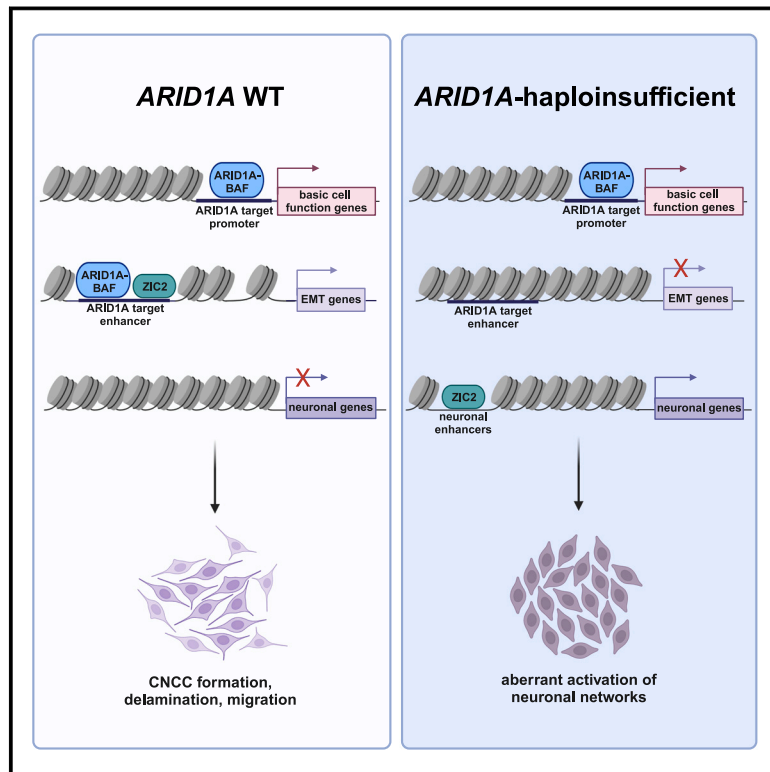


# ARID1A-BAF coordinates ZIC2 genomic occupancy for epithelial-to-mesenchymal transition in cranial neural crest specification

## Graphical abstract



## Authors

Samantha M. Barnada,  
Aida Giner de Gracia,  
Cruz Morenilla-Palao, ...,  
Steven B. McMahon, Eloísa Herrera,  
Marco Trizzino

## Correspondence

[e.herrera@umh.es](mailto:e.herrera@umh.es) (E.H.),  
[m.trizzino@imperial.ac.uk](mailto:m.trizzino@imperial.ac.uk) (M.T.)

**ARID1A modulates chromatin accessibility at enhancers of genes required for epithelial-to-mesenchymal transition, a process essential for cranial neural crest cell (CNCC) specification. Haploinsufficiency of ARID1A attenuates ZIC2 binding at these enhancers, resulting in impaired CNCC formation with an aberrant neuronal trajectory. This study reveals an ARID1A-ZIC2 axis essential for CNCC specification.**



# ARID1A-BAF coordinates ZIC2 genomic occupancy for epithelial-to-mesenchymal transition in cranial neural crest specification

Samantha M. Barnada,<sup>1</sup> Aida Giner de Gracia,<sup>3,10</sup> Cruz Morenilla-Palao,<sup>3,10</sup> Maria Teresa López-Cascales,<sup>3,10</sup> Chiara Scopa,<sup>4,10</sup> Francis J. Waltrich, Jr.,<sup>5</sup> Harald M.M. Mikkers,<sup>6</sup> Maria Elena Cicardi,<sup>4</sup> Jonathan Karlin,<sup>1</sup> Davide Trotti,<sup>4</sup> Kevin A. Peterson,<sup>7</sup> Samantha A. Bruggmann,<sup>8</sup> Gijs W.E. Santen,<sup>9</sup> Steven B. McMahon,<sup>1</sup> Eloísa Herrera,<sup>3,\*</sup> and Marco Trizzino<sup>1,2,\*</sup>

## Summary

The BAF chromatin remodeler regulates lineage commitment including cranial neural crest cell (CNCC) specification. Variants in BAF subunits cause Coffin-Siris syndrome (CSS), a congenital disorder characterized by coarse craniofacial features and intellectual disability. Approximately 50% of individuals with CSS harbor variants in one of the mutually exclusive BAF subunits, *ARID1A/ARID1B*. While *Arid1a* deletion in mouse neural crest causes severe craniofacial phenotypes, little is known about the role of ARID1A in CNCC specification. Using CSS-patient-derived *ARID1A*<sup>+/-</sup> induced pluripotent stem cells to model CNCC specification, we discovered that *ARID1A*-haploinsufficiency impairs epithelial-to-mesenchymal transition (EMT), a process necessary for CNCC delamination and migration from the neural tube. Furthermore, wild-type ARID1A-BAF regulates enhancers associated with EMT genes. ARID1A-BAF binding at these enhancers is impaired in heterozygotes while binding at promoters is unaffected. At the sequence level, these EMT enhancers contain binding motifs for ZIC2, and ZIC2 binding at these sites is ARID1A-dependent. When excluded from EMT enhancers, ZIC2 relocates to neuronal enhancers, triggering aberrant neuronal gene activation. In mice, deletion of *Zic2* impairs NCC delamination, while *ZIC2* overexpression in chick embryos at post-migratory neural crest stages elicits ectopic delamination from the neural tube. These findings reveal an essential ARID1A-ZIC2 axis essential for EMT and CNCC delamination.

## Introduction

The mammalian SWI/SNF complex, referred to as BRG1/BRM associated factor (BAF), is an ATP-dependent chromatin remodeler.<sup>1</sup> BAF plays global roles in lineage specification,<sup>2-4</sup> pluripotency,<sup>5</sup> tumorigenesis,<sup>6</sup> and basic cellular processes<sup>7-9</sup> by modulating chromatin accessibility and interacting with transcription factors<sup>10</sup> to impact gene regulation. BAF comprises three main modules: (1) the ATPase module that hydrolyzes ATP for catalytic activity to alter nucleosome positioning, (2) the ARP module, which aids in ATPase module functions, and (3) the core module required for complex assembly, stabilization, and DNA binding.<sup>11,12</sup> The three BAF subtypes, canonical BAF, polybromo-associated BAF (PBAF), and non-canonical (GLTSCR1/1L-containing) BAF (GBAF)<sup>5,13</sup> include all the main modules in varying configurations that contain interchangeable subunits that confer context-dependent functions.<sup>14-16</sup> Canonical BAF (hereafter “BAF”) is the only subtype that contains the large, mutually exclusive AT-rich interaction domain 1A and 1B (*ARID1A/ARID1B*) subunits in its core module.<sup>11,12,15</sup>

ARID1A and ARID1B are necessary for BAF complex formation, stability, and DNA binding.<sup>12,15,16</sup> *De novo* alterations in *ARID1A* can occur at all stages from germline mutations to somatic mosaicism in embryonic development<sup>17,18</sup> and adult tissue.<sup>6,19</sup> In cancer, ARID1A is reported to have context-dependent oncogenic<sup>20,21</sup> and/or tumor suppressive functions<sup>22-25</sup> as variants in *ARID1A* occur in ~10% of all tumors<sup>6,26</sup> and ~50% of specific cancer types.<sup>24,27</sup> In a developmental setting, heterozygous loss-of-function variants in *ARID1A* (as well as variants in other BAF subunits) are causative of Coffin-Siris syndrome (CSS)<sup>28,29</sup> due to haploinsufficiency. CSS is a rare developmental disorder characterized by systemic congenital anomalies, including distal limb phenotypes, intellectual disability, and coarse craniofacial features.<sup>30-32</sup> These dysmorphic craniofacial features, including a depressed nasal bridge, short nose, averted nares, broad philtrum, and wide mouth,<sup>32</sup> result from impaired craniofacial development.

The process of forming the craniofacial skeleton relies on complex spatiotemporal regulation and patterning and

<sup>1</sup>Department of Biochemistry and Molecular Biology, Thomas Jefferson University, Philadelphia, PA, USA; <sup>2</sup>Department of Life Sciences, Imperial College London, London, UK; <sup>3</sup>Instituto de Neurociencias de Alicante (Consejo Superior de Investigaciones Científicas- Universidad Miguel Hernández, CSIC-UMH). Campus San Juan, Avd. Ramón y Cajal s/n, 03550 San Juan de Alicante, Spain; <sup>4</sup>Jefferson Weinberg ALS Center, Vickie and Jack Farber Institute for Neuroscience, Department of Neuroscience, Thomas Jefferson University, Philadelphia, PA, USA; <sup>5</sup>Department of Pharmacology, Physiology, and Cancer Biology, Thomas Jefferson University, Philadelphia, PA 19107, USA; <sup>6</sup>Department of Cell & Chemical Biology, Leiden University Medical Center, Leiden, the Netherlands; <sup>7</sup>The Jackson Laboratory, 600 Main St, Bar Harbor, ME 04609, USA; <sup>8</sup>Division of Developmental Biology, Department of Pediatrics at Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA; <sup>9</sup>Department of Clinical Genetics, Leiden University Medical Center, Leiden, the Netherlands

<sup>10</sup>These authors contributed equally

\*Correspondence: e.herrera@umh.es (E.H.), m.trizzino@imperial.ac.uk (M.T.)

<https://doi.org/10.1016/j.ajhg.2024.07.022>

© 2024 The Author(s).

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).



includes derivatives of all three germ layers and neural crest cells (NCCs).<sup>33</sup> NCCs are a transient cell population that arise between weeks 3 and 4 of human embryonic development during neurulation.<sup>34</sup> Post-gastrulation, the dorsal portion of the ectoderm begins to specify into neuroectoderm, forming the neural plate that is flanked by the neural plate border.<sup>34</sup> During neurulation, the neural plate invaginates and separates from the dorsal non-neural ectoderm at the neural plate border to form the neural tube.<sup>34</sup> At this neural plate border zone, NCCs undergo epithelial-to-mesenchymal transition (EMT), delaminate, and migrate to populate their terminal sites in the developing embryo.<sup>35–37</sup>

Distinctively, NCCs are multipotent and will form characteristic ectodermal derivatives, such as peripheral neurons, as well as form cell types that are typically mesodermal derived such as cardiac tissue and craniofacial bone and cartilage.<sup>38</sup> Cranial neural crest cell (CNCC)-derived bone and cartilage form the frontal, nasal, zygomatic, maxillary, mandibular, and dentary bones, as well as the bones of the inner ear and the hyoid.<sup>37,39</sup> Moreover, malformations of these bones underlie many of the craniofacial anomalies seen in individuals with CSS, including a depressed nasal bridge, short nose, averted nares, broad philtrum, and wide mouth.<sup>32</sup> Similarly, *Arid1a* conditional knockout models in the neural crest of developing mouse embryos<sup>40</sup> recapitulate these craniofacial phenotypes that are characteristic of CSS. Together, the clinical manifestations of individuals with *ARID1A*-CSS and the craniofacial anomalies in conditional *Arid1a*<sup>+/-</sup> mouse models demonstrate the indispensable role of ARID1A-BAF in neural crest formation and function. However, the molecular pathways regulated by ARID1A during craniofacial development and CNCC specification remain poorly understood.

Herein, we sought to uncover the role of ARID1A-BAF in the specification and migration of CNCCs in the context of craniofacial development. Our data suggests that ARID1A-BAF is required to activate gene networks associated with EMT. Using induced pluripotent stem cells (iPSCs) derived from individuals with *ARID1A*<sup>+/-</sup> CSS, we discovered impaired *ARID1A* correlates with the upregulation of neuronal networks at the expense of EMT networks. Motifs matching the binding site for the ZIC2 transcription factor were highly enriched at the ARID1A-bound genes linked to EMT, and *in vitro* and *in vivo* experiments suggest ARID1A-mediated induction of EMT occurs via ZIC2. These findings highlight an essential axis involving an ARID1A-ZIC2 interaction required for NCC EMT and suggests a pathogenic mechanism that may underlie the *ARID1A*-associated CSS phenotypes.

## Material and methods

### Human iPSC collection, reprogramming, and culture

The CTRL1 iPSC line was obtained from the iPSC Core at the University of Pennsylvania (SV20, male, age 43). Skin fibroblasts were obtained from two pediatric individuals with CSS (CSS1, teenage

male; CSS2, female) by the team of Dr. Gijs Santen at Leiden University (LUMC).

We complied with all relevant ethical regulations for work with human participants and informed consent was obtained by LUMC, which approved the protocol under the coordination of Dr. Gijs Santen. The study was conducted in accordance with the criteria set by the Declaration of Helsinki. The protocol to make iPSCs was approved by the institutional review board (IRB) at LUMC. Collecting patient-derived material and establishing iPSCs were all performed according to local (LUMC) IRB protocols.

The fibroblasts were reprogrammed into iPSCs with the polycistronic lentiviral vector LV.RRL.PPT.SF.hOKSM.idTomato.-preFRT<sup>41</sup> as described elsewhere.<sup>42</sup> Multiple clones per line were derived and one *ARID1A* mutant clone per line was used for this study: CSS1 = LUMC0081iARID02 or 81-2 and CSS2 = LUMC0064iARID04 or 64-4.

Since CSS2 exhibited somatic mosaicism, *ARID1A* wild-type iPSC clones were also isolated from CSS2 (CTRL2; LUMC0064iARID09 or 64-9). For each clone per line, pluripotency was assessed by immunofluorescence microscopy using antibodies against NANOG, OCT3/4, SSEA4, and Tra-1-81 under maintenance conditions and antibodies against TUBB3, AFP, and CD31 after spontaneous differentiation into the three germ layers as described elsewhere.<sup>41,43</sup> Clones with proper pluripotent characteristics were selected for downstream usage. Karyotyping by G banding was assessed for all three lines by the LUMC, and short tandem repeat profiling was performed by the LUMC, and all four lines were profiled at the Stem Cell and Regenerative Neuroscience Center at Thomas Jefferson University. The iPSC lines were expanded in mTeSR1 medium (STEMCELL Technologies) on Matrigel (BD Pharmingen) or Geltrex<sup>TM</sup> LDEV-Free hESC-qualified Reduced Growth Factor Basement Membrane Matrix (Fisher Scientific). Small cell clusters (50–200 cells) were passaged ~1:10 at 80% confluency using 0.5 mM EDTA (STEMCELL Technologies).

### RNA extraction followed by sequencing

Cells were lysed in Tri-reagent (Zymo) and total RNA was extracted using Direct-zol RNA Miniprep kit (Zymo) according to the manufacturer's instructions. RNA was quantified, and RNA integrity was checked by TapeStation 4150 (Agilent). Only samples with an RNA integrity number above 8.0 were used for transcriptome analysis. RNA libraries were prepared using 1 µg of total RNA input using NEB-Next Poly(A) mRNA Magnetic Isolation Module, NEBNext Ultra<sup>TM</sup> II Directional RNA Library Prep Kit for Illumina and NEBNext Multiplex Oligos Dual Index Primers for Illumina (New England BioLabs) according to the manufacturer's instructions (New England Biolabs). Libraries were sequenced on a NextSeq2K (Illumina) generating single-end ~138-bp reads. Each cell line at the specified timepoint(s) were collected from three separate differentiations for emphasis on reproducibility.

### RNA-sequencing analysis

All raw fastq files were analyzed via FastQC (<https://github.com/s-andrews/FastQC>), and adapters were removed with TrimGalore (<https://github.com/FelixKrueger/TrimGalore>). Kallisto<sup>44</sup> was used to quantify reads mapping to each gene. Differential gene expression levels were determined using DESeq2.<sup>45</sup> Briefly, abundance for each replicate of each sample at the specified timepoint were converted to both raw gene count and transcript per million (TPM). The raw gene count for each sample (including replicates) was used to determine differential gene expression with DESeq2

( $p$  value < 0.05 and FDR < 5%) and were grouped based on condition (either CTRL or CSS). The TPMs were used for the violin plots comparing expression levels of individual genes. All statistical analyses were performed using the latest versions of BEDTools,<sup>46</sup> deepTools,<sup>47</sup> R, and Prism.

### **In vitro immunohistochemistry**

Immunohistochemistry (IHC) of iPSCs and CNCCs was performed in  $\mu$ -Slide 4 Well Glass Bottom (IBIDI 80426). Upon fixation (4% formaldehyde for 10 min), cells were permeabilized in blocking solution (0.1% Triton X-100, 1 $\times$  phosphate buffered saline [PBS], 5% normal donkey serum) and then incubated with the antibody of interest. The total number of cells in each field was determined by counterstaining cell nuclei with 4,6-diamidino-2-phenylindole dihydrochloride (DAPI; Sigma-Aldrich; 50 mg/mL in PBS for 15 min at room temperature). Immunostained cells were analyzed via confocal microscopy using a Nikon A1R+. Images were captured with a 4 $\times$  objective for iPSCs and a 60 $\times$  objective for CNCCs with a pinhole of 1.0 Airy unit. Analyses were performed in sequential scanning mode to rule out cross bleeding between channels. Fluorescence intensity and counting quantifications were performed with Fiji (<https://fiji.sc>) and the NIS-Elements AR software ([https://www.microscope.healthcare.nikon.com/en\\_AOM/products/software/nis-elements/nis-elements-advanced-research](https://www.microscope.healthcare.nikon.com/en_AOM/products/software/nis-elements/nis-elements-advanced-research)). In detail, to analyze the net intensity of the cells, we created a mask around each nucleus using DAPI intensity as the base criterion. This mask served as the reference for measuring the fluorescence intensity or number of cells positive for nuclear proteins. To analyze cytoplasmic proteins, the previously drawn nucleus mask was expanded to encompass the entire cytoplasm. All antibodies are listed in Table S1. All two-tailed unpaired t tests were performed in Prism.

