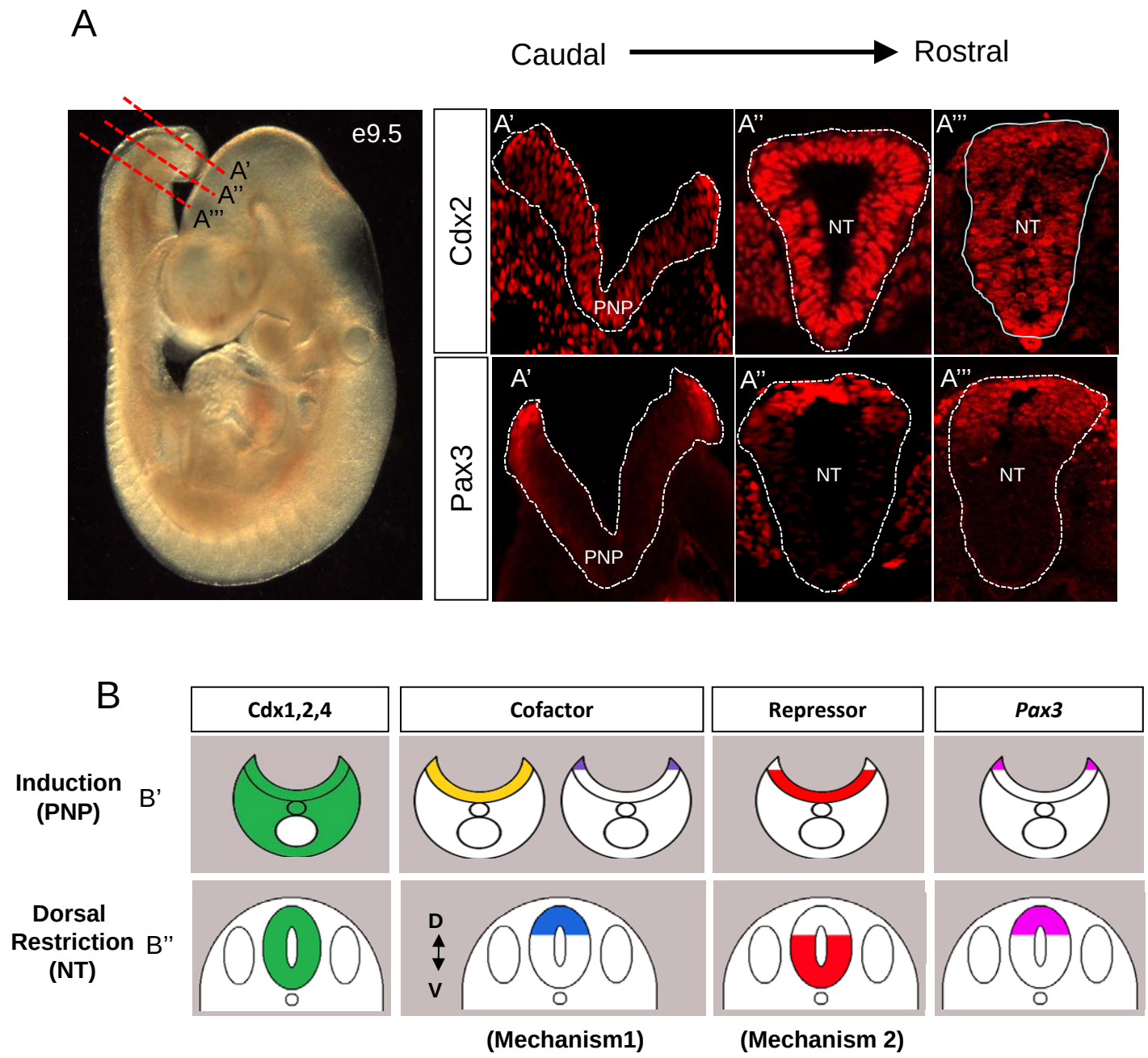
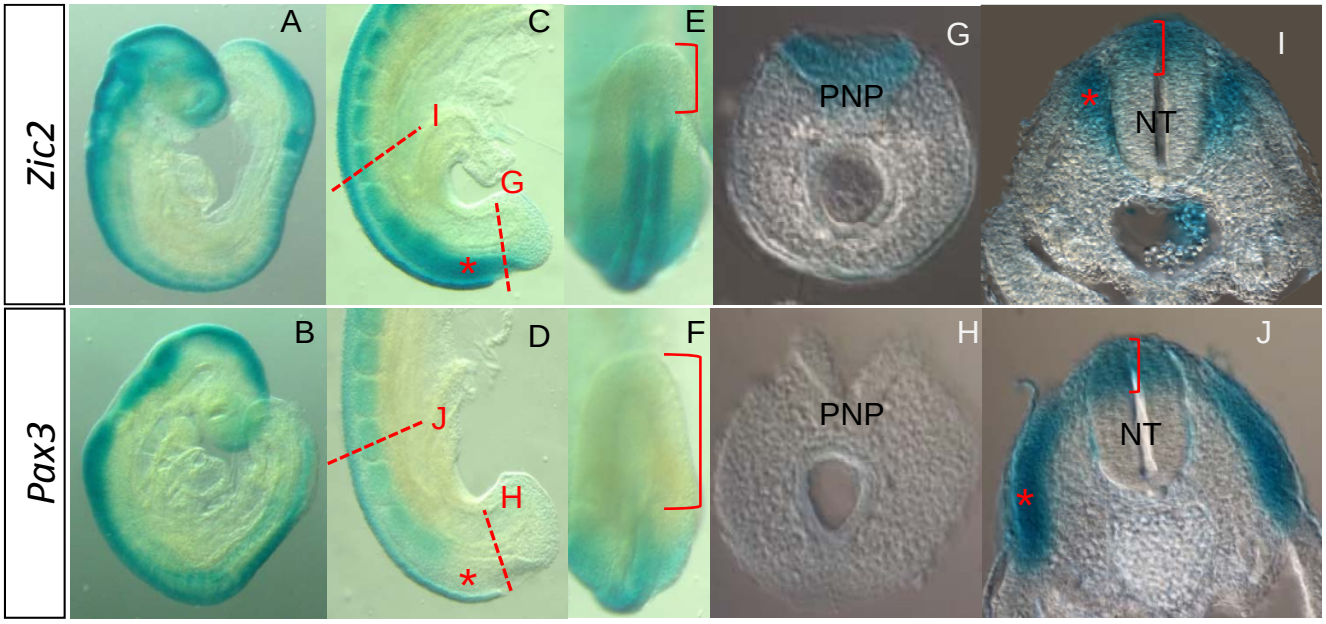

