The Role of Long-Range Super-Enhancers in Craniofacial Morphogenesis

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

SANDRA KESSLER

Basel, 2023

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von

Erstbetreuer: Prof. Dr. Filippo M. Rijli
Zweitbetreuer: Prof. Dr. Antoine H.F.M. Peters
Externe Expertin: Prof. Dr. Licia Selleri

Basel, den 21.02.2023

Prof. Dr. Marcel Mayor

The Dean of Faculty

Abstract

One of the most defining characteristics that sets individuals apart from each other is the shape of their face. Craniofacial morphogenesis is a fascinating and highly complex process that involves the coordination of multiple genetic, cellular, and environmental factors. A central component of this process are cranial neural crest cells (CNCCs), a population of cells that temporarily possess stem cell-like characteristics, with the potential to differentiate into various types of cells. Furthermore, growing evidence suggests that distal *cis*-regulatory elements called enhancers are major contributors to craniofacial development. In particular, large clusters of enhancers known as super-enhancers have been associated with genes that play crucial roles in differentiation and cell identity determination. However, while many putative craniofacial enhancers have been identified, only a few have been functionally validated. Furthermore, the role of super-enhancers (SEs) in craniofacial morphogenesis has not been explored so far. Thus, we decided to focus on the identification of putative craniofacial super-enhancers.

Using a combination of ChIP-seq and Promoter-Capture Hi-C (PCHi-C) data we identified 2,322 putative long-range SEs of which 177 specifically interact with transcription factor genes involved in the establishment of CNCCs' positional identity in craniofacial development. Out of this list of SEs, five SEs targeting *Hoxa2*, a master regulator of second pharyngeal arch (PA2) CNCCs fate, specifically in PA2-derived CNCCs stood out due to their extremely long genomic distance to *Hoxa2* and their ability to interact with *Hoxa2* across a topologically associating domain (TAD) boundary. This configuration struck us since TADs are known to spatially insulate regulatory elements and preferentially show interactions within themselves, thereby creating discrete functional and structural blocks that promote intra-domain enhancer-promoter interactions, while limiting contacts between regulatory elements across TADs. According to their genomic distance to each other and the distribution of H3K27ac, we separated the *Hoxa2* SEs into *Hoxa* Inter-TAD Regulatory Element 1 and 2 (HIRE1 and HIRE2). Both of these large regulatory elements are highly conserved within eutherian mammals and upon deletion in mice, we observed varying phenotypes. Homozygous deletion of HIRE1 resulted in a phenocopy of the full homeotic *Hoxa2* knock-out phenotype and abnormalities in PA3- and PA4-derived structures, which correlated a strong downregulation in gene expression of *Hoxa2* and *Hoxa3* in CNCCs and hindbrain rhombomeres. On the other hand, deletion of HIRE2 only caused mild alterations of CNCC-derived skeletal structures and a small external ear on a *Hoxa2* haploinsufficient sensitized background.

These phenotypes are particularly striking because both *Hoxa2* and *Hoxa3* were previously believed to be regulated by proximal enhancers in CNCCs and the hindbrain. A transcription factor motif analysis in combination with the analysis of PCHi-C data of PA2 CNCCs from Hoxa2 knockout mice and published ChIP-seq data for Hoxa2, Pbx, and Meis from PA2 CNCCs suggest that Hoxa2 is at least partially involved in its own regulation together with its co-factors Meis and Pbx.

Taken together, the findings presented in this thesis give new insights into the importance of long-range SEs in controlling key developmental transcription factors in craniofacial morphogenesis, which could have implications for the study of disease and rare congenital disorders. Additionally, they change our understanding of the regulation of *Hoxa2* in CNCCs from a model that solely relies on multiple proximal enhancers, to a model that is based on multiple long-range inter-TAD SEs.

Acknowledgements

I am deeply grateful to all of those who have supported me throughout the journey of completing this project.

First and foremost, I would like to express my heartfelt thanks to my supervisor, Prof. Filippo M. Rijli, for his expert guidance, unwavering support, and constant encouragement throughout my PhD. His expertise and experience provided invaluable direction for my research and I am very grateful for the time and effort he invested in me. Thank you for always believing in me.

I would also like to express my sincerest appreciation to my colleague and advisor Maryline Minoux. Her expertise and perspectives were instrumental in shaping my work and I am grateful for her willingness to always take the time to review and provide feedback on my progress. I am also grateful for the many stimulating discussions and brainstorming sessions that we had throughout the course of my PhD.

I would like to thank the members of my thesis committee, Antoine H.F.M. Peters, Licia Selleri, and Michael Stadler, for their valuable input, constructive feedback, and invaluable expertise. I would like to express special thanks to Michael for his availability and help with computational analyses. His positive attitude and friendly demeanor have made working with him a pleasure and I truly appreciate the time and effort he has put into helping me. Special thanks also go to Licia for her kindness, and willingness to listen, offer advice, and provide support, which have been invaluable to me and have made a significant impact on my well-being.

I would also like to extend my gratitude to the technicians of the Rijli lab, Nathalie Vilain and Sebastien Ducret. They are the backbone of the lab and without them nothing would work. I am also grateful to all other present and past members of the Rijli lab, namely Fiona Ross, Adwait Salvi, Gabriele Pumo, Yousra Ben Zouari, Guoliang Lyu, Onkar Joshi, Leslie Bargsted, Dania Machlab, Upasana Maheshwari, Taro Kitazawa, and Joachim Wolff, for their support, encouragement, and collaboration throughout my PhD. I loved working with every single one of them and greatly appreciated our scientific and non-scientific discussions and all the laughs and fun times we had together.

I would also like to acknowledge the support of the FMI and all its facilities, without which this work would not have been possible. In particular, I would like to extend my gratitude to Charlotte from the computational biology platform, Sebastien, Sirisha, Stéphane, and Eliza from the functional genomics facility, all members of the media kitchen and animal facility, and

Contents

Abbreviations

3C	Chromosome Conformation Capture
3D	Three dimensional
AP	Anteroposterior
ATAC − seq	Assay for Transposase-Accessible Chromatin with high-throughput sequencing
CGI	CpG island
ChIP − seq	Chromatin Immunoprecipitation Sequencing
CNCCs	Cranial neural crest cells
DNA	Deoxyribonucleic acid
EMT	Epithelial-to-mesenchymal transition
ESCs	embryonic stem cells
FNP	Frontonasal process
GRN	Gene regulatory network
H3K27ac	Histone 3 Lysine 27 acetylation
H3K27me3	Histone 3 Lysine 27 tri-methylation
H3K4me1	Histone 3 Lysine 4 mono-methylation
H3K4me2	Histone 3 Lysine 4 di-methylation
H3K4me3	Histone 3 Lysine 4 tri-methylation
H3K9me	Histone 3 lysine 9 methylation
Hi − C	High-throughput Chromosome Conformation Capture
HICE	Highly interconnected enhancers
HIRE1	*Hoxa* inter-TAD regulatory element 1
HIRE2	*Hoxa* inter-TAD regulatory element 2
Kb	Kilobases
Mb	Megabases
Md	Mandibular process
MET	Mesenchymal-to-epithelial transition
Mx	Maxillary process
NC	Neural crest
NCCs	Neural crest cells
PA	Pharyngeal arch

PcG	Polycomb group
PCHi − C	Promoter-Capture Hi-C
PG	Paralogue group
PolII	RNA Polymerase II
*PRC*1	Polycomb repressive complex 1
*PRC*2	Polycomb repressive complex 2
R	Rhombomere
RA	Retinoic acid
RNA − seq	RNA-sequencing
RNA	Ribonucleic acid
ROSE	rank ordering of super-enhancers
SE	Super-enhancer
T F	Transcription factor
T SS	Transcription start site

Introduction

How an organism, made up of hundreds to billions of cells, is formed from just a single cell has fascinated scientists for many years now. In mammals, this process begins with the exponential cell division of the zygote into two, four, eight, and so on cells, until the zygote has formed a sphere of cells known as the blastocyst. Post implantation into the uterine wall, the development of the various tissues and organs of the body starts through a combination of cell division, cell migration, and cell differentiation.

One of the most intricate processes in embryogenesis is the development of the face. Across different species of vertebrates, there is a wide range of facial shapes, facilitating different ways of feeding, communicating, and performing specialized functions. The process of craniofacial morphogenesis begins early in embryonic development and is a highly complex, multi-step process precisely controlled at the genomic, molecular, and cellular level. Central to this process is a transient population of cells with stem cell-like properties known as neural crest cells (NCCs), capable of giving rise to a wide range of cell types. The precise coordination of transcription in both time and space is crucial for this process and plays a vital role in the development of the whole organism. Such spatiotemporal gene expression programs are established by the coordinated action of *cis*-regulatory elements known as enhancers that engage in physical contact with their target promoters. Additionally, transcription is intimately linked to the three-dimensional (3D) chromatin landscape, which is heavily influenced by epigenetic mechanisms, such as DNA methylation and histone modifications. It is believed that morphological intra- and inter-species variability of structures such as the vertebrate head are caused by changes in the enhancer sequence and activity. Moreover, disruption of these programs by genomic alterations, including deletions, duplications, inversions, insertions, and translocations, not only in genes but also in *cis*-regulatory elements have been shown to cause malformations and disease. Thus, the study of craniofacial enhancers would provide major new insights into the regulation of the morphogenesis of facial structures and could help in determining the etiology behind facial abnormalities and congenital diseases.

1.1   Mechanisms of gene expression regulation in development

Over the course of development, a single genome produces diverse cell types with unique characteristics and functions mediated by differential gene expression patterns that are critical for cell fate decisions, lineage commitment, organ formation, and the development of the overall body plan. The 3D genome organization as well as transcriptional and epigenetic regulation are key mechanisms that control robust and precise regulation of transcription in space and time in order to facilitate these intricate gene expression changes in cells throughout embryogenesis and organogenesis. Transcriptional regulation involves the binding of transcription factors (TFs) to specific DNA sequences called *cis*-regulatory elements, such as promoters and enhancers, which then recruit the RNA polymerase II (PolII) machinery to transcribe the DNA into RNA. 3D chromatin folding mediates the positioning of the chromosomes in the nucleus and is required for enhancer-driven long-range gene regulation. Epigenetic regulation refers to changes in gene expression that occur without changes to the underlying DNA sequence, but rather through modifications to the structure of chromatin, such as DNA methylation and histone modifications.

Promoters are regulatory elements in close proximity to the transcription start site (TSS) of the gene they regulate and are sufficient to initiate transcription, containing binding sites for the PolII enzyme and TFs. On the other hand, enhancers are *cis*-acting DNA elements that can be located up to several megabases (Mb) away from their target gene promoter. They play a key role in the regulation of precise and robust gene expression by physically interacting with the promoter region of their target genes via a process called loop extrusion. Taken together, gene promoters and their enhancers form gene regulatory landscapes that shape the complex and dynamic patterns of transcriptional activity required for development. Such regulatory landscapes specifically fold in the nucleus, forming complex 3D structures that influence enhancer-promoter communication and drive cell type-specific gene expression. As such, the 3D organization of the genome also plays a crucial role in the regulation of gene expression by mediating the physical proximity of enhancers and promoters. Nevertheless, how exactly enhancers find their target genes over large genomic distances in the nuclear space and what prevents them from aberrantly activating other genes in their immediate vicinity remains poorly understood and is an active field of research (Andrey & Mundlos, 2017; Long et al., 2016; Robson et al., 2019; Schoenfelder & Fraser, 2019; Shlyueva et al., 2014). Disruptions in the enhancer-promoter circuitry through chromosomal rearrangement, genetic variation within enhancers, or epigenetic modulation have been linked to an increasing number of diseases and developmental malformations, highlighting the key role of enhancers in

regulating transcription (Claringbould & Zaugg, 2021; Spielmann et al., 2018; Zaugg et al., 2022).

In addition, epigenetic modifications can influence gene expression by controlling the accessibility of promoters and enhancers to the transcription machinery (e.g. TFs and PolII). Most of these modifications are dynamic, as they can change over time in response to environmental cues and internal signals. For example, different cell types have distinct epigenetic signatures that reflect their specific gene expression patterns. Certain epigenetic marks have been associated with active or inactive chromatin, and thus are commonly used to predict the location of putative promoters and enhancers.

1.1.1 The Epigenome

The epigenome refers to the complete set of epigenetic modifications that occur on the genome, including chemical modifications to the DNA molecule such as DNA methylation, and modifications to the histone proteins around which the DNA is wrapped, such as acetylation and methylation. By regulating the organization of chromatin and the accessibility of DNA, these chemical modifications affect how the genome functions in various developmental stages, tissue types, and disease states (Bernstein et al., 2007; Zhou et al., 2011). Therefore, it is of importance to shortly review some of the organizational layers of chromatin as well as the techniques to detect them (Fig. 1a): Epigenetic modifications of nucleosomes mapped by ChIP-seq (chromatin immunoprecipitation followed by sequencing), chromatin accessibility identified by ATAC-seq (Assay for Transposase-Accessible Chromatin with high-throughput sequencing) and DNA methylation identified by bisulfite sequencing.

Nucleosomes are fundamental building blocks of eukaryotic chromatin, which is formed by coiling approximately 146 base pairs of DNA around a core made up of a histone octamer (Luger et al., 1997). Chromatin is divided into two main structural domains, euchromatin and heterochromatin. Euchromatin is an open and accessible structure that allows for the expression of genes, while heterochromatin is a tightly packed structure that represses gene expression. Histone modifications such as acetylation and methylation can influence the formation of these domains and in turn, the accessibility of genes to the transcription machinery (Tamaru, 2010). These modifications can be mapped genome-wide by ChIP-seq. These ChIP-seq maps have linked specific modifications with either the activation or repression of genes, as well as different genomic features such as promoters, transcribed regions, enhancers, and insulators (Fig. 1b)(Zhou et al., 2011). For instance, acetylation of Histone 3 (H3) and methylation of H3 lysine 4 (H3K4me) are associated with chromatin accessibility and transcriptional activity and are consequently found at most active
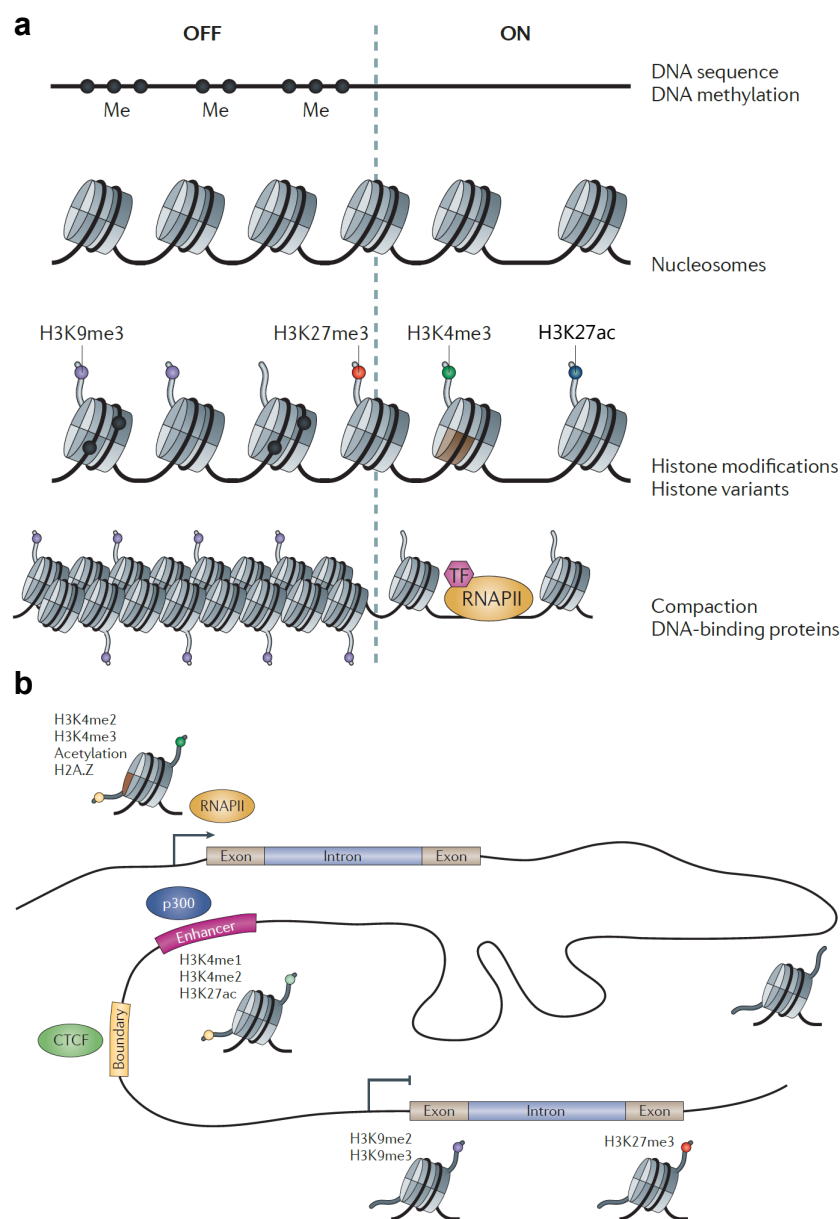
Figure 1. Epigenetic layers of chromatin involved in transcription regulation. a There are multiple epigenetic features at different levels of chromatin organization associated with inactive (off) or active (on) transcription. Inactive chromatin is characterized by DNA methylation, tightly packed nucleosomes, and repressive histone modifications. Active chromatin is open and accessible, allowing the transcription machinery to access the DNA, and marked by histone modifications associated with active transcription. **b** Promoters, enhancers, and boundary elements are indicated on a schematic genomic region. Active promoters are commonly marked by histone H3 lysine 4 dimethylation (H3K4me2), H3K4me3, histone H3 lysine 27 acetylation (H3K27ac), and H2A.Z. Enhancers are marked by H3K4me1, H3K4me2, and H3K27ac and are often bound by the histone acetyltransferase p300. Repressed genes are generally embedded in domains of histone H3 lysine 9 methylation (H3K9me2/3) or histone H3 lysine 27 tri-methylation (H3K27me3). CTCF binds many sites in convergent orientation to act as boundary element for cohesin-mediated loop extrusion, thereby creating insulated regulatory domains called topologically associated domains (TADs). The histone variant H2A.Z is colored in brown and the different histone modifications in **a** and **b** are indicated by colored dots. Figure adapted from Zhou et al., 2011.

promoters and enhancers. Specifically, H3K4 trimethylation (H3K4me3) mainly occurs at active promoters, H3K4 monomethylation (H3K4me1) at active enhancers, whereas H3K4 dimethylation (H3K4me2) and H3 lysine 27 acetylation (H3K27ac) marks both active promoters and enhancers (Bernstein et al., 2007; Zhou et al., 2011). In contrast, methylated H3 lysine 9 (H3K9me) has a clear role in heterochromatin formation and gene silencing (Bernstein et al., 2007; Zhou et al., 2011). The Trimethylation of H3 lysine 27 (H3K27me3) is probably one of the most well-studied histone modifications and has been associated with a wide range of mostly repressive regulatory mechanisms, including chromatin compaction, bivalency, and 3D genome architecture (Blanco et al., 2020). The mark is catalyzed by the Polycomb repressive complex 2 (PRC2) and often covers expansive genomic regions (Bernstein et al., 2007; Zhou et al., 2011). The term bivalency or bivalent domain was introduced by Bernstein and colleagues (Bernstein et al., 2006) to describe the seemingly contradictory co-occurrence of H3K4me3 and H3K27me3 at highly conserved non-coding elements associated with promoters of developmentally expressed TFs in mouse embryonic stem cells (mESCs). In addition, they proposed that promoter bivalency represses gene expression and keeps them poised for activation upon differentiation (Bernstein et al., 2006). Bivalent promoters are frequently located in CpG-rich regions, which are believed to play a role in recruiting both Polycomb group (PcG) and Trithorax group proteins in ESCs (Blanco et al., 2020). Later it was found that also at enhancers H3K4me1/H3K4me2 and H3K27me3 marks co-occur in mESCs and human ESCs (hESCs) (Rada-Iglesias et al., 2011; Zentner et al., 2011). Similar to promoters, bivalent enhancers are poised for activation during development and lack the active H3K27ac mark (Blanco et al., 2020). Promoter and enhancer bivalency is not a feature specific to ESCs. For example, it has been shown in mice that bivalency at promoters and enhancers of positional genes, involved in craniofacial development, is required for proper craniofacial morphogenesis (Minoux et al., 2017). It should be pointed out that, despite the presence of PcG proteins, bivalent domains are situated in regions of open chromatin and are associated with the active euchromatin (Blanco et al., 2020). Finally, PcG-bound regions have been found to form a spatial network in ESCs, which consists of both inter- and intra-chromosome interactions including genes with bivalent promoters that are poorly expressed, such as the HOX cluster genes, which are among the most strongly Polycomb-occupied and H3K27me3-enriched regions in ESCs (Denholtz et al., 2013; Freire-Pritchett et al., 2017; Kundu et al., 2017; Schoenfelder, Sugar, et al., 2015). Upon differentiation, bivalent genes that become active lose bound PcG proteins and consequently leave the repressive Polycomb network (Bonev et al., 2017; Joshi et al., 2015). In addition, it has been proposed that interactions between bivalent promoters and poised enhancers, which depend on PcG

binding, are pre-formed in ESCs in order to provide robust gene activation during differentiation (Cruz-Molina et al., 2017).

Chromatin accessibility is the extent to which nuclear macromolecules, such as TFs, can make physical contact with DNA that is packaged in chromatin and is determined by the presence and organization of nucleosomes and other chromatin-binding factors that restrict access to DNA. The arrangement of nucleosomes across the genome is not uniform; while densely located within facultative and constitutive heterochromatin, histones are less present at regulatory loci, including enhancers, insulators, and the gene body of transcribed genes (Klemm et al., 2019). Accessibility of chromatin is regulated by various modifications to the DNA and associated histone proteins, such as methylation and acetylation, which can either compact or loosen the chromatin structure. The majority of TFs bind to open chromatin and actively compete with histones and other chromatin-binding proteins to regulate nucleosome occupancy, thereby promoting local access to DNA (Klemm et al., 2019). In turn, the accessibility of DNA in a specific cell type influences TF binding (Klemm et al., 2019). On the other hand, a special class of TFs, called pioneer factors, have the unique ability to open closed chromatin to activate gene expression (Balsalobre & Drouin, 2022). Thus, chromatin accessibility is a reflection of the combined TF binding and thereby the regulatory capability of a specific genetic location. Techniques such as ATAC-seq(Buenrostro et al., 2015) can be used for establishing genome-wide enhancer footprints.

DNA methylation is the process by which a methyl group is added to the cytosine of a CpG dinucleotide in DNA and is typically associated with transcriptional repression and heterochromatin (Isbel et al., 2022). In vertebrates, this CpG methylation is essential and covers most CpGs in the genome except active regulatory elements (Isbel et al., 2022). Gene regulation can be affected by DNA methylation, specifically at CpG islands, which are a class of CpG-rich regions that constitute the majority of promoters in vertebrate genomes (Bird et al., 1985). DNA methylation represses transcription by directly or indirectly preventing the binding of transcription factors to promoters and enhancers (Isbel et al., 2022). This epigenetic mark can be experimentally detected at base pair resolution by bisulfite sequencing.

1.1.2   3D genome organization

The chromatin is organized within the 3D nuclear space into hierarchical layers, which are thought to be structural and functional building blocks of genome organization. These include compartments, topologically associating domains (TADs), and chromatin loops (Fig. 2). Our current knowledge of the 3D genome structure is mainly derived from experiments based on microscopy

and 3C (chromosome conformation capture) based techniques, such as 4C, 5C, and Hi-C (high-throughput chromosome conformation capture)(Lieberman-Aiden et al., 2009; McCord et al., 2020). 3C-type methods are proximity-ligation assays, meaning they rely on formaldehyde fixation to crosslink chromosomal regions in close spatial proximity, followed by restriction enzyme digestion of the crosslinked chromatin. After this step, proximal overhangs created by the restriction enzymes are ligated. The interaction rate between two genomic loci can be determined by the frequency at which the ligation product resulting from that interaction is observed, which can be measured by sequencing. The more regular the interaction between two loci, the more frequent the resulting ligation product will be. Each of the 3C techniques can be used to address different scientific questions: 3C measures the interaction between two specific genomic loci (referred to as 'one-to-one'), while 4C maps all interactions that one locus has with the surrounding linear genome ('one-to-many'). 5C is known as the 'many-to-many' technique giving a two-dimensional contact map of all interactions within a region, which then was extended by high-throughput sequencing to Hi-C ('all-to-all')(Schoenfelder & Fraser, 2019). In particular, Hi-C has enabled the genome-wide measurement of physical contacts within and across chromosomes at high-resolution (McCord et al., 2020; Pachano et al., 2022; Rowley & Corces, 2018; Spielmann et al., 2018).

On a large scale, chromosomes are spatially segregated into two compartments called A and B compartments. The A compartment is generally associated with accessible, transcriptionally active euchromatin, whereas the B compartment is associated with condensed, transcriptionally silent heterochromatin (Lieberman-Aiden et al., 2009; Spielmann et al., 2018). Furthermore, A compartments are positioned near nuclear speckles, known to be favorable for transcription, while B compartments are mainly located at the nuclear envelope or nucleolus, which are unfavorable for transcription (S. S. Rao et al., 2014; Robson et al., 2019). These compartments are further subdivided into submegabase-sized regions called topologically associated domains (TADs) that preferentially interact within themselves to create discrete functional and structural blocks (Dixon et al., 2012; Nora et al., 2012; Sexton et al., 2012). Many TAD structures are conserved throughout cell differentiation and across cell types (Dixon et al., 2012; Zhan et al., 2017), and a very prominent model for their formation is by the loop extrusion activity mediated by CTCF and cohesin (de Wit et al., 2015; Guo et al., 2015; S. S. Rao et al., 2014). The cohesin complex, which forms a ring structure, is able to extrude DNA loops until it reaches a barrier, which usually consists of CTCF bound to DNA in opposing orientation (I. F. Davidson et al., 2019; I. F. Davidson et al., 2022; Fudenberg et al., 2016; Sanborn et al., 2015). Indeed CTCF binding sites are enriched at TAD boundaries (Rowley & Corces, 2018; Spielmann et al., 2018) and depletion of either cohesin
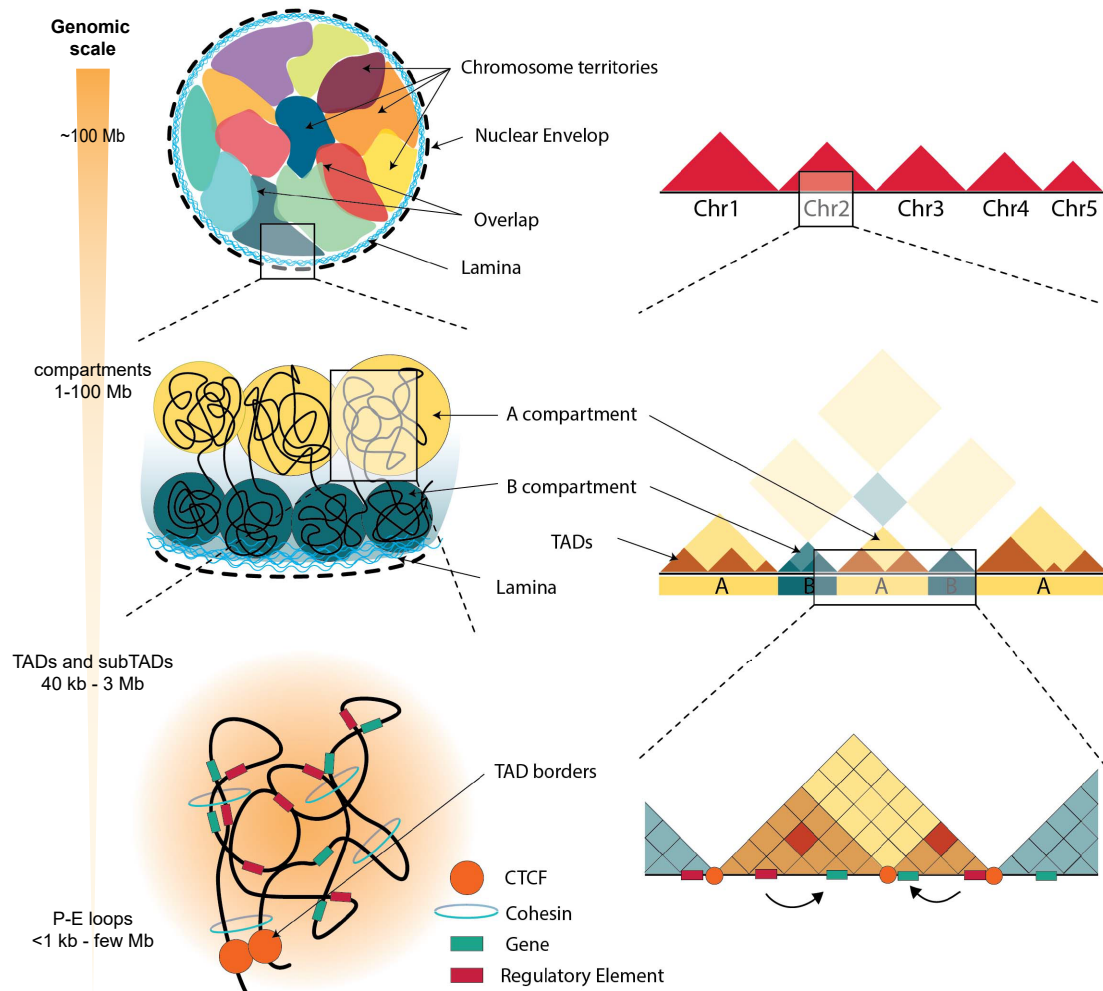
Figure 2. Layers of chromatin organization - from the 3D nucleus to the linear genome. At the level of the nucleus each chromosome resides in a specific territory and interactions across chromosomes are rare occurrences. At the chromosomal level, open and closed chromatin separates into multi-megabase-sized A (yellow) and B (blue) compartments associated with transcriptionally active and inactive genomic regions, respectively. B compartments are predominantly located at the nuclear lamina and A and B compartments preferentially interact with other A and B compartments, respectively, as illustrated in the schematic Hi-C heatmap. At shorter genomic length scales, high-resolution Hi-C revealed that chromatin folds into topologically associating domains (TADs) and subTADs which preferentially interact within themselves (indicated by darker shades on schematic Hi-C heatmaps) and arise from cohesin-mediated interactions between paired CTCF proteins. Within each TAD, dynamic cohesin-mediated loops contribute to chromatin folding. Promoter-enhancer (P-E) contacts occur via loop formation mainly within TADs (indicated by dark spots on schematic Hi-C heatmap) at varying genomic distances. Figure adapted from Boltsis et al., 2021.

or CTCF disrupts TAD structure and favors interactions within the active and repressive compartments (Nora et al., 2017; S. S. P. Rao et al., 2017; Schwarzer et al., 2017), highlighting the importance of these two factors in TAD and loop formation. TADs can be further subdivided into smaller subTADs and loops, all of which bring distant *cis*-regulatory elements, such as promoters and enhancers, in physical contact (Furlong & Levine, 2018; McCord et al., 2020; Rowley & Corces, 2018; Spielmann et al., 2018). Similar to TADs the formation of subTADs and enhancer-promoter loops is believed to be mainly facilitated by proteins such as CTCF, cohesin, and Mediator (Spielmann et al., 2018). However, in contrast to TADs, which are mostly conserved across cell types, most enhancer-promoter loops dynamically change during differentiation and are cell type-specific (Javierre et al., 2016).