### **CNCC specification**

The iPSC lines were specified into CNCC as previously described.<sup>3,48</sup> In brief, iPSCs were cultured with CNCC specification media: 1:1 Neurobasal medium/D-MEM F-12 medium (Invitrogen), 1 $\times$  penicillin-streptomycin solution, 1 $\times$  Glutamax supplement (100 $\times$  stock, Invitrogen), 0.5 $\times$  B-27 supplement with vitamin A (50 $\times$  stock, Invitrogen), 0.5 $\times$  N-2 supplement (100 $\times$  stock, Invitrogen), 5  $\mu$ g/mL bovine insulin (Sigma-Aldrich), 20 ng/mL EGF (epidermal growth factor; Sigma-Aldrich), and 20 ng/mL hFGF (human fibroblast growth factor; Biolegend) for 6 days. Medium was changed every 1–2 days. At day ~7 when early migratory CNCCs first appeared, the cells were transitioned to CNCC early maintenance media: 1:1 Neurobasal medium/D-MEM F-12 medium (Invitrogen), 1 $\times$  penicillin-streptomycin solution, 1 $\times$  Glutamax supplement (100 $\times$  stock, Invitrogen), 0.5 $\times$  B-27 supplement with vitamin A (50 $\times$  stock, Invitrogen), 0.5 $\times$  N-2 supplement (100 $\times$  stock, Invitrogen), 1 mg/mL bovine serum albumin, serum replacement grade (Gemini Bio-Products # 700-104 P), 20 ng/mL EGF (Sigma-Aldrich), and 20 ng/mL hFGF (Biolegend) for 7 days. After 14 full days of CNCC specification, the cells were maintained in an “early CNCC” state for up to day 21 (seven additional days) for subsequent experiments, or immediately transitioned to CNCC late maintenance media: 1:1 Neurobasal medium/D-MEM F-12 medium (Invitrogen), 1 $\times$  penicillin-streptomycin solution, 1 $\times$  Glutamax supplement (100 $\times$  stock, Invitrogen), 0.5 $\times$  B-27 supplement with vitamin A (50 $\times$  stock, Invitrogen), 0.5 $\times$  N-2 supplement (100 $\times$  stock, Invitrogen), 1 mg/mL bovine serum albumin, serum replacement grade (Gemini Bio-Products # 700-104 P), 20 ng/mL EGF (Sigma-Aldrich), 20 ng/mL hFGF (Biolegend), 3  $\mu$ M ChIRON 99021 (Selleck Chem

S1263), and 50 ng/mL bone morphogenetic protein 2 (BMP2) (Peprotech 120-02). CNCCs transitioned to late maintenance media can be maintained for up to two additional weeks.

### **Flow cytometry**

CTRL and CSS patient-derived iPSCs were treated with accutase for 5 min to obtain a single-cell suspension. Cells were then washed with cold 1 $\times$  PBS + 2% fetal bovine serum (FBS), and live cells were counted and resuspended in 1 $\times$  PBS to 1  $\times$  10<sup>6</sup> cells/mL. LIVE/DEAD aqua stain (ThermoFisher L34965) was prepared following the manufacturer’s instructions, and 1  $\mu$ l of prepared stain was added per 1  $\times$  10<sup>6</sup> cells and incubated for 30 min at 4°C in the dark. Cells were then washed twice in 1 $\times$  PBS + 2% FBS. Cells were resuspended in up to 100  $\mu$ l 1 $\times$  PBS +2% FBS and 2  $\mu$ l of respective antibodies (1:50 dilution; Table S1) and incubated for 15 min at 4°C in the dark. The antibodies were removed, and the cells were washed once in 1 $\times$  PBS + 2% FBS. Stained cells were resuspended in 300  $\mu$ l of PBS + 2% FBS and strained through a 35  $\mu$ m strainer. Flow cytometry data were acquired using a BD Celesta and analyzed with FlowJo Software v10.10.

### **Western blotting**

Cells were washed three times in 1 $\times$  PBS and lysed in radioimmunoprecipitation assay buffer (RIPA buffer; 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% Igepal, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulfate [SDS] 500  $\mu$ M DTT) with protease inhibitors. Whole-cell lysate was loaded in Novex WedgeWell 4%–20% Tris-Glycine Gel (Invitrogen) and subject to SDS-PAGE. Proteins were then transferred to an Immobilon-P polyvinylidene fluoride or Nitrocellulose membrane (Thermo Fisher) for antibody probing. Membranes were blocked in a 10% BSA/Tris-buffered saline with Tween (TBST) solution for 30 min then incubated with diluted primary antibodies (Table S1) in a 5% BSA/TBST solution. Membranes were then washed with TBST and incubated with diluted secondary antibodies (Table S1). Chemiluminescent signal was detected using the KwikQuant Western Blot Detection Kit (Kindle BioSciences) or the Amersham ECL Prime Western Blotting Detection Reagents (Cytiva) and a KwikQuant Imager.

### **Transwell assay**

The transwell assay was performed using the Cell Migration Assay Kit (abcam ab235694) following the manufacturer’s protocol. Briefly, CNCCs were cultured to ~80% confluency and growth factors (EGF and FGF) were removed from the CNCC early maintenance media for 24 h to prevent proliferation. The bottom chamber of the migration plate was prepared with growth-factor-free CNCC early media and 10% FBS, and the top chamber was assembled. Cells were harvested and counted, and 200,000 cells were resuspended in growth-factor-free CNCC early maintenance media and seeded in the top chamber. Cells were incubated in the migration chamber at 37°C for 48 h. After 48 h, the cells that invaded through the migration membrane were dissociated and incubated in cell dye for 1 h at 37°C. One set of standards per cell type was created via serial dilution and incubated in cell dye for 1 h at 37°C. Absorbance for each sample and standard was read at Ex/Em-530/590 nm using a PolarStar Optima plate reader (BMG LabTech). Percent migration was calculated from the linear curve of the plotted standards.

### **Scratch wound assay**

CNCCs were cultured to ~80% confluency, and growth factors (EGF and FGF) were removed from the CNCC early maintenance

media for 24 h to prevent proliferation. Cells were harvested and counted, and 100,000 cells were resuspended in growth-factor-free CNCC early maintenance media and seeded in an IncuCyte ImageLock 96-well plate (Sartorius 4379) in triplicate and incubated for 48 h at 37°C. The IncuCyte 96-Well Woundmaker Tool was used to create identical wounds in each well. The wells were washed twice to remove the cells dislodged from the Woundmaker Tool, and the media was refreshed. The cells were incubated at 37°C, and migration was assayed for 96 h with repeat scanning every 2 h via the IncuCyte ZOOM Live-Cell Analysis System. Average percent confluence and wound density was quantified and normalized relative to each individual cell count per replicate.

### ChIP sequencing

All samples from different conditions were processed together to prevent batch effects. Approximately 13 million cells were cross-linked with 1% formaldehyde for 5 min at room temperature, quenched with 125 mM glycine, harvested, and washed twice with 1× PBS. The fixed cell pellet was resuspended in chromatin immunoprecipitation (ChIP) lysis buffer (150 mM NaCl, 1% Triton X-100, 0.7% SDS, 500 μM DTT, 10 mM Tris-HCl, 5 mM EDTA), and chromatin was sheared to an average length of 200–900 base pairs using a Covaris S220 Ultrasonicator at 5% duty factor between 6 and 9 min. The chromatin lysate was diluted with SDS-free ChIP lysis buffer. 15 μg of antibody was used for ARID1A and ZIC2 and 3 μg of antibody for H3K27ac (Table S1). The antibody was added to at least 5 μg of sonicated chromatin along with Dynabeads Protein G magnetic beads (Invitrogen) and incubated with rotation at 4°C overnight. The beads were washed twice with each of the following buffers: Mixed Micelle Buffer (150 mM NaCl, 1% Triton X-100, 0.2% SDS, 20 mM Tris-HCl, 5 mM EDTA, 65% sucrose), Buffer 200 (200 mM NaCl, 1% Triton X-100, 0.1% sodium deoxycholate, 25 mM HEPES, 10 mM Tris-HCl, 1 mM EDTA), LiCl detergent wash (250 mM LiCl, 0.5% sodium deoxycholate, 0.5% NP-40, 10 mM Tris-HCl, 1 mM EDTA), and a final wash was performed with cold 0.1× Tris-EDTA (TE). Finally, beads were resuspended in 1× TE containing 1% SDS and incubated at 65°C for 10 min to elute immunocomplexes. The elution was repeated twice, and the samples were incubated overnight at 65°C to reverse cross linking along with the input (5% of the starting material). The DNA was digested with 0.5 mg/mL proteinase K for 1 h at 65°C and then purified using the ChIP DNA Clean & Concentrator kit (Zymo) and quantified with QUBIT. Bar-coded libraries were made with the NEBNext Ultra II DNA Library Prep Kit for Illumina using NEBNext Multiplex Oligos Dual Index Primers for Illumina (New England BioLabs) and sequenced on NextSeq2K (Illumina) producing ~138 bp single-end reads.

### ATAC sequencing

For assay for transposase-accessible chromatin with sequencing (ATAC-seq) experiments, 50,000 cells per condition were processed as previously described.<sup>49</sup> Briefly, 50,000 cells were collected, washed, and lysed. The chromatin was subjected to transposition/library preparation via a Tn5 transposase using the Tagment DNA Enzyme and Buffer Kit (Illumina 20034197) and incubated at 37°C for 30 min with slight agitation. Transposed DNA was purified using a MinElute PCR Purification Kit (Qiagen 28004). Transposed DNA fragments were then amplified using a universal and barcoded primer.<sup>49</sup> Thermal cycling parameters were set as follows: 1 cycle of 72°C for 5 min and 98°C for 30 s, followed by 5 cycles of 98°C for 10 s, 63°C for 30 s, and 72°C for

1 min. The amplification was paused, and 5 μL of the partially amplified, transposed DNA was used for a qPCR side reaction, including the universal and sample-specific barcoded primers,<sup>49</sup> PowerUp SYBR Green Master Mix (Applied Biosystems), NEBNext High-Fidelity 2× PCR Master Mix, and nuclease-free water. The qPCR side reaction parameters were set as follows: 1 cycle of 72°C for 5 min and 98°C for 30 s, followed by 40 cycles of 98°C for 10 s, 63°C for 30 s, and 72°C for 1 min. The Rn vs. cycle was plotted to determine the remaining number of PCR cycles needed where 1/3 of the maximum fluorescent intensity corresponds to the additional cycle number. The remaining partially amplified transposed DNA was fully amplified using the previous parameters with the additional cycle number determined from the qPCR side reaction. The amplified, transposed DNA was purified using AMPure XP beads (Beckman Coulter A63881) and sequenced using an Illumina NextSeq2K, generating ~138 bp single-end reads.

### ChIP- and ATAC-seq analysis

After removing the adapters with TrimGalore, the sequences were aligned to the human reference genome hg19 using the Burrows-Wheeler Alignment tool with the MEM algorithm.<sup>50</sup> Aligned reads were filtered based on mapping quality (MAPQ > 10) to restrict our analysis to higher quality and uniquely mapped reads, and PCR duplicates were removed. For ATAC-seq analysis, all mapped reads were offset by +4 bp for the forward strand and –5 bp for the reverse strand. MACS2<sup>51</sup> was used to call peaks using the default parameters at 5% FDR. All statistical analyses were performed using the latest versions of BEDTools,<sup>46</sup> deepTools,<sup>47</sup> R, and Prism. Peaks were visualized in Integrative Genomics Viewer (<https://igv.org>) on genome build hg19.

### Motif analysis

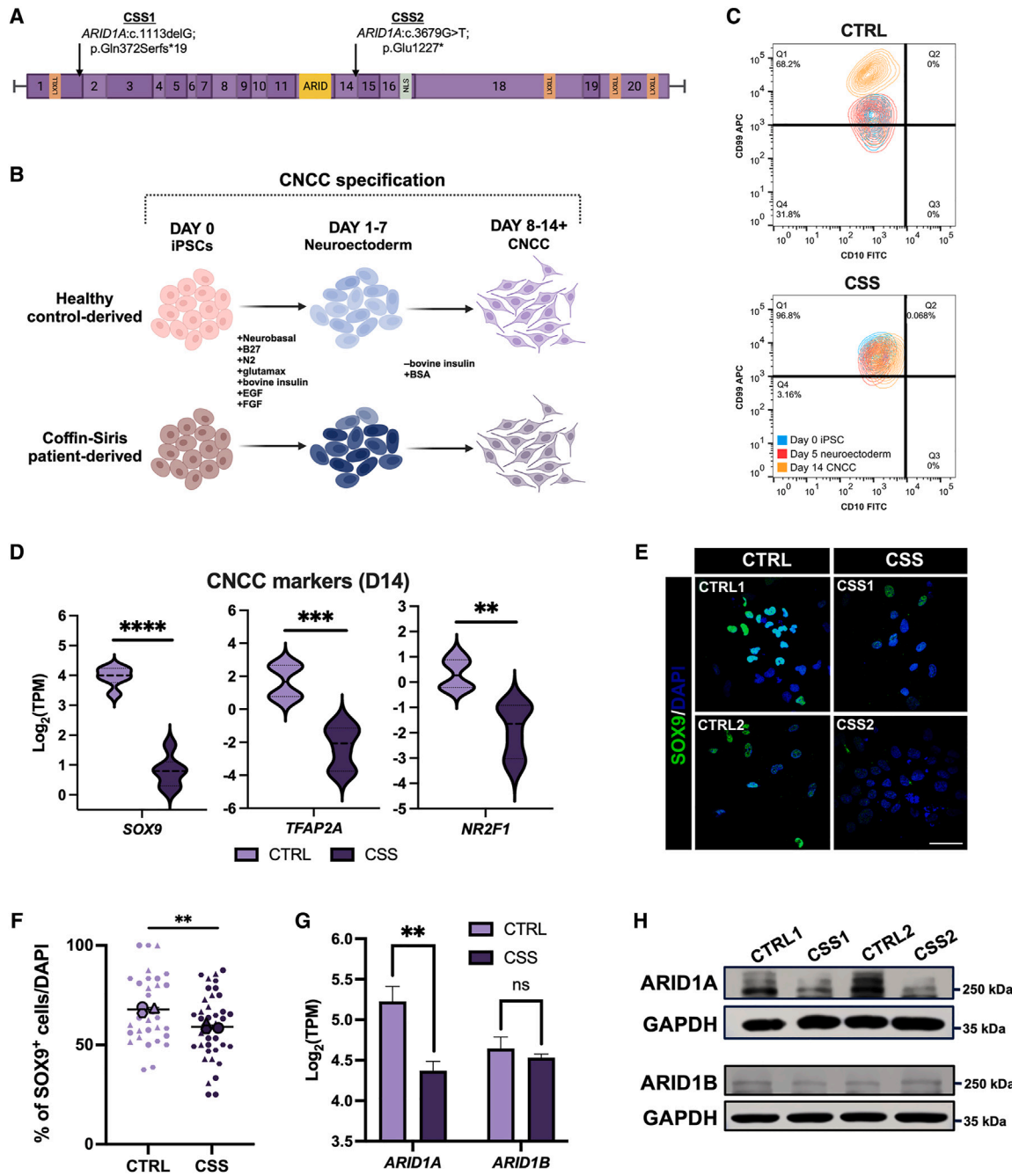
Fasta files for the coordinates of interest were produced using BEDTools.<sup>46</sup> Motif analysis of all lost ARID1A-bound regions, ARID1A-bound promoters, ARID1A “gained” regions, and ZIC2 relocated regions was performed using HOMER.<sup>52</sup> Shuffled input sequences were used as background. E values <0.001 were used as a threshold for significance. Transcription factor motif quantification was performed using FIMO<sup>53</sup> (The MEME Suite) on a background of lost ARID1A-bound regions.

### Mouse lines

The B6;FVB-Tg(*Zic2*-EGFP)HT146Gsat/Mmucd line (identification number RP23-158G6)<sup>54</sup> was generated by GENSAT<sup>55</sup> and obtained from the Mutant Mouse Regional Resource Center ([https://www.mmrc.org/catalog/sds.php?mmrc\\_id=30037](https://www.mmrc.org/catalog/sds.php?mmrc_id=30037)). Hereafter, this line is known as Tg(*Zic2*<sup>EGFP</sup>). The B6.129S4-*Zic2*<sup>tm1Jaru</sup>/JaruRbrc strain (Mouse Genome Informatics: 2156825; RBRC00165)<sup>56</sup> was obtained from the RIKEN repository. Hereafter, this line is known as *Zic2*<sup>kd</sup>. *Zic2*<sup>+/+</sup>;Tg(*Zic2*<sup>EGFP</sup>) and *Zic2*<sup>kd/kd</sup>;Tg(*Zic2*<sup>EGFP</sup>) embryos were obtained from breeding *Zic2*<sup>+/-</sup>;Tg(*Zic2*<sup>EGFP</sup>) mice and are referred to as *Zic2*-WT and *Zic2*-mutant, respectively. The day a vaginal plug appeared was considered embryonic day E0.5. All mouse lines were congenic on a C57BL/6J background and were kept in a timed pregnancy-breeding colony at the Instituto de Neurociencias. The animal protocols were approved by the Institutional Animal Care and Use Committee and met European and Spanish regulations.