A
Caudal
Rostral
e9.5
A'
A''
A'''
Cdx2
Pax3
PNP
NT
B
Cdx1,2,4
Cofactor
Repressor
Pax3
Induction (PNP)
B'
Dorsal Restriction (NT)
B''
D
V
(Mechanism1)
(Mechanism 2)

FIGURE 1


Zic2
Pax3
A
B
C
D
E
F
G
H
I
J
PNP
PNP
NT
NT
*
*
*
*

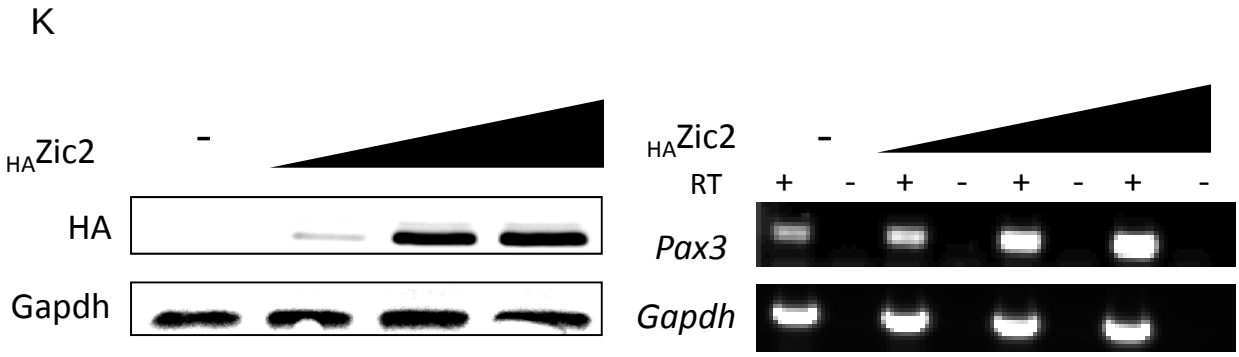

K
HAZic2
-
HA
Gapdh
HAZic2
-
RT
+ - + - + - + -
Pax3
Gapdh

FIGURE 2

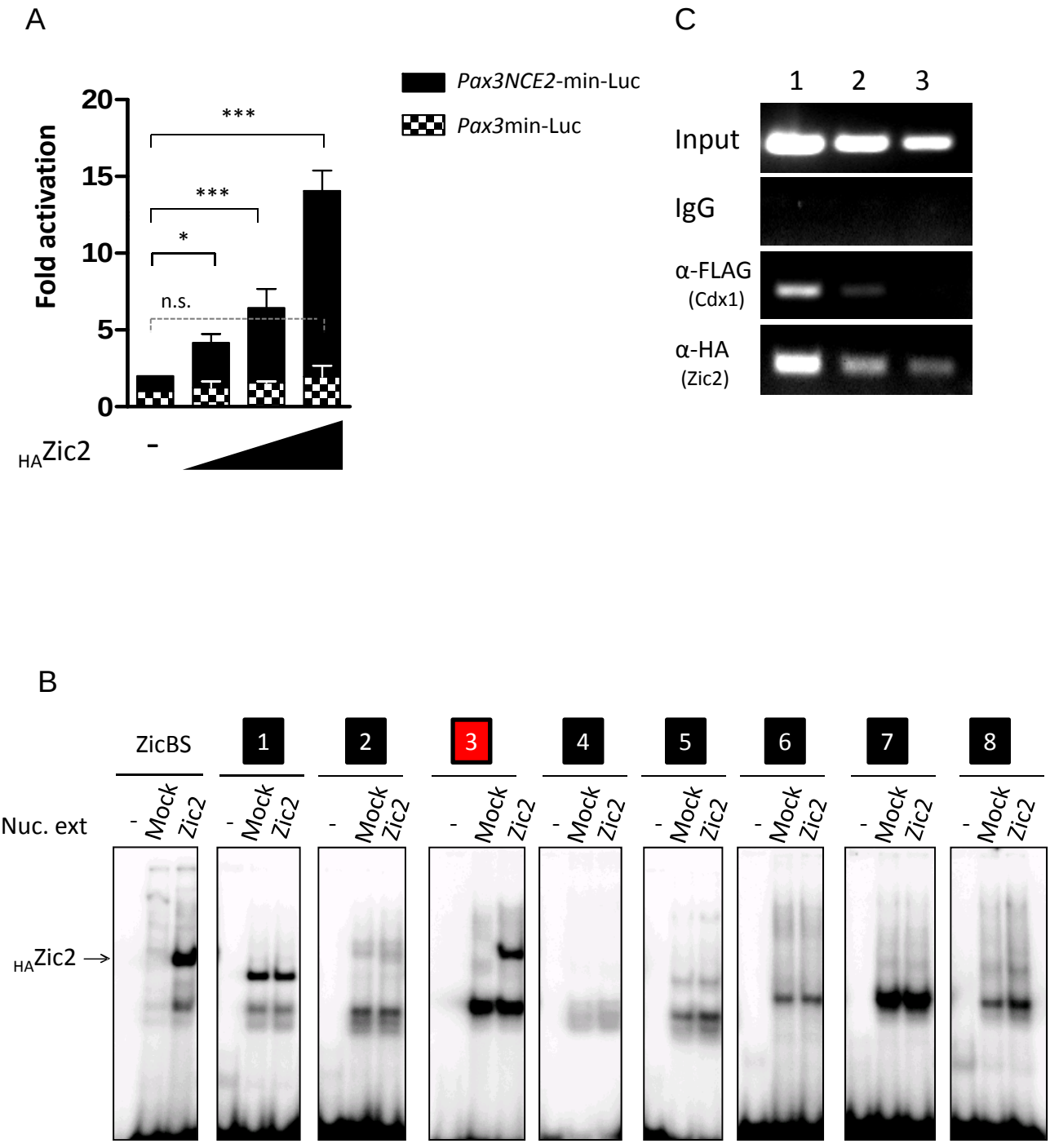

A
Pax3NCE2-min-Luc
Pax3min-Luc
Fold activation
20
15
10
5
0
***
***
*
n.s.
HAZic2
-
C
1
2
3
Input
IgG
α-FLAG
(Cdx1)
α-HA
(Zic2)
B
ZicBS
1
2
3
4
5
6
7
8
Nuc. ext
-
Mock
Zic2
HAZic2