Changes in the 3D genome structure are believed to play a key role in gene expression regulation by altering contact frequencies between promoters and enhancers. The majority of enhancer–promoter interactions occur within TADs (Spielmann et al., 2018). In this regard, TADs have been proposed to spatially insulate regulatory elements, thereby promoting intra-domain enhancer-promoter interactions, while limiting communication between regulatory elements across TADs to prevent abnormal gene activation (Robson et al., 2019; Schoenfelder et al., 2018). This idea is supported by several lines of evidence: For instance, enhancers located in one TAD preferentially interact with promoters within the same TAD, as opposed to those in neighboring TADs (Symmons et al., 2014). Furthermore, perturbations of TADs by deletion, inversion, or duplication of TAD boundaries in mice have been shown to result in ectopic promoter-enhancer interactions causing developmental defects (Franke et al., 2016; Guo et al., 2015; Lupiáñez et al., 2015), activation of oncogenes (Hnisz et al., 2016), or loss of developmental robustness (Despang et al., 2019). However, some studies do not see great effects on gene expression upon deletion of CTCF sites or TAD boundaries (Ghavi-Helm et al., 2019; Kane et al., 2022; Williamson et al., 2019) and upon global depletion of CTCF or cohesin only a fraction of genes exhibit changes in expression despite the loss of TAD structure (Nora et al., 2017; S. S. P. Rao et al., 2017; Schwarzer et al., 2017). A more detailed analysis employing promoter-capture Hi-C found that the relatively mild effects of CTCF or cohesin depletion on steady-state transcription can be attributed to the fact that many interactions between promoters and active enhancers are maintained, and only a small number of new promoter-enhancer contacts are formed in these conditions (Thiecke et al., 2020). In addition, cohesin-mediated loop extrusion was recently found to be only necessary for gene activation from long-range (>400kb) enhancers, but not to facilitate close (100 kb) enhancer-promoter contacts at the *Shh* locus (Kane et al., 2022). Thus, CTCF/cohesin-mediated TADs and loops appear to

be particularly important for long-range gene regulation of a subset of genes and compaction of large regulatory domains, while cohesin-independent promoter-enhancer loops must be regulated by alternative mechanisms (Furlong & Levine, 2018; Kane et al., 2022; McCord et al., 2020; Schoenfelder & Fraser, 2019; Thiecke et al., 2020).

In addition, it should be noted that TAD boundary insulation is not absolute. Interactions within a TAD may be favored by loop extrusion (Dekker & Mirny, 2016), but TADs and CTCF loops dynamically fold and unfold (Gabriele et al., 2022; Mach et al., 2022) and inter-TAD contacts may occur, depending on boundary strength and chromatin composition (Bonev et al., 2017; Paulsen et al., 2019; Szabó & Mayor, 2018). In human hematopoietic cells about one-third of significant interactions between promoters and putative enhancers have been shown to cross a TAD boundary, which however is less than expected to occur at random (Javierre et al., 2016). Moreover, expression quantitative trait loci (eQTLs) have been found to interact with genes across TADs (Javierre et al., 2016). In fact, a recent study proposed that TAD boundaries and CTCF interactions do not mediate absolute transcriptional insulation, instead, the level of transcriptional insulation varies based on factors such as the strength of enhancers, the strength of the boundary, and the linear genomic distance of an enhancer to the promoter (Zuin et al., 2022).

1.1.3 Enhancers

Enhancers are defined as 100 - 2,000 bp long non-coding DNA sequences with the ability to control cell-type specific spatiotemporal gene expression programs largely independent of their relative distance, location, or orientation to their cognate promoter (Furlong & Levine, 2018; Kvon et al., 2021; Long et al., 2016; Schoenfelder & Fraser, 2019). Additionally, they are believed to play a significant role in causing phenotypic variation, driving evolutionary changes, and contributing to the development of disease (Claringbould & Zaugg, 2021; Pachano et al., 2022; Zaugg et al., 2022). They are known to activate their target gene through physical contact via loop formation, mediated by direct contact between factors bound to the regulatory sequences, as has been shown by studies engineering artificial loops between a gene and distal enhancer sequences (Bartman et al., 2016; Chen et al., 2018; Deng et al., 2012; Fukaya et al., 2016; Kim et al., 2019). Moreover, enhancers are characterized by concentrated clusters of TF binding motifs, which mediate the cooperative binding of multiple TFs in close proximity. These, in turn, recruit co-factors, such as p300, Mediator, and BRG1 to activate or increase transcription of the target gene (Kvon et al., 2021; Long et al., 2016; Schoenfelder & Fraser, 2019).

The functional status of an enhancer can be classified through their DNA accessibility

status, determined by techniques such as ATAC-seq or DNAse-seq, and post-translational histone modifications on nucleosomes in the vicinity of the enhancer via ChIP-seq: active (H3K4me1/H3K4me2 and H3K27ac), poised (marked by H3K4me1/H3K4me2 and H3K27me3), and primed/neutral (marked by H3K4me1/H3K4me2 only) (Creyghton et al., 2010; Rada-Iglesias et al., 2011; Shlyueva et al., 2014). Hence, it has become common practice to use histone marks, DNA accessibility, and co-factor binding for genome-wide predictions of putative enhancers in various cell types (Shlyueva et al., 2014). To prevent the wrongful assignment of an enhancer to its target gene, it is recommended to validate their interaction using advanced microscopy techniques or 3C-type methods (Schoenfelder & Fraser, 2019).

Predictions of putative enhancers have revealed that enhancers by far outnumber genes in mammalian genomes (Anderson et al., 2014; Gorkin et al., 2020; Kundaje et al., 2015; Shen et al., 2012). As such, it does not come as a surprise that in the majority of cases multiple enhancers with overlapping or partially overlapping spatiotemporal activities cooperate to regulate their target gene (Andrey & Mundlos, 2017; Kvon et al., 2021). This is a fundamental feature in gene regulation, as it adds multiple layers of regulatory complexity. This allows for a wide range of gene expression patterns and in turn a broad range of cellular states using a relatively small set of genes (Long et al., 2016). For example, the combined actions of at least 11 enhancers control the expression of the *Sonic hedgehog* (*Shh*) gene in various tissues, such as the central nervous system, epithelial linings, and limbs, during the development of mouse embryos (Anderson et al., 2014). In rare cases, the loss of a single enhancer can lead to the loss of gene expression selectively in the tissue, where the enhancer is active, causing severe phenotypes (Fig. 3a) (Anderson et al., 2014; Kragesteen et al., 2018; Lettice et al., 2003; Sagai et al., 2005; Yanagisawa et al., 2003). More often than not, deletion of a single enhancer results in no or fairly mild phenotypes due to minimal effects on target gene expression (Attanasio et al., 2013; Long et al., 2020; Osterwalder et al., 2018). On the other hand, double enhancer deletions often result in severe phenotypes, mimicking complete gene loss in the related tissues (Antosova et al., 2016; Dickel et al., 2018; Hay et al., 2016; Osterwalder et al., 2018). These findings suggest that there is a redundancy between both enhancers, as under normal conditions one of them is sufficient to drive gene expression. However, in a sensitized genetic background, the deletion of a single enhancer can lead to abnormal phenotypes, suggesting that enhancer redundancy can provide robustness against genetic disruptions (Long et al., 2020; Osterwalder et al., 2018). This suggests that although enhancers can be redundant in terms of their activity pattern, they are not necessarily functionally identical.
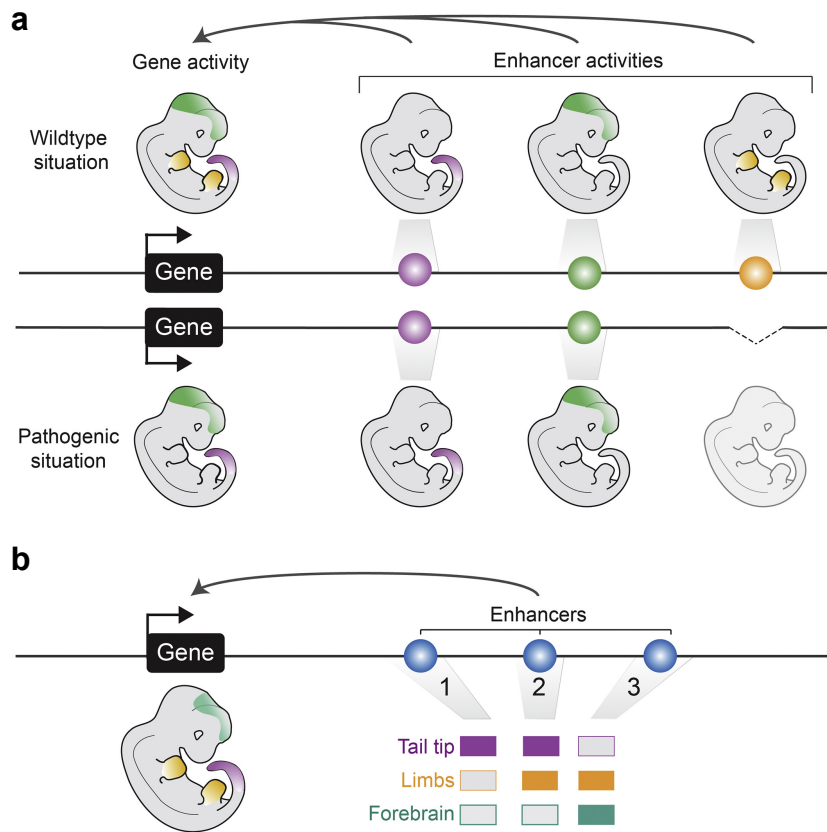
Figure 3. Enhancer tissue-specificity and redundancy. **a** One gene's activity can be regulated by different enhancers in different tissues. The loss of one enhancer can result in tissue-specific reduction or loss of gene expression. **b** Multiple 'shadow enhancers' with overlapping or partially overlapping spatiotemporal activities cooperate to regulate their target gene. This enhancer redundancy provides robustness against genetic and environmental perturbations during development. Figure taken from Robson et al., 2019.

Enhancer Redundancy

In recent years the term 'shadow enhancer' has increasingly been used to describe such sets of redundant enhancers targeting the same gene with overlapping or partially overlapping activities (Fig. 3b)(Kvon et al., 2021). They have been found to be associated especially with developmentally regulated genes and act as a safeguard for crucial developmental processes, mitigating the risks of severe outcomes caused by genetic or environmental factors (Antosova et al., 2016; Frankel et al., 2010; Gorkin et al., 2020; Osterwalder et al., 2018). The exact mechanisms by which shadow enhancers confer resistance against genetic perturbations are not fully understood and are the subject of ongoing research. One proposition is that each enhancer can independently drive gene expression to levels adequate for normal development, meaning multiple shadow enhancers increase the probability that the target gene is properly activated (Kvon et al., 2021). However, a study in *D. melanogaster* points towards another potential mechanism, where each enhancer responds to a different set of TFs (Wunderlich et al., 2015). The independent regulation of shadow enhancers via different combinations of TFs could be a widespread mechanism to confer robustness (Cannavò et al., 2016; Kvon et al., 2021).

Various modes of interaction between shadow enhancers have been proposed to be re-

sponsible for fine-tuning their target gene's expression pattern. The classical mode of enhancer interaction is described as them being arranged in a highly modular and functionally autonomous fashion, where each enhancer within the regulatory landscape of a gene individually adds to the total expression pattern of a gene (Kvon et al., 2021; Long et al., 2016; Schoenfelder & Fraser, 2019; Visel et al., 2009). For instance, *Indian hedgehog* (*Ihh*) is regulated by a multi-component enhancer ensemble that comprises at least nine enhancers with specific tissue combinations that act in an additive way (Will et al., 2017). Furthermore, the -globin gene locus, which is regulated by a cluster of enhancer elements known as a super-enhancer, has shown that each enhancer acts independently and additively without any clear evidence of synergistic or higher-order effects (Hay et al., 2016). In contrast, in *Drosophila* embryos it has been shown that in addition to the additive mode of interaction, shadow enhancers can also behave in a synergistic manner, significantly increasing target gene expression beyond what would be expected from their individual activities, or sub-additively, resulting in lower target gene expression than expect (Bothma et al., 2015; Scholes et al., 2019). Sub-additivity between enhancers suggests that there is competition among shadow enhancers for promoter occupancy (Kvon et al., 2021; Long et al., 2016). Furthermore, a very recent study discovered that four individually weak enhancers, of which three form a super-enhancer, collaborate in a super-additive fashion to regulate *Fgf5* expression in mESCs upon exit from naive pluripotency (Thomas et al., 2021). In other cases, shadow enhancers have been shown to repress each other's activity fully or in part, resulting in a reduction or loss of target gene expression (Bothma et al., 2015; Dunipace et al., 2011; Perry et al., 2011). Taken together, these findings indicate that gene regulatory landscapes composed of multiple enhancers are regulated in an extremely complex manner with many case-to-case variations.

Enhancer-Promoter cross-talk

The functional importance of enhancers has been recognized more than 30 years ago. Nevertheless, how enhancers target genes over extremely large distances, sometimes up to several megabases, within the nucleus and the mechanisms that prevent enhancers from activating other nearby genes remain largely unknown. Major technological advances within the last couple of years have lead to a better understanding of enhancer function and spatial genome organization, including genome-wide mapping of enhancer-promoter contacts at high resolution in various cell types (Freire-Pritchett et al., 2017; Javierre et al., 2016; Mifsud et al., 2015; Novo et al., 2018; Rubin et al., 2017; Schoenfelder, Furlan-Magaril, et al., 2015; Siersbaek et al., 2017; Wilson et al., 2016), large-scale transgenic reporter assays (Kvon, 2015; Pennacchio et al., 2006; Visel et al.,

2007), genome and epigenome editing techniques (Diao et al., 2017; Fulco et al., 2016; Klann et al., 2017; Sanjana et al., 2016; Simeonov et al., 2017), engineering of enhancer-promoter interactions (Bartman et al., 2016; Deng et al., 2012; Deng et al., 2014) and advanced microscopy techniques (Bintu et al., 2018; Boettiger et al., 2016; Chen et al., 2018; Gu et al., 2018). These techniques not only enabled researchers to accurately assign enhancer-promoter pairs but also brought forward various theories to explain how enhancers and promoters interact, including tracking, linking, looping, phase-separated condensates, and combinations of these mechanisms, as reviewed in various sources (Furlong & Levine, 2018; Hnisz et al., 2017; Robson et al., 2019; Schoenfelder & Fraser, 2019; Zabidi & Stark, 2016). Out of these different models, cohesin-driven loop formation has become the most supported one within the field (Furlong & Levine, 2018). In support of this, recent studies suggest that long-range regulation by distal enhancers might rely on transient interactions within the constraints of a TAD (Gabriele et al., 2022; Mach et al., 2022). Another set of recent studies has further demonstrated that enhancer function is not independent of the linear distance to their target gene promoters, instead, gene expression levels and distances between a promoter and enhancer are anti-correlated in a non-linear fashion (Rinzema et al., 2022; Zuin et al., 2022). For instance, a developmental enhancer is only able to activate the tissue-specific gene *Gria1*, which has a CpG-poor promoter, if it is placed in close proximity to the gene (Pachano et al., 2022). In contrast, cohesin-mediated loop extrusion appears to be able to negate the effects of linear distance between enhancers and promoters (Kane et al., 2022).

Interestingly, it has been reported that tissue-specific genes with CpG-poor promoters are regularly located in close proximity to their *cis*-regulatory elements, whereas developmental genes, most of which have CpG-rich promoters in vertebrates, are often regulated by distal enhancers over quite long distances of up to over 1 Mb (Bahr et al., 2018; Benko et al., 2009; Herranz et al., 2014; Lettice et al., 2003; Long et al., 2020). Thus, the ability of distal enhancers to act on their target genes despite their linear distance in the genome may be due to specific additional mechanisms (Batut et al., 2022; Kane et al., 2022; Rinzema et al., 2022). In *Drosophila* short DNA sequences located in enhancers and/or promoters with the ability to facilitate interactions between both have been defined as tethering elements (Pachano et al., 2022). These tethering elements mediate long-range enhancer-promoter contacts and thereby ensure the specific timing of developmental gene transcription during the early stages of *Drosophila* embryogenesis (Batut et al., 2022). Similar elements with the ability to facilitate long-range enhancer-promoter interactions have been described in mammals (Pachano et al., 2022). For example, CTCF has recently been shown to be responsible for the tethering of certain promoters to their distal enhancers, but it remains unclear if CTCF has

to be bound to both of the regulatory elements or just one of them (Kubo et al., 2021; Oh et al., 2021). Apart from CTCF binding sites, CpG islands (CGIs) have recently been identified as potent tethering elements that aid in both physical and functional communication between developmental genes and enhancers that are rich in CpG (poised enhancers)(Pachano et al., 2021). The interactions between poised enhancers and their target genes in pluripotent cells are likely facilitated by PcG complexes, which are brought to CGIs found at both enhancers and target gene promoters (Crispatzu et al., 2021; Cruz-Molina et al., 2017; Pachano et al., 2022). Lastly, specific TFs such as YY1 may be able to act as tethering elements and promote long-range enhancer-promoter contacts (Pachano et al., 2022). It is likely, that there are further tethering elements that still need to be identified. Moreover, the different types of tethering elements may collaborate in order to increase specificity and/or gene expression levels (Pachano et al., 2022).

High-resolution Promoter-Capture Hi-C technology at varying developmental time points and in different cell types has revealed that certain enhancer-promoter contacts are pre-formed while others are gained *de novo* in an activity-dependent manner (Andrey et al., 2017; Robson et al., 2019; Schoenfelder & Fraser, 2019). Pre-formed interactions in ESCs are mediated by PRC1 and PRC2 complexes, which keep the enhancer-promoter pair in a poised state, allowing enhancers to readily activate their target genes during differentiation (Cruz-Molina et al., 2017). In other cases, pre-established looping between enhancers and promoters appears to be mediated by CTCF and cohesin, as is the case for the contact between the *Shh* promoter and the ZRS enhancer (Paliou et al., 2019). Consequently, spatial proximity between enhancers and promoters does not necessarily lead to induction of gene expression, instead proteins like CTCF and cohesin, or PRC1 and PRC2, keep inactive enhancers and promoters in close proximity to "prime" genes for activation (Robson et al., 2019; Schoenfelder & Fraser, 2019). While some contacts are pre-formed during development, cell fate changes often involve dynamic changes in enhancer-promoter interactions linked to active transcription and chromatin modifications (Andrey et al., 2017; Bonev et al., 2017; Javierre et al., 2016). For instance, via a topological switch the *Pen* enhancer, active in both embryonic hind- and forelimbs, specifically activates *Pitx1* only in the hindlimb (Kragesteen et al., 2018). In conclusion, these data indicate that both pre-formed and facultative contacts are present in regulatory landscapes to accommodate the diverse and specific demands of gene functions (Robson et al., 2019).

3C experiments demonstrated that shadow enhancers are regularly in physical contact with each other and their target gene (Allahyar et al., 2018; Beagrie et al., 2017; Hay et al., 2016; Hughes et al., 2014; Jiang et al., 2016; Madsen et al., 2020; Proudhon et al., 2016), providing

evidence for simultaneous promoter activation. On the other hand, a study in *D. melanogaster* using quantitative live imaging demonstrated that a single enhancer can simultaneously activate two proximal promoters, resulting in synchronized transcriptional bursts (Fukaya et al., 2016). Collectively, the studies indicate that multiple shadow enhancers can work together to regulate the expression of a single target promoter and that enhancer-promoter interactions can involve multiple DNA elements (Kvon et al., 2021).

In contrast to the enhancer-promoter looping model stands the concept of dynamic "transcriptional hubs" or phase-separated "transcriptional condensates", which could potentially explain the regulation of one promoter by multiple enhancers (J. M. Alexander et al., 2019; Benabdallah et al., 2019; Cho et al., 2018; Hnisz et al., 2017; Montavon et al., 2011; Sabari et al., 2018; Tsai et al., 2019; Zabidi & Stark, 2016). Very high concentrations of TFs, co-activators such as Mediator and BRD4, and PolII have been found to be involved in the formation of such hubs (Cho et al., 2018; Hnisz et al., 2017; Sabari et al., 2018). The hub model may also explain why an enhancer does not always need to be physically connected to its target promoter to activate it (J. M. Alexander et al., 2019; Benabdallah et al., 2019). Furthermore, transcriptional co-activators have been shown to form phase-separated condensates at super-enhancer associated genes (Hnisz et al., 2017; Sabari et al., 2018), which supports the idea that multiple shadow enhancers and their target promoter can exist together in the same microenvironment, creating a multi-enhancer hub (Kvon et al., 2021). In conclusion, not one single model applies to all enhancer-promoter interactions, instead, each of the different mechanisms contributes to varying degrees to the cross-talk between enhancers and promoters.

Super-Enhancers

Over the years, genomic regions that spread out over large distances containing multiple enhancers in close proximity have been recognized by several studies as substantially different from a single/typical enhancer and were thus given their own terminology. The term most used today is "super-enhancer" and was coined in 2013 by the Young lab (Hnisz et al., 2013; Loven et al., 2013; Whyte et al., 2013). They defined a super-enhancer (SE) as a cluster of enhancers that have an exceptionally strong enrichment of the active chromatin mark H3K27ac, much denser binding of transcriptional co-activators like Mediator, BRD4 and cell type–specific master TFs, and a stronger ability to activate transcription in comparison to typical enhancers (Hnisz et al., 2013; Loven et al., 2013; Whyte et al., 2013). Moreover, they created an algorithm, named rank ordering of super-enhancers (ROSE) to identify SEs on a genome-wide scale using ChIP-seq data from active

chromatin marks like H3K27ac, co-activators, or binding of cell type–specific TFs (Hnisz et al., 2013; Loven et al., 2013). In short, the location of enhancers in a given cell type is predicted using e.g. H3K27ac or Mediator complex ChIP-seq data. Single putative enhancers with a distance of less than 12.5 kb from each other are grouped together, generating larger genomic fragments. All enhancers and grouped fragments are ranked according to their H3K27ac enrichment, a cut-off is mathematically defined and all regions above this point are defined as super-enhancers. Called SEs that are located within a ±2000 bp window flanking an annotated TSS are excluded (Hnisz et al., 2013; Loven et al., 2013; Whyte et al., 2013). This same algorithm has recently been used to predict SEs in early human craniofacial development and found that developmental genes are enriched near these SEs (Wilderman et al., 2018). It should be kept in mind that the SE prediction algorithm used in these studies does not consider any information about the physical arrangement of chromatin, and instead assigns SEs to their target genes based solely on their location on the linear genome, which can result in incorrect enhancer-promoter pair assignments. In addition, the 12.5 kb window deciding if two or more enhancers are grouped, is set arbitrarily. Thus, it would be advisable to validate and functionally assess SE predictions with high-resolution enhancer-promoter maps via 3C-based methods or advanced microscopy techniques and genome editing, as has been done in various recent studies (Hay et al., 2016; Hnisz et al., 2016; Huang et al., 2018; Jiang et al., 2016; Proudhon et al., 2016; Sabari et al., 2018; Thomas et al., 2021).

The main differences between a SE and a typical enhancer have been shown to lie in their size, their increased capability to activate transcription, the density and composition of enhancer-associated factors, and their high enrichment in disease-associated genetic variation (Hnisz et al., 2013; Whyte et al., 2013). Initially, it was described in mESCs that SEs commonly occur in proximity to important pluripotency genes, including the ESC master regulators Oct4, Sox2, and Nanog which also bind to the SEs. The binding levels of these three factors were similar at typical enhancers and SEs (Whyte et al., 2013). In contrast, two other TFs associated with stem cell identity, namely Klf4 and Esrrb, had a significantly stronger presence at SEs compared to typical enhancers (Whyte et al., 2013). Thus, the genes that code for master TFs in ESCs appear to be controlled by SEs, which in turn are regulated by the binding of these TFs, resulting in an auto-regulatory feedback loop where master TFs regulate their own expression (Whyte et al., 2013).

SEs are prevalent not only in stem cells but also in many other cell types, typically located close to genes that play crucial roles in determining cell identity (Hnisz et al., 2013; Siersbæk et al., 2014; Whyte et al., 2013). In support of the claim that SEs are cell type-specific, it has been shown that most SEs only occur in very few or single cell types (Wang et al., 2019). Furthermore,

multiple recent studies observed that the single enhancer elements forming a SE, called constituent enhancers, are physically contacting each other and their target promoters to jointly control target gene expression (Beagrie et al., 2017; Dowen et al., 2014; Hnisz et al., 2016; Huang et al., 2018; Ji et al., 2016; Jiang et al., 2016; Madsen et al., 2020; Proudhon et al., 2016). Interestingly, not all constituent enhancers are functionally equivalent in a given cell type and some constituent enhancers show significantly higher TF binding density than others, referred to as "epicenters" (Adam et al., 2015), "hotspots" (Siersbæk et al., 2014), or "enhancer hubs" (Huang et al., 2018). For example, Hnisz et al. (2015) generated deletions of single constituent enhancers, which in most of the cases resulted in decreased expression levels of the SE-associated gene, but in at least one case the expression of the associated gene increased. Moreover, they found that some SE constituents have a higher regulatory potential than others, as their loss has a particularly strong effect on the target gene transcription and additionally affects the chromatin state of the entire SE. Similar observations were made by other studies upon deletion of single constituent enhancer (Huang et al., 2018; Jiang et al., 2016; Proudhon et al., 2016; Shin et al., 2016; Siersbæk et al., 2014). In addition, the loss of one constituent enhancer has also been shown to reduce the physical contacts between the other constituents of a SE leading to reduced SE activity (Jiang et al., 2016; Proudhon et al., 2016). Another mode of function for SEs has been presented by Thomas et al. (2021). They found that three intergenic constituent enhancers contribute to *Fibroblast growth factor 5* (*Fgf5*) induction at distinct time points during exit from naive pluripotency, while a third poised intragenic enhancer amplifies the expression of *Fgf5* in a super-additive manner at each time point (Thomas et al., 2021). Taken together, these findings suggest that there is a highly complex functional and structural interdependence between individual enhancer elements of some SEs, required to control the expression of their target genes. However, there are also instances where constituent enhancers appear to act independently in an entirely additive fashion (Hay et al., 2016) or are partially redundant without synergistic effects (Moorthy et al., 2017). Nevertheless, this does not contradict the concept that many collaborating enhancers confer complexity and robustness to gene regulatory landscapes mitigating the effects of the loss of single constituent enhancers on gene expression.

Interactions between SEs and their target genes have been found to mainly occur in so-called insulated neighborhoods formed by the looping of two convergent CTCF sites co-occupied by cohesin (Dowen et al., 2014; Ji et al., 2016). Considering the extremely high activation potential of SEs, it may be necessary to keep SEs and their target genes insulated to prevent SEs from aberrantly activating other genes. Indeed, deletion of a CTCF binding site at either border of a SE

domain resulted in a substantial upregulation of genes immediately outside the normally insulated domain (Dowen et al., 2014; Hnisz et al., 2016; Ji et al., 2016). In contrast to the looping model, it has been proposed that active SEs form phase-separated condensates mediated by the high density of co-activators, including Mediator and BRD4, found at SEs (Hnisz et al., 2017). In support of this model, two recent studies found that BRD4 and Mediator as well as Mediator and PolII form condensates that co-localized with SEs (Cho et al., 2018; Sabari et al., 2018).

Tumorigenesis and developmental defects have been increasingly implicated with mis-regulated SEs (Wang et al., 2019). For instance, single nucleotide polymorphisms (SNPs) that are linked to diseases are more commonly found in SEs than in typical enhancers and cancer cells can gain SEs driving oncogene expression (Chapuy et al., 2013; Hnisz et al., 2013; Loven et al., 2013). One of the mechanisms that can lead to the activation of oncogenes by SEs is the disruption of insulated neighborhoods containing SEs (Hnisz et al., 2016), and another is the establishment of novel SEs through somatic mutation (Mansour et al., 2014).

1.2   The Neural Crest and Craniofacial Morphogenesis

The neural crest (NC) was first described in 1868 by Wilhelm His (His, 1868) as a group of cells located between the neural tube and the non-neural ectoderm giving rise to spinal and cranial ganglia in chicken embryos. Since then major advances in the research of this vertebrate synapo-morphic cell population have been made, from grafting and early cell lineage tracing experiments (e.g. N. M. Le Douarin & Teillet, 1973; Noden, 1983; P. Trainor & Krumlauf, 2000) to system-level and genomic approaches (e.g. Gammill & Bronner-Fraser, 2002; Lignell et al., 2017; Minoux et al., 2009; Schorle et al., 1996), and most recently next generation sequencing techniques (e.g. Lukoseviciute et al., 2018; Minoux et al., 2017; Soldatov et al., 2019; Zalc et al., 2021), revealing the stem cell-like properties, extensive migratory abilities and differentiation capabilities of NC cells (NCCs) (Bronner & LeDouarin, 2012; Martik & Bronner, 2017, 2021; Murillo-Rincón & Kaucka, 2020; Sauka-Spengler & Bronner-Fraser, 2008; Simões-Costa & Bronner, 2015; P. A. Trainor & Krumlauf, 2000). Being a key defining trait in the evolution of vertebrates, NCCs have been of particular interest in evolutionary and clinical research (Green et al., 2015; Sauka-Spengler et al., 2007). Alterations in the development of these cells are associated with a high proportion of birth defects and a variety of syndromes termed neurocristopathies, of which many include cranio-facial abnormalities (Bolande, 1974; Crane & Trainor, 2006; Siismets & Hatch, 2020; Takahashi et al., 2013). Furthermore, there is a rapidly growing understanding of the cellular programs and regulatory mechanisms that play a role in the biology of NCCs. As a result, they are also a highly
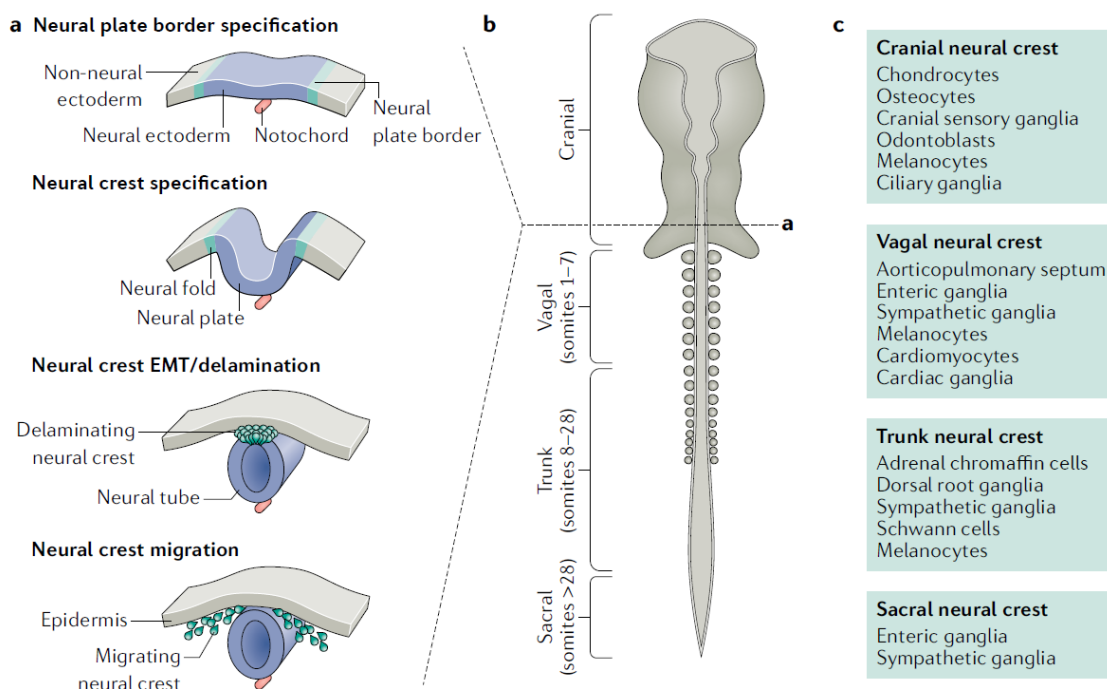
Figure 4. Development of the neural crest. a Step-wise formation of the neural crest: 1. formation of the neural plate border, 2. neural crest specification, 3. delamination of neural crest cells from the dorsal part of the forming neural tube, 4. neural crest cell migration to distant locations within the embryo. **b** The neural crest is divided into four segments along the anteroposterior body axis: cranial, vagal, trunk, and sacral neural crest. The dashed line indicates the location of the schematic sections depicted in **a**. **c** Migrated neural crest cells differentiate into unique cell types, depending on their origin and final axial positioning within the embryo. Figure taken from Martik and Bronner, 2021.

relevant system for studying gene regulatory interactions and networks (Parker et al., 2018).