### Constructs and *in vivo* electroporation

Fertilized White Leghorn chicken eggs were incubated at 38°C until the desired developmental stage, according to Hamburger and



**Figure 1. *ARID1A*-haploinsufficient cells fail to specify as CNCCs**

(A) Schematic of the *ARID1A* gene (GenBank: NM\_006015.6) highlighting the two nonsense mutations in each CSS cell line. Each number represents a corresponding exon in *ARID1A*. LXXLL, protein-interaction domain; ARID, AT-rich interaction domain or DNA-binding domain; NLS, nuclear localization signal. Made with [BioRender.com](https://BioRender.com).

(B) Graphical illustration of 2D *in vitro* CNCC specification. EGF, epidermal growth factor; FGF, fibroblast growth factor; BSA, bovine serum albumin. Made with [BioRender.com](https://BioRender.com).

(C) CSS lines at day 14 (D14) fail to show a distinctive increase in the CD99 surface marker representative of successful CNCC specification. All populations (D0 iPSC, D5 neuroectoderm, D14 CNCC) are CD10<sup>+</sup> with a slight increase in the CD10 surface marker at D14, which is characteristic of preparation of a differentiation trajectory toward the mesenchymal lineage.

(D) Violin plots displaying  $\log_2(\text{TPM})$  of CNCC markers *SOX9*, *TFAP2A*, and *NR2F1* at D14 between CTRL and CSS cell lines. A two-tailed unpaired t test was performed and  $p < 0.05$  was considered significant; *SOX9*: \*\*\*\* $p < 0.0001$ , *TFAP2A*: \*\*\* $p = 0.0001$ , and *NR2F1*: \*\* $p = 0.0012$ .

(E and F) (E) Representative images and (F) quantification of a *SOX9* immunofluorescence performed at D14 in CTRL and CSS cell lines. (E) DAPI staining on nuclei in blue. Merged images shown taken at 60 $\times$  magnification; scale bar, 50  $\mu\text{m}$ . (F) SuperPlot quantification of percentage of *SOX9*<sup>+</sup> cells per DAPI.  $n = 3$  represented by a distinctive shape with each small data point representing a captured image. The larger data points correspond to the average values of each replicate. A two-tailed unpaired t test performed on the average values between CTRL and CSS cell lines;  $p < 0.05$  was considered significant; \*\* $p = 0.0032$ .

(legend continued on next page)

Hamilton.<sup>57</sup> Platinum electrodes were used with 5 × 10 ms pulses at 25 V generated with a TSS20 Ovodyne electroporator. The following plasmids were injected at the indicated concentrations: pCAG-ZIC2 plasmid (1 µg/µL) and CAG-GFP (0.5 µg/µL). Due to technicalities underlying electroporation, all analyses of electroporated chicken embryos were confined to the trunk region.

### In vivo IHC and iDISCO

For IHC, chick embryos fixed overnight at 4°C, and mouse embryos fixed for 4–5 h (E8.5–E9.5) with 4% paraformaldehyde (PFA)/PBS and were washed with PBS. Cryostat (40 µm) and vibratome (70 µm) sections or whole embryos were incubated with the corresponding primary and secondary antibodies, according to standard protocols. For immunolabeling-enabled three-dimensional imaging of solvent-cleared organs (iDISCO), embryos were fixed overnight at 4°C with 4% PFA/PBS, and immunolabeling was performed before clarification. The iDISCO protocol was performed as previously published.<sup>58</sup> IHC images were captured with an Olympus FV1000 confocal IX81 microscope/FV10-ASW Software. Images of iDISCO-clarified embryos were obtained with an Ultramicroscope II (LaVisionBiotec). All antibodies are listed in Table S1.

### In vivo IHC analysis

The number of Sox10<sup>+</sup> or FoxD3<sup>+</sup> cells per slice was quantified and normalized by neural tube length and the ratio of positive cells inside the tube compared with the total number of cells in each cryosection (30 µm). A region of interest was delineated in the migration zone outside the neural tube, and the fluorescence intensity was measured (IntDen parameter in Fiji/ImageJ) and normalized to total fluorescence. Two complementary regions of interest were delineated along the migratory stream (one in the ventromedial area and another in the dorsolateral area). GFP fluorescence intensity was measured, and the total fluorescence area was calculated as integrated density (fluorescence of the selected area – mean background fluorescence) in Fiji. All two-tailed unpaired t tests were performed in Prism.

### Statistical and genomic analysis

All statistical analyses were performed using the latest versions of BEDTools,<sup>46</sup> deepTools,<sup>47</sup> R, and Prism. Superplot quantification represents at least  $n = 3$  represented by a distinctive shape with each small data point representing a captured image. The larger data points correspond to the average values of each replicate. A two-tailed unpaired t test performed on the average values. A value of  $p < 0.05$  was considered significant; \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ; \*\*\*\* $p < 0.0001$ ; and n.s., not significant. Error bars represent standard error of the mean. Pathway analysis was performed with Ingenuity-Pathway Analysis Suite (IPA; QIAGEN Inc., <https://digitalinsights.qiagen.com/IPA>).

## Results

### Control and CSS patient-derived iPSCs are pluripotent

We leveraged control and *ARID1A* variant-containing, human-derived iPSCs, to investigate the role of *ARID1A* in the

BAF chromatin remodeling complex during cranial neural crest specification and craniofacial development. To generate these iPSCs, skin fibroblasts were collected from two individuals with CSS each harboring distinct, previously reported heterozygous, *de novo*, nonsense mutations in *ARID1A*<sup>17</sup> (MIM: 614607). CSS patient 1-derived cells (CSS1) originated from a teenage male harboring a *ARID1A*:c.1113delG (p.Gln372Serfs\*19) variant and CSS patient 2-derived cells (CSS2) originated from a female harboring a *ARID1A*:c.3679G>T (p.Glu1227\*) variant (Figure 1A; GenBank: NM\_006015.6). Both variants were classified as pathogenic based on the criteria set forth by the American College of Medical Genetics and Genomics.<sup>59</sup> A subsequent *ARID1A* wild-type iPSC line was used as a control (CTRL1). While germline mutations in *ARID1A* lead to a severe CSS subtype with multiple congenital anomalies and reduced survival, mosaic variants are often observed in individuals with a more moderate CSS phenotype.<sup>60</sup> The *ARID1A*-haploinsufficient CSS fibroblasts (CSS1 and CSS2) were reprogrammed into iPSCs (Figure S1A), and due to the associated mosaicism, it was possible to isolate *ARID1A* wild-type iPSCs (CTRL2) in addition to the *ARID1A*-haploinsufficient iPSCs from CSS2 for an isogenic study system. Genotypes of the resulting iPSC lines were verified via deep sequencing.

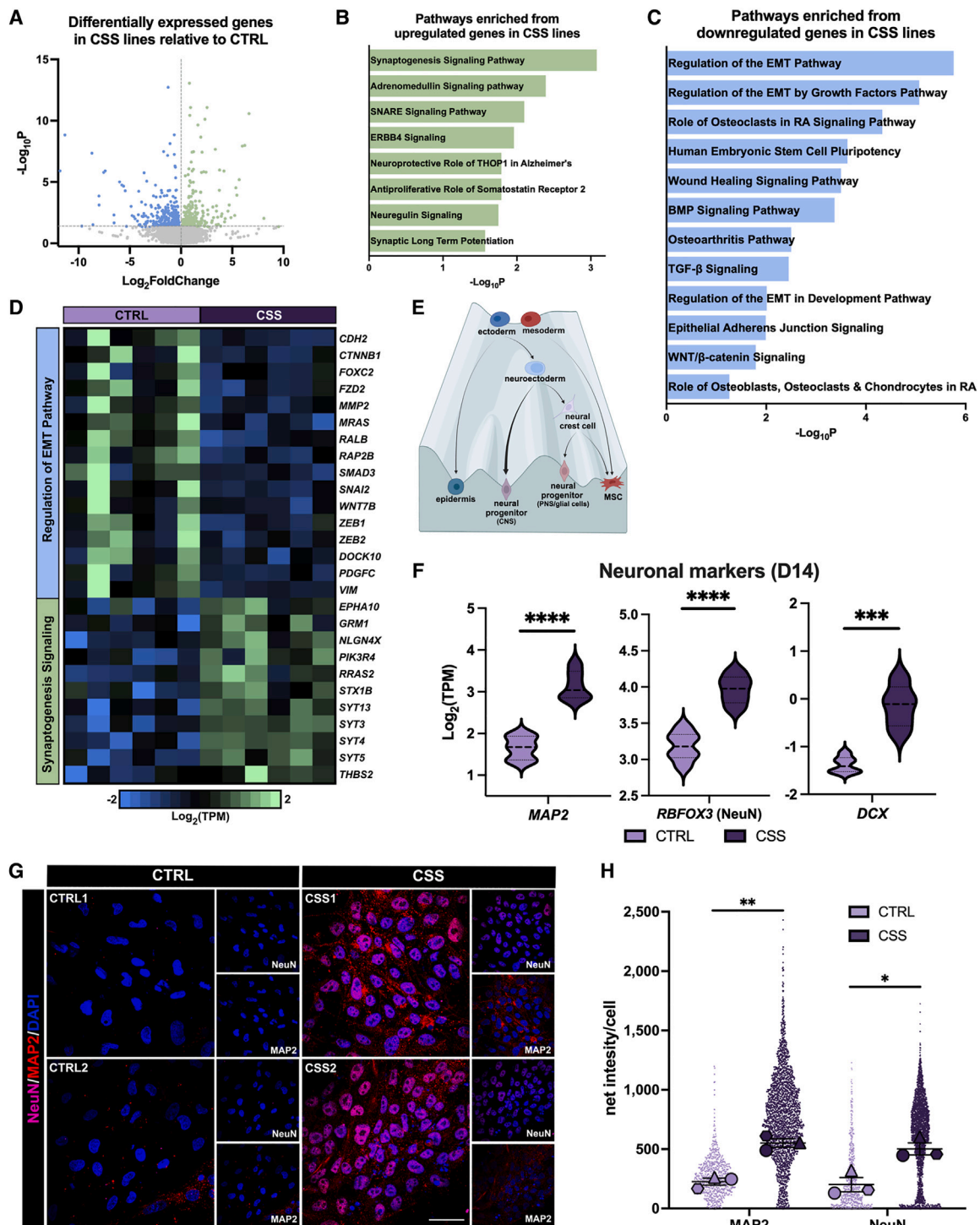
Previous studies from our lab<sup>3</sup> and others<sup>61,62</sup> have shown that *ARID1A*-containing BAF is present in pluripotent stem cells (esBAF; Figure S1B). Additionally, GBAF (non-canonical BAF) is an alternate embryonic stem cell-specific configuration of BAF that lacks *ARID* subunits and instead incorporates *GLTSCR1/GLTSCR1L* and *BRD9*.<sup>5,13</sup> Notably, GBAF is also present in embryonic stem cells where it is implicated in regulating pluripotency.<sup>5</sup> Consistent with the lack of *ARID* subunits in GBAF, the *ARID1A*-haploinsufficient iPSCs display regular iPSC morphology, including small cells making up distinct, circular, compact colonies with well-defined colony borders (Figure S1C) and broadly express *OCT4*, *SOX2*, and *NANOG* mRNA with comparable protein levels (Figures S1D–S1F). RNA-seq of iPSCs identified only 121 differentially expressed genes between CTRL and CSS cells ( $p < 0.05$ ; FDR < 5%;  $\log_2(\text{fold change}) \pm 1$ ; Table S2). Notably, genes involved in pluripotency regulation such as *OCT4*, *SOX2*, and *NANOG* were not differentially expressed (Table S2). These data show that CSS patient-derived iPSCs maintain pluripotency and *ARID1A*-haploinsufficiency did not impact pluripotent gene expression in this cell type.

### CSS cranial NCCs exhibit migratory defects

As premature loss of pluripotency did not appear to be the pathological mechanism of *ARID1A* haploinsufficiency, we

---

(G and H) (G) Log<sub>2</sub>(TPM) bar plot and (H) immunoblot of *ARID1A* and *ARID1B* at day 10 of CNCC specification between CTRL and CSS cell lines. Reduction of *ARID1A* transcript and protein levels correspond to the *ARID1A*-haploinsufficient genotype while *ARID1B* levels are consistent across CTRL and CSS cell lines. A two-tailed unpaired t-test was performed in (G) and  $p < 0.05$  was considered significant; *ARID1A*: \*\* $p = 0.0029$ , *ARID1B*: ns  $p = 0.4592$ . Error bars represent the standard error of the mean.



**Figure 2. CSS patient-derived cell lines do not undergo EMT and default to the neuronal lineage**

(A) Volcano plot of differentially expressed genes in CSS cell lines relative to CTRLs at day 10 of CNCC specification as determined by DESeq2.<sup>45</sup> Green points represent upregulated genes while blue points represent downregulated genes.  $n = 652$ .

(B) Neuronal pathways are enriched from differentially upregulated genes in CSS cell lines at day 10 of CNCC specification and the respective  $-\log_{10} p$  value as determined by IPA.

(C) Pathways enriched from differentially downregulated genes in CSS cell lines at day 10 of CNCC specification and the respective  $-\log_{10} p$  value as determined by IPA. Identified downregulated pathways are essential for successful CNCC specification, including epithelial-to-mesenchymal transition (EMT), BMP, TGF- $\beta$ , and Wnt signaling. RA, rheumatoid arthritis.

(D) Heatmap of genes involved in the most significant upregulated and downregulated pathways presented in (B) and (C) between CTRL and CSS cell lines. CTRL and CSS columns represent three individual replicates from three separate CNCC specifications for CTRL1/CTRL2 and CSS1/CSS2, respectively.

(legend continued on next page)



next investigated the role of ARID1A in cranial neural crest specification. CTRL and CSS iPSCs were specified to CNCCs as per previously published protocols<sup>48,63,64</sup> (Figure 1B). RNA-sequencing analysis comparing CTRL iPSCs (Table S3) and derived CNCCs (Table S4) revealed a transcriptional signature in which CNCC markers were highly expressed while pluripotency genes were downregulated (Figure S1G), supporting successful CNCC specification (Figure S1H).

Previous work from our group demonstrated a bimodal switch between ARID1A and ARID1B in the BAF complex during the CNCC specification process<sup>3</sup> (Figure S1B). At the onset of neuroectoderm formation, ARID1A protein levels are downregulated and nearly undetectable while ARID1B protein levels increase<sup>3</sup> (Figure S1B). ARID1B levels remain elevated throughout the duration of the neuroectodermal stage (days 1–7) but are then downregulated during the onset of EMT and CNCC formation. At this stage, ARID1A protein levels are re-upregulated. This latter switch occurs at day 8–9 of our CNCC specification protocol, and ARID1A is retained in the BAF complex as migratory CNCCs are formed<sup>3</sup> (Figure S1B). Notably, this time point immediately precedes EMT and delamination of the CNCCs as they acquire a migratory phenotype.

To investigate the impact ARID1A haploinsufficiency has on CNCC identity, we harvested specified CNCCs on day 14. At this stage, CTRL cells displayed characteristic single-cell CNCC morphology, while CSS patient-derived cells adopted a more elongated and clustered cellular morphology (Figure S1I). By day 14, CSS cell lines failed to form the distinct CD99<sup>+</sup> CNCC population, which is characteristic of successful specification (Figures 1C and S2) as observed by flow cytometry. Contrary to results observed in iPSCs, RNA sequencing performed in CTRL and CSS-derived CNCCs at day 14 revealed over 3,800 differentially expressed genes ( $p < 0.05$ ; FDR < 5%;  $\log_2(\text{fold change}) +/ - 1$ ; Table S5). Notably, genes associated with CNCC specification and migration (i.e., *SOX9*, *TFAP2A*, and *NR2F1*) were significantly decreased in the CSS cells relative to CTRL CNCCs (Figure 1D). Immunostaining at day 14 validated the transcriptomic data showing a reduced number of SOX9<sup>+</sup> CSS cells relative to CTRLs (Figures 1E and 1F).

As ARID1A levels are upregulated by day 8–9 of our *in vitro* CNCC specification program,<sup>3</sup> we chose day 10 for comprehensive genomic studies, focused on the role of ARID1A during CNCC formation. First, we wanted to elucidate if other ARID proteins (i.e., ARID1B or the PBAF member ARID2) could compensate for ARID1A haploinsufficiency. As expected, ARID1A protein levels were significantly diminished in the CSS lines at day 10 of CNCC formation (Figures 1G and 1H). However, neither ARID1B nor ARID2 were upregulated at the transcript (Figures 1G and S1J) or protein level (Figure 1H) at the same time point, suggesting other BAF configurations could not compensate for the partial loss of ARID1A during CNCC formation. Together, these data implicated ARID1A-BAF as a key regulator of CNCC formation.

#### ARID1A-haploinsufficiency results in decreased expression of EMT genes and increased expression of neuronal genes

In addition to the 3,800 genes differentially expressed at the end of the CNCC specification protocol (day 14; Table S5), RNA-sequencing analyses performed at day 10 of CNCC specification (i.e., at the onset of ARID1A re-upregulation) identified 652 differentially expressed genes in the CSS lines compared to CTRLs (Figure 2A; Table S6). Of these 652 genes, 349 were downregulated and 303 were upregulated in CSS lines. Downregulated genes were significantly enriched in pathways associated with EMT while the upregulated genes were significantly enriched in neuronal pathways (Figures 2B–2D). Notably, both neural progenitors and NCCs are derived from the neuroectoderm (Figure 2E). We previously demonstrated that this developmental precursor was regulated by a specific version of BAF incorporating ARID1B rather than ARID1A.<sup>3</sup> However, the ARID1A-haploinsufficient lines failed to upregulate EMT and CNCC genes while aberrantly upregulating neuronal genes. Thus, the partial loss of ARID1A in CSS cells appeared to impair formation of the multipotent CNCC lineage from neuroectoderm, triggering a default differentiation to the neuronal lineage (Figure 2E).