FIGURE 3

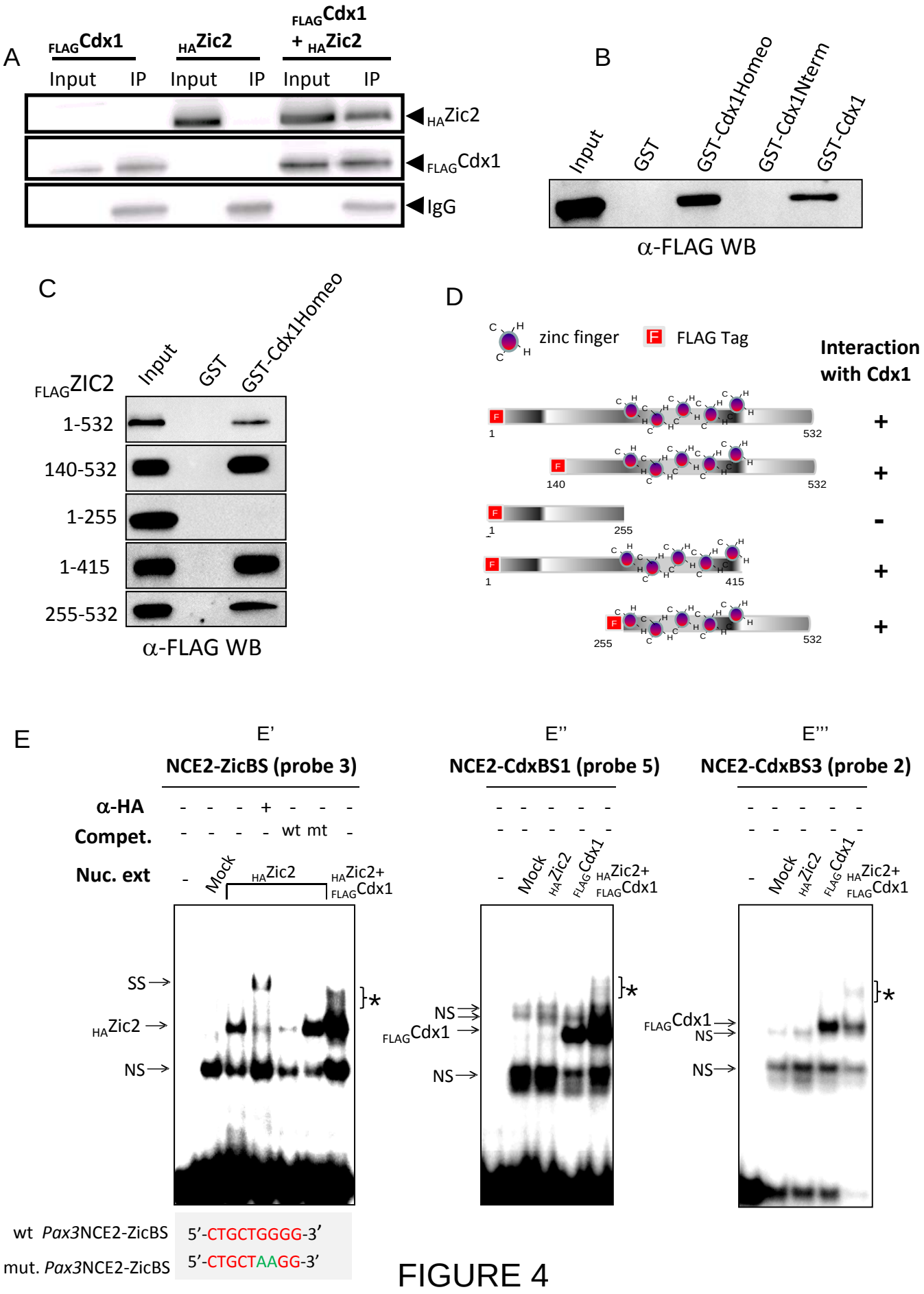

A
FLAGCdx1
HAZic2
FLAGCdx1 + HAZic2
Input
IP
Input
IP
Input
IP
HAZic2
FLAGCdx1
IgG
B
Input
GST
GST-Cdx1Homeo
GST-Cdx1Nterm
GST-Cdx1
α-FLAG WB
C
FLAGZIC2
Input
GST
GST-Cdx1Homeo
1-532
140-532
1-255
1-415
255-532
α-FLAG WB
D
zinc finger
FLAG Tag
Interaction with Cdx1
1
532
+
140
532
+
1
255
-
1
415
+
255
532
+
E
E'
NCE2-ZicBS (probe 3)
α-HA
Compet.
wt mt
Nuc. ext
Mock
HAZic2
HAZic2+ FLAGCdx1
SS
HAZic2
NS
E''
NCE2-CdxBS1 (probe 5)
Mock
HAZic2
FLAGCdx1
HAZic2+ FLAGCdx1
NS
FLAGCdx1
NS
E'''
NCE2-CdxBS3 (probe 2)
Mock
HAZic2
FLAGCdx1
HAZic2+ FLAGCdx1
FLAGCdx1
NS
NS
wt Pax3NCE2-ZicBS 5'-CTGCTGGGG-3'
mut. Pax3NCE2-ZicBS 5'-CTGCTAAGG-3'