Many of the molecular mechanisms driving NC development have been uncovered in recent years, leading to the establishment of a gene regulatory network (GRN) that governs each step of the developmental process (Martik & Bronner, 2017; Sauka-Spengler & Bronner-Fraser, 2008; Simões-Costa & Bronner, 2015). In short, GRNs are made up of modules of regulatory information that contain the genes and interactions that function at different stages of development. The regulatory interactions present in a cell at any given time represent its regulatory state. GRNs further describe the regulatory interactions that occur between transcription factors and cis-regulatory elements at different stages of development in a particular cell type (E. H. Davidson, 2010). The formation of neural crest cells is controlled by a feedforward GRN that is based on data from numerous organisms and cell types, including basal vertebrates, mice, and human pluripotent stem cells (Martik & Bronner, 2017; Sauka-Spengler & Bronner-Fraser, 2008; Simões-Costa & Bronner, 2015). Despite some species-specific differences, the overall architecture and composition of the GRN controlling NC development is highly similar across all vertebrates (Sauka-Spengler et al., 2007). In recent years multiple studies have shown that next to regulation through transcrip-

tion factors, epigenetic mechanisms are also crucial for proper NC development, particularly in controlling the timing of gene expression at different stages of development (N. Hu et al., 2014; Minoux et al., 2017). Knowledge about the conservation and changes in gene regulatory logic across species is central to our understanding of how the neural crest evolves, how morphological traits such as variable craniofacial features emerge, and how NCCs are able to differentiate in such a wide range of cell types.

For the formation of the neural crest, a series of regulatory events have to take place early in development: induction, specification, delamination, migration, and differentiation (Fig.4a). The formation of the neural plate border and NC is initiated during gastrulation by environmental signals and a cascade of transcription factors. WNT, BMP, and FGF pathways combinatorially refine the border between the forming neural and non-neural ectoderm, driving the expression of neural plate border specifier genes (Garcıa-Castro et al., 2002; Groves & LaBonne, 2014; Khudyakov & Bronner-Fraser, 2009; Martik & Bronner, 2021; Monsoro-Burq et al., 2005; Stuhlmiller & García-Castro, 2012b).

After the formation of the neural plate border, the expression of NC specifier genes, including *Foxd3, Soxe, Snai1/2* and *Tfap2* (Martik & Bronner, 2017, 2021; Simões-Costa & Bronner, 2015), is induced by the concerted action of neural plate border specifiers and signaling pathways (e.g. *Msx1, Zic1, Tfap2* and *Pax3/7*) (Basch et al., 2006; Garnett et al., 2012; Monsoro-Burq et al., 2005; Nikitina et al., 2008; Plouhinec et al., 2014; Sato et al., 2005; Stuhlmiller & García-Castro, 2012a). Once specified, the NCCs undergo an epithelial-to-mesenchymal transition (EMT) and separate from the dorsal part of the neural tube to migrate extensively throughout the embryo (Szabó & Mayor, 2018; Williams et al., 2019). Upon reaching their final location within the embryo NCCs transition from moving streams of cells to aggregates within complex structures. This process often involves a reverse of the EMT, called the mesenchymal-to-epithelial transition (MET). Based on their anteroposterior (AP) location along the body axis the NCCs will differentiate into distinct derivatives (Martik & Bronner, 2017, 2021; Simões-Costa & Bronner, 2015). While some derivatives are common between the cranial, vagal, trunk, and lumbosacral subpopulations of NCCs, certain NC-derived structures are unique to a particular subpopulation (Fig. 4c). As such the cranial NCCs (CNCCs) have the unique ability to produce chondrogenic and skeletal tissue (N. Le Douarin & Kalcheim, 1999; Santagati & Rijli, 2003).
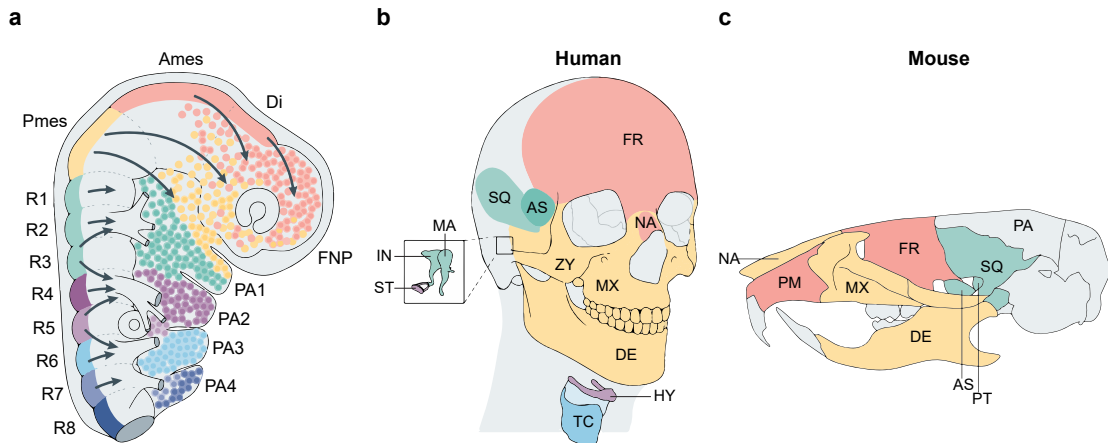
Figure 5. Skeletal fate of cranial neural crest cells in human and mouse. a During development, the cranial neural crest cells (CNCC) migrate (arrows) in discrete streams from the diencephalon (Di), anterior and posterior mesencephalon (Ames and Pmes), and rhombomeres (R1-R8) to colonize the frontonasal process (FNP) and anterior pharyngeal arches (PA1-PA4), as indicated by the color code. **b** and **c** Depicted are craniofacial skeletal structures in human (**b**) and mouse (**c**) that are derived from distinct CNCC subpopulations. Structures are color-coded to match the CNCC subpopulations in panel **a** to indicate their origin. Panel **a** and **b** adapted from Martik and Bronner, 2021. Panel **c** adapted from Santagati and Rijli, 2003. AS, alisphenoid bone; DE, dentary bone; FR, frontal bone; HY, hyoid bone; IN, incus; MA, malleus; MX, maxillary bone; NA, nasal bone; PA, parietal bone; PM, premaxillary bone; PT, pterygoid bone; SQ, squamosal bone; ST, stapes; ZY, zygomatic bone.

1.2.1 The Cranial Neural Crest

The development of the vertebrate head is one of the most complex processes in embryogenesis, orchestrated by the coordinated interplay between a series of molecular and morphogenetic events and extrinsic position-specific environmental signals (Minoux et al., 2017; Minoux & Rijli, 2010; Santagati & Rijli, 2003). Central to this process are the CNCCs, which arise from the developing diencephalon, mesencephalon, and hindbrain region and are the primary source of skeletogenic head mesenchyme (Fig. 5). Accordingly, the frontonasal skeleton is formed by rostral CNCCs, which also significantly contribute to the membranous bones of the skull. More posterior CNCCs of the hindbrain region, which is transiently subdivided into neuroepithelial segments called rhombomeres (R), colonize the pharyngeal arches (PA), paired segmental bulges that form on the lateral surface of the head during the early stages of embryonic development (Graham & Richardson, 2012). From there they give rise to the cartilages and bones of the upper and lower jaws, middle ear ossicles, outer ear, hyoid, and thyroid structures (Fig. 5b and c)(Gammill & Bronner-Fraser, 2003; N. M. Le Douarin et al., 2004; Minoux & Rijli, 2010; Parker et al., 2018; Santagati & Rijli, 2003). Additionally, CNCCs contribute to forming peripheral neurons and glia and the connective tissue associated with mesoderm-derived head muscles (P. A. Trainor & Krumlauf, 2000). The NCC subpopulations targeting the individual PAs migrate in discrete streams (Fig. 5a), which is

of importance for proper craniofacial patterning (Minoux & Rijli, 2010). As such, specific rhombomeres have different contributions to these streams. For instance, NCCs that migrate into the first pharyngeal arch (PA1), which is subdivided into the mandibular (Md) and maxillary process (Mx), are mainly derived from the posterior mesencephalon, rhombomere 1 (R1), and R2, while those that migrate into the second (PA2) and third pharyngeal arch (PA3) mostly come from R4 and R6, respectively (Birgbauer et al., 1995; Kulesa & Fraser, 2002; Lumsden et al., 1991). These CNCC subpopulations express different Hox genes along the AP axis, which is essential for their AP patterning. As such, CNCCs that contribute to the frontonasal process (FNP) and the first pharyngeal arch (PA1) do not express any *Hox* genes, while CNCCs of the second and more posterior arches express different combinations of *Hox* genes, starting with *Hoxa2* as the single *Hox* gene expressed in PA2 (T. Alexander et al., 2009; Minoux et al., 2009; Minoux & Rijli, 2010; P. Trainor & Krumlauf, 2000). An important factor defining the competence of CNCCs to form facial skeletal elements is related to the presence or absence of *Hox* gene expression.

1.2.2 Cranial neural crest cell plasticity

In the past there has been a long-standing debate about the CNCC plasticity versus pre-patterning or in other words the question of whether CNCCs are fate committed before migration, thereby being pre-patterned, or retain multipotency throughout migration able to respond to environmental cues at their final destination (N. M. Le Douarin et al., 2004; Santagati & Rijli, 2003; P. A. Trainor & Krumlauf, 2000, 2001). This issue arose from the fact that the posterior and anterior cephalic NC domains differ with regards to their expression of *Hox* genes, which has an impact on their competence. The pharyngeal *Hox* code and its significance in craniofacial morphogenesis will be discussed in detail in section 1.2.3. In short, the more posterior CNCCs colonizing PA2 to PA4 express a combination of *Hox* genes, while the anterior CNCCs populating the FNP and PA1 do not express any *Hox* genes (Couly et al., 1996; Hunt, Gulisano, et al., 1991; Prince & Lumsden, 1994). *Hox*-negative and *Hox*-positive CNCCs differ in their ability to generate pharyngeal arch skeletal derivatives and *Hox* expression has been demonstrated to prevent the formation of skeletal derivatives of the frontonasal process and first pharyngeal arch (Couly et al., 1998; Creuzet et al., 2002). In fact, only *Hox*-negative CNCCs can produce the membrane bones of the facial skeleton (Couly et al., 2002), whereas both *Hox*-positive and *Hox*-negative CNCCs can produce cartilage derivates.

In the 1980s, the concept of a pre-patterned cranial NC was prevalent, mainly due to the results of a NC grafting experiment in the avian embryo by Drew Noden in 1983 (N. M.

Introduction

Le Douarin et al., 2004; Santagati & Rijli, 2003; P. A. Trainor & Krumlauf, 2000, 2001). In this study, the transplantation of presumptive first arch NC primordia, which is *Hox*-negative, to a more posterior position in the neural tube (roughly R4-R5), from where normally the CNCCs of the second arch originate, led to a partial duplication of lower jaw skeletal elements (Noden, 1983). In the following years, this concept was challenged by conflicting results from different studies (N. M. Le Douarin et al., 2004; Santagati & Rijli, 2003; P. A. Trainor & Krumlauf, 2000, 2001). For instance, contrary to Noden Couly et al. (1998) did not observe a duplication of first arch elements upon transposition of the *Hox*-negative NC domain to the more posterior domain of the presumptive second arch NC in the chick. They further found that *Hox*-positive neural fold tissue transplanted into the anterior *Hox*-negative domain is not able to produce any skeletal tissue (Couly et al., 1998; Couly et al., 2002). These results pointed towards a lack of pre-patterning information in *Hox*-negative CNCCs progenitors, giving them a high degree of plasticity and the ability to readily respond to local patterning cues, while expression of *Hox* genes appeared to provide an AP positional identity to CNCCs pre-migration. This was supported by the finding that the entire *Hox*-negative NC domain acts as a group of equivalence when it comes to building the facial skeleton, as implantation of just small fragments of the anterior NC, following the removal of the entire domain, leads to the formation of a complete face (Couly et al., 2002). Additionally, a study in the mouse showed that when grafting cells from R3, R4 or R5 into R2 the majority of the grafted cells maintained their *Hox* gene AP identity (P. Trainor & Krumlauf, 2000). However, some single cells which separated from the primary transplanted tissue and mixed with the local cell population, failed to maintain their original *Hox* gene expression patterns, thus losing their identity (P. Trainor & Krumlauf, 2000). Similarly, in zebrafish moving individual hindbrain cells from R2 to R6 or vice versa resulted in a total reversal of Hox gene expression (Schilling et al., 2001). The movement of these cells to a new location was accompanied by changes in their fate, which took on characteristics of their new environment. The extent to which this plastic transformation occurred depended on the timing and size of the transplant (Schilling et al., 2001). Moreover, in both mouse and zebrafish, the transplanted cells contributed to pharyngeal derivatives appropriate to their new AP position (Schilling et al., 2001; P. Trainor & Krumlauf, 2000).

Coming back to the findings of Noden (1983), a later study demonstrated that the duplications of first arch derivatives observed by Noden were related to the inclusion of the isthmus in the transplantations (P. A. Trainor et al., 2002). The isthmus expresses *Fgf8* which silences *Hoxa2* (Irving & Mason, 2000), so repression of *Hoxa2* in the presumptive second arch NCC was the likely cause for the development of ectopic first arch skeletal structures. In summary, all these findings

indicated that the AP positional identity of neither the *Hox*-positive nor *Hox*-negative CNCCs is fixed in the progenitors or passively inherited from the hindbrain to the pharyngeal arches, and revealed a high degree of plasticity in these cells, which is inconsistent with the pre-patterning model. Instead, it appears that NCCs rely on specific cues in the pharyngeal arch environment as they migrate to develop their proper regional identity. The size of the cell community also appears to be a significant factor, suggesting that a complex interplay of genetics and cellular interactions plays a role in hindbrain and NC patterning (Santagati & Rijli, 2003; P. A. Trainor & Krumlauf, 2000). Nevertheless, they do not fully resolve the debate over the balance between multipotency and early restriction in neural crest cell fate

More recent studies using a wide range of experimental approaches, including chromatin profiling via ChIP-seq, single-cell transcriptomics, single-molecule fluorescent in situ hybridization, and modern clonal analysis tools, have given new insights regarding this matter. For instance, Minoux et al. (2017) have demonstrated the importance of epigenetic regulation in the maintenance of positional plasticity in CNCCs and in controlling axial regional identity of postmigratory CNCCs. In pre-migratory CNCCs positional genes and enhancers are kept in a poised chromatin state, meaning they are accessible and bivalent, ready to be activated upon receipt of the right environmental inputs. Interestingly, promoters and enhancers remain in a poised-for-action chromatin state in post-migratory CNCC subpopulations, in which they are not active (Minoux et al., 2017). Hence, epigenetic poising may enable CNCCs to swiftly adjust their response to differences in local signaling, possibly helping them handle unexpected changes in location due to migration errors. Furthermore, it indicates that post-migratory CNCCs retain a certain level of plasticity. Such a model could explain past observations such as the growth of an ectopic upper beak upon the transplantation of the frontonasal ectodermal zone (FEZ) to a different region of the FNP (D. Hu et al., 2003), or the formation of supernumerary jaw elements upon grafting of chick endodermal tissue (Couly et al., 2002). In addition, the findings of Minoux et al. (2017) could explain differences in facial shape between individuals. Another study in avian embryos identified a neural crest stem cell niche in the dorsal neural tube with high expression of neural crest genes, pluripotency factors, and lineage markers using multiplex single-molecule fluorescence in situ hybridization (Lignell et al., 2017). In agreement with other studies, they also found that migratory CNCCs have the ability to take on multiple fates and that cell fate commitment is a gradual process resulting from the competition between gene expression programs (Baggiolini et al., 2015; Soldatov et al., 2019). In fact, using single-cell transcriptomics and bioinformatics to analyze the branching trajectories of murine NCCs revealed new insights into CNCC multipotency, EMT, and fate specification (La Manno et

al., 2018; Soldatov et al., 2019; Zalc et al., 2021). Specifically, Zalc et al. (2021) demonstrate that during delamination CNCCs lose the expression of neuroepithelial positional genes and adopt a uniform transcriptional signature, resulting in an equivalent CNCC population with the ability to readily adapt to future migratory and postmigratory environments (Zalc et al., 2021). These results could explain the previously observed ability of premigratory CNCCs to adapt to a new position upon transplantation at a different axial level (Couly et al., 2002). It is furthermore consistent with the presence of Polycomb-dependent bivalent chromatin at the promoters of facial patterning genes in early CNCCs, which is resolved into transcriptionally active states in response to environmental cues (Minoux et al., 2017). Moreover, CNCC progenitors appear to be able to transiently reactive pluripotency factors suggesting that these cells are naturally reprogrammed (Zalc et al., 2021). Complementary to these findings is the work of Soldatov et al. (2019), who also observe the erasure of the AP neuroepithelial transcriptional signature in delaminating NCCs. Moreover, they present a novel model for the process of cell fate commitment, defined by the progressive limitation of cell fates through sequential phases of coactivation, biasing, and resolution of competing fates (Soldatov et al., 2019). Taken together, these recent findings highlight the plasticity of CNCCs.

1.2.3 The Role of *Hox* Gene Expression in Cranial Neural Crest Cells

The *Hox* gene family is a group of homeodomain transcription factors that are evolutionarily conserved and play fundamental roles in the regulation of axial patterning in the development of various tissues in animals (T. Alexander et al., 2009; Mallo et al., 2010). In mammals, these genes are organized into four distinct clusters (*Hoxa-Hoxd*) that arose through duplication events early in the evolution of vertebrates, and can be divided into 13 paralogue groups (PGs) based on their sequence characteristics (e.g. *Hoxa1* and *Hoxb1* belong to PG1). Moreover, *Hox* genes exhibit temporal and spatial colinearity along their clusters, meaning that their timing and spatial domains of expression along the anteroposterior (AP) axis are correlated with their relative gene order within each chromosomal cluster (Deschamps & Duboule, 2017; Kmita & Duboule, 2003). In the rhombomeres and cranial neural crest, these nested domains of *Hox* gene expression provide a combinatorial code that specifies regional properties in the developing head (Hunt, Gulisano, et al., 1991; Hunt, Whiting, et al., 1991; P. A. Trainor & Krumlauf, 2000). To understand the functional importance of colinear *Hox* gene expression in the establishment of CNCCs positional identity, it is important to consider how they pattern the AP axis of the hindbrain and pharyngeal arches.

Hox gene patterning in the hindbrain and cranial neural crest

For the establishment of ordered and nested gene expression domains of the nervous system, signaling pathways such as retinoic acid (RA), FGF, and WNT generate opposing gradients along the anteroposterior axis of the nervous system in the developing embryo (Parker et al., 2018). These gradients initiate a regulatory cascade of interactions between transcription factors (e.g. *Hoxa1, Krox20, Kreisler* and *Cdx1*), which generates tightly registered stripes of gene expression that will eventually become rhombomere boundaries (T. Alexander et al., 2009; Parker et al., 2016; Tümpel et al., 2009). *Hox* genes play a key role in this cascade. Through the presence of multiple RA response elements (RARE) within the different Hox clusters, they are able to directly respond to RA signaling creating nested expression domains. Additionally, they interact with other segmentally expressed genes, in order to control rhombomere formation, and are involved in extensive auto- and cross-regulatory circuits required for the specification of segmental identities and maintenance of segmental gene expression (Parker et al., 2019). A key characteristic of this process is the formation of lineage-restricted segments with distinct *Hox* gene expression patterns, leading to the formation of an AP *Hox* "code" that shapes regional variations in the hindbrain (Fig. 6).
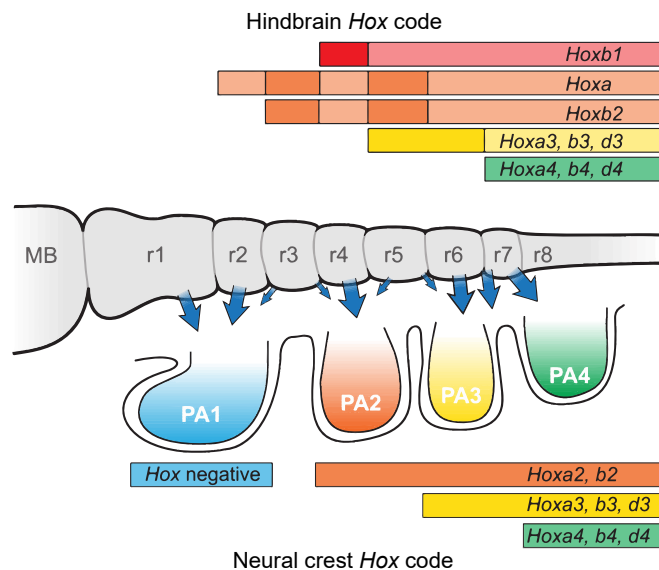

Figure 6. *Hox* **code in the hindbrain and cranial neural crest cells.** Schematic illustration *Hox* gene expression patterns in the mouse hindbrain (top) and cranial neural crest cells (CNCCs) (bottom) at pharyngula stage from the lateral viewpoint. Rhombomeres (r1-r8) and pharyngeal arches (PA1-PA4) are annotated. CNCCs migrate in distinct streams from the rhombomeres to the pharyngeal arches (dark blue arrows). The anteroposterior (AP) *Hox* code in the hindbrain and pharyngeal arches is indicated by different colors. Darker color shading within the hindbrain *Hox* expression domains (for instance, dark red for Hoxb1 in r4) reflects stronger levels of gene expression. Figure taken from Parker et al., 2018.

Hoxa1 and *Hoxb1* are the first *Hox* genes to be expressed in the neuroepithelium prior to the segmentation of the hindbrain into rhombomeres at around embryonic day 7.5 (E7.5) (Tümpel et al., 2009). Their early expression is driven by RA signaling, with *Hoxa1* extending anteriorly up to the presumptive R3 territory (Makki & Capecchi, 2010) and *Hoxb1* extending up to the presumptive R4 region (Studer et al., 1994; Wilkinson et al., 1989). Neither of the two *Hox* genes is expressed in migrating NCCs (Fabik et al., 2021). However, all NCCs that emanate from R4 are

derived from the *Hoxa1* lineage (Makki & Capecchi, 2010). Further downstream of the regulatory cascade, between E8.0 and E8.5, other *Hox* genes from PG1–4 are activated, once the expression of transcription factors *Krox20* and *Kreisler*, which is limited to certain rhombomeres, has been established. For instance, at first *Hoxa2* and *Hoxa3* are directly upregulated by *Krox20* in R3 and R5 via paralogous enhancers 5' to each gene (Nonchev, Maconochie, et al., 1996), followed by expression in R4 where *Hoxa2* is regulated by *Hoxb1* via an intronic enhancer and *Hoxb2* by a 5' enhancer (M. K. Maconochie et al., 1997; Tumpel et al., 2007). *Hoxa2* is expressed in R2 driven by the activity of a *cis*-regulatory element located in its own coding region (Tumpel et al., 2008). Lastly, PG3 *Hox* gene expression is initiated in R5 by Krox20 and Kreisler (Manzanares et al., 1999; Manzanares et al., 1997; Manzanares et al., 2002), whereas PG4 *Hox* genes are activated posteriorly to the R6/R7 boundary driven by RA signalling (Gould et al., 1997; Morrison et al., 1996; Nolte et al., 2006).

The *Hox* code from the AP axis of the neural tube is transposed to the AP axis of the head and pharyngeal arch mesenchyme, as the NCCs that originate from the hindbrain delaminate and migrate in distinct dorsolateral streams (Fig. 5a and Fig. 6) (Hunt & Krumlauf, 1991; Lumsden et al., 1991; Sechrist et al., 1993; P. A. Trainor & Krumlauf, 2001). However, the CNCCs *Hox* code is not simply a copy of the expression patterns found in the neural tube. In fact, the expression of *Hox* genes in migratory NCCs and developing rhombomeres is controlled by independent *Hox* gene enhancers, which may cause differences in expression between NCC subgroups and their corresponding rhombomeres of origin (Hunt et al., 1995; Hunt, Gulisano, et al., 1991; M. Maconochie et al., 1999; P. A. Trainor & Krumlauf, 2000; Tumpel et al., 2008). For instance, the expression of Hoxb1 is shut off in migrating R4-derived NCCs. On the other hand, R4-derived NCCs retain the expression of both *Hoxa2* and *Hoxb2* upon migration to PA2, whereas NCCs derived from the *Hoxa2* positive R2 domain are devoid of *Hox* gene expression after delamination (Hunt, Gulisano, et al., 1991; Hunter & Prince, 2002; Prince & Lumsden, 1994). By the time that all CNCCs have migrated to their final positions a pharyngeal Hox code of nested gene expression can be observed, where PA1 is Hox negative, PA2-PA4 express Hoxa2 and Hoxb2, PA3-PA4 express varying levels of Hox PG3 genes, and Hoxd4 is highly expressed in PA4 (Fig. 6).

Hox genes in craniofacial morphogenesis

The importance of accurate *Hox* gene expression along the AP axis of the pharyngeal arches for craniofacial morphogenesis has been emphasized by a multitude of *Hox* gene knock-out and ectopic over-expression experiments (Chisaka & Capecchi, 1991; Chojnowski et al., 2016; Condie

& Capecchi, 1994; Gendron-Maguire et al., 1993; Kitazawa et al., 2015; Kitazawa et al., 2022; Manley & Capecchi, 1997; Minoux et al., 2009; Minoux et al., 2013; Rijli et al., 1993; Santagati et al., 2005). The targeted inactivation of *Hoxa2* in the mouse (Gendron-Maguire et al., 1993; Rijli et al., 1993) provided the first piece of evidence for the involvement of *Hox* genes in establishing the AP positional identity of CNCCs, resulting in the homeotic transformation of PA2 skeletal derivatives into a subset of PA1-like structures and death within 24h post birth (Fig. 7b). More specifically, the lesser horns of the hyoid bone, styloid process, and stapes, PA2-derived structures, are absent in *Hoxa2*-null mice and replaced by a mirror image duplication of the incus, malleus, gonial bone, tympanic ring and part of the Meckel's cartilage (Fig. 7b)(Gendron-Maguire et al., 1993; Rijli et al., 1993), which are normally derived from PA1 CNCCs (Kontges & Lumsden, 1996). Additionally, the PA2-derived external ear or pinna is completely absent (Fig. 7b), whereas the PA1-derived external auditory canal is duplicated (Minoux et al., 2013; Rijli et al., 1993; Santagati et al., 2005). Similar results were obtained in *Xenopus* (Baltzinger et al., 2005) and zebrafish (Hunter & Prince, 2002) upon downregulation of *Hoxa2*. These observations suggest that *Hoxa2* plays a key role across species in the specification of PA2 CNCCs identity and the morphogenesis of PA2-derived elements. Further support for this was provided by Santagati et al. (2005) through the temporal inactivation of *Hoxa2* in mice demonstrating that *Hoxa2* is required for PA2 CNCC morphogenetic program post-migration.

Conversely, ectopic expression of *Hoxa2* in *Hox*-negative CNCCs resulted in varying morphological phenotypes. For example, in the study of Kitazawa et al. (2015) ectopic expression of *Hoxa2* in mice suppresses the formation of PA1-derived jaw structures and repatterns them into PA2-derived structures, thereby duplicating PA2 derivatives (Fig. 7c). A range of studies in *Xenopus*, zebrafish, and chick observed very similar phenotypes upon ectopic expression of *Hoxa2* (Grammatopoulos et al., 2000; Hunter & Prince, 2002; Pasqualetti et al., 2000). Additionally, the conditional ectopic expression of *Hoxa2* in CNCCs is by itself enough to transform proximal PA1 into PA2-like structures, including a duplication of the pinna (Kitazawa et al., 2015; Minoux et al., 2013). However, another study found in their chick model system that altered expression of *Hoxa2* in *Hox*-negative CNCCs led to the lack of the entire facial skeleton (Creuzet et al., 2002). This apparent discrepancy between studies has been resolved very recently by Kitazawa and colleagues (2022) demonstrating that the severity of hypoplasia/malformation in craniofacial structures is linked to the degree of *Hoxa2* overexpression. Overall these findings further highlight the evolutionarily conserved function of *Hoxa2* in craniofacial morphogenesis and provide evidence for a model in which CNCCs from PA1 and PA2 share an underlying *Hox*-free ground state. Interest-
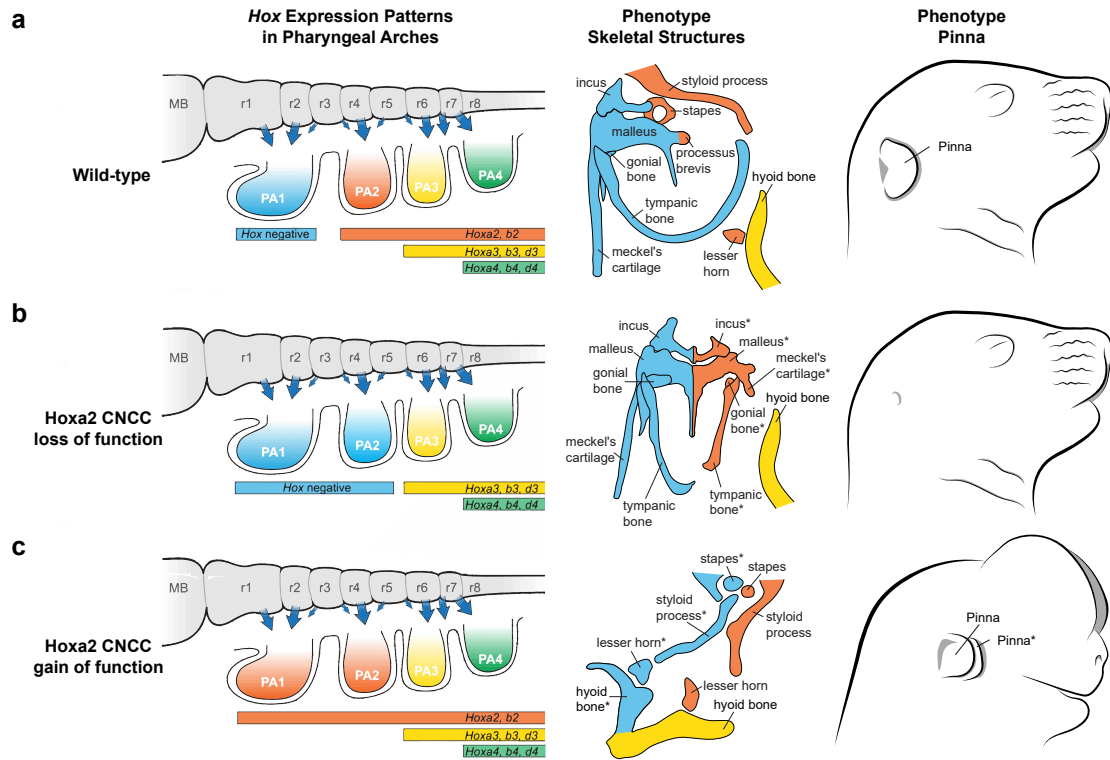
Figure 7. Effects of the loss or gain of function of *Hoxa2* gene expression in cranial neural crest cells. a Schematic representation of the CNCCs *Hox* code in the pharyngeal arches (PA1-PA4) at pharyngula stage (as in Fig. 6)(left), the middle ear ossicles and hyoid bone skeletal structures at E18.5 (middle), and the pinna at E18.5 (right) of wild-type mouse embryos. The colors of individual PAs reflect their anteroposterior *Hox* code. The colors of the skeletal structures are matched with the colors of the PAs indicating from which PA each structure is derived (for instance, the stapes is derived from PA2 CNCCs). **b** Schematic representation as in **a** showing the effects of loss of function of *Hoxa2* in CNCCs, namely a homeotic transformation of PA2 to PA1-like skeletal structures at the level of the middle ear, loss of the lesser horn of the hyoid bone, and loss of the PA2-derived pinna. **c** Schematic representation as in **a** showing the effects of gain of function of *Hoxa2* in CNCCs (based on Kitazawa et al., 2015), namely a homeotic transformation of PA1 into PA2-like structures at the level of the middle ear, a duplication of the hyoid bone including the lesser horn, and a duplication of the pinna. * marks duplicated structures. Panel **a** taken from Parker et al., 2018.

ingly, similar severe craniofacial defects could be observed in the mouse upon conditional deletion of *Ezh2* in CNCCs, which encodes a subunit of PRC2 that catalyzes H3K27 trimethylation (Minoux et al., 2013; Schwarz et al., 2014). Indeed, the four *Hox* clusters and positional genes are embedded in repressive H3K27me3 domains, while their promoters are usually in a bivalent state (Minoux et al., 2017; Schwarz et al., 2014). Upon deletion of *Ezh2* in CNCCs H3K27me3 is lost genome-wide leading to the ectopic expression of normally bivalent genes, including the *Hox* genes and other positional genes, in the anterior CNCCs of the FNP and PA1 (Minoux et al., 2017; Schwarz et al., 2014). Thus, *Hox* genes and other positional genes are epigenetically regulated in CNCCs and this regulation is essential in defining the distinct positional molecular identities of CNCCs.