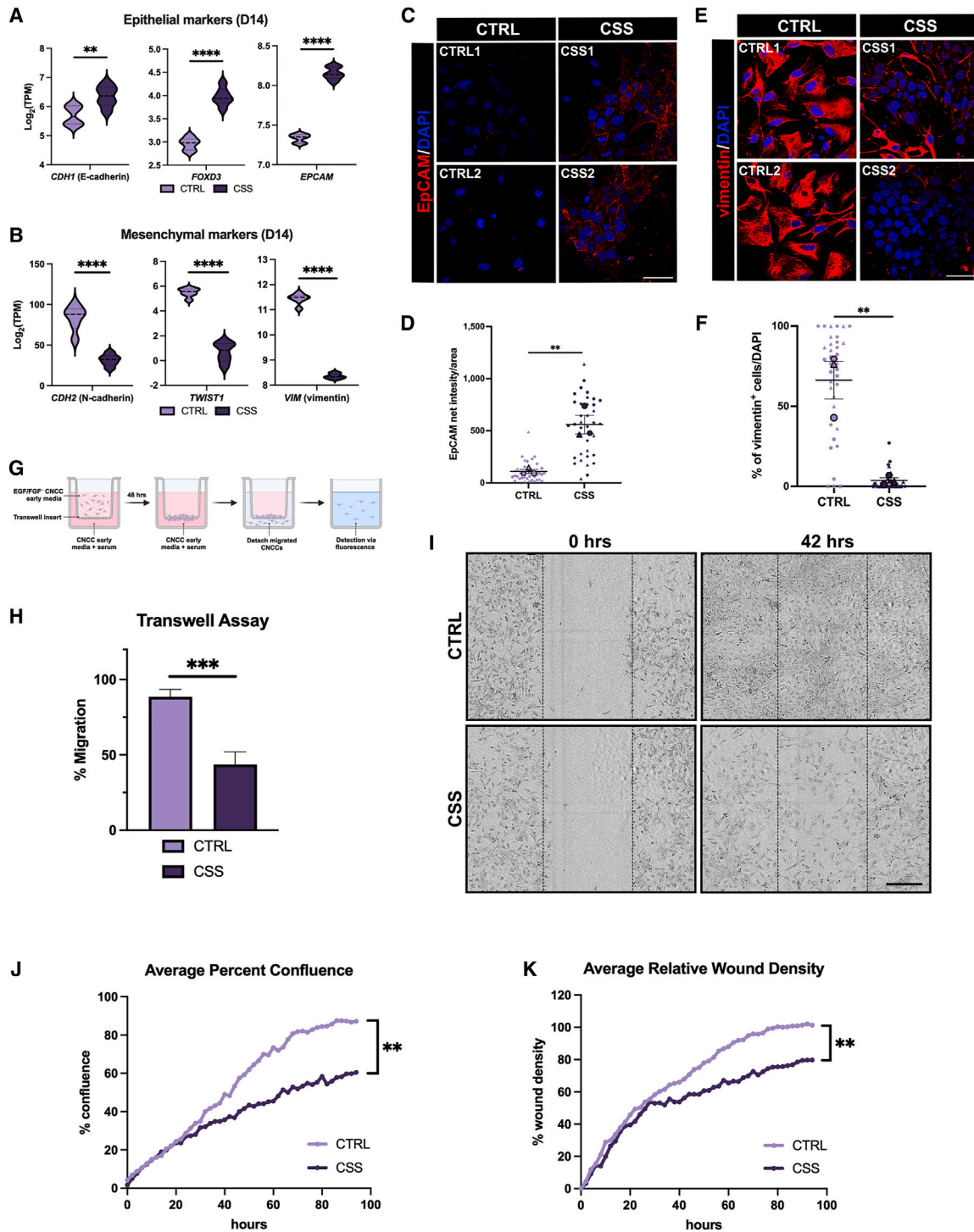
Further, we examined transcript levels of neuronal markers at the conclusion of the CNCC specification

---

(E) Adapted schematic of Waddington's epigenetic landscape<sup>65</sup> depicting restrictive cell fate commitment in the context of neural crest formation. The mesoderm gives rise to MSC (mesenchymal stem cells), and the ectoderm differentiates into the epidermis and neuroectoderm. Neuroectodermal fate defaults to the neural progenitor lineage of the central nervous system (CNS) represented by a thick arrow when neural crest fate cannot be specified. Neuroectoderm also gives rise to multipotent NCCs, which will form various derivatives, including the craniofacial bones and cartilage and neural progenitors and glia of the peripheral nervous system (PNS). Other derivatives of these cell types are not shown in the schematic. Made with [BioRender.com](https://www.biorender.com).

(F) Violin plots displaying  $\log_2(\text{TPM})$  of neuronal markers *MAP2*, *RBFOX3* (NeuN), and *DCX* at day 14 between CTRL and CSS cell lines. A two-tailed unpaired t test was performed, and  $p < 0.05$  was considered significant; *MAP2*: \*\*\*\* $p < 0.0001$ , *RBFOX3* (NeuN): \*\*\*\* $p < 0.0001$ , and *DCX*: \*\*\* $p = 0.0001$ .

(G and H) (G) Representative images and (H) quantification of MAP2 and NeuN immunofluorescence performed at day 14 in CTRL and CSS cell lines. (G) DAPI staining on nuclei in blue. Images taken at 60 $\times$  magnification; scale bar, 50  $\mu\text{m}$ . (H) SuperPlot quantification of net intensity of MAP2 and NeuN per cell.  $n = 3$  represented by a distinctive shape with each small data point representing a captured image. The larger data points correspond to the average values of each replicate. A two-tailed unpaired t test performed on the average values between CTRL and CSS cell lines;  $p < 0.05$  was considered significant; *MAP2* \*\* $p = 0.0021$  NeuN \* $p = 0.0184$ . Error bars represent the standard error of the mean.



**Figure 3. Impaired EMT compromises the specification of *ARID1A*-haploinsufficient cells into migratory CNCCs**

(A) Violin plots displaying  $\log_2(\text{TPM})$  of epithelial/pre-migratory CNCC markers *CDH1* (E-cadherin), *FOXD3*, and *EPCAM* at day 14 between CTRL and CSS cell lines. A two-tailed unpaired t test was performed and  $p < 0.05$  was considered significant; *CDH1* (E-cadherin):  $**p = 0.0079$ , *FOXD3*:  $****p < 0.0001$ , and *EPCAM*:  $****p < 0.0001$ .

(B) Violin plots displaying  $\log_2(\text{TPM})$  of mesenchymal/migratory CNCC markers *CDH2* (N-cadherin), *TWIST1*, and *VIM* (vimentin) at day 14 between CTRL and CSS cell lines. A two-tailed unpaired t test was performed and  $p < 0.05$  was considered significant; *CDH2* (N-cadherin):  $****p < 0.0001$ , *TWIST1*:  $****p < 0.0001$ , and *VIM* (vimentin):  $****p < 0.0001$ .

(C and D) (C) Representative images and (D) quantification of an EpcAM immunofluorescence performed at day 14 of CNCC specification in CTRL and CSS cell lines. (C) DAPI staining on nuclei in blue. Merged images shown taken at  $60\times$  magnification; scale bar,  $50\mu\text{m}$ . (D) SuperPlot quantification of EpcAM net intensity per area.  $n = 3$  represented by a distinctive shape with each small data point representing a captured image. The larger data points correspond to the average values of each replicate. A two-tailed unpaired t-test

(legend continued on next page)

process (day 14). The expression of neuronal genes *MAP2*, *RBFOX3* (NeuN), and *DCX* were significantly upregulated in CSS lines compared to CTRLs (Figure 2F). Moreover, immunofluorescence for MAP2 and NeuN revealed a massive increase in both markers in the CSS cell lines (Figures 2G and 2H). Notably, these cells featured a level of heterogeneity correlated with the observed cellular morphology (Figure S1I). CSS cells that displayed a clustered and elongated morphology (Figure S1I) were positive for these neuronal markers (Figures 2G and 2H). Few isolated single cells formed in the CSS lines by day 14; however, these cells did not exhibit this aberrant neuronal protein accumulation by immunofluorescence. These data suggest that the *ARID1A*-haploinsufficient lines were unable to commit toward the CNCC lineage, while aberrantly upregulating a mix of neural progenitor and neuronal markers.

### Functional EMT is impaired in *ARID1A*-haploinsufficient CNCCs

CNCCs are a multipotent population with the potential to differentiate into both ectodermal and mesenchymal derivatives<sup>38</sup> (Figure 2E). To migrate and differentiate into the terminal derivatives, CNCCs must undergo EMT allowing for delamination and migration from the neural tube.<sup>35–37</sup> Key mechanisms involved in developmental EMT include regulation of cadherin switching whereby E-cadherin is downregulated, and N-cadherin is upregulated.<sup>66–68</sup> This switch is accompanied by increased levels of essential mesenchymal factors vimentin, SNAIL, SLUG, and the TWIST, as well as ZEB transcription factor families.<sup>66</sup>

Notably, we found that the expression of genes defining epithelial/neuroectodermal cells, including *CDH1* (E-cadherin) and *EPCAM*, and pre-migratory CNCCs (*FOXD3*), were all upregulated in the *ARID1A*-haploinsufficient lines compared to CTRL CNCCs (day 14; Figure 3A). Conversely, the expression of mesenchymal and migratory CNCC genes, such as *CDH2* (N-cadherin), *TWIST1*, and *VIM* (vimentin), were markedly downregulated in the *ARID1A*-haploinsufficient lines compared to CTRL CNCCs (Figure

3B). Immunostaining for the epithelial marker EpCAM (Figures 3C and 3D) and the mesenchymal marker vimentin (Figures 3E and 3F) further supported the transcriptomic data, strongly indicating that the CSS cell lines remained in a neuroepithelial state with some degree of heterogeneity (Figure 3E; CSS1), while the CTRL cells successfully underwent EMT.

To functionally assess the migratory potential of CTRL and CSS cell lines *in vitro*, we performed transwell (Figures 3G and 3H) and scratch wound assays (Figures 3I–3K). After CNCCs were specified, growth factors were removed from the media to allow for an accurate assessment of cell migration. CSS cells had approximately 50% reduced migration than CTRLs through the transwell membrane (Figure 3H). Similarly, the CTRL scratch wound was visually confluent while the CSS cell lines exhibited minimal migration (Figure 3I). A significant decrease in average percent confluence in the scratch wound as well as relative wound density was observed in CSS lines compared to CTRLs (Figures 3J and 3K). Collectively, these data suggest that EMT was affected in CSS lines as they failed to acquire a functional migratory phenotype.

### *ARID1A*-BAF regulates accessibility at EMT enhancers

To investigate the mechanism underlying impaired EMT in the CSS lines, we performed chromatin immunoprecipitation followed by sequencing (ChIP-seq) for *ARID1A* to assess alterations in binding at day 10 (i.e., *ARID1A* re-upregulation onset). Simultaneously, we performed an ATAC-seq to interrogate chromatin accessibility at the same time point. We detected 21,724 shared *ARID1A* ChIP-seq peaks across CSS and CTRL cells (Figure S3A), 3,398 new *ARID1A* peaks gained in CSS cell lines, and 19,401 *ARID1A* peaks lost in CSS lines (Figures 4A and 4B). Similarly, there were ~46,065 conserved ATAC-seq peaks between CSS and CTRL cell lines (Figure 3SB), 9,163 peaks exclusive to CSS cells, and 20,216 peaks exclusive to CTRL cells (Figures S3C and S3D). Consistent with *ARID1A* loss being heterozygous, average profiles indicate that both *ARID1A* binding and chromatin accessibility at the ~20,000 sites were nearly half of the peak height of the CTRL levels

---

performed on the average values between CTRL and CSS cell lines;  $p < 0.05$  was considered significant;  $**p = 0.0078$ . Error bars represent the standard error of the mean.

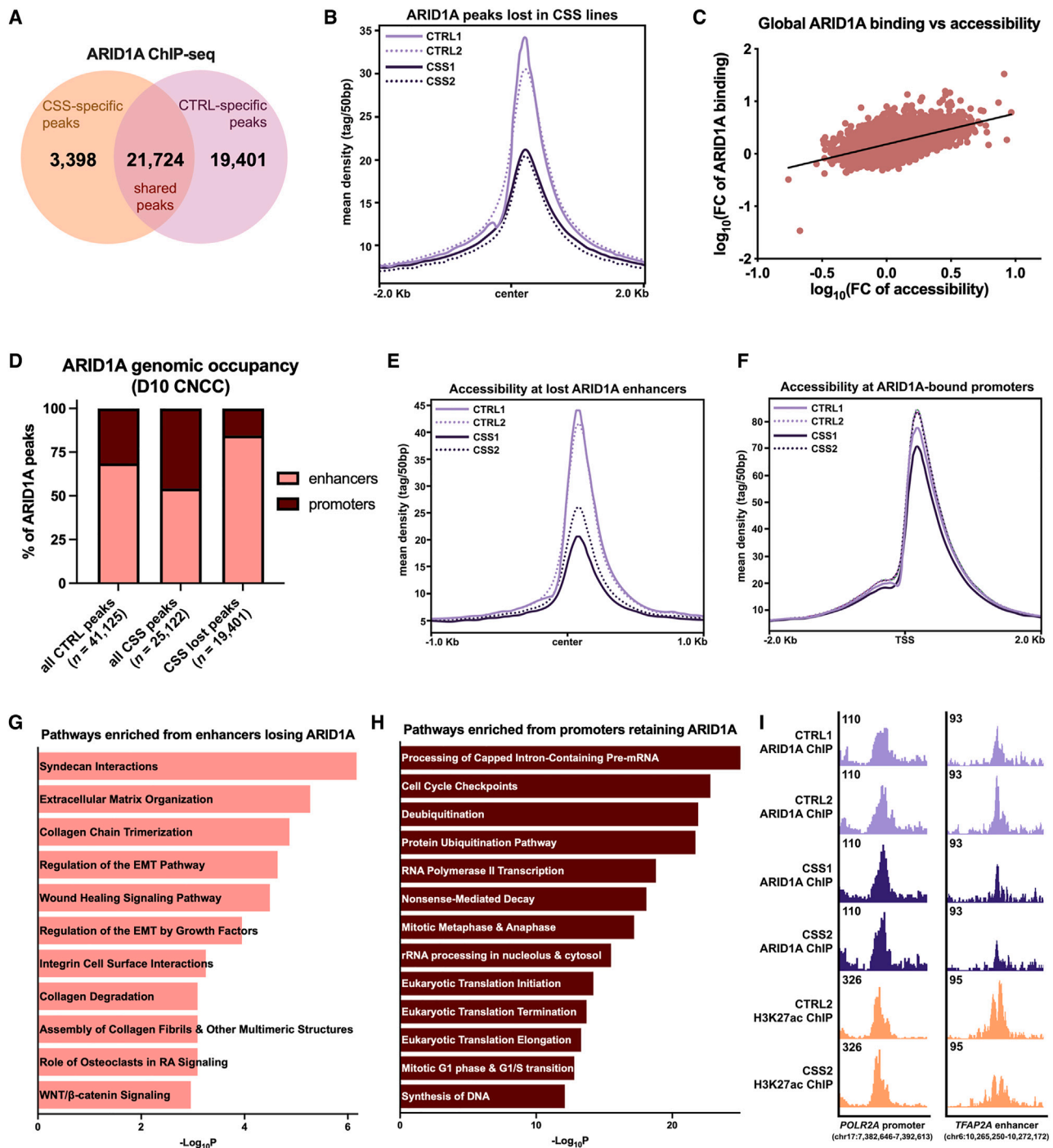
(E and F) (E) Representative images and (F) quantification of a vimentin immunofluorescence performed at D14 of CNCC specification in CTRL and CSS cell lines. (E) DAPI staining on nuclei in blue. Merged images taken at 60 $\times$  magnification; scale bar, 50 $\mu$ m. (F) SuperPlot quantification of percentage of vimentin-positive cells per DAPI.  $n = 3$  represented by a distinctive shape with each small data point representing a captured image. The larger data points correspond to the average values of each replicate. A two-tailed unpaired t test performed on the average values between CTRL and CSS cell lines;  $p < 0.05$  was considered significant;  $**p = 0.0062$ . Error bars represent the standard error of the mean.

(G) Graphical illustration of a transwell assay. Made with [BioRender.com](https://www.biorender.com).

(H) Bar plot depicting average percent migration of CTRL and CSS cells through a transwell membrane. A two-tailed unpaired t test performed between CTRL and CSS cell lines;  $p < 0.05$  was considered significant;  $***p = 0.0002$ . Error bars represent the standard error of the mean.

(I) Representative brightfield images of a scratch wound assay (Incucyte) at 0 h and 42 h post-scratch between CTRL and CSS cell lines. 42 h represents the earliest time point of a visually closed wound in CTRL cells. Scale bar, 300  $\mu$ m, 10 $\times$  magnification.