FIGURE 4

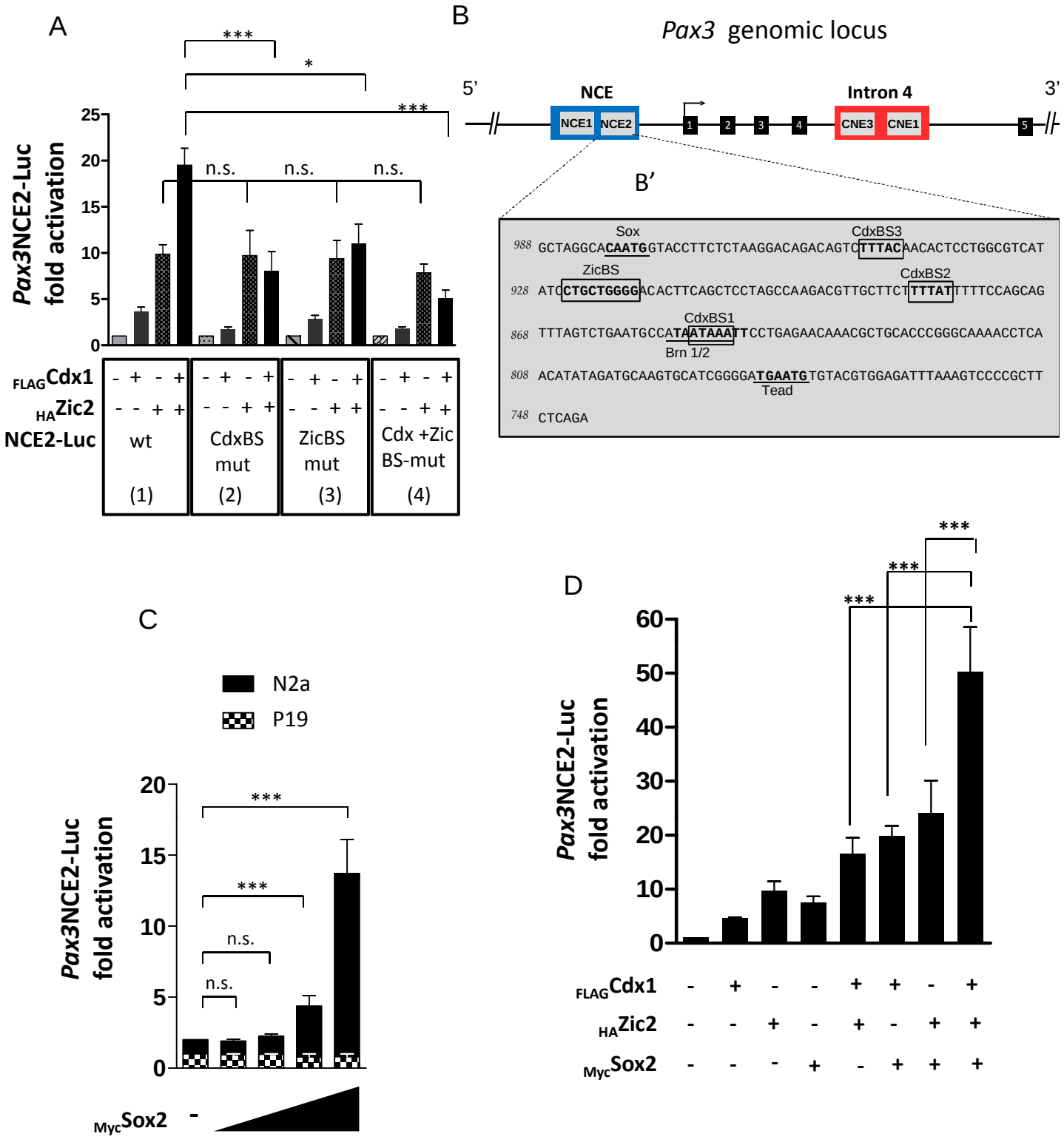