Malformations of PA3 and PA4 CNCC-derived skeletal elements, but no homeotic transformations, have been described upon inactivation of single *Hox* PG3 and PG4 genes or combinations of them (Boulet & Capecchi, 1996; Chisaka & Capecchi, 1991; Chojnowski et al., 2014; Chojnowski et al., 2016; Condie & Capecchi, 1994; Manley & Capecchi, 1995, 1997). For instance, inactivation of *Hoxa3* in mice causes defects in the development of certain pharyngeal organs and PA3-derived craniofacial structures, such as the throat cartilages (Chisaka & Capecchi, 1991; Manley & Capecchi, 1995). Namely, the lesser horns of the hyoid bone are reduced or missing and the greater horn of the hyoid is fused with the superior horn of the thyroid cartilage (Fig x) (Chojnowski et al., 2016; Condie & Capecchi, 1994; Manley & Capecchi, 1997). Progressively more severe phenotypes can be observed in double and triple *Hox* PG3 mutants (Condie & Capecchi, 1994; Manley & Capecchi, 1997). Remarkably, deletion of the entire *HoxA* cluster in mouse NCCs results in a partial homeotic transformation of PA3 and PA4 derivatives toward a PA1-like morphology along with the PA2 homeosis induced by the absence of *Hoxa2* (Minoux et al., 2009). This NCC-specific *HoxA* cluster deletion has revealed that *Hoxa2* also plays a role in the patterning of PA3 and PA4 which is only apparent in the absence of *Hoxa3*, supporting the idea that *Hoxa3* has a more significant role than *Hoxa2* in patterning PA3 and PA4 (Minoux et al., 2009). Taken together, these studies strongly support the idea that a common *Hox*-free ground patterning program underlies the development of pharyngeal arches and that *Hox* genes from different PGs (such as *Hoxa2* and *Hoxa3*) act synergistically to pattern the derivatives of PA3 and PA4 (Minoux et al., 2009; Minoux & Rijli, 2010). Further phenotypic mouse data analysis indicates that the *HoxA* cluster genes are key in giving CNCCs their axial identity, while the *HoxB* cluster helps to fine-tune the patterning of the derivatives (Vieux-Rochas et al., 2013).

Despite all the phenotypic data on *Hox* mutants, there is still a lack of understanding of the specific mechanisms through which *Hox* genes influence the development of the head and face. This is a significant gap in our knowledge in the field of craniofacial morphogenesis. Studies on the role of *Hoxa2* provide an example of this lack of understanding. While it is clear that *Hoxa2* expression is required in CNCCs from PA2, the exact role of *Hoxa2* in this process is not fully understood. Previous research has indicated that *Hoxa2* expression is necessary for the proper differentiation of these cells (Santagati et al., 2005), but it is not yet known if *Hoxa2* also plays a role during the migration of these cells (Parker et al., 2018). Post-migration *Hoxa2* has been shown to modulate the response of CNCCs in PA2 to the local environmental signals thereby altering the competence of these cells (Bobola et al., 2003). Consequently, *Hoxa2* directly or indirectly represses different components of the ossification pathway which are normally active in

Introduction

PA1, including *Pitx1*, *Sox9*, *Lhx6*, *Alx4*, *Six2* and *Runx2* (Bobola et al., 2003; Kanzler et al., 1998; Kutejova et al., 2005; Kutejova et al., 2008). Furthermore, a recent study by Minoux et al. (2013) has shown that *Hoxa2* plays a role in the development of the external ear through its influence on the BMP signaling pathway and the expression of *Eya1*. Even more direct targets of *Hoxa2* could be identified through recent advancements in genomic technologies. As such, ChIP-seq data from mouse PA2 cells has revealed that *Hoxa2* has a positive feedback relationship with its cofactors *Meis1/2* (Amin et al., 2015), which plays a role in downstream Wnt signaling (Donaldson et al., 2012), leads to a transient activation of *Meox1* (Kirilenko et al., 2011), helps specify *Hmx1* expression in the lateral facial mesenchyme through an evolutionarily conserved enhancer (Rosin et al., 2016). Besides, these ChIP-seq experiments have found that *Meis*, *Pbx*, and *Hoxa2* bind to the area upstream of *Hoxa2* in PA2, which may suggest that self-regulation plays a role in its expression in neural crest tissue (Amin et al., 2015; Donaldson et al., 2012). Lastly, *Hoxa2* and *Hoxa3* have the ability to dimerize, which may play an important role in the patterning of PA3 and PA4 derivatives (Mallen et al., 2021).

Hoxa2 regulation in the hindbrain and cranial neural crest

Due to the fact that *Hoxa2* is one of the most important factors in specifying CNCCs AP identity, particularly of the second pharyngeal arch, and required for proper craniofacial morphogenesis, it is crucial to understand how *Hoxa2* gene expression is regulated along the AP axis both in the hindbrain and NCCs originating from it. However, there is limited information available on the specific regulatory elements that control the activation and maintenance of *Hoxa2* gene expression and whether the regulatory influences from these elements are shared or distinct in both tissues. In particular, the mechanisms that regulate *Hoxa2* expression in the NC are not as well understood as the current knowledge of rhombomeric *Hox* regulation (Parker et al., 2018).

Over the years a wide range of proximal enhancers likely to be involved in the regulation of *Hoxa2* gene expression have been identified. Mainly via transient enhancer activity assays, most of these regulatory elements have been shown to have non-overlapping activity patterns in different regions of the hindbrain and CNCCs, pointing towards distinct regulatory mechanisms between these tissues. To date we have knowledge about a total of five putative NC-specific enhancers (NC1-NC5) for *Hoxa2* (M. Maconochie et al., 1999; McEllin et al., 2016), which partially overlap with eight separate R3/R5 enhancers (RE1-RE5, *Krox20*, *Sox*) (M. K. Maconochie et al., 2001; Nonchev, Maconochie, et al., 1996; Nonchev, Vesque, et al., 1996; Tümpel et al., 2006), five R4 enhancers (Lampe et al., 2008; Tumpel et al., 2007) and five R2 enhancers (Frasch et al.,
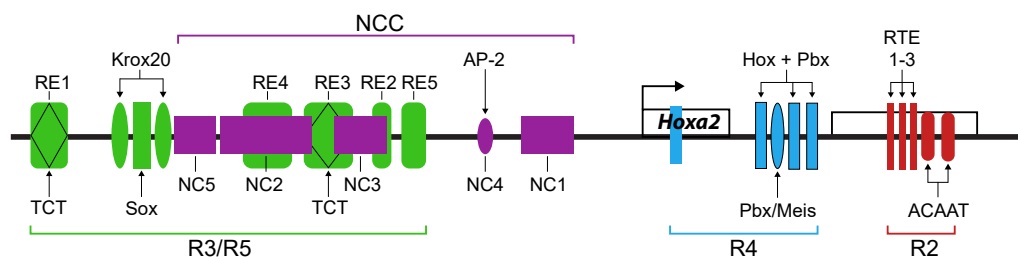
Figure 8. *Hoxa2* regulatory landscape. Schematic illustration of the enhancers regulating *Hoxa2* expression in neural crest cells (NCCs) and rhombomeres (R2-R5).The exons of *Hoxa2* are represented by white boxes and the transcription start site by an arrow. The colored boxes depict the enhancers: R2 elements in red, R4 elements in blue, R3/R5 elements in green and NCC elements in purple (not drawn to scale). Known direct inputs from transcription factors and transcription factor motifs in these enhancers are depicted by arrows. *Hoxa2* is regulated by independent enhancers in the four different tissues. Figure adapted from Parker et al., 2019; Tümpel et al., 2006.

1995; Tumpel et al., 2008) in the mouse (Fig. 8 x)(Parker et al., 2019; Parker et al., 2018; Tümpel et al., 2009). The enhancers regulating *Hoxa2* expression in R4 consist of three intronic Hox/Pbx binding sites, one intronic Prep/Meis binding site (Tumpel et al., 2007), and one exonic Hox/Pbx binding site (Lampe et al., 2008), required for regulatory activity. This entire R4 regulatory module responds to the expression of both *Hoxa2* and *Hoxb1*, indicating that *Hoxa2* is first activated in R4 as a direct target of *Hoxb1* and then sustains its own expression via an autoregulatory feedback loop (Lampe et al., 2008; Tumpel et al., 2007). Hox proteins achieve specific binding and transcriptional activity through interactions with the Pbx and Prep/Meis family of Hox cofactors (Bobola & Sagerström, 2022; Moens & Selleri, 2006).

Hoxa2 is the only *Hox* gene expressed in R2, where five elements embedded in the second exon of *Hoxa2* collaborate to direct its expression. Two of the enhancers contain binding sites for Sox proteins that are essential for the activity of the enhancers (Frasch et al., 1995; Tumpel et al., 2008). Then, in R3 and R5 expression of *Hoxa2* is in part mediated by multiple *Krox20* binding sites in close proximity to a BoxA motif (Sox site) upstream of the *Hoxa2* promoter (M. K. Maconochie et al., 2001; Nonchev, Maconochie, et al., 1996; Nonchev, Vesque, et al., 1996). In addition the *cis*-regulatory elements RE1-RE5, upstream of *Hoxa2*, are required for the activity of *Hoxa2* in R3/R5, of which RE1 (5' of the *Krox20* motifs) and R3 contain TCT motifs crucial for R3/R5 expression, while RE2, RE4 and RE5 contain no obvious motifs or shared common elements (M. K. Maconochie et al., 2001; Tümpel et al., 2006).

When it comes to the five enhancers (NC1-NC5) that are required for *Hoxa2* expression in NCCs, there is very little knowledge about how their activity is regulated. In the mouse, one of the elements, NC4, contains a binding site for transcription factor AP-2a/Tfap2a that is essential for enhancer activity (M. Maconochie et al., 1999). Furthermore, both NC2 and NC3 overlap with

33

RE2-RE4 (Fig. 8) and show enhancer activity in both NCCs and R3 and R5 of the hindbrain. A recent study found Meis and Hox/Pbx binding sites in the NC3 enhancer (Parker et al., 2019). The Meis site is involved in mediating enhancer activity in both R3/R5 and NCCs, whereas the Hox/Pbx site only contributes to enhancer activity in NCCs (Parker et al., 2019). Taken together, these findings suggest that Meis and Pbx, in addition to their already known fundamental functions in tissue patterning (Machon et al., 2015; Moens & Selleri, 2006), may play an important role in the regulation of *Hoxa2* gene expression in the hindbrain and NC in concert with or independent of Hox proteins (Parker et al., 2018). Moreover, they show that *Hoxa2* expression in the hindbrain and NCCs is largely regulated by independent enhancer elements. It should be pointed out that none of the above studies analyzed the effect of deleting single or multiple of these enhancers on *Hoxa2* expression *in vivo*. Hence, there remains a level of uncertainty regarding the regulatory function of each of these putative enhancers on *Hoxa2* gene expression.

1.3 Thesis objective

The goal of this thesis was to gain a deeper understanding of the mechanisms that control craniofacial development. To address this we focused on the systematic identification of putative enhancers, in particular super-enhancers, that might play a role in the regulation of key developmental genes involved in face formation. To this end, we employed PCHi-C, Hi-C, ChIP-seq, ATAC-seq, and RNA-seq leading to the discovery of multiple SEs required for the expression of *Hoxa2*, a master regulator of PA2 CNCCs fate.

Results

This project was a collaboration with Maryline Minoux, Onkar Joshi, Yousra Ben-Zouari, and the other authors listed in the manuscript. My contributions included the performance of most of the experiments, computational data analysis, and interpretation of experimental data. Namely, I performed all experiments on pinna-derived CNCCs at E12.5 and E14.5. In addition, I carried out all the Hi-C experiments, except in mESC which was done by Adwait Salvi, performed RNA-seq on PA2 and PA3 of WT and homozygous HIRE1 mutant embryos, did all whole mount in-situ hybridization experiments, and analyzed mouse skeletons with the help of Maryline Minoux and Filippo M. Rijli. Furthermore, I checked the conservation of HIRE1 and HIRE2 across species and performed a transcription factor motif analysis in collaboration with Michael Stadler. Lastly, I performed the computational quality control of ChIP-seq, ATAC-seq and RNA-seq data.
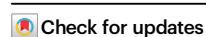
nature communications

A multiple super-enhancer region establishes inter-TAD interactions and controls *Hoxa* function in cranial neural crest

Sandra Kessler [1,2,4], Maryline Minoux[1,3,4], Onkar Joshi[1,4], Yousra Ben Zouari [1,4], Sebastien Ducret[1], Fiona Ross[1,2], Nathalie Vilain[1], Adwait Salvi[1,2], Joachim Wolff[1], Hubertus Kohler[1], Michael B. Stadler [1] & Filippo M. Rijli [1,2] ✉

Enhancer-promoter interactions preferentially occur within boundary-insulated topologically associating domains (TADs), limiting inter-TAD interactions. Enhancer clusters in linear proximity, termed super-enhancers (SEs), ensure high target gene expression levels. Little is known about SE topological regulatory impact during craniofacial development. Here, we identify 2232 genome-wide putative SEs in mouse cranial neural crest cells (CNCCs), 147 of which target genes establishing CNCC positional identity during face formation. In second pharyngeal arch (PA2) CNCCs, a multiple SE-containing region, partitioned into *Hoxa* Inter-TAD Regulatory Element 1 and 2 (HIRE1 and HIRE2), establishes long-range inter-TAD interactions selectively with *Hoxa2*, that is required for external and middle ear structures. HIRE2 deletion in a *Hoxa2* haploinsufficient background results in microtia. HIRE1 deletion phenocopies the full homeotic *Hoxa2* knockout phenotype and induces PA3 and PA4 CNCC abnormalities correlating with *Hoxa2* and *Hoxa3* transcriptional downregulation. Thus, SEs can overcome TAD insulation and regulate anterior *Hoxa* gene collinear expression in a CNCC subpopulation-specific manner during craniofacial development.

The establishment of complex spatiotemporal gene expression programs, controlling appropriate patterning events during development, involves regulatory DNA enhancer elements engaged in physical contact with their target gene promoters, sometimes over long distances[1]. The importance of these *cis*-regulatory elements during development is highlighted by the fact that their disruption can lead to disease and congenital disorders in humans[2–4]. Furthermore, a number of studies underline the relevance of enhancers in morphogenesis[5–11].

Among the most complex processes during embryonic development is the morphogenesis of the craniofacial skeleton, which derives from the cranial neural crest cells (CNCCs), a multipotent cell population that arises dorsally in the forming neural tube and migrates to colonize the facial and pharyngeal prominences[12–15]. To generate

craniofacial skeletal and cartilage structures with proper shape, size, and orientation, CNCCs need to acquire specific positional identities and patterning information. This is achieved through the coordinated action of transcription factors, whose proper expression in the distinct facial and pharyngeal CNCC subpopulations is established in response to position-specific environmental signals[13,15,16].

Hox genes encode conserved homeodomain transcription factors that in mammals are organized into four clusters (*Hoxa-d*)[17–19]. While the CNCCs colonizing the frontonasal process and first pharyngeal arch (PA1) do not express *Hox* genes, *Hox* genes are required to provide rostrocaudal positional identity to the hindbrain rhombomeres (R)[20,21] and CNCCs of PA2-PA4, providing each CNCC subpopulation with unique patterning information[13,15,20]. Among the four clusters,

[1]Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058 Basel, Switzerland. [2]University of Basel, Basel, Switzerland. [3]INSERM UMR 1121, Université de Strasbourg, Faculté de Chirurgie Dentaire, 8, rue Sainte Elisabeth, 67 000 Strasbourg, France. [4]These authors contributed equally: Sandra Kessler, Maryline Minoux, Onkar Joshi, Yousra Ben Zouari. ✉e-mail: filippo.rijli@fmi.ch

Hoxa genes play a predominant role in patterning skeletogenic CNCCs[22]. *Hoxa2* is the only *Hoxa* gene expressed in CNCCs of PA2[23,24]. In the mouse, inactivation of *Hoxa2* induces homeotic transformation of the PA2-derived skeletal elements into a subset of PA1-like, *Hox*-negative, structures[25–28]. *Hoxa2* is also necessary and sufficient for external ear morphogenesis and *Hoxa2/HOXA2* hypomorph/haploinsufficient mutations in both mouse and humans cause microtia[26–31]. Moreover, *Hoxa2* and *Hoxa3* synergistically pattern PA3 and PA4 CNCC derivatives[22]. However, very little is known about how *Hoxa2* expression is regulated in PA2 CNCCs during craniofacial development, with only a few proximal regulatory elements identified in mouse[20] for which no functional data are available to date.

At the genome-wide level, in vivo epigenomic mapping and transgenic assays have identified putative enhancers that are active in the craniofacial prominences of embryonic day (E) 10.5 and 11.5 mouse embryos[5,16]. However, the target genes transcriptionally regulated by these putative enhancers remain largely unknown. This is partly due to the lack of comprehensive maps of 3D chromatin organization coupling these enhancers to their target promoters. Chromosome Conformation Capture (3C)-based techniques[32] have allowed the characterization of 3D chromatin interaction networks leading to the identification of specific long-range regulatory elements for developmentally important genes[33–39]. Moreover, these techniques have revealed that enhancer-promoter interactions preferentially occur within topologically associating domains (TADs), evolutionary and developmentally largely invariant 3D chromatin structures[40–42].

Very few putative long-range craniofacial enhancers have been functionally tested by in vivo targeted deletion so far, which resulted in quite mild variations in the shape and size of the affected structures[5,7,9]. This indicated that regulatory redundancy may contribute to buffer potentially deleterious phenotypic effects of single enhancer mutations[10]. Enhancer clusters, also termed super-enhancers (SEs), have been identified genome-wide as large, highly active regulatory domains containing clusters of active enhancers in close linear proximity, ensuring high expression levels of target genes[43,44]. However, little is known about SEs and their target genes active during craniofacial development, as well as about their potential functional impact on in vivo gene regulation in the context of 3D chromatin topology.

Here, we aimed to systematically identify SEs that might play a role in craniofacial morphogenesis by controlling the transcriptional regulation of key genes involved in establishing CNCC subpopulation-specific transcriptional programs[16]. Using Hi-C[45] and Promoter-Capture Hi-C (PCHi-C)[37] assays, in association with chromatin immunoprecipitation followed by sequencing (ChIP-seq) and Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq) we identified 2232 putative SEs. 147 of these targeting transcription factor-coding genes involved in establishing post-migratory CNCC positional identities[16]. We then focused on a large genomic region containing two subregions of 175 kb (*Hoxa* Inter-TAD Regulatory Element 1, HIRE1) and 39 kb (*Hoxa* Inter-TAD Regulatory Element 2, HIRE2) composed of multiple SEs, localized about 1.07 and 1.33 Mb away from the *Hoxa2* locus, respectively. We show that HIRE1 and HIRE2 are highly conserved in mammals, including humans, and establish inter-TAD long-range interactions with *Hoxa2* selectively in PA2 but not PA1 CNCCs, skipping the 3′ (centromeric) TAD neighboring the anterior *Hoxa* cluster. CRISPR-mediated targeted deletion of HIRE1 in the mouse phenocopied the full homeotic *Hoxa2* knockout phenotype in PA2 CNCCs and additionally induced malformations in PA3 and PA4 CNCC-derived skeletal structures, correlating with transcriptional downregulation of both *Hoxa2* and *Hoxa3*. In contrast, targeted deletion of HIRE2 did not yield major alterations of CNCC-derived skeletal structures, suggesting functional redundancy, but nonetheless resulted in microtic (i.e., small and malformed) pinnae in adult mice when put on a *Hoxa2* haploinsufficient sensitized background.

Thus, a multiple SE-containing region can overcome TAD insulation and provide very long-range transcriptional regulation in a CNCC subpopulation-specific manner and ensure robust target gene expression levels during craniofacial development.

Results

### Genome-wide identification of super-enhancers in post-migratory CNCC subpopulations

To identify SEs that might play regulatory roles in CNCC subpopulations of distinct developing facial prominences, we analyzed H3K27ac ChIP-seq datasets from E10.5 mouse CNCCs of *Hox*-negative frontonasal process (FNP), maxillary (Mx) and mandibular (Md) components of PA1, and of *Hoxa2*-expressing PA2[16] (Fig. 1a). We first merged the H3K27ac ChIP-seq reads from these four CNCC subpopulations for peak calling and then largely followed the workflow of the ROSE algorithm to identify SEs[44,46] (Methods). Briefly, individual H3K27ac peaks within 12.5 kb or less were merged into larger regions, and, in each CNCC subpopulation, the H3K27ac signal was quantified to distinguish SEs from typical enhancers[44] (Supplementary Fig. 1). We excluded merged regions that overlapped with promoters, to identify only distal enhancer elements. In total, we found 2232 putative SEs that were active in at least one of the four CNCC subpopulations (Supplementary Data 1).

We next focused on the 237 transcription factor-coding genes differentially expressed between FNP, Mx, Md, and PA2 CNCCs (i.e., "positional" transcription factors, previously identified in ref. 16) and assessed if they were targeted by one or more SE(s). Among the putative SEs described above, for each CNCC subpopulation, we selected the SEs that were connected by at least one significant interaction to the promoters of positional transcription factor coding genes that displayed expression levels of RPKM >2 (Methods). To this aim, we performed PCHi-C[37] in duplicate for each of the four CNCC subpopulations (Supplementary Fig. 2a). Briefly, we used biotinylated RNA bait probes targeting promoter regions (Methods) to selectively enrich for all distal genome-wide sequences interacting with promoters from a pool of 'all-to-all' genomic interactions generated by Hi-C, followed by high throughput paired-end sequencing and statistical analysis. For CNCC collection, we micro-dissected E10.5 FNP, Mx, Md, and PA2 prominences and isolated red fluorescent protein (RFP)-expressing CNCCs by fluorescence-activated cell sorting (FACS) (Figs. 1a, 2a and Supplementary Fig. 3) (Methods).

In total, 62 out of 237 positional gene promoters showed, in the CNCC(s) where they were expressed, at least one significant interaction with at least one restriction fragment overlapping with a putative SE region (Fig. 1b and Supplementary Data 2). Among these targets, we found genes coding for transcription factors whose mutations result in craniofacial abnormalities in humans and/or are involved in prominence-specific CNCC patterning in mouse, including *Msx1/2*, *Tfap2b*, *Pax3*, *Alx4*, *Six1/2*, *Alx1*, *Hoxa2*, *Pitx1*, *Barx1*, *Meis1/2*, *Dlx3*, and *Hand2* (Fig. 1b and Supplementary Data 2). Some promoters were linked to multiple SEs, while one SE interacted with two gene promoters, namely those of *Hes6* and *Twist2*. Overall, we found 147 putative SEs associated with a positional transcription factor-coding gene promoter in at least one CNCC subpopulation (Fig. 1b and Supplementary Data 2).

Identification of inter-TAD super-enhancers targeting *Hoxa2*

Hoxa2 is the only *Hoxa* gene expressed in the CNCCs of PA2[23,24] (Fig. 2b) and is required to pattern all PA2-derived skeletal and cartilaginous structures, including the pinna[25–28]. Notably, PCHi-C revealed that 5 putative SEs, SE1–5, selectively targeted *Hoxa2* in E10.5 PA2 CNCCs, where this gene is highly expressed, unlike in the other CNCC subpopulations (Figs. 1b, 3a). SE1–5 were all located in a genomic region at a very large distance (>1 Mb) 3′ (centromeric) from the *Hoxa2* promoter (Figs. 1b, 3a). To assess temporal SE targeting dynamics and
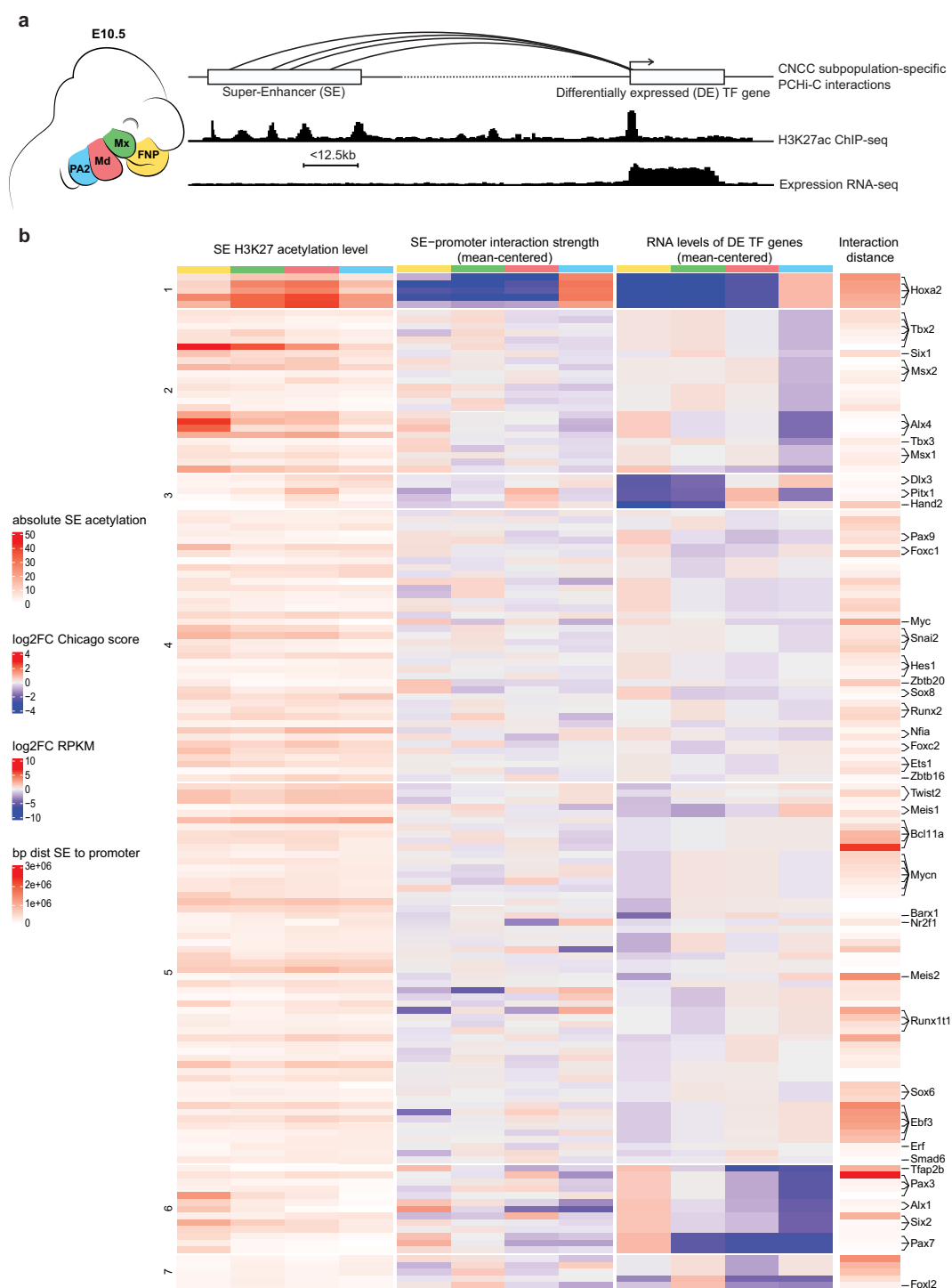
correlate it with the *Hoxa2* expression pattern in PA2 CNCCs and their derivatives, we analyzed PA2-derived pinna CNCCs at E12.5 and E14.5, in addition to E10.5 PA2 CNCCs. We micro-dissected E12.5 and E14.5 pinnae, isolated RFP-expressing CNCCs by FACS (Fig. 2a and Supplementary Fig. 3; Methods), and processed them for Hi-C, PCHi-C (Supplementary Fig. 2b), RNA-seq, ATAC-seq, and ChIP-seq assays. *Hoxa2* is highly expressed in E10.5 PA2 CNCCs[16] and its expression level further

increased in the E12.5 developing pinna (logFC = 0.753; FDR = 8.89E-07), while decreasing from E12.5 to E14.5 (logFC = −0.455; FDR = 5.71E-05) (Supplementary Data 3). Accordingly, the *Hoxa2* locus remained accessible and was enriched with active H3K27ac and H3K4me2 histone marks from E10.5 to E14.5, while the remainder of the transcriptionally silent *Hoxa* cluster was blanketed by the Polycomb (Pc)-dependent repressive H3K27me3 mark (Fig. 2b)[16]. In contrast, in *Hox-*

Fig. 1 | Super-enhancer calling and assignment to positional transcription factor-coding genes. a (Left) Schematic of mouse facial prominences at embryonic day 10.5 (E10.5). Cranial neural crest cells (CNCCs) of the frontonasal process (FNP), maxillary (Mx), mandibular (Md), and second pharyngeal arch (PA2) are depicted in yellow, green, red, and blue, respectively. (Right) Each CNCC subpopulation was subjected to RNA-seq, H3K27ac ChIP-seq, and PCHi-C. Promoter distal ChIP-seq peaks with a maximum distance of 12.5 kb from each other were used for super-enhancer (SE) calling. Patterns of ChIP-seq and RNA-seq signals at one SE-promoter pair, where both elements are active, are shown in the style of genome browser tracks. Links of SEs to their target genes were identified with PCHi-C focusing on positional transcription factor (TF) coding genes[16] and are represented as arcs. **b** Heatmap of the SEs assigned to positional transcription factor-coding genes. Each heatmap row represents a SE-promoter pair. In each CNCC subpopulation, the SEs

were linked to promoters of positional transcription factor coding genes[16] if there was at least one significant interaction between the two elements, and if the gene was expressed (>2 RPKM). 147 SEs were linked to 62 different genes, with 148 unique pairings. The row annotation highlights transcription factor coding genes involved in craniofacial development and/or malformations, if mutated. From left to right, the heatmap displays the H3K27 acetylation level of each SE for each CNCC subpopulation, the mean CHiCAGO score of all interactions from a promoter bait of a positional transcription factor coding gene to a restriction fragment that overlaps with a SE, the expression of the positional transcription factor coding genes and the distance between the two elements. The interaction strength and gene expression are given as log2 FC for each CNCC subpopulation in relation to the mean across all four CNCC subpopulations. The rows were grouped by k-means clustering on the gene expression levels (cluster numbers are indicated on the left).

free E10.5 Md CNCCs, the whole *Hoxa* cluster was embedded into a Polycomb H3K27me3 repressive domain (Fig. 2b)[16]. As in E10.5 PA2, in E12.5 and E14.5 pinna CNCCs, SE1–5 were accessible and active, as indicated by ATAC-seq peaks and H3K27ac enrichment (Fig. 3a and Supplementary Fig. 4a).