(J and K) (J) Average percent confluence of the wound area and (K) average relative wound density over 96 h between CTRL and CSS cell lines from the scratch wound assay performed in (I). A two-tailed unpaired t test performed between CTRL and CSS cell lines;  $p < 0.05$  was considered significant; percent confluence:  $**p = 0.0014$ ; percent wound density:  $**p = 0.0077$ .



**Figure 4. Chromatin dynamics and ARID1A genomic occupancy in CTRL and CSS cell lines**

(A) Venn diagram displaying the number of CSS-specific peaks (ARID1A peaks present only in CSS cell lines or “gained” in CSS cell lines; 3,398), CTRL-specific peaks (ARID1A peaks present only in CTRLs or “lost” in CSS cell lines; 19,401), and shared ARID1A peaks between CTRL and CSS cell lines (conserved ARID1A binding; 21,724) via ARID1A ChIP-seq performed at day 10 of CNCC specification.

(B) Average profile of ARID1A peaks lost in CSS cell lines at day 10 ( $n = 19,401$ ). Center represents the average overlapping global binding of ARID1A at CTRL-specific peaks.

(C) Scatterplot and linear regression of the  $\log_{10}$ (fold-change) of ARID1A peak coverage versus the  $\log_{10}$ (fold-change) of ATAC peak coverage displaying a significant Pearson correlation ( $r = 0.4940$ , two-tailed \*\*\*\* $p < 0.0001$ ) between global ARID1A binding and global accessibility.

(D) Stacked bar plot depicting the percentage of ARID1A peaks enriched at enhancers ( $>1$  kb from the closest transcription start site or TSS) and promoters ( $<1$  kb from the closest TSS) in all CTRL ARID1A peaks (including peaks shared with CSS lines), all CSS ARID1A peaks (including peaks shared with CTRL lines), and all ARID1A peaks that are lost in CSS lines (CSS lost peaks). ARID1A peaks in CSS cell lines are enriched at promoter regions relative to CTRLs; 54% at enhancers and 46% at promoters in CSS cell lines versus 69% at enhancers and 31% at promoters in CTRLs. Approximately 85% of lost ARID1A peaks in the CSS lines were lost at enhancer regions.

(legend continued on next page)

(Figures 4B and S3E). Moreover, the loss of ARID1A binding and accessibility were correlated (Figure 4C).

Analyzing the genomic occupancy of ARID1A, approximately 70% of the CTRL ARID1A ChIP-seq peaks were at enhancer regions (>1 kb from the closest transcription start site [TSS]) and 30% at promoters (<1 kb from the closest TSS; Figure 4D). Conversely, ARID1A ChIP-seq peaks in the CSS lines were enriched equally between enhancers and promoters (Figure 4D). The majority of the ARID1A peaks lost in the CSS cell lines were at enhancer regions (85% at enhancers and 15% at promoters), while most of the ARID1A peaks at promoters were retained (Figure 4D). This suggests that BAF binding at promoter regions may be essential, while BAF binding at enhancers may be dispensable. Consistent with this, accessibility and active histone marks at ARID1A-bound enhancers was significantly diminished in CSS lines (Figures 4E and S3F), while accessibility and active state at ARID1A-bound promoter regions was largely retained (Figures 4F and S3G) via ATAC-seq and H3K27ac ChIP-seq respectively. Pathway analysis on differentially expressed genes associated with enhancers that lost ARID1A binding revealed enrichment for EMT, wound healing, extracellular and collagen dynamics, and integrin interactions (Figure 4G). However, pathway analysis on genes associated with ARID1A-regulated promoters revealed enrichment for basic cellular processes including cell cycle, transcription, translation, and DNA synthesis (Figure 4H). For example, the promoter of *POLR2A*, encoding the main subunit of RNA polymerase II, harbors conserved ARID1A occupancy, while a *TFAP2A* enhancer displayed loss of both ARID1A binding and H3K27ac ChIP-seq signal (Figure 4I). Notably, ARID1A gained peaks were closest to aberrantly upregulated genes that were enriched for neuronal processes and cell death pathways (Figure S3H). These data suggested that ARID1A-BAF regulates EMT by opening the chromatin at EMT-associated enhancers, and the function of these enhancers is impaired in *ARID1A*-haploinsufficiency.

#### ARID1A-BAF regulates the accessibility at enhancers of EMT genes to coordinate ZIC2 binding

To predict potential interactions between ARID1A and transcription factors in the regulation of EMT enhancers during CNCC specification, we performed a motif analysis

on the ARID1A-bound enhancers and promoters. Regions that gained ARID1A binding and promoters that retained ARID1A binding both harbored conventional promoter motifs (Figures S4A and S4B), while the EMT enhancers that lose ARID1A binding in CSS cells revealed motifs for CNCC lineage-specific transcription factors<sup>4,48,69–75</sup> (Figure 5A). Strikingly, nearly 40% of these enhancers harbored a ZIC family motif. *ZIC1* is not expressed at this developmental stage,<sup>76,77</sup> while *ZIC2* and *ZIC3* are both highly expressed in day 10 CNCCs<sup>56,78–81</sup> (Table S7). Reportedly, *ZIC2* variants cause holoprosencephaly type 5 in humans,<sup>82</sup> a developmental syndrome that shares many phenotypic features<sup>83</sup> with CSS.<sup>30–32</sup> Therefore, we investigated *ZIC2* as a potential interactor of ARID1A in the regulation of EMT.

To determine whether *ZIC2* requires ARID1A to bind to EMT enhancers, we performed a ChIP-seq for *ZIC2* in our isogenic CTRL and CSS lines. In *ARID1A*-haploinsufficient conditions, we observed attenuation of *ZIC2* binding at EMT enhancers normally regulated by ARID1A (Figures 5B, 5C and S4C) that also exhibited loss of accessibility (Figure S4D). Similar to ARID1A, *ZIC2* also preferentially bound enhancers in control conditions, and upon the loss of *ARID1A*, the global enhancer occupancy of *ZIC2* was attenuated (Figures S4E and S4F). Specifically, enhancers for essential mesenchymal and CNCC factors, *CDH2* (N-cadherin) and *GLI2* displayed decreased ARID1A and *ZIC2* binding in the CSS line of the isogenic pair (Figure 5D).

We further examined *ZIC2* genome-wide occupancy and identified ~5,500 *ZIC2* peaks exclusive to the CSS line. These sites did not exhibit ARID1A binding even in CTRL conditions, indicating that *ARID1A*-haploinsufficiency triggered genomic relocation of *ZIC2* to non-ARID1A regulated sites (Figures 5E and S4G). These sites of relocated *ZIC2* binding were enriched for promoter motifs (Figure S4H) and displayed aberrant accessibility likely regulated by another BAF configuration or chromatin remodeler (Figures 5F and S4I).

Notably, *ZIC2* was bound at regulatory regions for ~21% of all the differentially upregulated genes in the CSS cell line. Genes associated with aberrantly bound *ZIC2* enhancers (Figure 5G) were enriched for neuronal pathways (Figure 5H), while *ZIC2*-bound promoters were associated

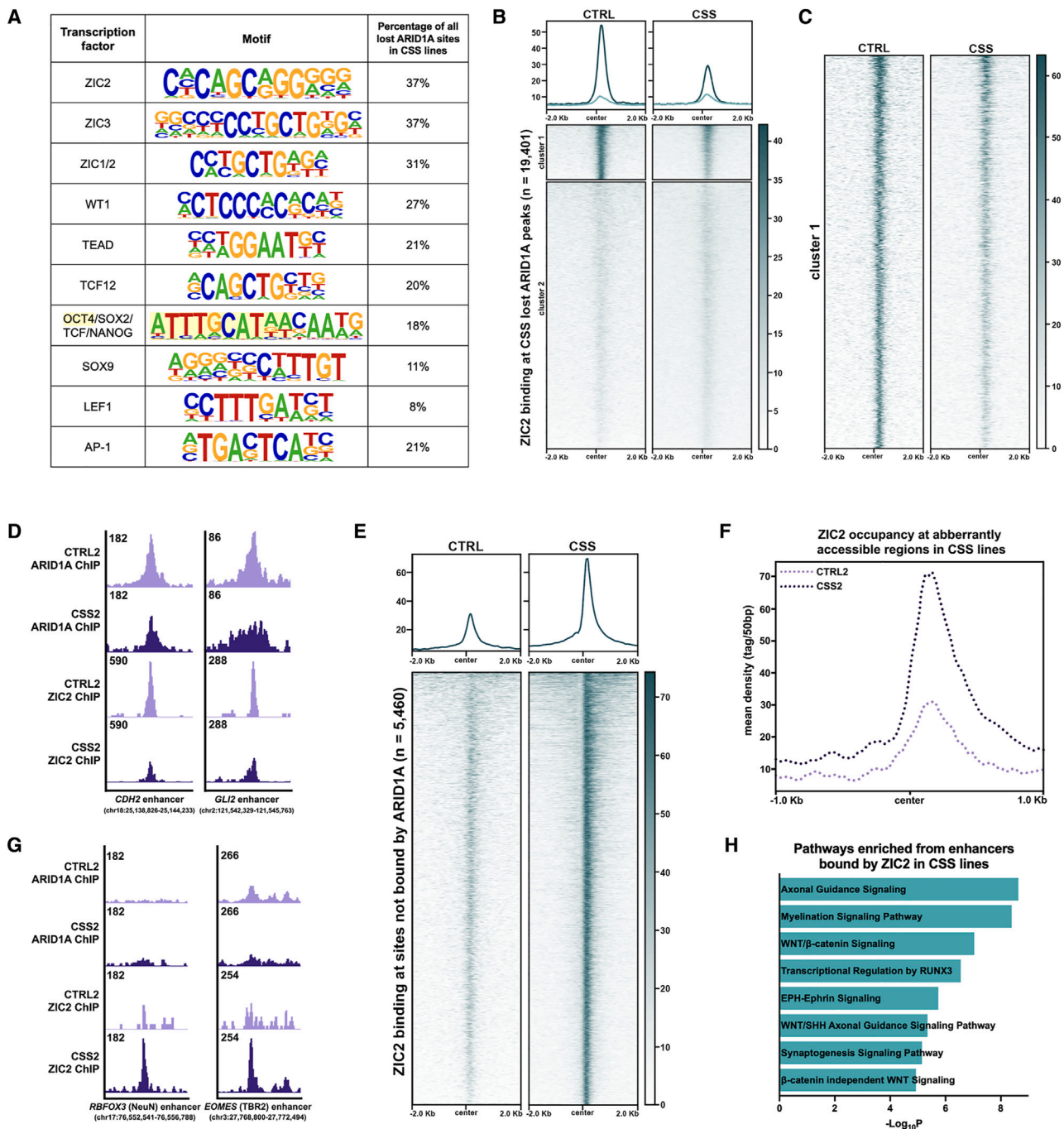
(E) Average profile of ATAC-seq peaks at enhancer regions that lose ARID1A binding and accessibility in the CSS cell lines at day 10 of CNCC specification ( $n = 3,203$ ). Center represents the average overlapping lost CSS ATAC/ARID1A peaks at enhancer regions across the genome.

(F) Average profile of ATAC-seq peaks at promoter regions that retain ARID1A binding and accessibility in the CSS cell lines at day 10 of CNCC specification ( $n = 11,231$ ). TSS represents the average overlapping retained ATAC-seq peaks at ARID1A-bound transcription start sites throughout the genome.

(G) Pathways enriched from genes that lose ARID1A binding at their enhancers in CSS cell lines at day 10 of CNCC specification and the respective  $-\log_{10} p$  value as determined by IPA. Pathways include those necessary for migration, extracellular matrix dynamics, and EMT.

(H) Pathways enriched from genes that retain ARID1A binding at their promoters in CSS lines at day 10 of CNCC specification and the respective  $-\log_{10} p$  value as determined by IPA. Pathways include those necessary for basic cellular functions and survival including DNA synthesis, transcription, translation, and cell cycle.

(I) Integrative Genomics Viewer example of a retained ARID1A peak in the CSS cell lines at the promoter of *POLR2A*, encoding the main subunit of RNA polymerase II, and a lost ARID1A peak at an enhancer for *TFAP2A*, encoding an essential CNCC marker. Representative promoter and enhancer regions marked by H3K27ac. Regions annotated using hg19.



**Figure 5. ZIC2 regulates CNCC and neuronal identity by activating accessible regulatory networks**

(A) Chart of enriched motifs at ARID1A-bound regions that are lost in CSS cell lines identified by HOMER. The presence of the motifs at each lost region was quantified via FIMO (MEME Suite). Yellow highlighted region represents the individual OCT4 motif.

(B) Clustered heatmap of ZIC2 binding at regions that lose ARID1A binding in the CSS lines at day 10 of CNCC specification (CTRL-specific;  $n = 19,401$ ).

(C) Expanded heatmap of cluster 1 from (B).

(D) Integrative Genomics Viewer example of lost ARID1A and ZIC2 peaks in CSS2 relative to the isogenic CTRL2 at ARID1A target enhancers at day 10 of CNCC specification. *CDH2* and *GLI2* are necessary for mesenchymal identity and migration of CNCCs respectively. Regions annotated using hg19.

(E) Heatmap of ZIC2 binding at regions that are not bound by ARID1A displaying a genomic relocation of ZIC2 in *ARID1A*-haploinsufficient conditions at day 10 of CNCC specification ( $n = 5,460$ ).

(F) Average profile of ZIC2 peaks at aberrantly accessible regions in CSS2 at day 10 of CNCC specification ( $n = 244$ ). Center represents the average overlapping binding of ZIC2 at CSS-specific ATAC peaks across the genome.

(legend continued on next page)

with common cellular processes (Figure S4J). Notably, ZIC2 protein levels were higher in CSS cell lines at day 10 of CNCC specification compared to CTRL (Figure S4K). These findings were consistent with the upregulation of neuronal genes observed in our ARID1A-haploinsufficient CSS lines. Collectively, these data suggest that ARID1A-BAF opened the chromatin at EMT enhancers during CNCC specification, allowing for ZIC2 binding. Furthermore, in the partial absence of ARID1A, ZIC2 relocated to neuronal enhancers whose accessibility was likely modulated by other chromatin remodelers.

### Zic2 is required for NCC EMT and delamination *in vivo*

To further investigate the role of Zic2 in EMT and NCC formation *in vivo*, we characterized the spatiotemporal levels of Zic2 during murine neurulation using a previously described Zic2 reporter mouse line, Tg(Zic2<sup>EGFP</sup>).<sup>54</sup> At E8.0, prior to CNCC delamination and migration (similar to the pre-migratory stage *in vitro*), Zic2 signal is detectable all along the dorsal neural tube in cells positive for Foxd3, a marker of pre-migratory NCCs (Figure 6A). By E9.5, there were few Zic2<sup>+</sup> pre-migratory cells remaining in the dorsal region of the neural tube, indicating successful NCC delamination and migration (Figure 6B; right). Additionally, we observed migratory NCCs (Sox10<sup>+</sup>/EGFP) migrating throughout the dorsolateral and ventromedial paths in the mesenchyme (Figure 6B; right). This observation suggested that Sox10/EGFP<sup>+</sup> cells migrating in the mesenchyme (Figure 6B; right) were previously Zic2<sup>+</sup> pre-migratory cells originating from the dorsal region of the neural tube (Figure 6A).

In order to assess the impact of Zic2 loss-of-function in NCCs, we crossbred the Tg(Zic2<sup>EGFP</sup>) strain with the previously reported Zic2<sup>kd</sup> line<sup>56</sup> to obtain mutant [Zic2<sup>kd/kd</sup>; Tg(Zic2<sup>EGFP</sup>)] and control [Zic2<sup>+/+</sup>; Tg(Zic2<sup>EGFP</sup>)] embryos in the same litter. In Zic2-WT embryos, many EGFP-labeled cells exhibited a stereotypical migration pattern through the dorsal sub-epidermal region of the mesenchyme (Figures 7A–7C). However, in the Zic2-mutant embryos, very few migrating EGFP-labeled cells were detected (Figures 7A–7C). Instead, an aberrant accumulation of EGFP<sup>+</sup> cells was observed in the dorsal neural tube in the mutants compared to the controls (Figure 7B). Notably, there was a significant reduction of migratory Sox10<sup>+</sup> cells detected in the Zic2-mutant embryos (Figures 7D and 7E). The Sox10<sup>+</sup> cells were mostly restricted to the dorsal neural tube in the Zic2-mutant embryos indicating a lack of EMT, delamination, and migration in the absence of Zic2 (Figure 7D). Consistent with our *in vitro* data, these results indicate that NCCs failed to delaminate from the neural tube in Zic2-mutant embryos, demonstrating a crucial role for Zic2 in cell delamination via EMT.

### ZIC2 gain-of-function at pre-migratory stages elicits aberrant NCC delamination *in vivo*

To explore whether Zic2 was sufficient to trigger cell detachment through EMT *in vivo*, we conducted gain-of-function experiments by ectopically expressing human ZIC2 and/or GFP into the neural tube of HH9-11 chick embryos (pre-neurulation). As expected, GFP<sup>+</sup> cells were largely observed in the neural tube at developmental stage HH24-25 (Figure 7F). Notably, ectopic expression of ZIC2 at the same developmental stages triggered a significant increase in the number of cells delaminating from the neural tube (Figures 7F and 7G). This indicates that ectopic expression of ZIC2 was sufficient to induce delamination in NCCs. Collectively our *in vivo* gain-of-function data further support a key role for ZIC2 in EMT and delamination of NCCs from the neural tube.