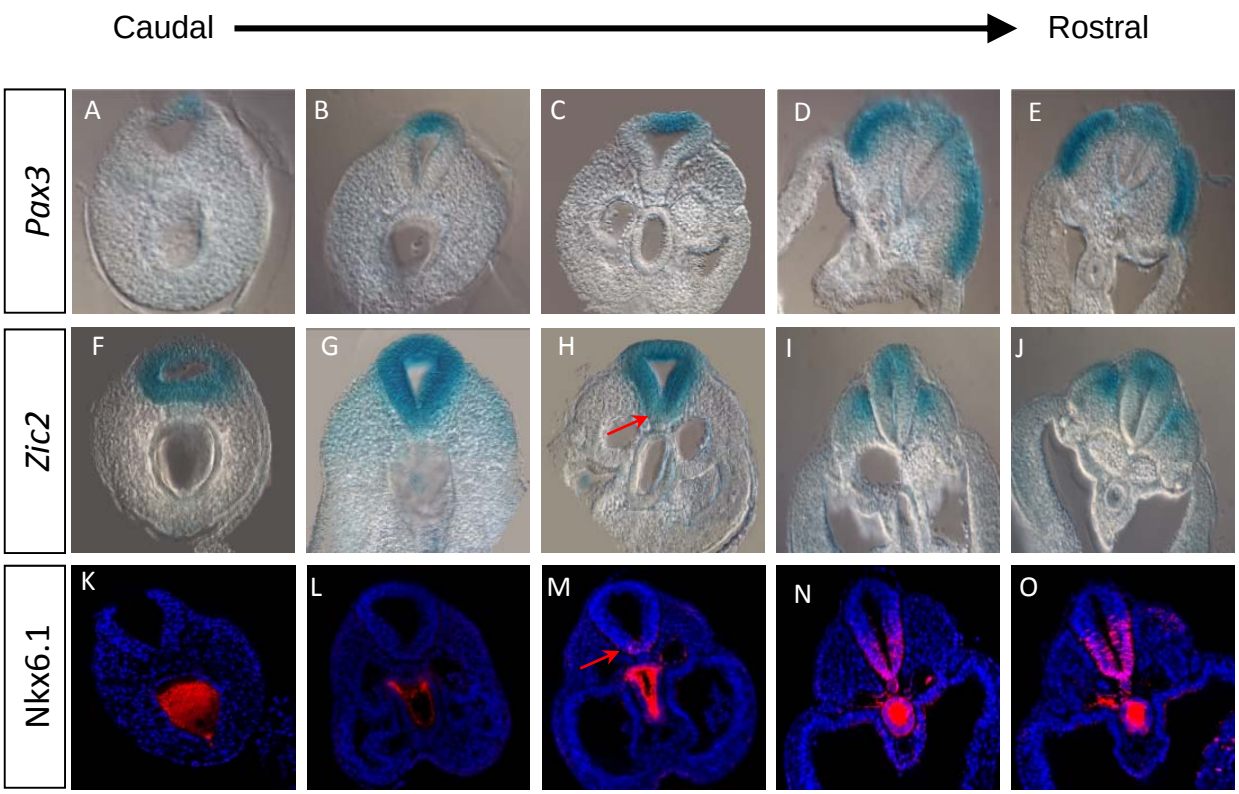
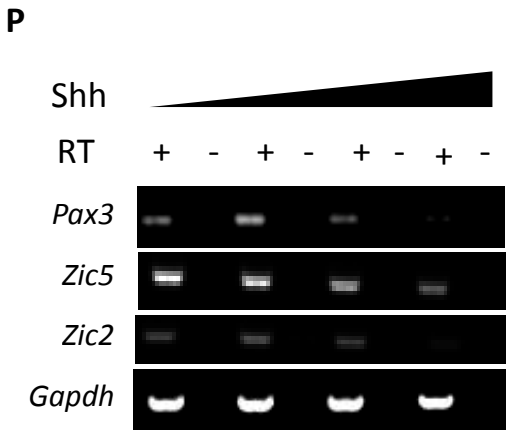