Hi-C profiles of mouse embryonic stem cells (mESCs) and E10.5 Md CNCCs, and of *Hoxa2*-expressing PA2 and pinna CNCCs at E10.5, E12.5, and E14.5 further showed that, in all these cell populations, the *Hoxa* cluster was embedded in a dense domain spanning the border between two adjacent 3′ and 5′ TADs[47] (blue box, Fig. 3b and Supplementary Fig. 2c). As assessed by CTCF binding profiles (Supplementary Fig. 2d) and TAD separation score (black arrow, Fig. 3b and Supplementary Fig. 2c; Methods), no difference among cell populations was observed in the position of a strongly predicted boundary, segregating the *Hoxa2* locus from the SE1–5-containing genomic region in distinct TADs (Fig. 3b and Supplementary Fig. 2c, d). SE1 partially overlapped this TAD boundary at its 5′ end, whereas SE2–5 covered a genomic region encompassing almost entirely the neighboring TAD across that boundary (Fig. 3b and Supplementary Fig. 2c).

Given their profile of interactions and the relative proximity of SE2–4 as compared to SE5, we further subdivided SE2–5 into two main regulatory subdomains. The first subdomain (chr6:50,913,170-51,087,888), hereafter referred to as '*Hoxa* inter-TAD regulatory element 1' (HIRE1) is located about 1.07 Mb away from *Hoxa2* and covers a 175 kb region encompassing SE2–4 (excluding the SE2 most 5′ end which does not interact with *Hoxa2*) (Fig. 3a, b and Supplementary Fig. 4). The second subdomain (chr6:50,789,172-50,828,639), referred to as HIRE2, is a 39 kb region encompassing SE5 and localized about 1.33 Mb away from *Hoxa2* (Fig. 3a, b and Supplementary Fig. 4). The SE1–5 interaction patterns with *Hoxa2* in PA2-derived CNCCs were similar from E10.5 through E14.5 (Fig. 3a, b). However, HIRE1 interacted more strongly with *Hoxa2* in PA2 and pinna CNCCs at E10.5 and E12.5, whereas HIRE2 had the strongest interaction in pinna CNCCs at E12.5 (Fig. 3c, d and Supplementary Fig. 5a). Notably, in *Hox*-negative Md, Mx, and FNP CNCC subpopulations, even though not interacting, HIRE1 and HIRE2 were also accessible and enriched with the active H3K27ac mark (Fig. 3a and Supplementary Fig. 4a), whereas the *Hoxa2* locus was maintained repressed by Polycomb-dependent H3K27me3 (Fig. 2b)[16].

The TAD containing HIRE1 and HIRE2 is a gene-poor chromosomal region (Fig. 3c). Virtual 4C plots from the HIRE1 or HIRE2 viewpoints confirmed the *Hoxa2*-specific interactions. Furthermore, they identified weak intra- and inter-TAD interactions with *Npvf* and *Hoxa1/Hoxa3*, respectively (Supplementary Fig. 5). None of these genes are expressed in the selected CNCC subpopulations, suggesting that these weak contacts are not functional nor specific, likely due to physical proximity of *Hoxa1/Hoxa3* to *Hoxa2* and *Npvf* to the HIRE1/2 SEs, respectively. Moreover, most Hi-C-based methods cannot resolve very proximal interactions (typically <30 kb) from the generally high background crosslinking frequency between genomic sequences over these distances. Thus, very long-range inter-TAD interactions were

established between *Hoxa2* and HIRE1/HIRE2 in E10.5 PA2 and E12.5-E14.5 pinna CNCCs. Remarkably, *Hoxa2* inter-TAD interactions were visible as an asymmetrical "architectural stripe"[48] on low-resolution Hi-C plots (arrowheads, Fig. 3b and Supplementary Fig. 5a) and confirmed at higher resolution by PCHi-C (Fig. 3a). Inter-TAD interactions were selective for PA2 and pinna CNCCs, which express *Hoxa2* (Figs. 2b, 3a, b and Supplementary Fig. 2c), as they were absent in *Hox*-free mESCs and E10.5 Md CNCCs (Fig. 3a, b and Supplementary Fig. 2c).

HIRE1 and HIRE2 are highly conserved in mammals

HIRE1 and HIRE2 sequences were highly conserved in eutherian mammals, whereas in more phylogenetically distant species, such as marsupials, birds, or fish, the degree of conservation was considerably lower (Fig. 4). Nevertheless, both HIRE2 and HIRE1 contained multiple 200–1000 base pair (bp) long elements showing high conservation in basewise analysis across 60 vertebrates. In HIRE1, 22 of such elements were conserved in bird species, with two elements notably conserved in fish as well (Fig. 4a). In HIRE2, six elements were conserved to some degree down to birds and two out of six showed conservation in the coelacanth fish (Fig. 4b). Furthermore, these highly conserved elements tended to overlap with ATAC-seq peaks in developing CNCCs and derivatives (Fig. 4, blue bars), suggesting a conserved functional role of HIRE1 and HIRE2 in vertebrates, in particular, Eutheria.

HIRE1 deletion phenocopies the full *Hoxa2* knockout phenotype, whereas HIRE2 deletion in a *Hoxa2* haploinsufficient background results in microtia

To assess their involvement in *Hoxa2* transcriptional regulation, we deleted HIRE1 and HIRE2 using the CRISPR-Cas9 system (Fig. 5a and Supplementary Fig. 4b; Methods). Mice lacking one copy of HIRE2 (*HIRE2^{del/wt}*) appeared phenotypically normal. Similarly, as compared to wild-type (WT), E18.5 *HIRE2^{del/del}* homozygous mutant fetuses ($n = 6$) did not display visible abnormalities of CNCC-derived pinna (Fig. 5b, c), middle ear, or hyoid structures (Fig. 6a–f and Supplementary Fig. 6a–d). To address the potential impact of the HIRE2 deletion on a haploinsufficient *Hoxa2* background, we generated trans-heterozygous mutants (Fig. 5a) by mating *HIRE2^{del/wt}* mice with *Hoxa2^{EGFP/wt}* mice, carrying a *Hoxa2* knockout allele[49]. While no clear defect of the pinna was visible at E18.5 in *HIRE2^{del/wt};Hoxa2^{EGFP/wt}* fetuses ($n = 4$; 8 pinnae) (Fig. 5d), adult trans-heterozygous animals displayed microtia, i.e., smaller and misshapen ears (Fig. 5e–g, $n = 5/5$; 10 pinnae). Mild skeletal abnormalities could be observed at E18.5 (Fig. 6g–i and Supplementary Fig. 6e, f), namely, the PA2-derived processus brevis of the malleus was reduced (asterisk, Fig. 6g, h, $n = 8/8$ sides) and an ectopic cartilage nodule was inconsistently present on the PA2-derived styloid process (white arrow, Fig. 6g, $n = 5/8$ sides). This suggests that HIRE2 is mostly functionally redundant, although still required to contribute to full *Hoxa2* expression levels (see below) on a sensitized *Hoxa2* haploinsufficient genetic background.

HIRE2 deletion might be mainly functionally compensated by HIRE1. Heterozygous mutant mice for HIRE1 deletion (*HIRE1^{del/wt}*)
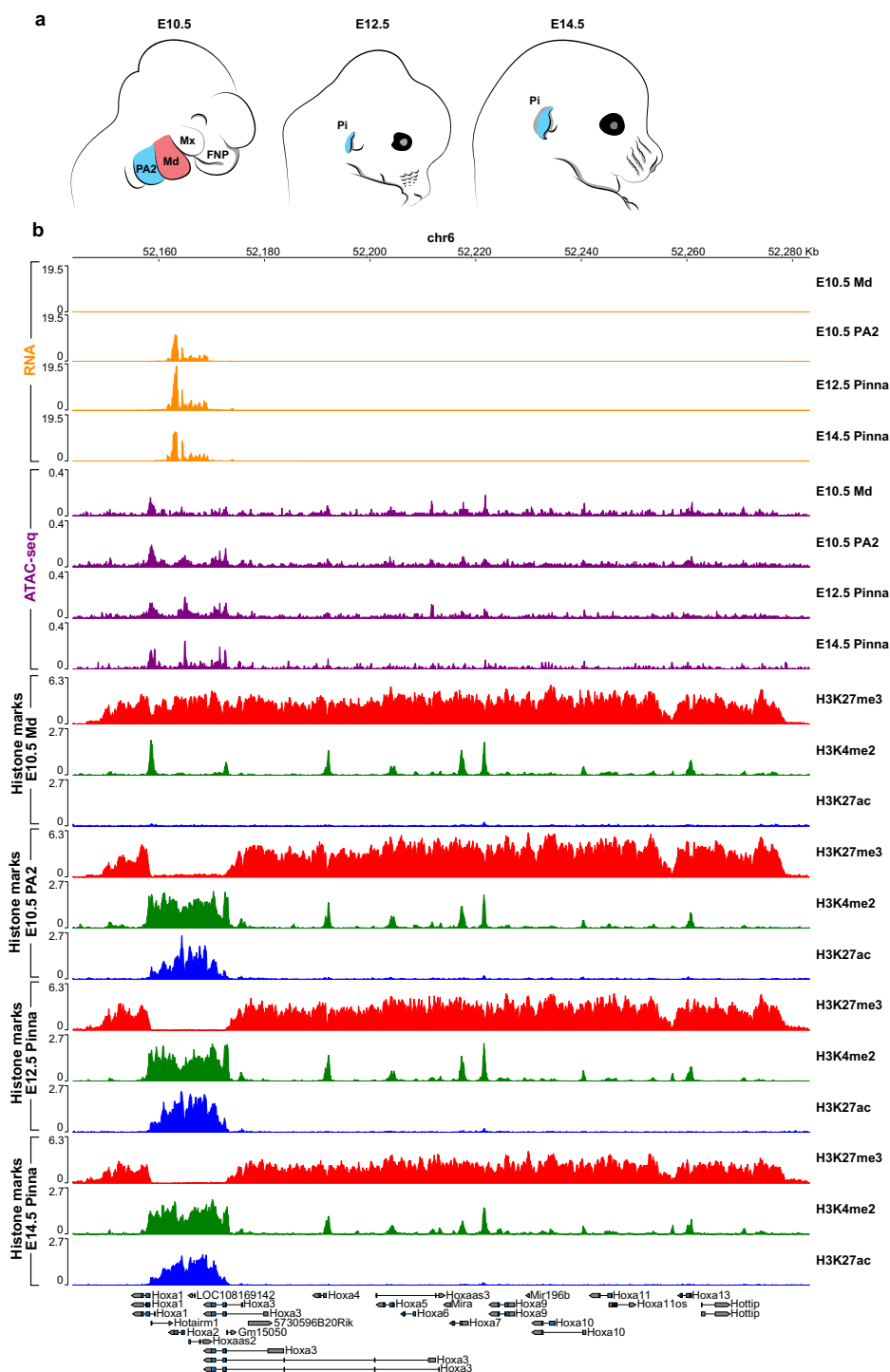
Fig. 2 | Chromatin state at *Hoxa2* locus. a Schematic of mouse developmental progression from E10.5 to E14.5. The mandibular process (Md) at E10.5 is highlighted in red. The second pharyngeal arch (PA2) at E10.5 and the PA2-derived pinna (Pi) at E12.5 and E14.5 are highlighted in blue. Mx, maxillary process of first pharyngeal arch; FNP, frontonasal process. **b** *Hoxa* cluster genome browser view with RNA (orange), chromatin accessibility (ATAC-seq, purple), and ChIP-seq profiles for H3K27me3 (red), H3K4me2 (green), and H3K27ac (blue) from E10.5 Md and PA2 cranial neural crest cells (CNCCs), and E12.5 and E14.5 pinna CNCCs.
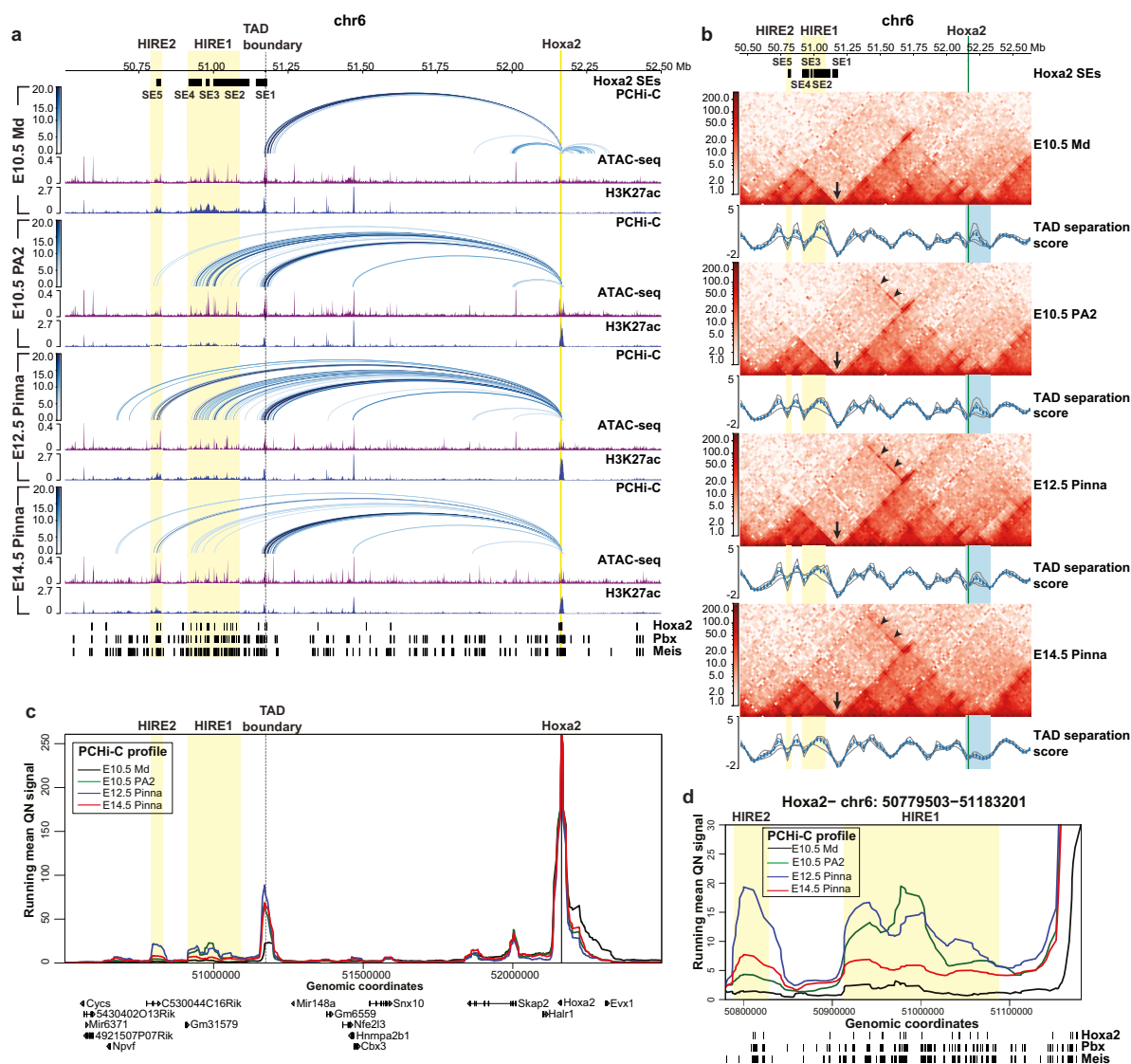
Fig. 3 | *Hoxa* long-range inter-TAD regulatory elements (HIREs). a Genome browser view of significant promoter capture Hi-C (PCHi-C) interactions (blue arcs) for *Hoxa2*, chromatin accessibility (ATAC-seq, purple), and ChIP-seq profile for H3K27ac (blue) in the mandibular process (Md) and second pharyngeal arch (PA2) cranial neural crest cells (CNCCs) at E10.5, and in E12.5 and E14.5 pinna CNCCs in a 2 Mb region of chromosome 6 (50,502,806–52,500,000 bp). PCHi-C interactions with a CHiCAGO score ≥5 are visualized. The color intensity of an arc indicates the CHiCAGO score, with a maximum value of 20 (i.e., interactions having a score ≥20 are shown in dark blue). Only interactions with *Hoxa2* are shown. Bottom, ChIP-seq binding sites for Hoxa2[57], Pbx, and Meis[58] in PA2 at E11.5. **b** Hi-C interaction heat-maps at 25 kb resolution in a 2.2 Mb region of chromosome 6, including the *Hoxa* cluster (50,442,417–52,636,150 bp) in Md and PA2 CNCCs at E10.5, and in pinna CNCCs at E12.5 and E14.5. TAD separation scores are called with HiCExplorer's hicFindTADs (shown as blue lines). Additional gray lines show the TAD scores for different window sizes. The blue highlight marks the domain encompassing the *Hoxa* cluster and *Evx1* (chr6:52,145,433–52,327,518). Arrows indicate the location of a TAD boundary between *Hoxa2* and HIRE1/HIRE2. Arrowheads highlight inter-TAD interactions between *Hoxa2* and HIRE1/HIRE2. **c** Virtual 4C profiles derived from PCHi-C of E10.5 Md (black) and PA2 (green) CNCCs and pinna CNCCs at E12.5 (blue) and E14.5 (red) on the *Hoxa2* promoter bait. The same chromosomal region is displayed for panels **a**, **c**. For better readability, genomic coordinates are only displayed below panel **c** and within the *Hoxa* cluster only show *Hoxa2* genomic position. **d** Zoom in on virtual 4C profiles of panel **c** at HIRE1 and HIRE2 (chr6:50779503 – 51183201). Bottom, ChIP-seq binding sites for Hoxa2[57], Pbx, and Meis[58] in PA2 at E11.5. In **a–d**, HIRE1 and HIRE2 are highlighted by yellow boxes. The *Hoxa2* locus is highlighted in yellow (**a**) or green (**b**). Black boxes at the top of panels **a** and **b** show the *Hoxa2* super-enhancer 1–5 (SE1–5) in PA2 CNCCs at E10.5.

appeared normal. In contrast, E18.5 *HIRE1^{del/del}* homozygous mutant (*n* = 4) and *HIRE1^{del/wt}*;*Hoxa2^{EGFP/wt}* trans-heterozygous mutant (*n* = 4) fetuses survived up to birth but died perinatally. All *HIRE1^{del/del}* and *HIRE1^{del/wt}*;*Hoxa2^{EGFP/wt}* mutants lacked the pinna, similarly to the full *Hoxa2^{−/−}* mutant phenotype[26] (Fig. 5h, i). Moreover, all *HIRE1^{del/del}* (*n* = 4/4, 8 sides) and *HIRE1^{del/wt}*;*Hoxa2^{EGFP/wt}* (*n* = 4/4, 8 sides) mutants

displayed skeletal malformations phenocopying the full *Hoxa2^{−/−}* phenotype[26]. Namely, the PA2 CNCC-derived stapes, styloid process and lesser horns of the hyoid bone were absent and replaced by a mirror image homeotic duplication of PA1-like structures, including a duplicated incus, malleus, tympanic bone, and a partially duplicated Meckel's cartilage (Fig. 6j–o). Like *Hoxa2^{EGFP/EGFP}* mutants[49] (*n* = 6/8
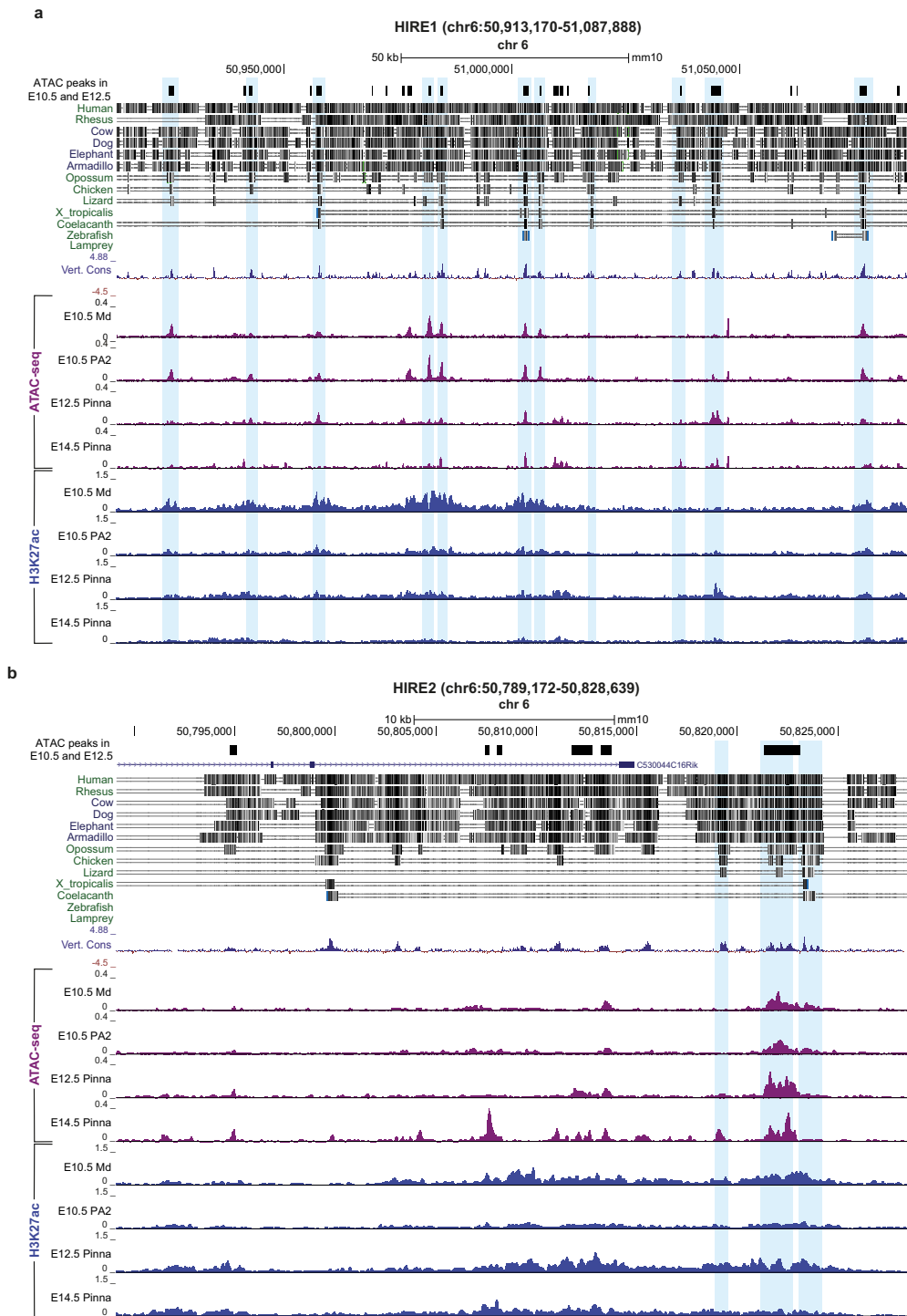
Fig. 4 | Conservation of HIRE1 and HIRE2 in Eutheria/Placentalia. Illustration of sequence conservation at HIRE1 (**a**) and HIRE2 (**b**). Top, combined ATAC-seq peaks called for E10.5 second pharyngeal arch (PA2) cranial neural crest cells (CNCCs) and E12.5 pinna CNCCs. The related sequence alignments of the indicated vertebrate species to the mouse genomic DNA sequence by Multiz program[134] are shown, as well as the 60 vertebrates Basewise Conservation (Vert. Cons) by PhyloP (blue/red), both extracted from UCSC genome browser. Corresponding chromatin accessibility (purple) and H3K27ac pattern (blue) in mandibular (Md) and PA2 CNCCs at E10.5 and pinna CNCCs at E12.5 and E14.5 are depicted. The blue boxes highlight regions with high levels of conservation between mouse and other vertebrate species at HIRE1 and HIRE2, which also show ATAC-seq peak overlap in CNCCs at different developmental stages.
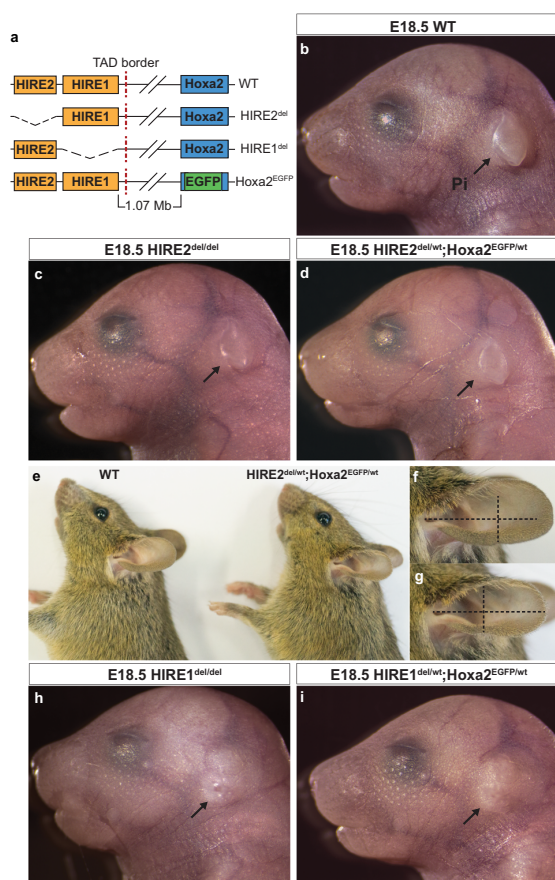
Fig. 5 | Effect of HIRE1 and HIRE2 deletions on pinna morphogenesis. a CRISPR/
Cas9 mediated deletion of HIRE1 and HIRE2 in vivo. In the *Hoxa2^EGFP* knockout allele,
Hoxa2 is replaced by *EGFP* knock-in[49]. **b–d** E18.5 external ear phenotype in wild-type
(WT) (**b**, representative of *n* = 4/4 fetuses), *HIRE2^del/del* homozygous mutant
(**c**, representative of *n* = 6/6 fetuses) and *HIRE2^del/wt*;*Hoxa2^EGFP/wt* trans-heterozygote
mutant (**d**, representative of *n* = 4/4 fetuses). **e** External ear phenotype in WT and
HIRE2^del/wt;*Hoxa2^EGFP/wt* adult mice. **f, g** Enlarged views of the pinna of WT
(**f**, representative of *n* = 4/4 animals) and *HIRE2^del/wt*;*Hoxa2^EGFP/wt* (**g**, representative of
n = 5/5 animals) adult mice in **e**. **h, i** E18.5 external ear phenotype in *HIRE1^del/del*
homozygous mutant (**h**, representative of *n* = 4/4 fetuses) and *HIRE1^del/wt*;*Hoxa2^EGFP/wt*
trans-heterozygous mutant (**i**, representative of *n* = 4/4 fetuses). Arrows show the
pinna (Pi), which has no visible abnormalities in **c** and **d** as compared to **b** but is
absent in **h** and **i** (phenocopying the *Hoxa2^−/−* phenotype[26]). The vertical and hor-
izontal dashed lines of equal length in panels **f** and **g** highlight the shape and size
differences of the external ear between WT and *HIRE2^del/wt*;*Hoxa2^EGFP/wt* adult mice.

fetuses), both *HIRE1^del/del* and *HIRE1^del/wt*;*Hoxa2^EGFP/wt* mutants also dis-
played a cleft secondary palate (*n* = 8/10 and 14/24 fetuses, respec-
tively) (Supplementary Fig. 7). Furthermore, additional typical features
of *Hoxa2^−/−* mutant fetuses could be observed in both *HIRE1^del/del* and
HIRE1^del/wt;*Hoxa2^EGFP/wt* mutants, such as partial duplications of the
pterygoid and squamosal bones, and bifurcation of the retrotympanic
process of the orthotopic squamosal bone (Supplementary Fig. 6g–j),
strongly suggesting that HIRE1 contributes to most of *Hoxa2* expres-
sion in PA2 CNCCs (see below).

We nonetheless observed variability in the extent of morpholo-
gical transformation of ectopic elements, as compared to *Hoxa2^−/−*
phenotype, namely in the shape of duplicated malleus and Meckel's
cartilage (Fig. 6k, n and Supplementary Fig. 8a–c; *n* = 8/8 sides for
HIRE1^del/del and *n* = 2/8 sides for *HIRE1^del/wt*;*Hoxa2^EGFP/wt* mutants). In
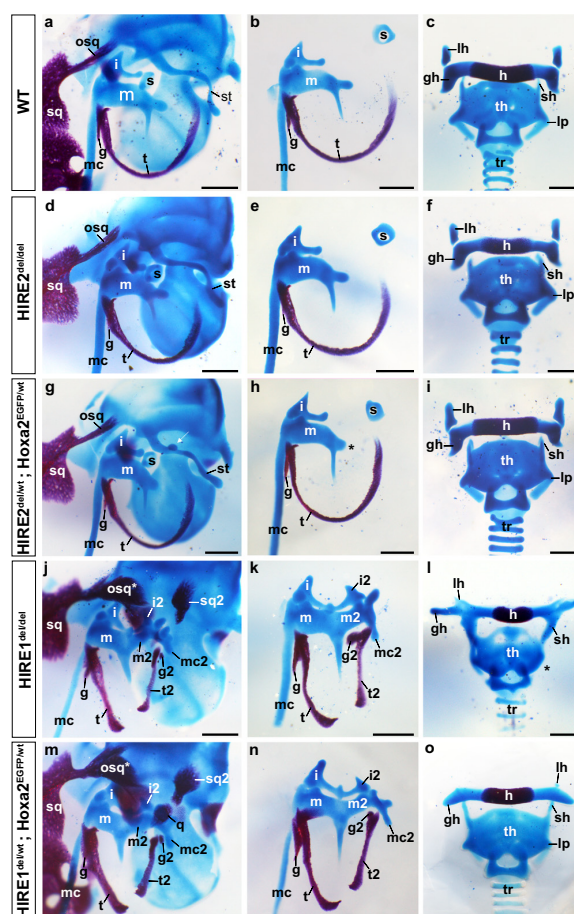some cases, there was no fusion between the incus and its duplicated

**Fig. 6 | Middle ear and hyoid skeletal changes in HIRE1 and HIRE2 mutant
fetuses.** Middle ear (**a, b, d, e, g, h, j, k, m, n**) and hyoid (**c, f, i, l, o**) skeletal
preparations from E18.5 wild-type (WT) (**a–c**), *HIRE2^del/del* homozygous (**d–f**),
HIRE2^del/wt;*Hoxa2^EGFP/wt* trans-heterozygous (**g-**), *HIRE1^del/del* homozygous (**j–l**), and
HIRE1^del/wt;*Hoxa2^EGFP/wt* trans-heterozygous (**m–o**) mutant fetuses. *HIRE2^del/del* skeletal
structures appear normal (**d–f**). **a–n** are representative images of WT (*n* = 8/8) (**a, b**),
HIRE2^del/del (*n* = 12/12) (**d, e**), *HIRE2^del/wt*;*Hoxa2^EGFP/wt* (*n* = 8/8) (**g, h**), *HIRE1^del/del* (*n* = 8/8)
(**j, k**) and *HIRE1^del/wt*;*Hoxa2^EGFP/wt* (*n* = 8/8) fetus sides. **c, f, i, l, o** are repre-
sentative images of WT (*n* = 4/4) (**c**), *HIRE2^del/del* (*n* = 6/6) (**f**), *HIRE2^del/wt*;*Hoxa2^EGFP/wt*
(*n* = 4/4) (**i**), *HIRE1^del/del* (*n* = 4/4) (**l**), and *HIRE1^del/wt*;*Hoxa2^EGFP/wt* (*n* = 4/4) (**o**) fetuses.
HIRE2^del/wt;*Hoxa2^EGFP/wt* mutant fetuses have a smaller processus brevis (*, **h**) and can
display a cartilage nodule on the styloid process (st) (white arrow, **g**, *n* = 5/8 fetus
sides) compared to WT (**a, b**). In *HIRE1^del/del* (**j, k**) and *HIRE1^del/wt*;*Hoxa2^EGFP/wt* (**m, n**)
middle ear homeotic duplications phenocopying those of conventional *Hoxa2^−/−*
mutant[26] are observed. In *HIRE1^del/wt*;*Hoxa2^EGFP/wt*, the lesser horns (lh) of the hyoid
bone (h) are absent and greater horns (gh) display an abnormal location, similar to
Hoxa2^−/− fetuses[26] (**o**). Superior horns (sh) of thyroid cartilage (th) are elongated,
and lateral processes (lp) of laryngeal cartilage are reduced (**o**). In *HIRE1^del/del*
mutants, lh are absent, gh extends dorsally, sh fuse with gh (**l**) and lp are absent (*, **l**).
g and g2 represent WT and duplicated gonial bones, respectively; i and i2, WT and
duplicated incus, respectively; m and m2, WT and duplicated mallei, respectively;
mc and mc2, WT and partially duplicated Meckel's cartilages, respectively; osq and
osq*, normal and modified retrotympanic (otic) process of squamosal bone; s,
stapes; sq, squamosal bone; sq2, ectopic squamosal bone; t and t2, WT and dupli-
cated tympanic bones, respectively; tr trachea. Scale bars represent 500 μm.

counterpart (Supplementary Fig. 8a, b; *n* = 1/8 sides for *HIRE1^del/del* and
n = 1/8 sides for *HIRE1^del/wt*;*Hoxa2^EGFP/wt*) and/or in most cases the gonial
bone was mirror-image duplicated (Fig. 6j, k, m, n and Supplementary
Fig. 8b, c; *n* = 8/8 sides for *HIRE1^del/del* and *n* = 6/8 sides for

HIRE1^{del/wt};*Hoxa2*^{EGFP/wt}) whereas in full *Hoxa2*^{−/−} mutants the duplicated gonial bone was fused to its normal counterpart into a single element[26] (Supplementary Fig. 8a). Moreover, the lesser horns of the hyoid bone were sometimes highly reduced, instead of absent (Fig. 6o and Supplementary Fig. 8d). Altogether, these mild phenotype variations could be explained by differences from the original genetic background and/or the presence of residual *Hoxa2* transcripts (see below).