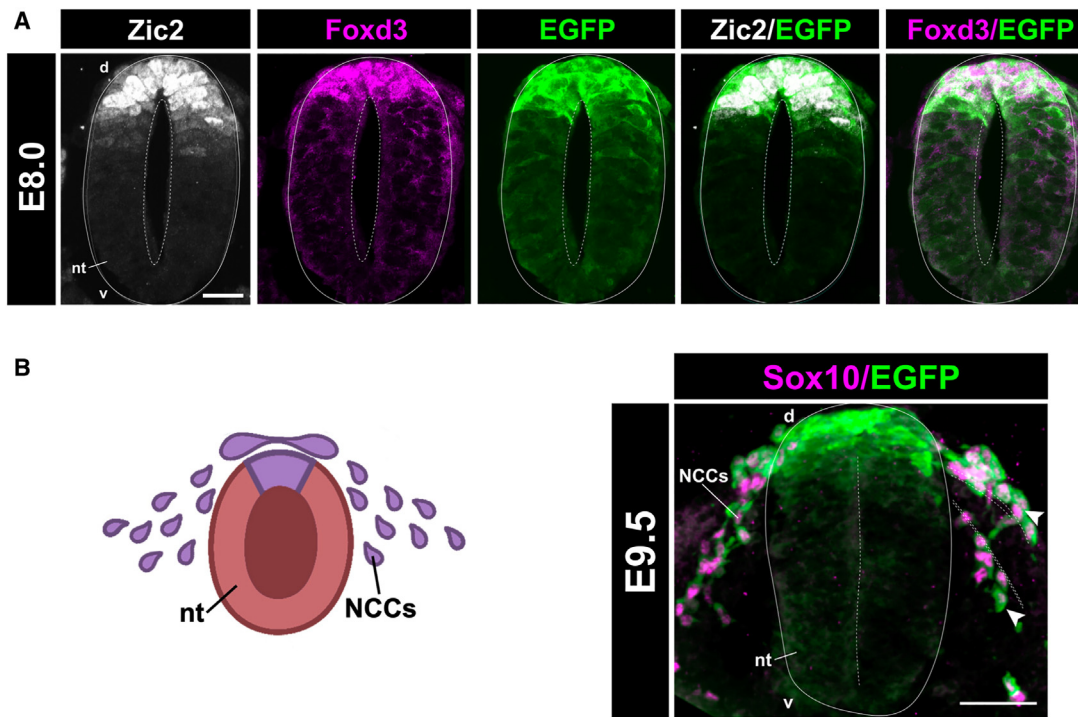
## Discussion

ARID1A is one of two integral DNA-binding subunits of the canonical BAF chromatin remodeling complex. While ARID1A-BAF has a wide-range of roles including regulating aberrant chromatin accessibility in cancer contexts,<sup>6,84</sup> its function in physiological craniofacial development is underappreciated. Variants in ARID1A are associated with CSS, a rare developmental disorder predominantly characterized by coarse craniofacial features.<sup>31,32</sup> The hyperplastic craniofacial skeletal features seen in CSS are derived from the cranial neural crest.<sup>31,33,37</sup>

Herein, we investigated the role of ARID1A-BAF in the molecular regulation of CNCC formation using ARID1A<sup>+/-</sup> patient-derived iPSCs. It is critical to highlight the considerable incidence of somatic mosaicism associated with ARID1A variants.<sup>17,18</sup> While *de novo* germline mutations have been reported, most individuals with CSS who harbor variants in ARID1A are typically mosaic.<sup>60</sup> Recently, it was shown that germline mutations in ARID1A result in a much more severe phenotype, often with multiple congenital anomalies, frequently leading to termination of pregnancy or spontaneous demise during or after birth.<sup>60</sup> In a research context, this mosaicism provides us with an isogenic cellular platform in which the wild-type and ARID1A<sup>+/-</sup> lines share the same genetic background. This offers the advantage of eliminating confounding genomic factors, ensuring all the cellular phenotypes observed are a direct result of ARID1A haploinsufficiency from a clinical setting.

We previously reported an essential ARID1 subunit switch between ARID1A and ARID1B in BAF that occurs at the transition of iPSC to neuroectoderm and neuroectoderm to CNCC.<sup>3</sup> Notably, subunit switching for distinct regulatory function is not restricted to BAF as this

(G) Integrative Genomics Viewer example of gained ZIC2 binding in CSS2 relative to the isogenic CTRL2 at sites not bound by ARID1A at day 10 of CNCC specification. RFX3 (NeuN) and EOMES (TBR2) are genes involved in neurogenesis. Regions annotated using hg19. (H) Pathways enriched from enhancers that are bound by ZIC2 in CSS2 at day 10 of CNCC specification and the respective  $-\log_{10} p$  value as determined by IPA. Pathways include those involved in neuronal signaling and processes.



**Figure 6. Zic2 activates neural crest EMT and delamination from the neural tube *in vivo***

(A) A transverse cross-section of the neural tube in an E8.0 Tg(*Zic2*<sup>EGFP</sup>) mouse embryo immunostained for Zic2 and Foxd3. Scale bar, 20  $\mu$ m, 40 $\times$  magnification.

(B) A schematic (left) and a transverse cross-section (right) of the neural tube in an E9.5 Tg(*Zic2*<sup>EGFP</sup>) mouse embryo immunostained for Sox10. White arrowheads highlight Sox10<sup>+</sup> migrating NCCs are also GFP<sup>+</sup> depicting migratory NCCs once expressed *Zic2* in a pre-migratory setting. Dotted lines show NCCs migrating in the dorsolateral/ventromedial paths. Scale bar, 50  $\mu$ m, 40 $\times$  magnification. d, dorsal; v, ventral; nt, neural tube; NCCs, neural crest cells. Schematic made with [BioRender.com](https://www.biorender.com).

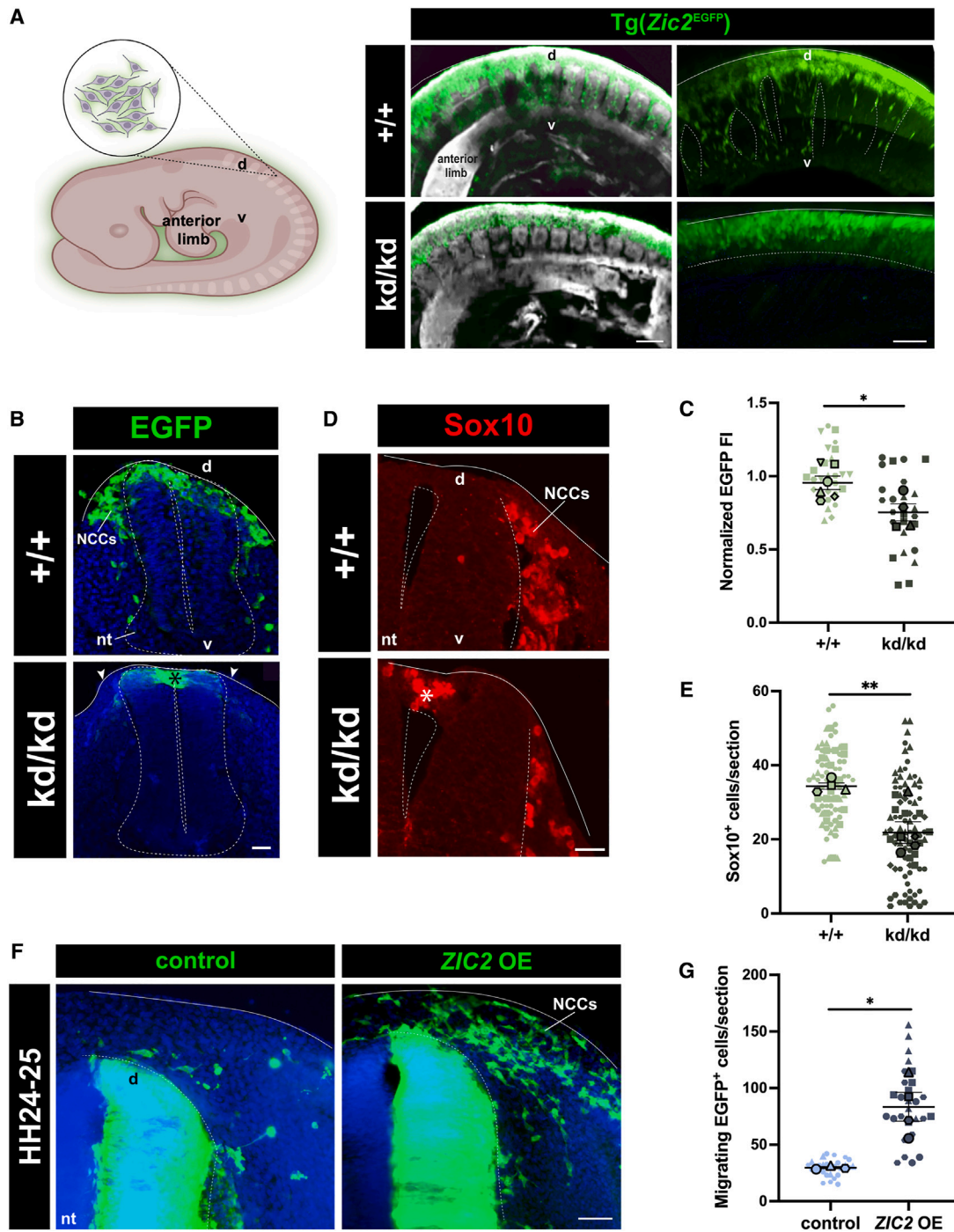
phenomenon occurs in other chromatin remodelers. For example, among the mutually exclusive ATPases in the NuRD complex, CHD3/4/5, each subunit coordinates specific sets of genes necessary for various stages of cortical development.<sup>85</sup> Ultimately, signaling pathways and interactors that regulate these ARID1 subunit switches, and its timing remain unknown. Similarly, recruitment of BAF to its target regions is globally regulated by the AP-1 transcription factor complex.<sup>71</sup> However, this recruitment process could also be guided by cell type-specific transcription factors essential for NCC formation<sup>48,72–74</sup> as depicted in our motif analysis of ARID1A-bound regions (Figure 5A).

In this study, we have unveiled an essential axis for CNCC formation, including ZIC2 and ARID1A-BAF. ZIC2 is a transcription factor of the GLI superfamily, which is known to regulate several aspects of neurodevelopment.<sup>86–88</sup> Variants of ZIC2 are associated with holoprosencephaly type 5,<sup>82</sup> a developmental syndrome that shares many craniofacial phenotypes<sup>83</sup> with CSS.<sup>32</sup> Individuals with holoprosencephaly harboring ZIC2 variants present with bitemporal narrowing, upslanting palpebral fissures, a short nose with anteverted nares, a broad and well demarcated philtrum, and large ears.<sup>83</sup> Notably, these specific craniofacial anomalies do not appear in individuals with pathogenic variants in other holoprosencephaly-associated genes.<sup>83</sup>

Our *in vitro* and *in vivo* data, encompassing human iPSCs, mouse models, and chick models, support a mechanism in which ARID1A opens the chromatin for ZIC2 at the onset of EMT to induce NCC delamination from the neural tube. With multiple *in vivo* models, we demonstrate that *Zic2* is indispensable for NCC EMT and subsequent migration, and validate the molecular foundation of this regulation through ARID1A-mediated accessibility at EMT enhancers *in vitro*. Supporting this model *in vivo*, *Zic2* is present in mouse pre-migratory NCCs and is required for their delamination from the neural tube. Conversely, ectopic expression of ZIC2 in chick embryos displays asynchronous delamination and EMT.

In *ARID1A*<sup>+/-</sup> conditions, accessibility at EMT enhancers is perturbed, preventing lineage-specific transcription factors, like ZIC2, from binding at these elements to promote EMT. In the partial absence of *ARID1A*, ZIC2 relocates to neuronal enhancers, suggesting ZIC2 plays a role in the aberrant activation of neuronal networks. As both NCCs and neurons are derived from the neuroectoderm,<sup>35,89–92</sup> and neuroectoderm-like cells are successfully specified in *ARID1A*-haploinsufficient conditions, our results potentially suggest a default trajectory toward the neuronal lineage when the neural crest lineage cannot be specified. However, further studies will be needed to validate this model.





**Figure 7.** *Zic2* perturbation prevents NCC migration, while *ZIC2* overexpression induces ectopic NCC delamination from the neural tube *in vivo*

(A) Schematic (left) and iDISCO (right) of the lateral view of E9.5 whole-mount control [*Zic2*<sup>+/+</sup>;Tg(*Zic2*<sup>EGFP</sup>)] and *Zic2*-mutant [*Zic2*<sup>kd/kd</sup>;Tg(*Zic2*<sup>EGFP</sup>)] mouse embryos displaying lack of migration into the mesenchyme in *Zic2*-mutant conditions. Scale bar, 100  $\mu$ m, 5 $\times$  magnification. Schematic made with BioRender.com.

(B) Representative images of a transverse cross-sections of the neural tube from E9.0 control [*Zic2*<sup>+/+</sup>; Tg(*Zic2*<sup>EGFP</sup>)] and *Zic2*-mutant [*Zic2*<sup>kd/kd</sup>;Tg(*Zic2*<sup>EGFP</sup>)] mouse embryos. Arrowheads highlight the absence of GFP<sup>+</sup> cells in the dorsal sub-epidermal area of a *Zic2*-mutant embryo, with GFP<sup>+</sup> cells accumulating inside the dorsal neural tube represented by an asterisk. Scale bar, 10  $\mu$ m, 40 $\times$  magnification.

(C) Superplot quantification of the fluorescent intensity (FI) of GFP<sup>+</sup> cells in the dorsal sub-epidermis of E9.0 [*Zic2*<sup>+/+</sup>; Tg(*Zic2*<sup>EGFP</sup>)] and [*Zic2*<sup>kd/kd</sup>; Tg(*Zic2*<sup>EGFP</sup>)] mouse embryos normalized to the control intensity.  $n = 4$ –6 independent experiments represented by a distinctive shape with each small data point representing at least four sections per embryo and 4–6 embryos per genotype. The larger data points correspond to the average values of each replicate. A two-tailed unpaired t test performed on the average values between *Zic2*-WT and *Zic2*-mutant embryos;  $p < 0.05$  was considered significant;  $*p = 0.0255$ . Error bars represent the standard error of the mean.

(legend continued on next page)

ZIC2 has multifaceted functions in neuronal development,<sup>86–88</sup> hindbrain patterning,<sup>78,93</sup> and axon guidance.<sup>54,94–96</sup> Zic2 controls axon midline repulsion, and this neurodevelopmental process has been recently proposed to be influenced by enhancers containing Zic2 binding motifs.<sup>97</sup> Thus, the genomic relocation of ZIC2 and aberrant upregulation of neuronal-associated genes is consistent with its known role in regulating neuronal processes. However, additional studies will be required to determine the molecular regulation and contribution from other lineage-specific transcription factors and chromatin remodelers underlying this relocation.

Notably, EMT is a main function of other non-developmental processes including wound healing<sup>98</sup> and fibrosis.<sup>99,100</sup> Therefore, it would be crucial to investigate whether ARID1A-BAF directs EMT in these settings and if distinctive cell-type-specific transcription factors are involved. Additionally, EMT extensively modulates cancer progression and metastasis.<sup>101</sup> ARID1A is mutated in ~10% of all tumors with postulated balancing oncogenic<sup>20,21</sup> and tumor suppressive functions.<sup>22–25,84</sup> However, ARID1A is predominantly reported to inhibit EMT in a cancer context.<sup>102–104</sup> This may suggest a physiological role for ARID1A-BAF in promoting EMT in embryonic development, but further studies will be needed to support this hypothesis.

In summary, this study provides crucial insight into how chromatin remodelers regulate neural crest specification, delamination, and migration, suggesting a potential pathological mechanism underlying severe craniofacial phenotypes in certain developmental syndromes.

## Data and code availability

The accession number for the RNA-seq, ChIP-seq, and ATAC-seq data reported in this paper is GEO: GSE261846.

## Acknowledgments

The authors are grateful to the families of the individuals with CSS for donating the samples and to Sandra Deliard and Alessandro Gardini for the use of the IPA software and for insightful comments. The authors thank the Genomic Facility at The Wistar

Institute (Philadelphia, PA) for the Next Generation Illumina Sequencing. For this work, M.T. was funded by the G. Harold and Leila Y. Mathers Foundation. E.H. was funded by MCIN (PID2022-138245NB-I00), the 'la Caixa' Foundation (HR21-00824), Generalitat Valenciana (PROMETEO/2020/007), and the Severo Ochoa Program for Centers of Excellence (CEX2021-001165-S). S.B.M. was funded by the National Institutes of Health (NIH R01 R01ES033197).

## Author contributions

M.T., E.H., and S.M.B. designed the project. S.M.B. performed all the human iPSC and CNCC experiments and analyzed all the genomic data. A.G.d.G., C.M.-P., and M.T.L.-C performed all the *in vivo* experiments. C.S. performed all the iPSC and CNCC immunofluorescence and contributed to data analysis. F.J.W., M.E.C., and J.K. contributed to some of the experiments. G.W.E.S. recruited the individuals with CSS and obtained the skin fibroblasts. H.M.M.M. reprogrammed the CSS patient-derived fibroblasts into iPSCs and performed quality checks. K.A.P., D.T., S.B.M., and S.A.B. contributed to data analysis and offered critical support. S.M.B. and M.T. analyzed the data and wrote the manuscript. All the authors read and approved the manuscript.