A
B
Pax3 genomic locus
5'
3'
NCE
Intron 4
NCE1
NCE2
CNE3
CNE1
B'
Sox
CdxBS3
ZicBS
CdxBS2
CdxBS1
Brn 1/2
Tead
988 GCTAGGCACAATGGTACCTTCTCTAAGGACAGACAGTCTTTACAACACTCCTGGCGTCAT
928 ATCCTGCTGGGGACACTTCAGCTCCTAGCCAAGACGTTGCTTCTTTTATTTTCCAGCAG
868 TTTAGTCTGAATGCCATAATAAATTCCTGAGAACAAACGCTGCACCCGGGCAAAACCTCA
808 ACATATAGATGCAAGTGCATCGGGGATGAATGTGTACGTGGAGATTTAAAGTCCCCGCTT
748 CTCAGA
Pax3NCE2-Luc fold activation
n.s.
FLAGCdx1
HAZic2
NCE2-Luc
wt
CdxBS mut
ZicBS mut
Cdx +Zic BS-mut
(1)
(2)
(3)
(4)
C
N2a
P19
MycSox2
D

FIGURE 5


Caudal
Rostral
Pax3
Zic2
Nkx6.1
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O


P
Shh
RT
+ - + - + - + -
Pax3
Zic5
Zic2
Gapdh

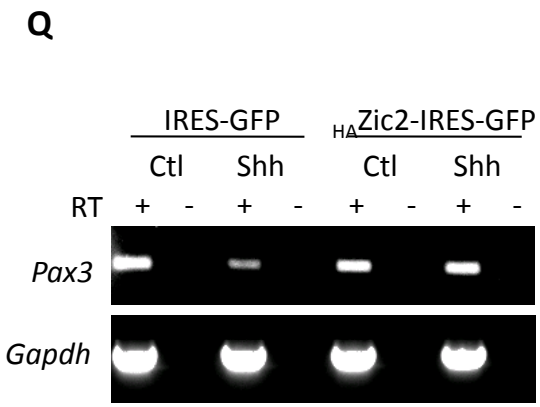