A notable difference between *HIRE1*^{del/del} and *HIRE1*^{del/wt};*Hoxa2*^{EGFP/wt} mutants concerned the ectopic atavistic palatoquadrate skeletal structure, described in *Hoxa2*^{−/−} mutants[26]. This structure was present with full penetrance in *HIRE1*^{del/wt};*Hoxa2*^{EGFP/wt} trans-heterozygous, though never in *HIRE1*^{del/del} mutants (Supplementary Fig. 6g–j). Moreover, in E18.5 *HIRE1*^{del/wt};*Hoxa2*^{EGFP/wt} trans-heterozygous mutants, the PA3-derived greater horns of the hyoid bone made an abnormal angle with the body of the hyoid bone, similar to the *Hoxa2*^{−/−} phenotype[26] (Fig. 6o and Supplementary Fig. 8d, g), the superior horns of the thyroid cartilage were elongated, and the lateral processes of the laryngeal cartilage were reduced (Fig. 6o and Supplementary Fig. 8d). In contrast, in all *HIRE1*^{del/del} homozygous mutants a fusion between the superior horns of the thyroid cartilage and the greater horns of the hyoid bone could be observed (Fig. 6l and Supplementary Fig. 8e), a phenotype reminiscent of *Hoxa3*^{−/−} knockout mutants[50,51]. In addition, the lateral process of the thyroid cartilage was absent (asterisk, Fig. 6l and Supplementary Fig. 8e) and the greater horns of the hyoid bone extended dorsally and fused with the dorsal part of the thyroid cartilage (*n* = 5/8) (Supplementary Fig. 8h, i). These malformations are reminiscent, while not identical, of those observed upon *Hoxa* cluster deletion in CNCCs[22]. As *Hoxa4* is not expressed at significant levels in PA3 and PA4[22,52], these malformations could therefore result from simultaneous downregulations of both *Hoxa2* and *Hoxa3* expression (see below). Hence, in addition to regulating *Hoxa2* expression in PA2 CNCCs, HIRE1 might be required to regulate *Hoxa2* and *Hoxa3* expression in PA3 and PA4 CNCCs. This is further emphasized by the finding that the observed malformations of PA3 and PA4-derived structures are stronger in *HIRE1*^{del/del} than in *HIRE1*^{del/wt};*Hoxa2*^{EGFP/wt} mutants, where the removal of only one allele of HIRE1 would result in a less severe reduction of *Hoxa3* expression.

The skeletal abnormalities in *HIRE1*^{del/del} and *HIRE1*^{del/wt};*Hoxa2*^{EGFP/wt} mutants were restricted to CNCC-derived structures. The axial and limb skeletons were normal. Moreover, while the otic capsule lacked the oval window as in full *Hoxa2*^{−/−} mutants[26], there was no otic skeletal phenotype reminiscent of *Hoxa1*^{−/−} mutants[53,54], suggesting that *Hoxa1* expression is not regulated by HIRE1.

HIRE1 is required for *Hoxa2* and *Hoxa3* expression in CNCCs

To assess the decrease in *Hoxa2* expression following the inactivation of HIRE1 or HIRE2, we performed a qRT-PCR analysis (Methods) in PA2 of E10.5 *HIRE2*^{del/del}, *HIRE2*^{del/wt};*Hoxa2*^{EGFP/wt}, *HIRE1*^{del/del} and *HIRE1*^{del/wt};*Hoxa2*^{EGFP/wt} mutant embryos (Supplementary Fig. 9). We collected PA2 of E10.5 WT and *Hoxa2*^{EGFP/wt} embryos as controls (Supplementary Fig. 9). In *Hoxa2*^{EGFP/wt} embryos, which do not display skeletal malformations, we detected about 50% of the normal *Hoxa2* transcript levels, as compared to WT (Supplementary Fig. 9). *HIRE2*^{del/del} and *HIRE2*^{del/wt};*Hoxa2*^{EGFP/wt} mutant embryos displayed 75% and 46% of the normal *Hoxa2* transcript levels, as compared to WT, respectively. This correlates with the absence of visible malformations of PA2-derived structures in E18.5 *HIRE2*^{del/del} mutant fetuses, and the presence of mild malformations in E18.5 *HIRE2*^{del/wt};*Hoxa2*^{EGFP/wt} mutant fetuses (Figs. 5, 6 and Supplementary Figs. 9, 10). By contrast, in *HIRE1*^{del/del} and *HIRE1*^{del/wt};*Hoxa2*^{EGFP/wt} mutant embryos, we found a drastic reduction of *Hoxa2* expression, with about 0–3% and 0–4% of *Hoxa2* transcripts, respectively, as compared to WT, thus explaining the finding that HIRE1 deletion phenocopied the full homeotic *Hoxa2* knockout phenotype in PA2 CNCCs (Figs. 5, 6 and Supplementary Figs. 9, 10).

Next, we carried out whole-mount and tissue section in situ hybridization for *Hoxa2* and *Hoxa3* in *HIRE1*^{del/del} embryos. At E10.5, *Hoxa2* expression was undetectable in PA2, in keeping with the qRT-PCR data (Supplementary Fig. 9), and highly reduced in more posterior pharyngeal arches (Fig. 7a, c, d, f and Supplementary Fig. 11a, b), as well as in the anterior hindbrain R2 to R5, as compared to WT (Fig. 7b, e). *Hoxa2* expression was, however, maintained normally in somites (Fig. 7a, d). *Hoxa3* expression was highly reduced in PA3 and more posterior pharyngeal arches (Fig. 7g, i, j, l and Supplementary Fig. 11c, d) but only slightly affected in the hindbrain of E10.5 *HIRE1*^{del/del} embryos (Fig. 7h, k) and maintained normally in somites (Fig. 7g, j). We did not detect changes in the expression patterns of *Hoxa1*, which at E8.5 is expressed in the hindbrain (Supplementary Figs. 11e–k, 8f–l).

Next, to quantify the transcriptional changes induced by HIRE1 deletion on both *Hoxa2* and *Hoxa3* and on their potential downstream targets in the pharyngeal region, we dissected PA2 and PA3 in E10.5 WT and *HIRE1*^{del/del} embryos and performed RNA-seq. In keeping with the in situ hybridization results, *Hoxa2* and *Hoxa2/Hoxa3* transcript levels were significantly reduced in PA2 and PA3, respectively, of *HIRE1*^{del/del} embryos as compared to WT (Fig. 8 and Supplementary Data 3). *Hoxa2* downregulation was, however, more severe in PA2 than in PA3 (Fig. 8a, c and Supplementary Data 3), while in PA3, *Hoxa2* relative transcript levels decreased less than *Hoxa3*, as compared to WT (Fig. 8a, c, e and Supplementary Data 3). Thus, the effect of HIRE1 on *Hoxa* gene regulation correlates with the gene linear position in the cluster. Secondly, confirming in situ hybridization experiments, our data indicate a stronger effect of HIRE1 on *Hoxa2* transcriptional regulation in its anterior-most domain of expression, both in the pharyngeal region and hindbrain.

Furthermore, in PA2 of E10.5 *HIRE1*^{del/del} mutants, 16 genes were upregulated and 41 downregulated, excluding *Hoxa2* itself (FDR <0.05 and log2 CPM ≥1), as compared to WT (Fig. 8a, b and Supplementary Data 3). In PA3, 8 genes were upregulated and 21 downregulated, excluding *Hoxa2* and *Hoxa3* (FDR <0.05 and log2 CPM ≥1) (Fig. 8c, d and Supplementary Data 3). Among the upregulated genes, we confirmed several genes known to be negatively regulated by Hoxa2, including *Gbx2*, *Pitx1*, *Alx4*, *Lhx6*, *Barx1*, and *Rspo2*[27,55–57] (Fig. 8a, b and Supplementary Data 3). Among the downregulated genes, we identified several known Hoxa2 targets, such as *Meis1*, *Meis2*, *Meox1*, *Fzd4*, *Zfp703*, and *Zfp503*[56–60] (Fig. 8a, b and Supplementary Data 3). These data confirm the specific role of HIRE1 for *Hoxa2* and *Hoxa3* transcriptional regulation in PA2 and PA3.

We next investigated if HIRE1 could also be required at the time of emergence and early stages of CNCC migration. In the mouse, *Hoxa2* expression in PA2 CNCCs is detected at an early migratory stage around E8.25-E8.5[23,24,61]. Notably, in *HIRE1*^{del/del} embryos, *Hoxa2* transcripts were already undetectable by in situ hybridization at E8.5 in the CNCCs arising from R4 and migrating into PA2 (Fig. 7m, q); no signal was detected at E9.0 and E9.5 as well (Fig. 7n–p, r–t), similarly to E10.5 *HIRE1*^{del/del} embryos (Fig. 7a–f). Moreover, we observed a strong *Hoxa2* downregulation in the anterior hindbrain of *HIRE1*^{del/del} embryos already at E8.5, from R2 to R5 (Fig. 7m, q), and even more so at E9.0 and E9.5 (Fig. 7n–p, r–t).

Altogether, our findings strongly suggest that HIRE1 is required for induction of high *Hoxa2* expression levels already at the emergence and earliest stages of CNCC migration and may also be involved in maintaining appropriate transcript levels through later developmental stages, in both hindbrain and CNCCs.

Identification of transcription factor binding motifs in HIRE1/HIRE2 and involvement of *Hoxa2* in its own long-range regulation

To investigate which transcription factors may be involved in binding HIRE1 and HIRE2 and potentially regulate long-range interactions with
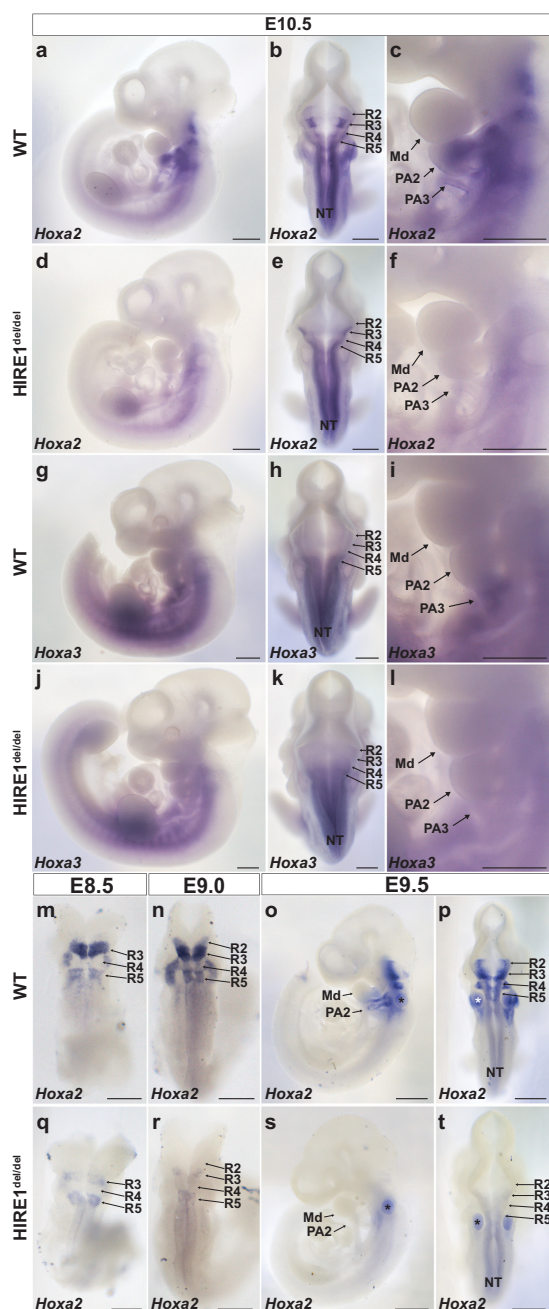
Fig. 7 | HIRE1 is required for *Hoxa2* and *Hoxa3* expression. Whole-mount in situ hybridization on E10.5 wild-type (WT) (**a–c, g–i**) and *HIRE1*$^{del/del}$ (**d–f, j–l**) embryos using *Hoxa2* (**a–f**) and *Hoxa3* (**g–l**) antisense probes (**a, c, d, f, g, i, j, l, o, s**) lateral and (**b, e, h, k, m, n, p, q, r, t**) dorsal views. **a–c** are representative images of E10.5 WT embryos (n = 6), **d–f** are representative images of E10.5 *HIRE1*$^{del/del}$ embryos (n = 6), **g–i** are representative images of E10.5 WT embryos (n = 3), and **j–l** are representative images of E10.5 *HIRE1*$^{del/del}$ embryos (n = 3). *Hoxa2* expression is not detectable in second (PA2) and third (PA3) pharyngeal arches of *HIRE1*$^{del/del}$ embryos (**d, f**) and is reduced in hindbrain rhombomeres 2–5 (R2–R5) (**e**) compared to WT (**a–c**). *Hoxa3* expression is strongly reduced in PA3 of *HIRE1*$^{del/del}$ embryos (**j, l**) and not affected in the hindbrain (**k**) compared to WT (**g–i**). Whole-mount in situ hybridization on WT (**m–p**) and *HIRE1*$^{del/del}$ (**q–t**) E8.5 (**m, q**), E9.0 (**n, r**), and E9.5 (**o, p, s, t**) embryos with *Hoxa2* antisense probe. **m–t** are representative images of n = 3 embryos for each stage and genotype. In E8.5 and E9.0 *HIRE1*$^{del/del}$ mutant embryos, *Hoxa2* expression is undetectable in R2 and R4 and severely reduced in R3 and R5 (**q, r**) compared to WT (**m, n**). At E9.5, *Hoxa2* expression is not detectable in PA2 of *HIRE1*$^{del/del}$ embryos (**s**) and is strongly reduced in hindbrain R2-5 (**t**) compared to WT (**o, p**). *, trapped dye in the otic capsule. Md, mandibular process. In **a–i, o, p, s, t** scale bars represent 500 μm. In **m, n, q, r** scale bars represent 250 μm.

Next, assuming that the observed accessibility changes are driven by differential transcription factor binding and that other genomic regions bound by the same transcription factors would show similar accessibility profiles, we ranked all ATAC peaks in E10.5 PA2 and E12.5 pinna CNCCs according to their similarity (Pearson's correlation coefficient) to the average accessibility profile of each of the three clusters (Supplementary Fig. 12c). We selected the 1000 peaks with the highest correlation to each cluster profile, resulting in three non-overlapping peak sets (Supplementary Fig. 12d). We then used these three sets of peaks and, as a control, a fourth set containing all residual ATAC peaks, and ran a motif enrichment analysis[62] resulting in a total of 382 significantly enriched motifs (Supplementary Fig. 12e and Supplementary Data 4). From these, we hierarchically clustered similar motifs and selected nine representative motifs with predicted transcription factor binding sites in the 31 ATAC-seq peaks overlapping with the HIRE2/HIRE1 region (Fig. 9a, Supplementary Data 5, 6, and Supplementary Figs. 12g, 13) (Methods).

For example, among the representative motifs, there were binding sites for Tal- and NFAT-related factors, potentially bound by Twist1, Twist2, and ZBTB18 and Nfatc1, Nfat5, and Nfatc3, respectively (Fig. 9a and Supplementary Fig. 12g). Each of these transcription factors is upregulated from E10.5 to E12.5 in CNCCs (Supplementary Data 3). Furthermore, we identified motifs for Hox-related factors, such as Hoxa2, and TALE-type homeodomain factors, such as Pbx and Meis. Both Meis and Pbx transcription factors are known Hox cofactors and form heterodimers with Hox proteins to bind to DNA[63,64]. Analysis of published ChIP-seq datasets for Hoxa2, Pbx, and Meis in PA2 at E11.5[57,58], revealed that these factors indeed showed enriched binding at HIRE1 and HIRE2 (Fig. 3a and Supplementary Fig. 4).

These latter findings indicated that Hoxa2 may play a role in long-range inter-TAD interactions with its own promoter. Notably, PCHi-C on E10.5 PA2 CNCCs in *Hoxa2*$^{EGFP/EGFP}$ knockout embryos[49] revealed that the strength of the interactions of *Hoxa2* with HIRE2 in E10.5 *Hoxa2*$^{EGFP/EGFP}$ knockout embryos was reduced by 27% (indeed, note that with CHiCAGO score ≥5, no arc is visible between *Hoxa2* and HIRE2 in *Hoxa2*$^{EGFP/EGFP}$ vs. WT; Fig. 9d and Supplementary Fig. 3). These data strongly suggest that Hoxa2 is partially involved in its own long-range transcriptional regulation likely with cofactors such as Pbx and Meis.

Discussion

Large clusters of enhancers in close genomic proximity and collectively bound by arrays of transcription factors, termed SEs, have been involved in the transcriptional control of cell identity during differentiation and disease[43,44,65–67]. Less is known about the potential

Hoxa2, we performed a motif enrichment analysis using the ATAC-seq data from PA2 and pinna CNCCs at E10.5 and E12.5, respectively. We first called the peaks from the datasets of both stages, merged them, and extracted putative enhancer peaks to obtain a total of 106,587 peaks. Thirty-one ATAC-seq peaks overlapped the region spanning from HIRE2 to HIRE1 (mouse GRCm38/mm10 chr6:50,789,172–51,087,888) (Supplementary Fig. 12a). We then clustered these 31 peaks according to their relative accessibility at E10.5 and E12.5 (cluster 1–3) (Supplementary Fig. 12b). Cluster 1 peaks were more accessible at E12.5, cluster 2 peaks were more accessible at E10.5, whereas cluster 3 peaks only showed minor accessibility differences between E10.5 and E12.5.
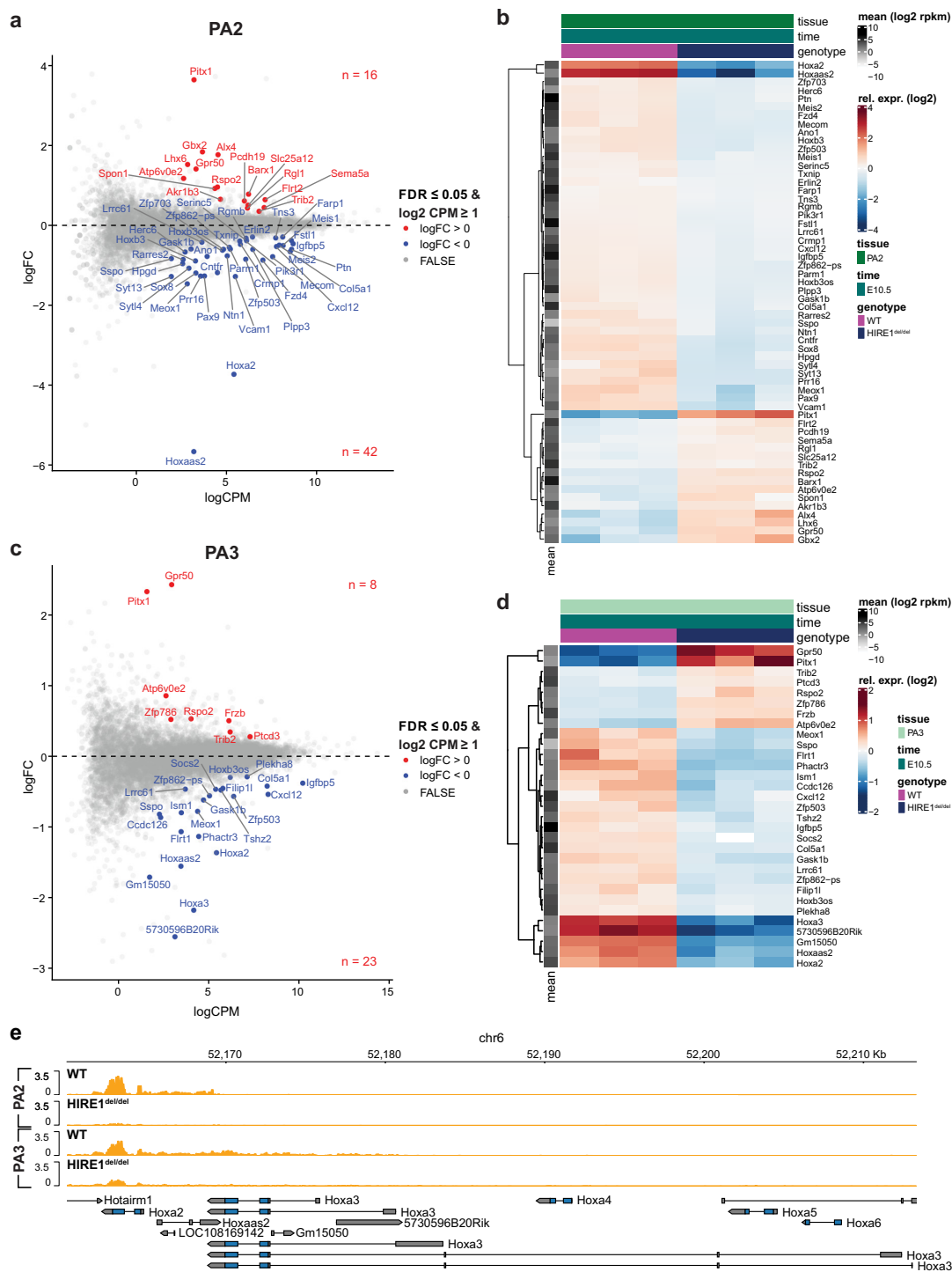
Fig. 8 | Genome-wide transcriptional effects of HIRE1 deletion. MA plots (**a**, **c**) showing the log2 fold-change (FC) versus the average log2 count per million (CPM) for the differential gene expression between wild-type and E10.5 *HIRE1^{del/del}* embryos, in pharyngeal arch 2 (PA2) (**a**) and 3 (PA3) (**c**). Each dot represents a single gene. Differentially expressed genes with FDR ≤0.05 and log2 CPM ≥1 are shown as red and blue dots. Positive log2 FC values correspond to genes upregulated (red dots) and negative log2 FC corresponds to genes downregulated (blue dots) in *HIRE1^{del/del}*. Heatmaps (**b**, **d**) of RNA levels relative to the mean over all wild-type and *HIRE1^{del/del}* replicates in PA2 (**b**) and PA3 (**d**) at E10.5. Differentially expressed genes with FDR ≤0.05 and log2 CPM ≥1 are shown. (**e**) Genome browser view at *Hoxa2* and *Hoxa3* loci showing RNA profiles (orange) in PA2 and PA3 of E10.5 wild-type (WT) and *HIRE1^{del/del}* embryos.
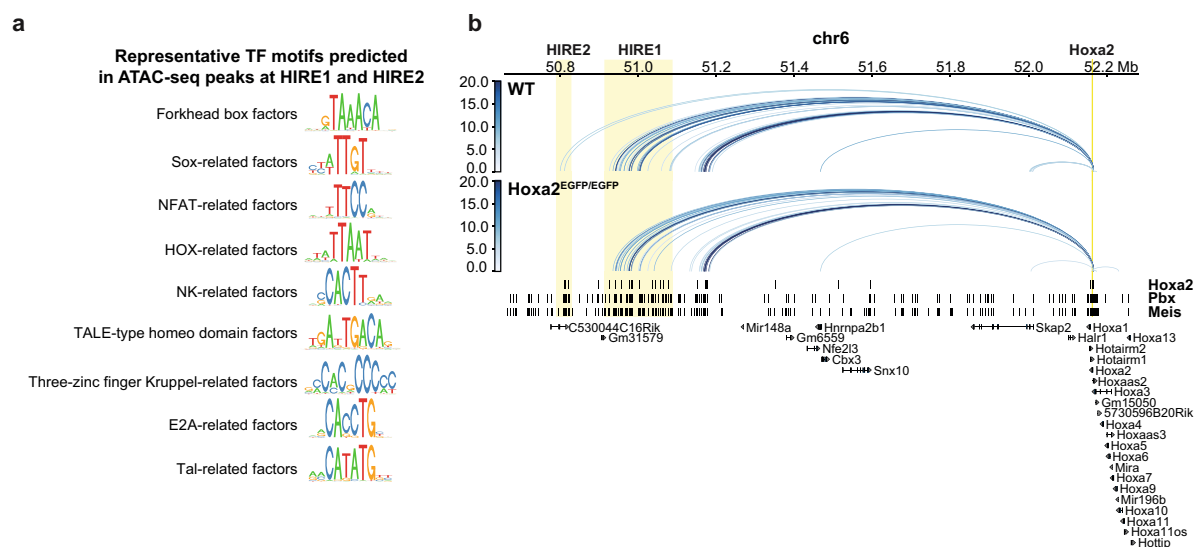
Fig. 9 | Involvement of Hoxa2 in inter-TAD long-range interactions between its promoter and HIRE2. a Representative transcription factor (TF) motifs predicted to be associated with ATAC-seq peaks of E10.5 PA2 and E12.5 pinna cranial neural crest cells (CNCCs) at HIRE1 and HIRE2. Each of the depicted motifs is representative of a cluster of similar motifs (Methods, Supplementary Figs. 12g, 13). **b** Significant promoter capture Hi-C interactions for *Hoxa2*, shown as blue arcs, in second pharyngeal arch (PA2) CNCCs of wild-type (WT) and *Hoxa2^EGFP/EGFP* embryos at E10.5, in a 1.7 Mb region of chromosome 6 (50,656,605–52,283,516). As in Fig. 3, only interactions with a CHiCAGO score ≥5 are visualized. The color intensity of an arc indicates the CHiCAGO score, with a maximum value of 20 (i.e., interactions having a score ≥20 are shown in dark blue). The visualized interactions are filtered to show only *Hoxa2* interactions. Significant interactions from *Hoxa2* to HIRE2 are lost in *Hoxa2^EGFP/EGFP*. Bottom, binding sites, identified by ChIP-seq, for Hoxa2[57], Pbx, and Meis[58] in PA2 at E11.5. HIRE1, HIRE2, and *Hoxa2* locus are highlighted by yellow boxes.

involvement of SEs in establishing the positional identities of specific cell populations during embryo morphogenesis. During craniofacial development, distinct migratory CNCC subpopulations contributing to the skeletal structures building the face acquire their positional identities and patterning information by the differential expression of key transcription factors induced by local signals and drive specific transcriptional programs in the different facial and pharyngeal prominences[15]. Here, we computationally identified 2232 putative CNCC subpopulation-specific SEs. Using PCHi-C, we found that 147 out of 2232 SEs selectively targeted 62 positional transcription factor-coding genes previously shown to be differentially expressed between FNP, Mx, Md, and PA2 CNCC subpopulations during craniofacial development[16]. Most interestingly, we identified very long-range (>1 Mb) inter-TAD SEs (SE1-SE5) that interact with *Hoxa2* in a CNCC subpopulation-specific manner. Based on their proximity, we grouped tandem SE2–4 into HIRE1, a 175 kb geneless genomic region, while SE5 encompassed a 39 kb region, termed HIRE2. HIRE1 and HIRE2 are highly conserved in mammals, including humans. Single (HIRE2, SE5) or multiple (HIRE1, SE2-SE4) functional deletions of inter-TAD SE elements in the mouse resulted in skeletal phenotypes with distinct severity and penetrance, partially or fully phenocopying *Hoxa2* knockout (Supplementary Fig. 10) as well as aspects of the *Hoxa3* mutant skeletal phenotypes. We further show that inter-TAD SEs selectively control the expression of *Hoxa2* and *Hoxa3* in neural derivatives, such as CNCCs and hindbrain rhombomeres, though not in mesoderm-derived tissues.

The putative craniofacial SEs, including *Hoxa2* SE1-SE5, were identified using the ROSE algorithm[44,46] (Methods). Similar to the original SE definition[44], we operationally used a maximum distance of 12.5 kb between individual active enhancers to consider them as part of a SE. This threshold is, however, arbitrary. Thus, even though SE2-SE4/HIRE1 and SE5/HIRE2 are separated by 84.4 kb without strong enrichment in H3K27ac, they might nevertheless be considered to belong to a single, extremely large (spanning 300–400 kb), distant SE regulatory

region targeting *Hoxa2* (and *Hoxa3*) in neural derivatives during development.

During facial morphogenesis, the transcriptional output of key genes needs to be tightly regulated, as certain structures may be sensitive even to small perturbations of gene dosage resulting in patterning abnormalities and disease (e.g., lower jaw[7]). *Hoxa2* provides positional identities and patterning information to all PA2 CNCCs derivatives and its function is highly conserved in vertebrate CNCCs[25,26,68–71]. In the mouse, *Hoxa2* is mainly required for morphogenesis of external (pinna) and middle ear structures[25,26,28,72]. Spatiotemporal control of *Hoxa2* expression levels is important to pattern distinct CNCC-derived skeletal elements[27,73,74], but how this is regulated at the transcriptional level is poorly understood. Enhancer clusters or SEs acting on a common target gene could provide a suitable regulatory landscape to control and fine-tune transcriptional output during the morphogenesis of facial elements. Previous work has shown that, in the presence of multiple enhancer clusters or SEs, individual enhancer constituents within clusters, or even individual enhancer clusters, may have overall weak activity on the target promoter and act in a partially redundant manner, when deleted; however, simultaneous deletion of multiple (clusters of) enhancers may result in synergistic combinatorial effects on target gene transcription levels[7,75–77]. The finding that *Hoxa2* expression is selectively regulated in the CNCCs of PA2 and derivatives by the activity of multiple clusters of inter-TAD long-range enhancers (SE1-SE5) prompted us to analyze their functional role by CRISPR-Cas9 mediated deletion.