## Declaration of interests

The authors declare no competing interests.

## Supplemental information

Supplemental information can be found online at <https://doi.org/10.1016/j.ajhg.2024.07.022>.

Received: April 25, 2024

Accepted: July 30, 2024

Published: September 2, 2024

## References

1. Hargreaves, D.C., and Crabtree, G.R. (2011). ATP-dependent chromatin remodeling: genetics, genomics and mechanisms. *Cell Res.* 21, 396–420. <https://doi.org/10.1038/cr.2011.32>.
2. Pagliaroli, L., and Trizzino, M. (2021). The Evolutionary Conserved SWI/SNF Subunits ARID1A and ARID1B Are Key Modulators of Pluripotency and Cell-Fate Determination. *Front. Cell Dev. Biol.* 9, 643361.

(D) Representative images of a transverse cross-sections of the neural tube from E10.5 control [*Zic2*<sup>+/+</sup>;Tg(*Zic2*<sup>EGFP</sup>)] and *Zic2*-mutant [*Zic2*<sup>kd/kd</sup>;Tg(*Zic2*<sup>EGFP</sup>)] mouse embryos immunostained for Sox10. The asterisk highlights an aberrant accumulation of Sox10<sup>+</sup> cells inside the neural tube at this developmental stage. Scale bar, 20 μm, 40× magnification.

(E) Superplot quantification of the number of Sox10<sup>+</sup> cells per section. *n* = 4–5 independent experiments represented by a distinctive shape with each small data point representing at least five sections per embryo and 4–5 embryos per genotype. The larger data points correspond to the average values of each replicate. A two-tailed unpaired t test performed on the average values between *Zic2*-WT and *Zic2*-mutant mouse embryos; *p* < 0.05 was considered significant; \*\**p* < 0.0074. Error bars represent the standard error of the mean.

(F) Transverse cross-sections of the developing neural tube at HH24-25 in chick embryos electroporated at HH9-11 with plasmids carrying *ZIC2*-GFP or GFP alone. OE, overexpression. Scale bar, 50 μm, 10× magnification.

(G) Superplot quantification of the number of migrating GFP<sup>+</sup> cells in sections of HH24-25 chick embryos electroporated at HH9-11 with plasmids carrying *ZIC2*-GFP or GFP alone. *n* = 3–4 independent experiments represented by a distinctive shape with each small data point representing at least five sections per embryo and 3–4 embryos per genotype. The larger data points correspond to the average values of each replicate. A two-tailed unpaired t test performed on the average values between control-GFP and *ZIC2*-GFP; *p* < 0.05 was considered significant; \**p* < 0.0161. Error bars represent the standard error of the mean. D, dorsal; v, ventral; nt, neural tube; NCCs, neural crest cells.

3. Pagliaroli, L., Porazzi, P., Curtis, A.T., Scopa, C., Mikkers, H.M.M., Freund, C., Daxinger, L., Deliard, S., Welsh, S.A., Offley, S., et al. (2021). Inability to switch from ARID1A-BAF to ARID1B-BAF impairs exit from pluripotency and commitment towards neural crest formation in ARID1B-related neurodevelopmental disorders. *Nat. Commun.* *12*, 6469. <https://doi.org/10.1038/s41467-021-26810-x>.
4. Alver, B.H., Kim, K.H., Lu, P., Wang, X., Manchester, H.E., Wang, W., Haswell, J.R., Park, P.J., and Roberts, C.W.M. (2017). The SWI/SNF chromatin remodelling complex is required for maintenance of lineage specific enhancers. *Nat. Commun.* *8*, 14648. <https://doi.org/10.1038/ncomms14648>.
5. Gatchalian, J., Malik, S., Ho, J., Lee, D.-S., Kelso, T.W.R., Shokhiev, M.N., Dixon, J.R., and Hargreaves, D.C. (2018). A non-canonical BRD9-containing BAF chromatin remodeling complex regulates naive pluripotency in mouse embryonic stem cells. *Nat. Commun.* *9*, 5139. <https://doi.org/10.1038/s41467-018-07528-9>.
6. Kadoch, C., Hargreaves, D.C., Hodges, C., Elias, L., Ho, L., Ranish, J., and Crabtree, G.R. (2013). Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nat. Genet.* *45*, 592–601. <https://doi.org/10.1038/ng.2628>.
7. Trizzino, M., Barbieri, E., Petracovici, A., Wu, S., Welsh, S.A., Owens, T.A., Licciulli, S., Zhang, R., and Gardini, A. (2018). The Tumor Suppressor ARID1A Controls Global Transcription via Pausing of RNA Polymerase II. *Cell Rep.* *23*, 3933–3945. <https://doi.org/10.1016/j.celrep.2018.05.097>.
8. Dykhuizen, E.C., Hargreaves, D.C., Miller, E.L., Cui, K., Korshunov, A., Kool, M., Pfister, S., Cho, Y.-J., Zhao, K., and Crabtree, G.R. (2013). BAF complexes facilitate decatenation of DNA by topoisomerase II $\alpha$ . *Nature* *497*, 624–627. <https://doi.org/10.1038/nature12146>.
9. Meisenberg, C., Pinder, S.I., Hopkins, S.R., Wooller, S.K., Benstead-Hume, G., Pearl, F.M.G., Jeggo, P.A., and Downs, J.A. (2019). Repression of Transcription at DNA Breaks Requires Cohesin throughout Interphase and Prevents Genome Instability. *Mol. Cell.* *73*, 212–223.e7. <https://doi.org/10.1016/j.molcel.2018.11.001>.
10. Barisic, D., Stadler, M.B., Iurlaro, M., and Schübeler, D. (2019). Mammalian ISWI and SWI/SNF selectively mediate binding of distinct transcription factors. *Nature* *569*, 136–140. <https://doi.org/10.1038/s41586-019-1115-5>.
11. He, S., Wu, Z., Tian, Y., Yu, Z., Yu, J., Wang, X., Li, J., Liu, B., and Xu, Y. (2020). Structure of nucleosome-bound human BAF complex. *Science* *367*, 875–881. <https://doi.org/10.1126/science.aaz9761>.
12. Varga, J., Kube, M., Luck, K., and Schick, S. (2021). The BAF chromatin remodeling complexes: structure, function, and synthetic lethality. *Biochem. Soc. Trans.* *49*, 1489–1503. <https://doi.org/10.1042/BST20190960>.
13. Alpsy, A., and Dykhuizen, E.C. (2018). Glioma tumor suppressor candidate region gene 1 (GLTSCR1) and its paralog GLTSCR1-like form SWI/SNF chromatin remodeling sub-complexes. *J. Biol. Chem.* *293*, 3892–3903. <https://doi.org/10.1074/jbc.RA117.001065>.
14. Wang, W., Côté, J., Xue, Y., Zhou, S., Khavari, P.A., Biggar, S.R., Muchardt, C., Kalpana, G.V., Goff, S.P., Yaniv, M., et al. (1996). Purification and biochemical heterogeneity of the mammalian SWI-SNF complex. *EMBO J.* *15*, 5370–5382. <https://doi.org/10.1002/j.1460-2075.1996.tb00921.x>.
15. Mashtalir, N., D'Avino, A.R., Michel, B.C., Luo, J., Pan, J., Otto, J.E., Zullo, H.J., McKenzie, Z.M., Kubiak, R.L., St. Pierre, R., et al. (2018). Modular Organization and Assembly of SWI/SNF Family Chromatin Remodeling Complexes. *Cell* *175*, 1272–1288.e20. <https://doi.org/10.1016/j.cell.2018.09.032>.
16. Mashtalir, N., Suzuki, H., Farrell, D.P., Sankar, A., Luo, J., Filipovski, M., D'Avino, A.R., St. Pierre, R., Valencia, A.M., Onikubo, T., et al. (2020). A Structural Model of the Endogenous Human BAF Complex Informs Disease Mechanisms. *Cell* *183*, 802–817.e24. <https://doi.org/10.1016/j.cell.2020.09.051>.
17. Santen, G.W.E., Aten, E., Vulto-van Silfhout, A.T., Pottinger, C., van Bon, B.W.M., van Minderhout, I.J.H.M., Snowdowne, R., van der Lans, C.A.C., Boogaard, M., Linsen, M.M.L., et al. (2013). Coffin–Siris Syndrome and the BAF Complex: Genotype–Phenotype Study in 63 Patients. *Hum. Mutat.* *34*, 1519–1528. <https://doi.org/10.1002/humu.22394>.
18. Wiczorek, D., Bögershausen, N., Beleggia, F., Steiner-Haldenstät, S., Pohl, E., Li, Y., Milz, E., Martin, M., Thiele, H., Altmüller, J., et al. (2013). A comprehensive molecular study on Coffin–Siris and Nicolaides–Baraitser syndromes identifies a broad molecular and clinical spectrum converging on altered chromatin remodeling. *Hum. Mol. Genet.* *22*, 5121–5135. <https://doi.org/10.1093/hmg/ddt366>.
19. Alfert, A., Moreno, N., and Kerl, K. (2019). The BAF complex in development and disease. *Epigenet. Chromatin* *12*, 19. <https://doi.org/10.1186/s13072-019-0264-y>.
20. Sun, X., Wang, S.C., Wei, Y., Luo, X., Jia, Y., Li, L., Gopal, P., Zhu, M., Nassour, I., Chuang, J.-C., et al. (2017). Arid1a Has Context-Dependent Oncogenic and Tumor Suppressor Functions in Liver Cancer. *Cancer Cell* *32*, 574–589.e6. <https://doi.org/10.1016/j.ccell.2017.10.007>.
21. Chandler, R.L., Damrauer, J.S., Raab, J.R., Schisler, J.C., Wilkerson, M.D., Didion, J.P., Starmer, J., Serber, D., Yee, D., Xiong, J., et al. (2015). Coexistent ARID1A–PIK3CA mutations promote ovarian clear-cell tumorigenesis through pro-tumorigenic inflammatory cytokine signalling. *Nat. Commun.* *6*, 6118. <https://doi.org/10.1038/ncomms7118>.
22. Cancer Genome Atlas Research Network, Kandoth, C., Schultz, N., Cherniack, A.D., Akbani, R., Liu, Y., Shen, H., Robertson, A.G., Pashtan, I., Shen, R., et al. (2013). Integrated genomic characterization of endometrial carcinoma. *Nature* *497*, 67–73. <https://doi.org/10.1038/nature12113>.
23. Varela, I., Tarpey, P., Raine, K., Huang, D., Ong, C.K., Stephens, P., Davies, H., Jones, D., Lin, M.-L., Teague, J., et al. (2011). Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma. *Nature* *469*, 539–542. <https://doi.org/10.1038/nature09639>.
24. Jones, S., Wang, T.-L., Shih, I.-M., Mao, T.-L., Nakayama, K., Roden, R., Glas, R., Slamon, D., Diaz, L.A., Jr., Vogelstein, B., et al. (2010). Frequent Mutations of Chromatin Remodeling Gene ARID1A in Ovarian Clear Cell Carcinoma. *Science* *330*, 228–231. <https://doi.org/10.1126/science.1196333>.
25. Wang, K., Kan, J., Yuen, S.T., Shi, S.T., Chu, K.M., Law, S., Chan, T.L., Kan, Z., Chan, A.S.Y., Tsui, W.Y., et al. (2011). Exome sequencing identifies frequent mutation of ARID1A in molecular subtypes of gastric cancer. *Nat. Genet.* *43*, 1219–1223. <https://doi.org/10.1038/ng.982>.
26. Fontana, B., Gallerani, G., Salamon, I., Pace, I., Roncarati, R., and Ferracin, M. (2023). ARID1A in cancer: Friend or foe?

- Front. Oncol. 13, 1136248. <https://doi.org/10.3389/fonc.2023.1136248>.
27. Wiegand, K.C., Shah, S.P., Al-Agha, O.M., Zhao, Y., Tse, K., Zeng, T., Senz, J., McConechy, M.K., Anglesio, M.S., Kalloger, S.E., et al. (2010). ARID1A Mutations in Endometriosis-Associated Ovarian Carcinomas. *N. Engl. J. Med.* 363, 1532–1543. <https://doi.org/10.1056/NEJMoa1008433>.
  28. Santen, G.W.E., Aten, E., Sun, Y., Almomani, R., Gilissen, C., Nielsen, M., Kant, S.G., Snoeck, I.N., Peeters, E.A.J., Hillhorst-Hofstee, Y., et al. (2012). Mutations in SWI/SNF chromatin remodeling complex gene ARID1B cause Coffin-Siris syndrome. *Nat. Genet.* 44, 379–380. <https://doi.org/10.1038/ng.2217>.
  29. Tsurusaki, Y., Okamoto, N., Ohashi, H., Kosho, T., Imai, Y., Hibi-Ko, Y., Kaname, T., Naritomi, K., Kawame, H., Wakui, K., et al. (2012). Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. *Nat. Genet.* 44, 376–378. <https://doi.org/10.1038/ng.2219>.
  30. Coffin, G.S., and Siris, E. (1970). Mental Retardation With Absent Fifth Fingernail and Terminal Phalanx. *Am. J. Dis. Child.* 119, 433–439. <https://doi.org/10.1001/archpedi.1970.02100050435009>.
  31. Schrier, S.A., Bodurtha, J.N., Burton, B., Chudley, A.E., Chiong, M.A.D., D'Avanzo, M.G., Lynch, S.A., Musio, A., Nyazov, D.M., Sanchez-Lara, P.A., et al. (2012). The Coffin-Siris syndrome: A proposed diagnostic approach and assessment of 15 overlapping cases. *Am. J. Med. Genet.* 158A, 1865–1876. <https://doi.org/10.1002/ajmg.a.35415>.
  32. Schrier Vergano, S., Santen, G., Wiczorek, D., Wollnik, B., Matsumoto, N., and Deardorff, M.A. (1993). In *Coffin-Siris Syndrome*, GeneReviews®, M.P. Adam, D.B. Everman, G.M. Mirzaa, R.A. Pagon, S.E. Wallace, L.J. Bean, K.W. Gripp, and A. Amemiya, eds. (University of Washington, Seattle).
  33. Gans, C., and Northcutt, R.G. (1983). Neural Crest and the Origin of Vertebrates: A New Head. *Science* 220, 268–273. <https://doi.org/10.1126/science.220.4594.268>.
  34. Gilbert, S.F. (2000). Formation of the Neural Tube. In *Developmental Biology*, 6th edition (Sinauer Associates).
  35. Ahlstrom, J.D., and Erickson, C.A. (2009). The neural crest epithelial-mesenchymal transition in 4D: a 'tail' of multiple non-obligatory cellular mechanisms. *Development* 136, 1801–1812. <https://doi.org/10.1242/dev.034785>.
  36. Morrison, J.A., McLennan, R., Wolfe, L.A., Gogol, M.M., Meier, S., McKinney, M.C., Teddy, J.M., Holmes, L., Semerad, C.L., Box, A.C., et al. (2017). Single-cell transcriptome analysis of avian neural crest migration reveals signatures of invasion and molecular transitions. *Elife* 6, e28415. <https://doi.org/10.7554/eLife.28415>.
  37. Martik, M.L., and Bronner, M.E. (2021). Riding the crest to get a head: neural crest evolution in vertebrates. *Nat. Rev. Neurosci.* 22, 616–626. <https://doi.org/10.1038/s41583-021-00503-2>.
  38. Gilbert, S.F. (2000). The Neural Crest. In *Developmental Biology*, 6th edition (Sinauer Associates).
  39. Santagati, F., and Rijli, F.M. (2003). Cranial neural crest and the building of the vertebrate head. *Nat. Rev. Neurosci.* 4, 806–818. <https://doi.org/10.1038/nrn1221>.
  40. Chandler, R.L., and Magnuson, T. (2016). The SWI/SNF BAF-A complex is essential for neural crest development. *Dev. Biol.* 411, 15–24. <https://doi.org/10.1016/j.ydbio.2016.01.015>.
  41. Warlich, E., Kuehle, J., Cantz, T., Brugman, M.H., Maetzig, T., Galla, M., Filipczyk, A.A., Halle, S., Klump, H., Schöler, H.R., et al. (2011). Lentiviral Vector Design and Imaging Approaches to Visualize the Early Stages of Cellular Reprogramming. *Mol. Ther.* 19, 782–789. <https://doi.org/10.1038/mt.2010.314>.
  42. Dambrot, C., van de Pas, S., van Zijl, L., Brändl, B., Wang, J.W., Schaliq, M.J., Hoeben, R.C., Atsma, D.E., Mikkers, H.M., Mummery, C.L., and Freund, C. (2013). Polycistronic lentivirus induced pluripotent stem cells from skin biopsies after long term storage, blood outgrowth endothelial cells and cells from milk teeth. *Differentiation* 85, 101–109. <https://doi.org/10.1016/j.diff.2013.01.001>.
  43. Themeli, M., Chhatta, A., Boersma, H., Prins, H.J., Cordes, M., de Wilt, E., Farahani, A.S., Vandekerckhove, B., van der Burg, M., Hoeben, R.C., et al. (2020). iPSC-Based Modeling of RAG2 Severe Combined Immunodeficiency Reveals Multiple T Cell Developmental Arrests. *Stem Cell Rep.* 14, 300–311. <https://doi.org/10.1016/j.stemcr.2019.12.010>.
  44. Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* 34, 525–527. <https://doi.org/10.1038/nbt.3519>.
  45. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. <https://doi.org/10.1186/s13059-014-0550-8>.
  46. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. <https://doi.org/10.1093/bioinformatics/btq033>.
  47. Ramírez, F., Dündar, F., Diehl, S., Grüning, B.A., and Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* 42, W187–W191. <https://doi.org/10.1093/nar/gku365>.
  48. Prescott, S.L., Srinivasan, R., Marchetto, M.C., Grishina, I., Narvaiza, I., Selleri, L., Gage, F.H., Swigut, T., and Wysocka, J. (2015). Enhancer Divergence and cis-Regulatory Evolution in the Human and Chimp Neural Crest. *Cell* 163, 68–83. <https://doi.org/10.1016/j.cell.2015.08.036>.
  49. Buenrostro, J.D., Wu, B., Chang, H.Y., and Greenleaf, W.J. (2015). ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr. Protoc. Mol. Biol.* 109, 21.29.1–21.29.9. <https://doi.org/10.1002/0471142727.mb2129s109>.
  50. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25, 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>.
  51. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nussbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based Analysis of ChIP-Seq (MACS). *Genome Biol.* 9, R137. <https://doi.org/10.1186/gb-2008-9-9-r137>.
  52. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol. Cell.* 38, 576–589. <https://doi.org/10.1016/j.molcel.2010.05.004>.
  53. Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: scanning for occurrences of a given motif. *Bioinformatics* 27, 1017–1018. <https://doi.org/10.1093/bioinformatics/btr064>.
  54. Escalante, A., Murillo, B., Morenilla-Palao, C., Klar, A., and Herrera, E. (2013). Zic2-Dependent Axon Midline Avoidance Controls the Formation of Major Ipsilateral Tracts in the