Q
IRES-GFP
HA Zic2-IRES-GFP
Ctl
Shh
Ctl
Shh
RT
+ - + - + - + -
Pax3
Gapdh

FIGURE 6

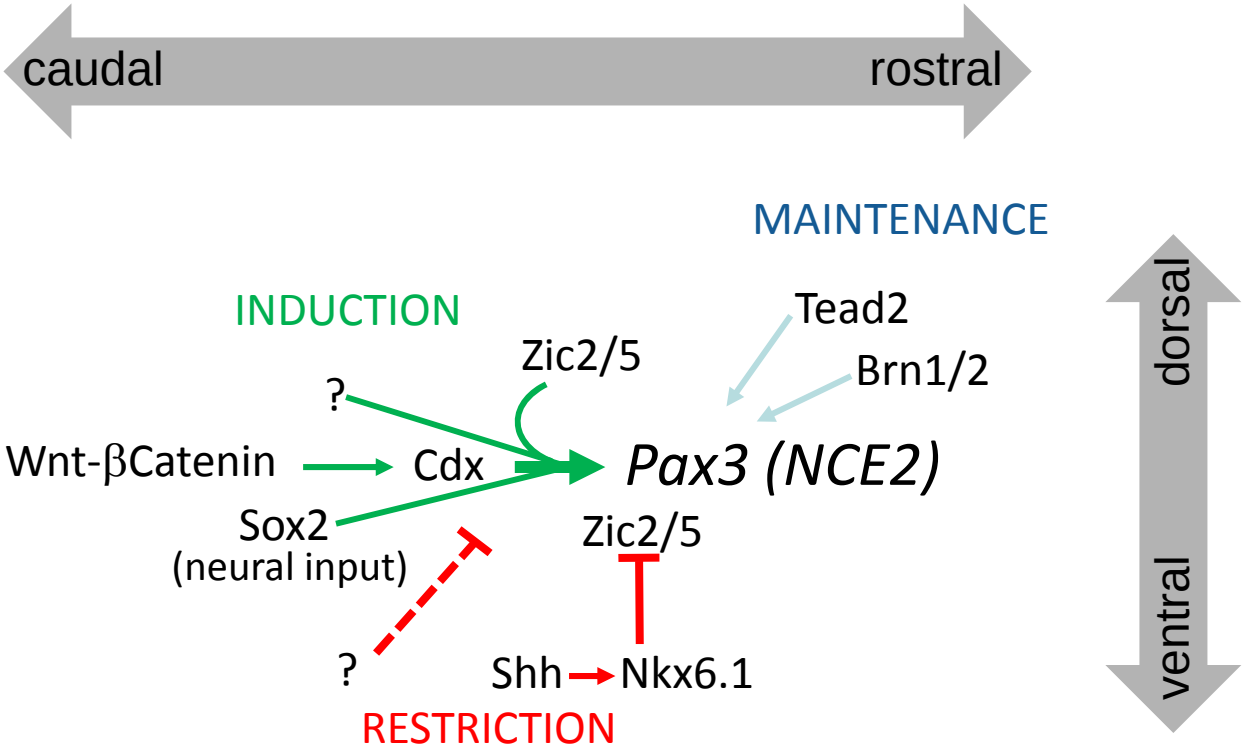

caudal
rostral
MAINTENANCE
INDUCTION
Tead2
Zic2/5
Brn1/2
?
Wnt-βCatenin
Cdx
Pax3 (NCE2)
Sox2
(neural input)
Zic2/5
?
Shh
Nkx6.1
RESTRICTION
dorsal
ventral

FIGURE 7

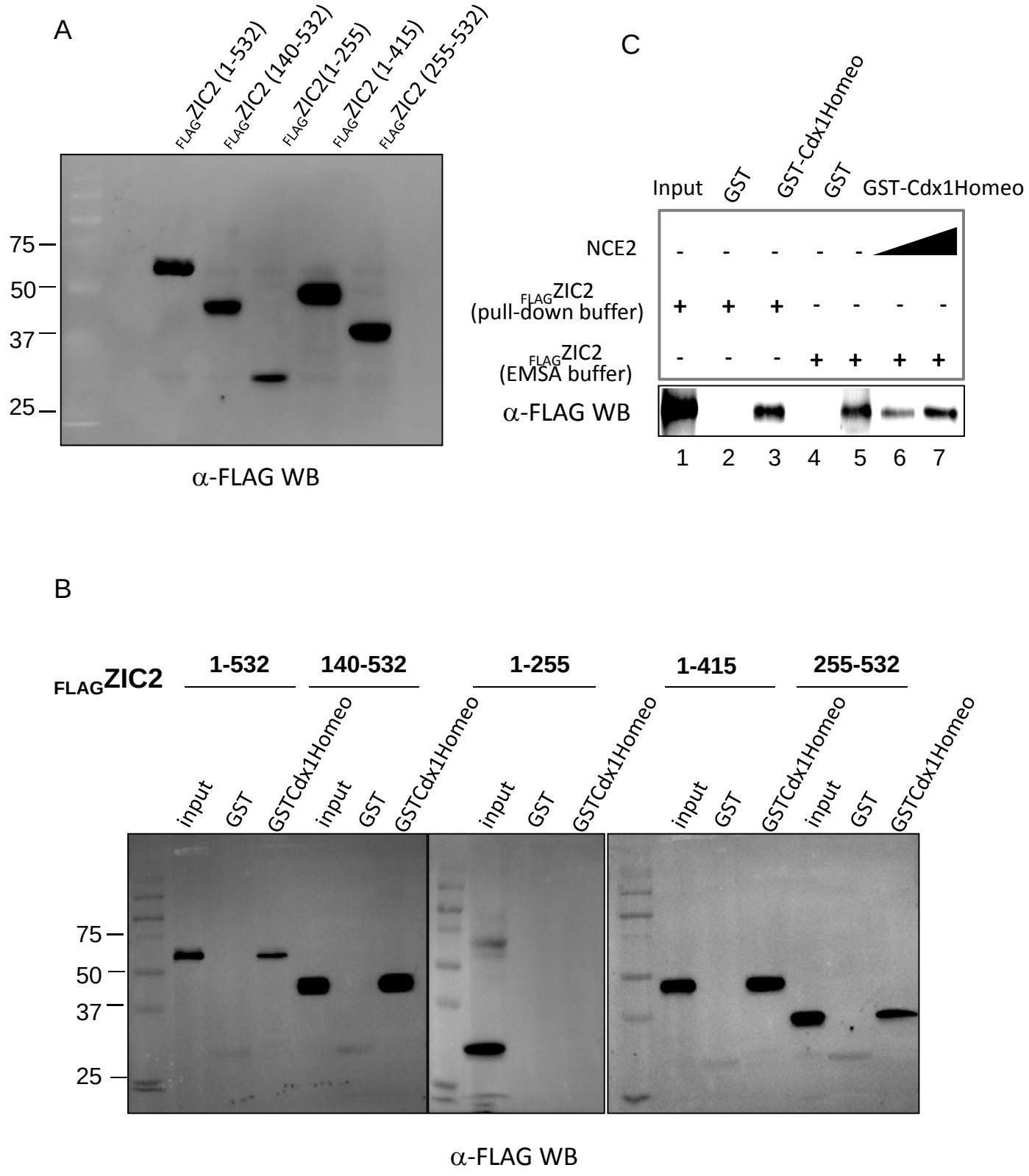

A
FLAGZIC2 (1-532)
FLAGZIC2 (140-532)
FLAGZIC2(1-255)
FLAGZIC2 (1-415)
FLAGZIC2 (255-532)
75
50
37
25
α-FLAG WB
C
Input
GST
GST-Cdx1Homeo
GST
GST-Cdx1Homeo
NCE2
FLAGZIC2 (pull-down buffer)
FLAGZIC2 (EMSA buffer)
α-FLAG WB
1 2 3 4 5 6 7
B
FLAGZIC2
1-532
140-532
1-255
1-415
255-532
input
GST
GSTCdx1Homeo
75
50
37
25
α-FLAG WB