Due to its partial overlapping with the TAD boundary, functional analysis of SE1 was not further pursued to avoid the potentially confounding effects of its deletion. HIRE2 homozygous inactivation did not result in a detectable *Hoxa2* mutant phenotype, indicating that its deletion could be fully compensated and redundant with other enhancers and/or identified SEs. In humans, HOXA2 haploinsufficiency causes microtic ears and hearing impairment[29–31], indicating that external ear morphogenesis is sensitive to HOXA2 dosage reduction

already at 50% of its normal levels. Remarkably, when put on a sensitized *Hoxa2* haploinsufficient genetic background, the deletion of HIRE2 was also associated with smaller, microtic, external ears in adult mice (Fig. 5e–g). Accordingly, HIRE2 is more enriched with H3K27ac and interacts more strongly with *Hoxa2* in pinna CNCCs at E12.5 (i.e., at the beginning of its formation) than in E10.5 PA2 (Fig. 3 and Supplementary Fig. 4). Thus, while HIRE2 homozygous deletion causes a modest reduction of *Hoxa2* expression (Supplementary Figs. 9, 10), which is nonetheless still compatible with full *Hoxa2* function, the HIRE2/SE5 contribution to *Hoxa2* transcription and function becomes critical below a 50% expression threshold, affecting the morphogenesis of dosage-sensitive structures such as the pinna (Supplementary Figs. 9, 10).

Strikingly, HIRE1 deletion phenocopied the full homeotic *Hoxa2* knockout phenotype, including the lack of pinna, and, in addition, resulted in *Hoxa3* knockout-like phenotypes in PA3-PA4-derived skeletal structures (Fig. 6j–o and Supplementary Figs. 6e–h, 7, 8). In *HIRE1*^{del/del} mutants, we confirmed strong downregulation of both *Hoxa2* and *Hoxa3* in CNCCs and anterior hindbrain (Figs. 7, 8), indicating that these distant SEs are involved in collinear regulation of *Hoxa2* and *Hoxa3* in neural derivatives and supporting the observation that SEs can coordinately regulate clustered genes[75,78]. The much stronger effect of HIRE1, as compared to HIRE2, most likely reflects the fact that HIRE1 comprises multiple SEs, namely SE2, SE3, and SE4, while HIRE2 is only composed of SE5. Even though HIRE1 deletion results in a complete lack of pinna, under dosage-sensitive conditions, HIRE2 does contribute to pinna morphogenesis as well (see above). This strongly indicates partial redundancy among SEs to achieve robustness against genetic or environmental perturbations, as well as cooperative or synergistic contributions to boost *Hoxa2* transcription and ensure reproducible expression patterns driving harmonious morphogenesis. These hypotheses will need to be tested by additional in vivo deletions and analysis.

Multiple cis-regulatory elements proximal to, or within, the *Hoxa2* locus have been identified, driving reporter genes expression in R2 and somites[79–81], R3/R5[61,82–84], R4[79,85,86], as well as in CNCCs[61]. However, the in vivo functional role of these proximal enhancers and their contribution to endogenous *Hoxa2* expression and patterning activity was not investigated. One of the major findings of this study is that most of the *Hoxa2* transcriptional output in neural crest and rhombomeres is dependent on multiple long-range, inter-TAD, interactions with clustered regulatory elements, HIRE1 and HIRE2, spanning a large genomic region of more than 1 Mb away from *Hoxa2*. To date, only a handful of very long-range regulatory sequences at more than 1 Mb genomic distance from their promoters have been identified, and they are all intra-TAD-located. Namely, these include the *Shh* ZRS, *Myc* BENC and MNE, and *SOX9* EC1.45 and EC1.25 enhancers[7,9,87–89]. This is the first time that an inter-TAD (super-)enhancer genomic region was identified and shown to be critical for the regulation of a key vertebrate developmental gene during face morphogenesis.

We generated E10.5, E12.5, and E14.5 CNCC subpopulation-specific Hi-C and PCHi-C data and CTCF ChIP-seq binding profiles and showed that the TAD organization flanking the *Hoxa* cluster is similar in Md and PA2 CNCCs at all stages analyzed and in ES cells (Fig. 3b and Supplementary Fig. 2c, d). Importantly, we identified a strong invariant TAD boundary with identical CTCF binding patterns in Md and PA2 CNCCs, that partitioned *Hoxa2* from HIRE1 and HIRE2 in distinct TADs in both cell populations (Fig. 3b and Supplementary Fig. 2d). This prompted the question of how these SEs can communicate in a subpopulation-specific manner with the *Hoxa2* promoter across this TAD boundary. Even though the functional relationship between TAD topology and gene regulation is debated[90], most of the enhancer-promoter (E-P) interaction pairs reside within the same TAD, suggesting that CTCF-bound TAD borders, while not absolutely required for intra-TAD E-P interactions[91], may provide intra-TAD transcriptional insulation thus

limiting inter-TAD E-P interactions[1,90]. On the other hand, TAD identification depends on Hi-C data resolution and the algorithm used for TAD calling[92]. Based on such computational methods, inter-TAD E-P interactions detected at a cell population level are indeed relatively rare[37,93,94]. However, single-cell approaches revealed greater than expected cell-to-cell heterogeneity and dynamic behavior of TADs, with only about 1.5–2.0-fold enrichment of intra-TAD vs. inter-TAD interactions in single cells[95]. Moreover, recent work[96] showed that enhancer strength, boundary strength, and distance all determine promoter sensitivity to CTCF-mediated transcriptional insulation at TAD boundaries. Thus, these findings suggest that TAD boundaries may not behave as absolute barriers to E-P interaction across them.

We found here that a strong TAD boundary may be overcome in vivo, in a cell population-specific manner, in the presence of multiple highly active enhancer clusters or SEs in tandem located in a different TAD than the promoter, and we further demonstrate that their SE-P interactions have strong transcriptional and in vivo functional impact. Intriguingly, HIRE1 and HIRE2 are active, i.e., enriched with H3K27ac, both in *Hoxa2*-expressing PA2 and *Hox*-free Md CNCCs, yet the contacts between HIRE1 and HIRE2 and *Hoxa2* only occur in PA2 CNCCs (Fig. 3a, b). In *Hox*-free CNCCs, the *Hoxa2* promoter may not be available for HIRE1/HIRE2 interaction since, together with the whole *Hoxa* cluster, it is embedded in a large repressive Polycomb domain[16] which may segregate in a repressive nuclear compartment distinct from the active HIRE1/HIRE2. Indeed, Polycomb binding at *Hoxa* promoters in developing limbs can prevent their interaction with active enhancers[97]. In PA2 CNCCs, local patterning signals may activate proximal enhancers so that the *Hoxa2* locus is "singled out" from the Polycomb repressed *Hoxa* cluster and transcriptionally induced, followed by removal of H3K27me3 and switched to H3K27ac deposition[16]; this might, in turn, allow rapid interaction with the inter-TAD SEs boosting *Hoxa2* expression to full transcriptional output. Similarly, *Hoxa3* might become collinearly connected to HIRE1/HIRE2 in PA3 CNCCs, where *Hoxa3* is transcriptionally induced and *Hoxa2* is expressed as well.

Moreover, tissue-specific 3D chromatin conformation can also contribute to enhancer activity and specificity[6]. For instance, the Pen enhancer shows activity in both developing forelimbs and hindlimbs, but it only controls *Pitx1* transcription in hindlimbs. This restricted enhancer activity is associated with a 3D chromatin configuration allowing Pen and *Pitx1* to interact only in hindlimbs, whereas enhancer and promoter are maintained physically separated in forelimbs[6]. Structural chromatin variants can however convert the inactive into an active 3D conformation, thereby inducing *Pitx1* misexpression in forelimbs[6]. Moreover, forced chromatin looping of strong enhancers to developmentally silenced promoters can be sufficient to stimulate transcription[98,99]. Thus, HIRE1/HIRE2-driven transcriptional regulation of *Hoxa2* may be allowed by a PA2-specific 3D chromatin configuration. Furthermore, HIRE1 and HIRE2 might be brought in proximity to the *Hoxa2* locus by a mechanism similar to the domain-skipping interactions described between Drosophila *Scr* and its distal enhancers T1, whereby the formation of an intervening TAD by boundary pairing is essential for distal, inter-TAD, E-P interaction[100]. Active *Hoxa2* and its SE region HIRE1/HIRE2 are both close to interacting TAD boundaries (Fig. 3b), suggesting that pairing between boundary elements might bring distant HIRE1/HIRE2 and its target promoter in proximity by domain-skipping chromatin folding in PA2 CNCCs.

Lastly, HIRE1 and HIRE2 could cooperate with proximal *Hoxa2* enhancers to allow for strong and precise *Hoxa2* expression in the hindbrain and CNCCs. Notably, motif enrichment analysis, and PCHi-C carried out in E10.5 *Hoxa2* full knockout embryos, showed that Hoxa2 itself is partially required to establish long-range interactions in PA2 CNCCs between its own promoter and HIRE2, likely with Pbx and Meis cofactors (Fig. 9). These data support the idea that the *Hoxa2* promoter must be active to recruit HIRE1/HIRE2 and are consistent with

the observation that interconnected autoregulatory loops often exist between SEs and their target promoters[44].

Methods

Mice and ethical statement

All animal experiments were approved by the Basel Cantonal Veterinary Authorities under permit number 2670 and conducted in accordance with the Guide for Care and Use of Laboratory Animals.

The generation of *Wnt1::Cre, ROSA-tdRFP (Rosa^RFP^)*, and *Hoxa2^EGFP^* were described elsewhere[49,101,102]. The HIRE1 and HIRE2 deletion lines were generated and characterized on the C57Bl/6J background as described in *Mouse genome editing using CRISPR/Cas9* section. All mice were maintained on a mixed background (C57Bl/6J; CD1). The mice were housed in a 12 h light:dark cycle and given ad libitum access to food and water for the duration of the study. The ambient temperature is $22 \pm 2\,°C$ and humidity is maintained at 45–65%.

For breeding, one or two female mice were introduced into a cage with a single male and monitored for timed pregnancies. Noon of the day of the vaginal plug was considered as E0.5. All mice used for breedings were at least 8 weeks old and not older than 6 months. To obtain E10.5, E12.5, and E14.5 *Wnt1::Cre;ROSA^RFP^* embryos (denoted as wild-type), we crossed *Wnt1::Cre* transgenic mice[101] with the *Rosa^RFP^* reporter[102] mouse line. To generate E10.5 *Hoxa2^EGFP^* mutant embryos, *Hoxa2^wt/EGFP^* transgenic mice[49] were crossed. To obtain E8.5, E9.5, E10.5, and E18.5 *HIRE1^del/del^* embryos, *HIRE1^wt/del^* mice were crossed. To obtain E18.5 *HIRE1^del/wt^;Hoxa2^EGFP/wt^* fetuses, *Hoxa2^EGFP/wt^* transgenic mice were crossed with *HIRE1^del/wt^* mice. To obtain E18.5 *HIRE2^del/del^* embryos, *HIRE2^del/wt^* mice were crossed. To obtain E18.5 and adult *HIRE2^del/wt^;Hoxa2^EGFP/wt^* specimen, the *Hoxa2^EGFP/wt^* transgenic mouse line and *HIRE2^del/wt^* mice were crossed.

As our study aims to uncover general molecular regulatory mechanisms during facial morphogenesis, sex was not considered in the study design, and findings apply to both sexes. For experiments, cell populations from embryos of different sexes were pooled to carry out molecular analysis, including sequencing.

All mouse lines are available upon request from the corresponding author.

Mouse genome editing using CRISPR/Cas9

Guide RNAs were designed using CRISPOR[103] to target upstream (Up) and downstream (Dw) the HIRE2 or HIRE1 genomic regions at the following protospacer sequences: HIRE2-Up: AGCACGTAGCACGTC AGTAG; HIRE2-Dw: TGTAGGGTATACTACTAGCC; HIRE1-Up: CACCC AGGAATAGGTGCGTC; HIRE1-Dw: GTGGCCCCCGACTAAACTAT. Deletions of the HIRE2 or HIRE1 genomic regions were performed by using the Alt-R CRISPR-Cas9 system from Integrated DNA Technologies (IDT) and ribonucleoprotein (RNP) complex delivery by mouse zygote electroporation. The RNP complex was composed of two crRNA (Up and Dw the HIRE2 or HIRE1 region), a tracrRNA and the Hifi Cas9 nuclease V3. Following IDT recommendations, a duplex was formed for each guide RNA by mixing an equal volume of crRNA (200 μM) and tracrRNA (200 μM), heated to 95 °C for 5 min, and allowed to cool down at room temperature for 10 min for annealing. The RNP complex was finally prepared by combining and incubating the two crRNA: tracrRNA duplexes (100 μM) and the Hifi Cas9 nuclease V3 (61 μM) in Opti-MEM medium (Gibco), respectively, at the final concentration of 3 μM and 1.2 μM, for 20 min at room temperature. Mouse zygotes (C57Bl/6J) were electroporated with the RNP complex using a NEPA21 electroporator and a CUY501P1-1.5 electrode (Nepa Gene) and re-implanted into foster females. Founder mice harboring the deletion were identified by PCR and sequencing. The mice were genotyped by PCR with the following primers for the HIRE2 deletion (297 bp wildtype and 377 bp deleted fragments): HIRE2-Fw1, 5′-CTTGGTTGGAGGC ATCCTTC-3′; HIRE2-Fw2, 5′-AGGGAGGTTAAAGTATTTAAGTAC-3′; HIRE2-Rv, 5′-CGCAAATTCAGTTCCCAGTAC-3′; and for the HIRE1

deletion (228 bp wildtype and 303 bp deleted fragments): HIRE1-Fw, 5′-ATAGCAGGCACTGAAGCCTG-3′; HIRE1-Rv1, 5′-GTCCTCGGTCCTCATC TCAG-3′; HIRE1-Rv2, 5′-ACCTTAGACAAGGCCTTATCTG-3′.

Culturing conditions

E14 mESCs, provided by L. Giorgetti's laboratory, were cultured on gelatin-coated culture plates in Glasgow minimum essential medium (Sigma-Aldrich, G5154) supplemented with 15% fetal calf serum (Eurobio Abcys), 1% L-glutamine (Thermo Fisher Scientific, 25030024), 1% sodium pyruvate (Thermo Fisher Scientific, 11360039), 1% MEM non-essential amino acids (Thermo Fisher Scientific, 11140035) 100 μM β-mercaptoethanol, 20 U/ml leukemia inhibitory factor (Miltenyi Biotec, premium grade) in 8% CO2 at 37 °C. The cells underwent mycoplasma contamination testing once a month, and no contamination was found.

Isolation of cranial neural crest cells (CNCCs)

To collect WT post-migratory CNCCs, we generated E10.5 *Wnt1::Cre;ROSA^RFP^* embryos and micro-dissected the frontonasal process (FNP), the maxillary (Mx) and mandibular (Md) components of the first pharyngeal arch (PA1) and the second pharyngeal arch (PA2). To collect CNCCs of the developing pinna, we generated E12.5 and E14.5 *Wnt1::Cre;ROSA^RFP^* embryos and micro-dissected the forming pinna prominence. The *Wnt1* promoter drives Cre expression in CNCC pre-migratory progenitors, resulting in permanent RFP reporter activity in their post-migratory progeny[101]. Fluorescence-activated cell sorting (FACS) was used to isolate the RFP-positive cranial CNCCs. Further processing of these cells was dependent on the downstream application (e.g., RNA-seq, ATAC-seq, ChIP-seq, Hi-C, or promoter capture Hi-C)—see below.

For the *Hoxa2* loss of function analysis, Md and PA2 from E10.5 *Hoxa2^EGFP/EGFP^* mutant embryos were micro-dissected and the GFP-positive CNCCs were isolated by FACS. The processing of these cells was adapted for promoter capture Hi-C.

To collect cells from *HIRE1^del^* and *HIRE2^del^*, we generated E10.5 embryos and micro-dissected PA2 and PA3. The processing of the cells was adapted for RNA-seq.

Sample preparation, RNA isolation and sequencing (RNA-seq)

Dissected tissue from E10.5 and E12.5 embryos was kept in 1× PBS on ice, then treated with 0.5% trypsin/1× EDTA at 37 °C for 10 min and immediately put on ice. Dissected tissue from E14.5 embryos was kept in 1× PBS on ice, then treated with papain digestion mix (10 mg ml^{-1} papain, 2.5 mM cysteine, 10 mM HEPES (pH 7.4), 0.5 mM EDTA, and 0.9× DMEM) for 7 min at 37 °C and immediately put on ice. The tissue was rinsed once in ice-cold 1× DMEM/10% FBS, followed by two washes in ice-cold 1× DMEM. The tissue was dissociated by pipetting. CNCCs from embryos with genotype *Wnt1::Cre;ROSA^RFP^* were filtered and collected by FACS. After sorting, CNCCs were pelleted by centrifugation for 5 min, 200×*g* at 4 °C. Total RNA was extracted using the Single Cell RNA Purification Kit (NORGEN, 51800) with genomic DNA digestion using an RNase-Free DNase I Kit (NORGEN, 25710) according to the manufacturer's protocol. For each sample, three independent biological replicates were prepared.

For sequencing of total RNA, unstranded RNA-Seq libraries were prepared using Illumina TruSeq RNA Library Prep Kit v2 according to manufacturers' instructions. Sequencing was performed on an Illumina HiSeq 2500 machine (50 bp read length, single-end) or an Illumina NovaSeq 6000 machine (50 bp read length, paired-end).

Sample preparation, chromatin immunoprecipitation and sequencing (ChIP-seq)

Tissue from E10.5, E12.5, and E14.5 embryos was dissociated as described for RNA-seq. Then the dissociated cells were cross-linked with 1% formaldehyde for 10 min at room temperature (RT) and

quenched with glycine (final concentration 125 mM) for 5 min at RT. Cells were spun down (500×*g*, 10 min, 4 °C). CNCCs with genotype *Wnt1::Cre;ROSA*^{RFP} were filtered and collected by FACS.

ChIP experiments with anti-H3K4me2 (Millipore, Cat. #07-030), anti-H3K27me3 (Millipore, Cat. # 07-449) anti-H3K27ac (Abcam, ab4729), or anti-CTCF (Cell Signaling, CST 2899) antibodies were performed as described in ref. 16.

ChIP libraries were generated using bar-coded adapters (NEB, E7335) in combination with the NEBNext® Ultra™ DNA Library Prep Kit for Illumina® (NEB, E7370) according to manufacturers' instructions. The quality of the libraries and size distribution was assessed on an Agilent 2100 Bioanalyzer (Agilent Technologies). All ChIP-seq experiments were performed with two independent biological replicates, except for CTCF, only one biological replicate was prepared per sample. Sequencing was performed on an Illumina HiSeq 2500 machine (50 bp read length, single-end).

Sample preparation, assay for transposase-accessible chromatin and sequencing (ATAC-seq)

Tissue from E10.5, E12.5, and E14.5 embryos was dissociated as described for RNA-seq. CNCCs with genotype *Wnt1::Cre;ROSA*^{RFP} were filtered and collected by FACS.

To identify open chromatin regions, we used the Assay for Transposase-Accessible Chromatin (ATAC-seq) protocol, performed according to ref. 104. with minor modifications as described previously in ref. 16. For each sample, at least two independent biological replicates were prepared.

Sequencing was performed on an Illumina NextSeq 500 (75 bp read length, paired-end).

Hi-C and PCHi-C sample preparation and sequencing

Tissue from E10.5, E12.5, and E14.5 embryos was dissociated as described for RNA-seq. E14 mESCs were collected with Accutase (Sigma-Aldrich, A6964) and resuspended in cold 1× PBS. Then the dissociated cells were cross-linked with 2% paraformaldehyde for 10 min at room temperature (RT) and quenched with glycine (final concentration 250 mM) for 5 min at RT. Cells were spun down (500×*g*, 10 min, 4 °C) and resuspended in cold 1× PBS. CNCCs with genotype *Wnt1::Cre;ROSA*^{RFP} or *Hoxa2*$^{EGFP/EGFP}$ were filtered and collected by FACS. Cells were spun down (500×*g*, 10 min, 4 °C), the supernatant was removed, and cell pellets were flash-frozen and stored at −80 °C.

Promoter-capture Hi-C was performed as described previously in ref. 105 with adaptations to our system. About 39021 biotinylated RNA probes were designed to enrich 22,225 annotated gene promoters in the mouse genome using PCHi-C. These probes were designed and used as described in refs. 37, 105 and purchased from Agilent Technologies. The input material per PCHi-C experiment was 2.5–5 × 10⁶ cells and per Hi-C experiment, 1–2 × 10⁶ cells. Cells were incubated in 1 mL (Hi-C: 500 μL) ice-cold lysis buffer (10 mM Tris-HCl pH 8, 0.2% (vol/vol) NP-40, 10 mM NaCl,1× protease inhibitor cocktail (Roche)) for 30 min with periodical mixing, followed by centrifugation to pellet nuclei (760×*g*, 5 min, 4 °C). Cell nuclei were resuspended in 358 μL (Hi-C: 179 μL) ice-cold 1.25× NEBuffer 2, 11 μL (Hi-C: 5.5 μL) of 10% SDS was added, and nuclei were incubated at 37 °C, shaking at 950 rpm, for 1 h. Then 75 μL (Hi-C: 37.5 μL) of 10% Triton X-100 was added and cell nuclei were incubated at 37 °C, shaking at 950 rpm, for 1 h. A 5 μL aliquot was taken as a control for undigested chromatin (stored at −20 °C). For HindIII digestion, 10 μL (Hi-C: 5 μL) of HindIII enzyme (NEB, cat. no. R0104T) was added, incubated 6 h at 37 °C shaking at 950 rpm, followed by the addition of another 10 μL (Hi-C: 5 μL) of HindIII and overnight incubation at 37 °C shaking at 950 rpm. A 5 μL aliquot was taken to test digestion efficiency (stored at −20 °C). To fill-in and biotinylate the restriction fragment overhangs 60 μL (Hi-C: 30 μL) of biotinylation master mix (6 μL 10× NEBuffer 2, 2 μL H₂O, 1.5 μL 10 mM dATP, 1.5 μL 10 mM dGTP, 1.5 μL 10 mM dTTP, 37.5 μL 0.4 mM biotin-14-

dATP, and 10 μL 5 U/μl Klenow (DNA polymerase I large fragment, NEB M0210L)) was added, incubated for 75 min at 37 °C mixing regularly by pipetting, and then placed on ice. Next, 454 μL (Hi-C: 227 μL) ligation master mix (100 μL 10× ligation buffer (NEB B0202), 10 μL 10 mg/mL BSA, 344 μL nuclease-free H₂O) and 25 μL (Hi-C: 12.5 μL) 1 U/μL T4 DNA ligase (Invitrogen 15224-024) was added for overnight blunt end ligation in a water bath at 16 °C. The next day ligated samples were incubated for a further 30 min at room temperature. The re-ligated chromatin products and test aliquots were de-cross-linked for at least 6 h by adding 100 μL (Hi-C: 50 μL) and 2.5 μL proteinase K (10 mg/mL), incubated at 65 °C. Afterward, 12 μL (Hi-C: 6 μL) or 0.5 μL of 10 mg/ml RNase A was added to the samples and test aliquots, respectively, and incubated for 60 min at 37 °C. Next, chromatin was precipitated by adding 1 volume of phenol:chloroform:isoamyl alcohol (25:24:1) (PC) to the samples and test aliquots, vigorously shaking them, followed by centrifugation at a maximum speed of a tabletop centrifuge (-16,500×*g*) at room temperature for 5 min. The upper phase containing the chromatin was transferred to a new 15 mL tube. To reduce loss of DNA, 1 mL of TLE buffer (10 mM Tris-HCl (pH 8.0), 0.1 mM EDTA) was added to the tube containing PC and residual chromatin, tube was vigorously mixed, followed by centrifugation at a maximum speed of a tabletop centrifuge (-16,500×*g*) at room temperature for 5 min. The upper phase containing the chromatin was transferred to the 15 mL tube. One-tenth volume of 3 M sodium acetate was added to the chromatin sample mixed by vortexing, followed by the addition of 3 volumes of 100% ethanol and vortexing to mix. The samples were frozen at −80 °C for at least 2 h. The precipitated chromatin was isolated by centrifugation at maximum speed for 45 min at 4 °C. The chromatin pellet was washed twice with 2 mL freshly prepared 70% ethanol and centrifuged at maximum speed for 15 min at 4 °C in between washes. Finally, the chromatin pellet was dried at room temperature and resuspended in 100 μL TLE buffer. To check the quality of the Hi-C library DNA, 1:50 dilution of the sample, and the undigested and digested controls were loaded on a 1.5% agarose gel. The DNA concentration of the Hi-C library was determined using Qubit™ dsDNA HS assay kit.

For Hi-C, 5 μg DNA and for PCHi-C, 10 μg DNA (split into two 5 μg aliquots) were used for library preparation. The next steps were performed per 5 μg aliquot. The DNA was sonicated using Covaris S220 (peak incidence power: 175 W; duty factor: 10%; cycles per burst: 200; time: 120 s) in a volume of 130 μL in a Covaris microTUBE to obtain 100–1000 bp long DNA fragments. End repair, dATP-tailing, and Adapter ligation were performed on beads. For this, 100 μL of Dynabeads™ MyOne™ Streptavidin T1 (Invitrogen) were washed twice with 400 μL Tween Buffer (TWB; 5 mM Tris-HCl pH 8.0, 0.5 mM EDTA, 1 M NaCl, 0.05% Tween 20) and resuspended in 400 μL 2× binding buffer (2×BB; 10 mM Tris-HCl pH 8.0, 1 mM EDTA, 2 M NaCl). The volume of the sonicated DNA was increased to 400 μL with TLE. DNA was mixed with washed beads (800 μL total volume) and incubated at RT for 45 min rotating at 5 rpm. Beads were reclaimed on a magnetic separation stand, the supernatant was removed, followed by a wash with 400 μL 1×BB, then with 100 μL 1×ligation buffer, and resuspended in 50 μL 1× ligation buffer. For end repair, each aliquot was mixed on ice with 50 μL 2.5 mM dNTP mix (12.5 μL of 10 mM of each dNTP), 18 μL of T4 DNA Polymerase (NEB M0203), 18 μL of T4 PNK (NEB M0201), 3.7 μL of DNA polymerase I large fragment (Klenow; NEB M0210), and 360.1 μL of H₂O, and then incubated at 20 °C for 60 min. Beads were reclaimed on a magnetic separation stand, and the supernatant was removed, followed by two washes with 500 μL 1×TWB, one wash with 500 μL 1×BB, one wash with 500 μL TLE and resuspended in 415 μL TLE. For dATP-tailing, each aliquot was mixed with 50 μL 10× NEBuffer 2, 5 μL 10 mM dATG, and 20 μL 5 U/μL Klenow Fragment (3′ → 5′ exo-) (NEB, M0212), and then incubated at 37 °C for 30 min. Beads were reclaimed on a magnetic separation stand, the supernatant was removed, followed by two washes with 500 μL 1×TWB, one wash with

500 μL 1×BB, one wash with 200 μL 1× ligation buffer, and resuspended in 100 μL 1× ligation buffer. For PCHi-C adapter ligation, 4 μL 15 μM pre-annealed PE adapters (described in ref. [105].) and 4 μL NEB T4 Ligase 400 U/mL (NEB, M0202) were added to each aliquot and incubated 2 h at room temperature. For Hi-C, 10 μL 15 μM TruSeq DNA Sgl index adapters (Illumina, Cat. No. 20016329) and 4 μL NEB T4 Ligase 400 U/mL (NEB, M0202) were added to each aliquot and incubated for 2 h at room temperature. Beads were reclaimed on a magnetic separation stand, the supernatant was removed, followed by two washes with 500 μL 1×TWB, one wash with 500 μL 1× BB, one wash with 200 μL 1×BB, one wash with 100 μL 1× NEBuffer 2, one wash with 50 μL 1× NEBuffer 2 and resuspended in 50 μL 1× NEBuffer 2. At this point, all aliquots per PCHi-C sample were pooled.

For on-bead PCHi-C and Hi-C library amplification, a reaction master mix was prepared, reaction volume per PCR tube was 25 μL (PCHi-C: 5 μL 5× Herculase II reaction buffer (Agilent), 0.25 μL 100 mM dNTPs, 2 μL PE PCR primer 1.0[105], 2 μL PE PCR primer 2.0[105], 0.5 μL Polymerase (Herculase, Agilent), 2.5 μL Hi-C DNA on beads, 12.75 μL nuclease-free H_2O; Hi-C: 12.5 μL KAPA HiFi HotStart ReadyMix (Roche), 2.5 μL Illumina primer mix (contained in KAPA HiFi HotStart Library Amplification Kit), 2.5 μL Hi-C DNA on beads, 7.5 μL nuclease-free H_2O) and PCR was run under the following conditions: 30 s at 98 °C; 6–7 cycles of 10 s at 98 °C, 30 s at 65 °C, 30 s at 72 °C; 7 min at 72 °C. Post-amplification PCR reactions per sample were pooled and DNA was cleaned up using AMPure XP beads and eluted in nuclease-free H_2O. Obtained DNA was used directly for Hi-C sequencing or for promoter-capture Hi-C as described in ref. [105]. Hi-C experiments on mESC and cells from E12.5 and E14.5 pinna were performed in biologically independent duplicates. Hi-C experiments on Md and PA2 at E10.5 have one replicate each. All promoter-capture Hi-C experiments have two biologically independent replicates. Sequencing was performed on an Illumina NextSeq 500 (75 bp read length, paired-end).

Quantitative real-time PCR (qRT-PCR)
Total RNA was extracted from PA2 of E10.5 embryos using the Single Cell RNA Purification Kit (NORGEN, 51800) with genomic DNA digestion using an RNase-Free DNase I Kit (NORGEN, 25710) according to the manufacturer's protocol. cDNA synthesis was performed using Superscript III (Invitrogen, ref#18080-044) according to the manufacturer's instructions and was then treated with RNAse H (NEB, M0297L) for half an hour at 37 °C. Samples were diluted 1:10 and qPCR was performed using StepOnePlus real-time PCR machine (Thermo Fisher) with SYBR Green PCR master mix (Thermo Fisher, ref#4309155) according to manufacturer's protocol.

The following primers were used: *Hoxa2* forward primer (FW) 'CAAGACCTCGACGC TTTCACAC', reverse primer (RV) 'CCTTCATCC AGGGATACTCAGGC'; *Gapdh* FW 'GAGAGGCCCTATCCCAACTC', RV 'GGTCTGGGATGGAAATTGTG'. The *Hoxa2* primers were specifically designed to only amplify from the *Hoxa2* wild-type allele but not the *Hoxa2* EGFP allele. Primer efficiencies of each primer pair were evaluated with a standard curve, and the occurrence of primer dimers was checked for with a melting curve and gel electrophoresis. Relative expression levels of *Hoxa2* were quantified using the ΔΔCt method, where *Hoxa2* Ct values were normalized to *Gapdh* levels and the average ΔCt of the wild-type samples. Statistical analysis was performed by one-way analysis of variance (ANOVA) followed by Dunnett's test.