- CNS. *Neuron* 80, 1392–1406. <https://doi.org/10.1016/j.neuron.2013.10.007>.
55. Gong, S., Zheng, C., Doughty, M.L., Losos, K., Didkovsky, N., Schambra, U.B., Nowak, N.J., Joyner, A., Leblanc, G., Hatten, M.E., and Heintz, N. (2003). A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. *Nature* 425, 917–925. <https://doi.org/10.1038/nature02033>.
  56. Nagai, T., Aruga, J., Minowa, O., Sugimoto, T., Ohno, Y., Noda, T., and Mikoshiba, K. (2000). *Zic2* regulates the kinetics of neurulation. *USA* 97, 1618–1623. <https://doi.org/10.1073/pnas.97.4.1618>.
  57. Hamburger, V., and Hamilton, H.L. (1951). A series of normal stages in the development of the chick embryo. *J. Morphol.* 88, 49–92. <https://doi.org/10.1002/jmor.1050880104>.
  58. Renier, N., Wu, Z., Simon, D.J., Yang, J., Ariel, P., and Tessier-Lavigne, M. (2014). iDISCO: A Simple, Rapid Method to Immunolabel Large Tissue Samples for Volume Imaging. *Cell* 159, 896–910. <https://doi.org/10.1016/j.cell.2014.10.010>.
  59. Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., et al. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* 17, 405–424. <https://doi.org/10.1038/gim.2015.30>.
  60. van der Sluijs, P.J., Joosten, M., Alby, C., Attié-Bitach, T., Gilmore, K., Dubourg, C., Fradin, M., Wang, T., Kurtz-Nelson, E.C., Ahlers, K.P., et al. (2022). Discovering a new part of the phenotypic spectrum of Coffin-Siris syndrome in a fetal cohort. *Genet. Med.* 24, 1753–1760. <https://doi.org/10.1016/j.gim.2022.04.010>.
  61. Ho, L., Jothi, R., Ronan, J.L., Cui, K., Zhao, K., and Crabtree, G.R. (2009). An embryonic stem cell chromatin remodeling complex, esBAF, is an essential component of the core pluripotency transcriptional network. *USA* 106, 5187–5191. <https://doi.org/10.1073/pnas.0812888106>.
  62. Ho, L., Ronan, J.L., Wu, J., Staahl, B.T., Chen, L., Kuo, A., Lesard, J., Nesvizhskii, A.I., Ranish, J., and Crabtree, G.R. (2009). An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. *USA* 106, 5181–5186. <https://doi.org/10.1073/pnas.0812889106>.
  63. Rada-Iglesias, A., Bajpai, R., Prescott, S., Bruggmann, S.A., Swigut, T., and Wysocka, J. (2012). Epigenomic Annotation of Enhancers Predicts Transcriptional Regulators of Human Neural Crest. *Cell Stem Cell* 11, 633–648. <https://doi.org/10.1016/j.stem.2012.07.006>.
  64. Bajpai, R., Chen, D.A., Rada-Iglesias, A., Zhang, J., Xiong, Y., Helms, J., Chang, C.-P., Zhao, Y., Swigut, T., and Wysocka, J. (2010). CHD7 cooperates with PBAF to control multipotent neural crest formation. *Nature* 463, 958–962. <https://doi.org/10.1038/nature08733>.
  65. Waddington, C.H. (1957). The Strategy of the Genes (Routledge). <https://doi.org/10.4324/9781315765471>.
  66. Cano, A., Pérez-Moreno, M.A., Rodrigo, I., Locascio, A., Blanco, M.J., del Barrio, M.G., Portillo, F., and Nieto, M.A. (2000). The transcription factor Snail controls epithelial–mesenchymal transitions by repressing E-cadherin expression. *Nat. Cell Biol.* 2, 76–83. <https://doi.org/10.1038/35000025>.
  67. Strobl-Mazzulla, P.H., and Bronner, M.E. (2012). A PHD12–Snail2 repressive complex epigenetically mediates neural crest epithelial-to-mesenchymal transition. *J. Cell Biol.* 198, 999–1010. <https://doi.org/10.1083/jcb.201203098>.
  68. Rogers, C.D., Saxena, A., and Bronner, M.E. (2013). Sip1 mediates an E-cadherin-to-N-cadherin switch during cranial neural crest EMT. *J. Cell Biol.* 203, 835–847. <https://doi.org/10.1083/jcb.201305050>.
  69. Cheung, M., and Briscoe, J. (2003). Neural crest development is regulated by the transcription factor Sox9. *Development* 130, 5681–5693. <https://doi.org/10.1242/dev.00808>.
  70. Simões-Costa, M., Tan-Cabugao, J., Antoshechkin, I., Sauka-Spengler, T., and Bronner, M.E. (2014). Transcriptome analysis reveals novel players in the cranial neural crest gene regulatory network. *Genome Res.* 24, 281–290. <https://doi.org/10.1101/gr.161182.113>.
  71. Vierbuchen, T., Ling, E., Cowley, C.J., Couch, C.H., Wang, X., Harmin, D.A., Roberts, C.W.M., and Greenberg, M.E. (2017). AP-1 Transcription Factors and the BAF Complex Mediate Signal-Dependent Enhancer Selection. *Mol. Cell.* 68, 1067–1082.e12. <https://doi.org/10.1016/j.molcel.2017.11.026>.
  72. Kim, S., Morgunova, E., Naqvi, S., Goovaerts, S., Bader, M., Koska, M., Popov, A., Luong, C., Pogson, A., Swigut, T., et al. (2024). DNA-guided transcription factor cooperativity shapes face and limb mesenchyme. *Cell* 187, 692–711.e26. <https://doi.org/10.1016/j.cell.2023.12.032>.
  73. Bhattacharya, D., Rothstein, M., Azambuja, A.P., and Simoes-Costa, M. (2018). Control of neural crest multipotency by Wnt signaling and the *Lin28/let-7* axis. *Elife* 7, e40556. <https://doi.org/10.7554/eLife.40556>.
  74. Hovland, A.S., Bhattacharya, D., Azambuja, A.P., Pramio, D., Copeland, J., Rothstein, M., and Simoes-Costa, M. (2022). Pluripotency factors are repurposed to shape the epigenomic landscape of neural crest. *Dev. Cell* 57, 2257–2272.e5. <https://doi.org/10.1016/j.devcel.2022.09.006>.
  75. Wang, J., Xiao, Y., Hsu, C.-W., Martinez-Traverso, I.M., Zhang, M., Bai, Y., Ishii, M., Maxson, R.E., Olson, E.N., Dickinson, M.E., et al. (2016). Yap and Taz play a crucial role in neural crest-derived craniofacial development. *Development* 143, 504–515. <https://doi.org/10.1242/dev.126920>.
  76. Hong, C.-S., and Saint-Jeannet, J.-P. (2007). The Activity of Pax3 and Zic1 Regulates Three Distinct Cell Fates at the Neural Plate Border. *Mol. Biol. Cell* 18, 2192–2202. <https://doi.org/10.1091/mbc.E06-11-1047>.
  77. Milet, C., Maczkowiak, F., Roche, D.D., and Monsoro-Burq, A.H. (2013). Pax3 and Zic1 drive induction and differentiation of multipotent, migratory, and functional neural crest in *Xenopus* embryos. *USA* 110, 5528–5533. <https://doi.org/10.1073/pnas.1219124110>.
  78. Elms, P., Siggers, P., Napper, D., Greenfield, A., and Arkell, R. (2003). *Zic2* is required for neural crest formation and hind-brain patterning during mouse development. *Dev. Biol.* 264, 391–406. <https://doi.org/10.1016/j.ydbio.2003.09.005>.
  79. Nakata, K., Nagai, T., Aruga, J., and Mikoshiba, K. (1997). *Xenopus Zic3*, a primary regulator both in neural and neural crest development. *USA* 94, 11980–11985. <https://doi.org/10.1073/pnas.94.22.11980>.
  80. Nakata, K., Nagai, T., Aruga, J., and Mikoshiba, K. (1998). *Xenopus Zic* family and its role in neural and neural crest development I. *Mech. Dev.* 75, 43–51. [https://doi.org/10.1016/S0925-4773\(98\)00073-2](https://doi.org/10.1016/S0925-4773(98)00073-2).
  81. Bellchambers, H.M., Barratt, K.S., Diamand, K.E.M., and Arkell, R.M. (2021). SUMOylation Potentiates ZIC Protein Activity to Influence Murine Neural Crest Cell Specification.

- Int. J. Mol. Sci. 22, 10437. <https://doi.org/10.3390/ijms221910437>.
82. Brown, S.A., Warburton, D., Brown, L.Y., Yu, C.Y., Roeder, E.R., Stengel-Rutkowski, S., Hennekam, R.C., and Muenke, M. (1998). Holoprosencephaly due to mutations in ZIC2, a homologue of Drosophila odd-paired. *Nat. Genet.* 20, 180–183. <https://doi.org/10.1038/2484>.
  83. Solomon, B.D., Lachawan, F., Mercier, S., Clegg, N.J., Delgado, M.R., Rosenbaum, K., Dubourg, C., David, V., Olney, A.H., Wehner, L.-E., et al. (2010). Mutations in ZIC2 in human holoprosencephaly: description of a Novel ZIC2 specific phenotype and comprehensive analysis of 157 individuals. *J. Med. Genet.* 47, 513–524. <https://doi.org/10.1136/jmg.2009.073049>.
  84. Wu, J.N., and Roberts, C.W.M. (2013). ARID1A Mutations in Cancer: Another Epigenetic Tumor Suppressor? *Cancer Discov.* 3, 35–43. <https://doi.org/10.1158/2159-8290.CD-12-0361>.
  85. Nitaraska, J., Smith, J.G., Sherlock, W.T., Hillege, M.M.G., Nott, A., Barshop, W.D., Vashisht, A.A., Wohlschlegel, J.A., Mitter, R., and Riccio, A. (2016). A Functional Switch of NuRD Chromatin Remodeling Complex Subunits Regulates Mouse Cortical Development. *Cell Rep.* 17, 1683–1698. <https://doi.org/10.1016/j.celrep.2016.10.022>.
  86. Aruga, J. (2004). The role of *Zic* genes in neural development. *Mol. Cell. Neurosci.* 26, 205–221. <https://doi.org/10.1016/j.mcn.2004.01.004>.
  87. Murillo, B., Ruiz-Reig, N., Herrera, M., Fairén, A., and Herrera, E. (2015). Zic2 Controls the Migration of Specific Neuronal Populations in the Developing Forebrain. *J. Neurosci.* 35, 11266–11280. <https://doi.org/10.1523/JNEUROSCI.0779-15.2015>.
  88. Brewster, R., Lee, J., and Ruiz i Altaba, A. (1998). Gli/Zic factors pattern the neural plate by defining domains of cell differentiation | *Nature* 393, 579–583. <https://doi.org/10.1038/31242>.
  89. Groves, A.K., and LaBonne, C. (2014). Setting appropriate boundaries: Fate, patterning and competence at the neural plate border. *Dev. Biol.* 389, 2–12. <https://doi.org/10.1016/j.ydbio.2013.11.027>.
  90. Khudyakov, J., and Bronner-Fraser, M. (2009). Comprehensive spatiotemporal analysis of early chick neural crest network genes. *Dev. Dynam.* 238, 716–723. <https://doi.org/10.1002/dvdy.21881>.
  91. Sauka-Spengler, T., and Bronner-Fraser, M. (2008). A gene regulatory network orchestrates neural crest formation. *Nat. Rev. Mol. Cell Biol.* 9, 557–568. <https://doi.org/10.1038/nrm2428>.
  92. Szabó, A., and Mayor, R. (2018). Mechanisms of Neural Crest Migration. *Annu. Rev. Genet.* 52, 43–63. <https://doi.org/10.1146/annurev-genet-120417-031559>.
  93. Drummond, D.L., Cheng, C.S., Selland, L.G., Hocking, J.C., Prichard, L.B., and Waskiewicz, A.J. (2013). The role of Zic transcription factors in regulating hindbrain retinoic acid signaling. *BMC Dev. Biol.* 13, 31. <https://doi.org/10.1186/1471-213X-13-31>.
  94. Herrera, E., Brown, L., Aruga, J., Rachel, R.A., Dolen, G., Mikoshiba, K., Brown, S., and Mason, C.A. (2003). Zic2 Patterns Binocular Vision by Specifying the Uncrossed Retinal Projection. *Cell* 114, 545–557. [https://doi.org/10.1016/S0092-8674\(03\)00684-6](https://doi.org/10.1016/S0092-8674(03)00684-6).
  95. García-Frigola, C., Carreres, M.I., Vegar, C., Mason, C., and Herrera, E. (2008). Zic2 promotes axonal divergence at the optic chiasm midline by EphB1-dependent and -independent mechanisms. *Development* 135, 1833–1841. <https://doi.org/10.1242/dev.020693>.
  96. Morenilla-Palao, C., López-Cascales, M.T., López-Atalaya, J.P., Baeza, D., Calvo-Díaz, L., Barco, A., and Herrera, E. (2020). A Zic2-regulated switch in a noncanonical Wnt/βcatenin pathway is essential for the formation of bilateral circuits. *Sci. Adv.* 6, eaaz8797. <https://doi.org/10.1126/sciadv.aaz8797>.
  97. Fernández-Nogales, M., López-Cascales, M.T., Murcia-Belmonte, V., Escalante, A., Fernández-Albert, J., Muñoz-Viana, R., Barco, A., and Herrera, E. (2022). Multiomic Analysis of Neurons with Divergent Projection Patterns Identifies Novel Regulators of Axon Pathfinding. *Adv. Sci.* 9, 2200615. <https://doi.org/10.1002/adv.202200615>.
  98. Haensel, D., and Dai, X. (2018). Epithelial-to-mesenchymal transition in cutaneous wound healing: where we are and where we are heading. *Dev. Dynam.* 247, 473–480. <https://doi.org/10.1002/dvdy.24561>.
  99. Barriere, G., Fici, P., Gallerani, G., Fabbri, F., and Rigaud, M. (2015). Epithelial Mesenchymal Transition: a double-edged sword. *Clin. Transl. Med.* 4, 14. <https://doi.org/10.1186/s40169-015-0055-4>.
  100. Kalluri, R., and Weinberg, R.A. (2009). The basics of epithelial-mesenchymal transition. *J. Clin. Invest.* 119, 1420–1428. <https://doi.org/10.1172/JCI39104>.
  101. Thiery, J.P. (2002). Epithelial-mesenchymal transitions in tumour progression. *Nat. Rev. Cancer* 2, 442–454. <https://doi.org/10.1038/nrc822>.
  102. Tomihara, H., Carbone, F., Perelli, L., Huang, J.K., Soeung, M., Rose, J.L., Robinson, F.S., Lissanu Deribe, Y., Feng, N., Takeda, M., et al. (2021). Loss of ARID1A Promotes Epithelial–Mesenchymal Transition and Sensitizes Pancreatic Tumors to Proteotoxic Stress. *Cancer Res.* 81, 332–343. <https://doi.org/10.1158/0008-5472.CAN-19-3922>.
  103. Somsuan, K., Peerapen, P., Boonmark, W., Plumworasawat, S., Samol, R., Sakulsak, N., and Thongboonkerd, V. (2019). ARID1A knockdown triggers epithelial-mesenchymal transition and carcinogenesis features of renal cells: role in renal cell carcinoma. *Faseb. J.* 33, 12226–12239. <https://doi.org/10.1096/fj.201802720RR>.
  104. Wilson, M.R., Reske, J.J., Holladay, J., Wilber, G.E., Rhodes, M., Koeman, J., Adams, M., Johnson, B., Su, R.-W., Joshi, N.R., et al. (2019). ARID1A and PI3-kinase pathway mutations in the endometrium drive epithelial transdifferentiation and collective invasion. *Nat. Commun.* 10, 3554. <https://doi.org/10.1038/s41467-019-11403-6>.