FIGURE S1

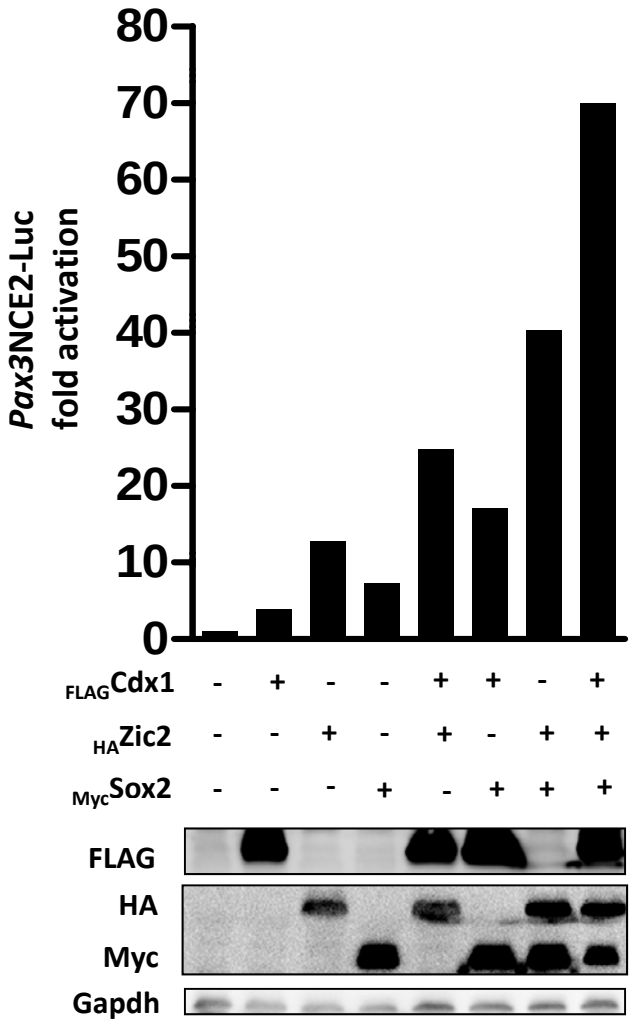
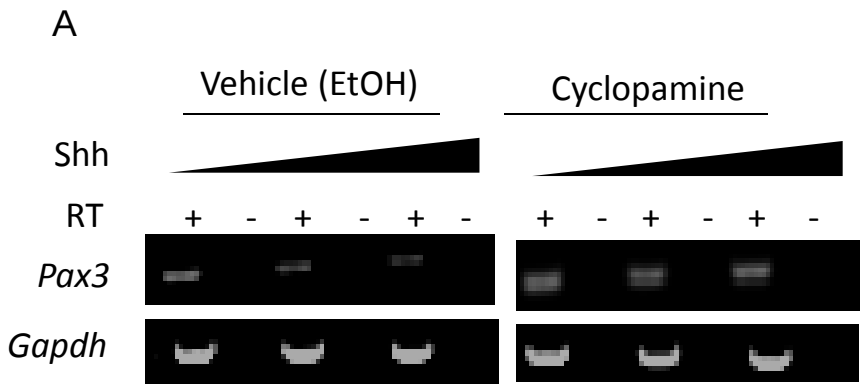

80
70
60
50
40
30
20
10
0
Pax3NCE2-Luc
fold activation
FLAGCdx1 - + - - + + - +
HAZic2 - - + - + - + +
MycSox2 - - - + - + + +
FLAG
HA
Myc
Gapdh

FIGURE S2


A
Vehicle (EtOH)
Cyclopamine
Shh
RT
+ - + - + - + - + - + -
Pax3
Gapdh

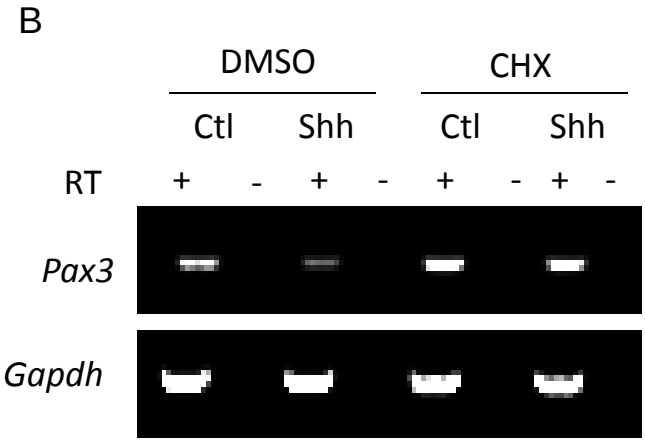

B
DMSO
CHX
Ctl
Shh
Ctl
Shh
RT
+ - + - + - + -
Pax3
Gapdh

FIGURE S3