Skeletal staining
E18.5 mouse embryos were skinned and eviscerated. Skeletal staining of the embryos was performed according to a previously described protocol[74]. Samples were fixed in 95% ethanol for 5–7 days. Subsequently, embryos were incubated with 0.015% alcian blue 8GS, 0.005% alizarin red S, and 5% acetic acid for 3 days with agitation at 37 °C. Samples were cleared in 1% KOH for several days and in 1% KOH/

glycerol series until surrounding tissues turned transparent. The samples were stored in glycerol for a longer term.

Whole-mount in situ hybridization (WISH)
In situ hybridization on whole-mount embryos was performed as previously described in ref. [106] with minor modifications. Embryos were dissected free of extraembryonic membranes in 1× PBS containing 0.1% Tween 20 (PBT) and fixed in 4% paraformaldehyde in 1× PBS for 2 h at room temperature or at 4 °C overnight. Throughout the procedure, the embryos were rocked gently on a mechanical rocking platform unless otherwise indicated. Embryos were washed three times with PBT, dehydrated through a methanol-PBT series into 100% methanol, and stored at −20 °C until further use. To summarize, embryos were rehydrated through a methanol-PBT series and washed three times in PBT, followed by a 1-h incubation in 6% hydrogen peroxide/PBT at room temperature and three washes in PBT. Next, embryos were treated with 5–10 μg/ml proteinase K in PBT for 1–2 min at room temperature (Proteinase K concentration and incubation time must be adjusted according to embryo size and a batch of Proteinase K), directly followed by refixation in 0.2% glutaraldehyde and 4% paraformaldehyde in 1× PBS for 20 min at room temperature and three washes with PBT. Embryos were pre-hybridized for 3–4 h at 70 °C in hybridization buffer (50% deionized formamide, 5× saline-sodium citrate (SSC) buffer pH 7.5, 100 μg/ml tRNA, 50 μg/ml heparin, 1% SDS). The hybridization buffer was replaced, RNA probes labeled with digoxigenin were added, and the embryos were incubated overnight at 70 °C. The embryos were rinsed once with Wash 1 (50% formamide, 4× SSC pH 7.5, 1% SDS), followed by two washes for 30 min each at 70 °C in Wash 1. The embryos were washed once at 70 °C for 10 min in a 1:1 mix of Wash 1 and Wash 2 (0.5 M NaCl, 10 mM Tris-HCl pH 7.5, 0.1% Tween 20), then three times in Wash 2 for 5 min at room temperature. The wash buffer was replaced by Wash 2 containing 100 μg/ml RNase A and incubated at 37 °C for 1 h. The embryos were washed once for 5 min with Wash 2, then once for 5 min at 70 °C with pre-heated Wash 3 (50% formamide, 2× SSC pH 7.5, 0.1% SDS). This was followed by two washes with Wash 3 at 70 °C for 1 h each. The embryos were washed three times with 1× maleic acid buffer containing 0.1% Tween 20 (MABT), then incubated for 2–3 h in MABT with 2% blocking reagent (Roche-11096176001). Sufficient embryo powder was heat-inactivated just before use in MABT with 2% blocking reagent 70 °C for 30 min, vortexed for 5 min, and placed on ice. Then, the required amount of anti-DIG AP FAB fragment (Roche-11093274910) was preabsorbed for 1 h at 4 °C with prepared embryo powder in MABT with 2% blocking reagent, followed by centrifugation at 4 °C for 10 min. The supernatant was diluted in MABT with 2% blocking reagent to obtain a final antibody dilution of 1:2000. Embryos were incubated with the preabsorbed antibody overnight at 4 °C. The embryos were washed three times at room temperature with MABT containing 2 mM levamisole, followed by eight washes for 1 h each at room temperature. The last wash was followed by three changes of NTMT (100 mM NaCl, 100 mM Tris-HCl pH 9.5, 50 mM MgCl2, 0.1% Tween 20) containing 2 mM levamisole. The color reaction was initiated by placing the embryos into NTMT containing 2 mM levamisole, 3.5 μL/mL NBT (Roche-11383213001), and 3.5 μL/mL BCIP (Roche-11383221001). Staining was allowed to proceed for multiple hours at room temperature in the dark. When staining was satisfactory, the embryos were washed three times with MABT, and left to wash in MABT at 4 °C until the background signal was sufficiently removed.

The following RNA probes were used: *Hoxa1*[107], *Hoxa2*[108], and *Hoxa3*[23,24].

Histological analysis and in situ hybridizations
Mouse embryos were collected and fixed in 4% PFA/1× PBS overnight at 4 °C. For cryostat sections, tissues were cryoprotected in 20% sucrose/1× PBS and embedded in 7.5% gelatin/10% sucrose/1× PBS before being

frozen at −80 °C. Cryostat sections (25 μm) were cut (CryoStar NX70) in sagittal orientation.

Frozen sections were fixed in 4% PFA/1× PBS, and washed in PBS. Acetylation was performed with acetic anhydride solution (1.35% (v/v) Triethanolamine, 0.175% (v/v) HCl, 0.25% (v/v) acetic anhydride) for 10 min at RT, before being pre-hybridized with Hybridization buffer (50% deionized formamide, 1× Salt, 10% dextran sulfate, 1 mg/ml tRNA, 1× Denhardt) for 1 h 30 at room temperature. Hybridization was performed with antisense riboprobes labeled with digoxigenin-11-d-UTP diluted 1:100 in hybridization buffer, for 16–18 h at 70 °C, in a humid chamber (humidified with 50% formamide/5× SSC). On day 2, Sections were rinsed with 5× SSC pH 4.5, followed by washing 2 h at 70 °C in 0.2× SSC pH 4.5, and 2× 10 min at RT in MABT buffer (0.1 M Maleic Acid, 0.15 M NaCl, 2 N NaOH, pH 7.5). After blocking 1h30 at RT in MABT/2% Blocking reagent (Roche Diagnostics), sections were incubated with anti-DIG antibody conjugated to alkaline phosphatase in MABT/2% Blocking reagent (1:5,000; Roche Diagnostics) overnight at RT. On day 3, slides were washed 8x In MABT (0.1 M Maleic Acid, 0.15 M NaCl, 2 N NaOH, pH 7.5), 2× 10 min in NTMT buffer (100 mM Tris-HCl, pH 9.5, 100 mM NaCl, 50 mM $MgCl_2$, 0.1% Tween 20, pH 9.5). The alkaline phosphatase activity was detected using NBT and BCIP diluted in NTMT buffer at 4 °C. To stop the reaction, slides were washed 2× 10 min in MABT. Slides were mounted with aqueous mounting media (Aquatex, Merck). Imaging was performed using Zeiss Axio Scan Z1, using 5× (NA 0.25), 10× (NA 0.45), and Zeiss upright Axio imager Z1 using 5× (NA 0.25).

The following RNA probes were used: *Hoxa2*[108] and *Hoxa3*[23,24].

Use of previously published datasets

Some datasets have been used from our previously published work[16], as indicated in the manuscript, and are published in the SuperSeries GSE89437. A number of RNA-seq, ChIP-seq, and ATAC-seq samples published in this series have been re-processed for this study (Supplementary Data 7).

Computational analysis

Reference genome and annotation. The mouse GRCm38/mm10 genome assembly was used as a reference. Gene annotation was obtained from the *TxDb.Mmusculus.UCSC.mm10.knownGene* Bioconductor package (https://doi.org/doi:10.18129/B9.bioc.TxDb.Mmusculus.UCSC.mm10.knownGene, version 3.4.7). For RNA-seq analysis, a single transcript start site (TSS) was selected randomly for each gene, and promoter regions were defined as a 2 kb window centered on the TSS.

For genome browser views, the number of alignments per 100 bp window in the genome and per million alignments in each sample were calculated, stored in BigWig format using the *qExportWig* function from the *QuasR* Bioconductor package (version 1.34.0)[109] and visualized as custom tracks in either UCSC genome browser (http://genome.ucsc.edu)[110] or pyGenomeTracks (version 3.7)[111].

RNA-seq data analysis

All RNA-seq experiments used in this study were sequenced single-end with 50 bp read length or paired-end 50 bp read length (Supplementary Data 7). The dataset containing single-end reads were not compared with the datasets containing paired-end reads. Thus, trimming the second read of the paired-end data was not necessary. Reads of single-end and paired-end datasets were aligned to the reference genome using the *QuasR* Bioconductor package (version 1.34.0)[109] by the *qAlign* function with parameters *aligner = "Rhisat2"* (version 1.12.0) and *splicedAlignment = TRUE*. Alignments overlapping genes from the *TxDb.Mmusculus.UCSC.mm10.knownGene* Bioconductor package (https://doi.org/doi:10.18129/B9.bioc.TxDb.Mmusculus.UCSC.mm10.knownGene, version 3.4.7) were quantified using the *qCount* function with parameter *orientation = "opposite"*.

The RNA-seq samples for E10.5 Md and E10.5 PA2 from public datasets (Supplementary Data 7) were processed in the same way for illustration as genome browser tracks.

Differential RNA-seq analysis

To identify differentially expressed genes, only expressed genes in both biological replicates are used (CPM > = 1). Differentially expressed genes are identified using *edgeR* (version 3.38.1)[112]. Two different models are fit for genes counts:

1/ ~ time point (E10.5 PA2; E12.5 Pinna; E14.5 Pinna)

2/ ~ genotype (PA2 WT and PA2 delta E3; PA3 WT and PA3 delta E3)

For each model, dispersions are estimated using *estimateDisp* and statistical significance is calculated using *glmQLFTest*. Differentially expressed genes are defined as genes with an FDR less than 0.05.

ChIP-seq data analysis

Reads were mapped to the reference genome using *bowtie2 w*ith default settings (version 2.4.2)[113,114] and converted to bam files using *samtools* (version 1.2)[115,116].

ATAC-seq data analysis

The adapter sequence CTGTCTCTTATACACA was trimmed from the 3' end of all samples using *cutadapt* (version 3.7)[117] with overlap = 1. The trimmed reads were aligned with *bowtie2* (version 2.4.2) with the options–*fr,–minins 0,–maxins 1000,–nodiscordant* and–*dovetail*, and converted to bam files using *samtools* (version 1.2)[115,116].

Hi-C data analysis

The Hi-C data has been mapped and quality controlled with *HiCUP* (version 0.6.1)[118], the interaction matrix was created with *Juicer* (version 1.6)[119], and was transformed into a *cool* matrix[120] with *hic2cool* (https://github.com/4dn-dcic/hic2cool, version 0.8.3) and *HiCExplorer's hicConvertFormat* (version 3.7.2)[121]. For the Hi-C on mESC, E12.5 pinna, and E14.5 pinna the *cool* matrices of biological replicates were merged into one *cool* matrix with *HiCExplorer's hicSumMatrices* (version 3.7.2)[122]. The merged *cool* matrix was normalized with *HiCExplorer's hicCorrectMatrix* (version 3.7.2)[122] using KR normalization. The TAD calling was applied by *HiCExplorer's hicFindTADs* (version 3.7.2)[122,123]. The Hi-C and TAD data was plotted with *pyGenomeTracks* (version 3.7)[111].

The differential interaction maps were generated by using the *FAN-C* suite of tools (Kruse et al., 2020). To derive the differential interactions maps, the normalized scores of one matrix were subtracted from the other using f*anc compare* tool, and the maps are plotted using *fancplot* tool.

Virtual 4C plots are generated by using *Virtual4CPlot* tool from *FAN-C*.

PCHi-C data analysis

Capture HiC libraries are first analyzed with *HiCUP* (version 0.6.1)[118], then significant interactions are identified using *CHiCAGO* (version 1.24.0)[124]. *CHiCAGO* pipeline is used with the recommended parameters for six cutter restriction enzymes[125]: *minFragLen = 150* and *maxFragLen = 40000, maxLBrownEst = 1500000* and *binsize = 20000*.

Interactions with a *CHiCAGO* score of 5 are considered as high-confidence interactions. Significant interactions are plotted as arcs linking baits to OE using the python api pygenometrack.

4C-like profiles of interactions are plotted using ChiCMaxima (version 1.0) browser[126]. The PCHi-C CHiCAGO interactions with a score ≥5 were plotted with *pyGenomeTracks* (version 3.7)[111].

The percentage of decrease of interactions of Hoxa2 promoter and HIRE2 SE in the PA2 E10.5 *Hoxa2*[EGFP/EGFP] compared to PA2 E10.5 WT is calculated using quantile normalized raw counts of Hoxa2 interactions with the HIRE2 region.

CTCF motif occurrences

To identify the orientation of CTCF motifs, we used the position frequency matrix of CTCF for Mouse downloaded from Jaspar 2022[127]. CTCF motif occurrences are identified by the *FIMO* package (version 5.4.1)[128]. Only motif sites with p-values <= 1.0E-4 are included. In total, 380951 CTCF motif occurrences were identified.

SE calling

SEs were identified computationally by referring to the ROSE algorithm[44,46]. First, bam files of the H3K27ac ChIP-seq alignments from all E10.5 samples (Supplementary Data 7) were merged. Then, peaks were called using MACS2 (version 2.1.3.3)[129] with the options *–nomodel,–shift 0,–extsize 141,–keep-dup all,–qvalue 0.001,–cutoff-analysis*. The extension size had been determined previously based on cross-correlation using the *csaw* Bioconductor package (version 1.28.0)[130]. Peaks that overlapped with promoter regions, i.e., 2500 bp upstream or downstream of the transcription start sites of the UCSC known genes (*TxDb.Mmusculus.UCSC.mm10.knownGene*, v. 3.10.0), or peaks overlapping with blacklisted regions[131] were removed. Only peaks located on chromosomes 1–19 or the X-chromosome were kept. Counts per million (CPM) were quantified for all remaining peaks for each sample individually using *QuasRs qCount* function with default parameters[109]. The quantified peaks were merged into larger regions if they were less than 12.5 kb apart. Subsequently, the CPMs within the individual peaks were summed. If the resulting stitched regions overlapped with the promoter regions, they were removed. The remaining regions were ranked by their summed counts for each sample. The threshold to distinguish between super- vs. non-SEs was determined as in ROSE. Lastly, only regions with at least 3 individual peaks were kept to select clusters of enhancers.

Calculation of interaction strengths between SEs and promoters

SEs were matched to differentially expressed transcription factor encoding genes if at least one significant interaction (Chicago score ≥5) was detected from the promoter bait to a restriction fragment overlapping with a SE region in at least one population. To ensure the interaction occurs in a population where both elements are active, the SE had to show acetylation levels above the relative threshold, and the gene had to be expressed with >2 RPKM in the same population.

To obtain the interaction strengths between SEs and associated transcription factor promoters, we quantified the mean Chicago score for interactions between all restriction fragments overlapping with a SE and the corresponding promoter for each sample.

To display the association between SE acetylation, link strength, and gene expression, a heatmap was generated with the *ComplexHeatmap* Bioconductor package (version 2.10.0)[132]. Rows were clustered using k-means clustering on the relative gene expression levels with seven centers.

Motif enrichment analysis

For transcription factor motif analysis, ATAC-seq data from E10.5 PA2 and E12.5 Pinna was used. Peaks were identified separately for each dataset using MACS2 (version 2.1.3.3)[129] with options *–nomodel,–shift –100,–extsize 200, -f BED,–gsize 1.87e9*, and *–qvalue 0.10* and a combined peak set was created by fusing peaks from E10.5 PA2 and E12.5 Pinna (*n* = 133,466). Since each ATAC sample was a mixture of mouse embryos of different sexes, only autosomal peaks were used for the analysis. Peaks were classified as "promoter" peaks, when the distance of a peak midpoint to the nearest TSS was smaller than 1 kb, or else as "non-promoter" peaks (hereon called enhancers). The ATAC-seq signal of E10.5 PA2 and E12.5 Pinna replicates was quantified individually at autosomal enhancers (*n* = 106,587) using *QuasR's qCount* function with default parameters[109] and normalized by dividing by the total number of alignments in each sample and multiplying by 1e6 to obtain counts

per million (CPM). For further analysis, the CPM values were log2-transformed with a pseudo-count of 1. Enhancers overlapping the genomic interval between HIRE2 and HIRE1 (chr6:50789172-51087888, mouse GRCm38/mm10) were extracted (*n* = 31). The average logCPM at each extracted enhancer was calculated and subtracted from the individual logCPM values of each sample to obtain the relative accessibility per enhancer and per sample. According to the relative accessibility values for each sample, the extracted enhancers were grouped into three clusters and heatmaps were drawn using the *ComplexHeatmap* Bioconductor package (version 2.6.2)[132]. The mean accessibility profile for each cluster was calculated and used to rank all enhancers by their similarity to each cluster profile (visualized using *vioplot* function from the *vioplot* package version 0.3.6). The top 1000 enhancers with the highest similarity in their accessibility profile compared to each cluster profile were grouped together (visualized using *upset* function from the *UpSetR* package version 1.4.0) and all residual peaks were grouped together as a control set. The *calcBinnedMotifEnr* function from the *monaLisa* package (version 0.1.40)[62] with *method = "R"* was used to identify what motifs are enriched in each of the three cluster profiles using the vertebrate list of motifs present in JASPAR (JASPAR2018)[133]. Motifs that had an FDR less than 0.001 in any of the three cluster profiles were selected. Heatmaps were drawn using the *ComplexHeatmap* Bioconductor package. JASPAR motifs were extracted from *JASPAR2018* package (version 1.1.1) using *getMatrixSet* function from the *TFBSTools* package (version 1.28.0) with *opts = list(tax_group = "vertebrates")*.

Selection of representative motifs

All motifs that had an FDR less than 0.001 in any of the three cluster profiles from the motif enrichment analysis were used to scan the sequences of the 31 putative enhancers in the region from HIRE2 to HIRE1 for motif hits using the *findMotifs* function from the *monaLisa* package (version 1.4)[62] with *min.score = 10.0* and *method = "matchPWM"*. Since the enhancers/ATAC peaks differ in lengths, the number of hits per peak per kb were calculated. For each motif, the average motif rate per kb across all enhancers was determined. To further summarize these motifs and select the most relevant ones, motifs were clustered based on the average motif hit rate per kb into two sets using *k-means* clustering. All the motifs in the cluster with the higher hit rate were used to perform a motif similarity analysis using the *motifSimilarity* function from the *monaLisa* package (version 1.4)[62] with *method = "R"*, which compares all pairs of motifs. The results of this analysis were grouped into ten clusters using *k-means* clustering (k = 10) and a heatmap was drawn using the *ComplexHeatmap* Bioconductor package (version 2.14.0)[132]. For each cluster, the motif with the least average distance to the other motifs in the cluster was selected as a representative motif.

Statistics and reproducibility

No statistical method was used to predetermine the sample size. No data were excluded from the analyses. The experiments were not randomized. The Investigators were not blinded to allocation during experiments and outcome assessment.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All raw sequencing data and processed data generated in this study have been deposited and are publicly available in the Gene Expression Omnibus (GEO) under GEO Series accession number GSE211904 (all data SuperSeries), GSE211899 (ATAC-seq), GSE211900(ChIP-seq), GSE211901 (Hi-C), GSE211902 (PCHi-C), and GSE211903 (RNA-seq). The following public sequencing datasets were published in ref. 16. and

available through GEO were used in this study: ChIP-seq, RNA-seq, and ATAC-seq data from mouse E10.5 cranial neural crest cell subpopulations (GSE89437). FACS gating strategies are presented in Supplementary Fig. 3. ChIP-seq peaks for Hoxa2 in PA2 at E11.5 were obtained from ref. 57, and ChIP-seq peak data for Pbx and Meis in PA2 at E11.5 were obtained from ref. 58. Further public databases used in this study are UCSC (mm10 reference genome assembly, gene annotation), JASPAR2018 (vertebrate transcription factor motifs)[133], JASPAR 2022 (vertebrate transcription factor motifs)[127], and the ENCODE Blacklist[131].

Code availability

Computational analyses were performed in R and Python using the publicly available packages described in Methods and Reporting Summary.

References

1. Schoenfelder, S. & Fraser, P. Long-range enhancer-promoter contacts in gene expression control. *Nat. Rev. Genet.* **20**, 437–455 (2019).
2. Claringbould, A. & Zaugg, J. B. Enhancers in disease: molecular basis and emerging treatment strategies. *Trends Mol. Med.* **27**, 1060–1073 (2021).
3. Smith, E. & Shilatifard, A. Enhancer biology and enhanceropathies. *Nat. Struct. Mol. Biol.* **21**, 210–219 (2014).
4. Spielmann, M., Lupianez, D. G. & Mundlos, S. Structural variation in the 3D genome. *Nat. Rev. Genet.* **19**, 453–467 (2018).
5. Attanasio, C. et al. Fine tuning of craniofacial morphology by distant-acting enhancers. *Science* **342**, 1241006 (2013).
6. Kragesteen, B. K. et al. Dynamic 3D chromatin architecture contributes to enhancer specificity and limb morphogenesis. *Nat. Genet.* **50**, 1463–1473 (2018).
7. Long, H. K. et al. Loss of extreme long-range enhancers in human neural crest drives a craniofacial disorder. *Cell Stem Cell* **27**, 765–783.e14 (2020).
8. Prescott, S. L. et al. Enhancer divergence and cis-regulatory evolution in the human and chimp neural crest. *Cell* **163**, 68–83 (2015).
9. Uslu, V. V. et al. Long-range enhancers regulating Myc expression are required for normal facial morphogenesis. *Nat. Genet.* **46**, 753–758 (2014).
10. Osterwalder, M. et al. Enhancer redundancy provides phenotypic robustness in mammalian development. *Nature* **554**, 239–243 (2018).
11. Dickel, D. E. et al. Ultraconserved enhancers are required for normal development. *Cell* **172**, 491–499.e15 (2018).
12. Le Douarin, N. & Kalcheim, C. *The Neural Crest* (Cambridge Univ. Press, 1999).
13. Santagati, F. & Rijli, F. M. Cranial neural crest and the building of the vertebrate head. *Nat. Rev. Neurosci.* **4**, 806–818 (2003).
14. Gammill, L. S. & Bronner-Fraser, M. Neural crest specification: migrating into genomics. *Nat. Rev. Neurosci.* **4**, 795–805 (2003).
15. Minoux, M. & Rijli, F. M. Molecular mechanisms of cranial neural crest cell migration and patterning in craniofacial development. *Development* **137**, 2605–2621 (2010).
16. Minoux, M. et al. Gene bivalency at Polycomb domains regulates cranial neural crest positional identity. *Science* **355**, eaal2913 (2017).
17. Duboule, D. The rise and fall of Hox gene clusters. *Development* **134**, 2549–2560 (2007).
18. Montavon, T. & Duboule, D. Chromatin organization and global regulation of Hox gene clusters. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **368**, 20120367 (2013).
19. Krumlauf, R. Hox genes, clusters and collinearity. *Int. J. Dev. Biol.* **62**, 659–663 (2018).
20. Parker, H. J., Pushel, I. & Krumlauf, R. Coupling the roles of Hox genes to regulatory networks patterning cranial neural crest. *Dev. Biol.* **444**, S67–S78 (2018).
21. Afzal, Z. & Krumlauf, R. Transcriptional regulation and implications for controlling Hox gene expression. *J. Dev. Biol.* **10**, 4 (2022).
22. Minoux, M., Antonarakis, G. S., Kmita, M., Duboule, D. & Rijli, F. M. Rostral and caudal pharyngeal arches share a common neural crest ground pattern. *Development* **136**, 637–645 (2009).
23. Hunt, P. et al. A distinct Hox code for the branchial region of the vertebrate head. *Nature* **353**, 861–864 (1991).
24. Hunt, P. et al. The branchial Hox code and its implications for gene regulation, patterning of the nervous system and head evolution. *Dev. Suppl.* **2**, 63–77 (1991).
25. Gendron-Maguire, M., Mallo, M., Zhang, M. & Gridley, T. Hoxa-2 mutant mice exhibit homeotic transformation of skeletal elements derived from cranial neural crest. *Cell* **75**, 1317–1331 (1993).
26. Rijli, F. M. et al. A homeotic transformation is generated in the rostral branchial region of the head by disruption of Hoxa-2, which acts as a selector gene. *Cell* **75**, 1333–1349 (1993).
27. Santagati, F., Minoux, M., Ren, S. Y. & Rijli, F. M. Temporal requirement of Hoxa2 in cranial neural crest skeletal morphogenesis. *Development* **132**, 4927–4936 (2005).
28. Minoux, M. et al. Mouse Hoxa2 mutations provide a model for microtia and auricle duplication. *Development* **140**, 4386–4397 (2013).
29. Alasti, F. & Van Camp, G. Genetics of microtia and associated syndromes. *J. Med. Genet.* **46**, 361–369 (2009).
30. Brown, K. K. et al. HOXA2 haploinsufficiency in dominant bilateral microtia and hearing loss. *Hum. Mutat.* **34**, 1347–1351 (2013).
31. Piceci, F. et al. Identification of a second HOXA2 nonsense mutation in a family with autosomal dominant non-syndromic microtia and distinctive ear morphology. *Clin. Genet.* **91**, 774–779 (2017).
32. Dekker, J., Rippe, K., Dekker, M. & Kleckner, N. Capturing chromosome conformation. *Science* **295**, 1306–1311 (2002).
33. Dryden, N. H. et al. Unbiased analysis of potential targets of breast cancer susceptibility loci by Capture Hi-C. *Genome Res.* **24**, 1854–1868 (2014).
34. Hughes, J. R. et al. Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-throughput experiment. *Nat. Genet.* **46**, 205–212 (2014).
35. Kolovos, P. et al. Targeted chromatin capture (T2C): a novel high resolution high throughput method to detect genomic interactions and regulatory elements. *Epigenetics Chromatin* **7**, 10 (2014).
36. Sahlen, P. et al. Genome-wide mapping of promoter-anchored interactions with close to single-enhancer resolution. *Genome Biol.* **16**, 156 (2015).
37. Schoenfelder, S. et al. The pluripotent regulatory circuitry connecting promoters to their long-range interacting elements. *Genome Res.* **25**, 582–597 (2015).
38. Mifsud, B. et al. Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. *Nat. Genet.* **47**, 598–606 (2015).
39. Andrey, G. et al. Characterization of hundreds of regulatory landscapes in developing limbs reveals two regimes of chromatin folding. *Genome Res.* **27**, 223–233 (2017).
40. Nora, E. P. et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381–385 (2012).
41. Sexton, T. et al. Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* **148**, 458–472 (2012).
42. Dixon, J. R. et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).

43. Hnisz, D. et al. Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934–947 (2013).

44. Whyte, W. A. et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013).

45. Lieberman-Aiden, E. et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293 (2009).

46. Loven, J. et al. Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* **153**, 320–334 (2013).

47. Berlivet, S. et al. Clustering of tissue-specific sub-TADs accompanies the regulation of HoxA genes in developing limbs. *PLoS Genet.* **9**, e1004018 (2013).

48. Vian, L. et al. The energetics and physiological impact of cohesin extrusion. *Cell* **175**, 292–294 (2018).

49. Pasqualetti, M. et al. A Hoxa2 knockin allele that expresses EGFP upon conditional Cre-mediated recombination. *Genesis* **32**, 109–111 (2002).

50. Condie, B. G. & Capecchi, M. R. Mice with targeted disruptions in the paralogous genes hoxa-3 and hoxd-3 reveal synergistic interactions. *Nature* **370**, 304–307 (1994).

51. Manley, N. R. & Capecchi, M. R. Hox group 3 paralogous genes act synergistically in the formation of somitic and neural crest-derived structures. *Dev. Biol.* **192**, 274–288 (1997).

52. Vieille-Grosjean, I., Hunt, P., Gulisano, M., Boncinelli, E. & Thorogood, P. Branchial HOX gene expression and human craniofacial development. *Dev. Biol.* **183**, 49–60 (1997).

53. Lufkin, T., Dierich, A., LeMeur, M., Mark, M. & Chambon, P. Disruption of the Hox-1.6 homeobox gene results in defects in a region corresponding to its rostral domain of expression. *Cell* **66**, 1105–1119 (1991).

54. Rossel, M. & Capecchi, M. R. Mice mutant for both Hoxa1 and Hoxb1 show extensive remodeling of the hindbrain and defects in craniofacial development. *Development* **126**, 5027–5040 (1999).

55. Bobola, N. et al. Mesenchymal patterning by Hoxa2 requires blocking Fgf-dependent activation of Ptx1. *Development* **130**, 3403–3414 (2003).

56. Kirilenko, P. et al. Transient activation of meox1 is an early component of the gene regulatory network downstream of hoxa2. *Mol. Cell Biol.* **31**, 1301–1308 (2011).

57. Donaldson, I. J. et al. Genome-wide occupancy links Hoxa2 to Wnt-beta-catenin signaling in mouse embryonic development. *Nucleic Acids Res.* **40**, 3990–4001 (2012).

58. Amin, S. et al. Hoxa2 selectively enhances Meis binding to change a branchial arch ground state. *Dev. Cell* **32**, 265–277 (2015).

59. Bridoux, L. et al. HOX paralogs selectively convert binding of ubiquitous transcription factors into tissue-specific patterns of enhancer activation. *PLoS Genet.* **16**, e1009162 (2020).

60. Mallen, J. et al. Molecular characterization of HOXA2 and HOXA3 binding properties. *J. Dev. Biol.* **9**, 55 (2021).

61. Maconochie, M. et al. Regulation of Hoxa2 in cranial neural crest cells involves members of the AP-2 family. *Development* **126**, 1483–1494 (1999).

62. Machlab, D. et al. monaLisa: an R/Bioconductor package for identifying regulatory motifs. *Bioinformatics* **38**, 2624–2625 (2022).

63. Mann, R. S., Lelli, K. M. & Joshi, R. Hox specificity unique roles for cofactors and collaborators. *Curr. Top. Dev. Biol.* **88**, 63–101 (2009).

64. Moens, C. B. & Selleri, L. Hox cofactors in vertebrate development. *Dev. Biol.* **291**, 193–206 (2006).

65. Hnisz, D. et al. Convergence of developmental and oncogenic signaling pathways at transcriptional super-enhancers. *Mol. Cell* **58**, 362–370 (2015).

66. Adam, R. C. et al. Pioneer factors govern super-enhancer dynamics in stem cell plasticity and lineage choice. *Nature* **521**, 366–370 (2015).

67. Parker, S. C. et al. Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. *Proc. Natl Acad. Sci. USA* **110**, 17921–17926 (2013).

68. Baltzinger, M., Ori, M., Pasqualetti, M., Nardi, I. & Rijli, F. M. Hoxa2 knockdown in Xenopus results in hyoid to mandibular homeosis. *Dev. Dyn.* **234**, 858–867 (2005).

69. Pasqualetti, M., Ori, M., Nardi, I. & Rijli, F. M. Ectopic Hoxa2 induction after neural crest migration results in homeosis of jaw elements in Xenopus. *Development* **127**, 5367–5378 (2000).

70. Grammatopoulos, G. A., Bell, E., Toole, L., Lumsden, A. & Tucker, A. S. Homeotic transformation of branchial arch identity after Hoxa2 overexpression. *Development* **127**, 5355–5365 (2000).

71. Hunter, M. P. & Prince, V. E. Zebrafish hox paralogue group 2 genes function redundantly as selector genes to pattern the second pharyngeal arch. *Dev. Biol.* **247**, 367–389 (2002).

72. Kitazawa, T. et al. Distinct effects of Hoxa2 overexpression in cranial neural crest populations reveal that the mammalian hyomandibular-ceratohyal boundary maps within the styloid process. *Dev. Biol.* **402**, 162–174 (2015).

73. Ohnemus, S., Bobola, N., Kanzler, B. & Mallo, M. Different levels of Hoxa2 are required for particular developmental processes. *Mech. Dev.* **108**, 135–147 (2001).

74. Kitazawa, T., Minoux, M., Ducret, S. & Rijli, F. M. Different ectopic Hoxa2 expression levels in mouse cranial neural crest cells result in distinct craniofacial anomalies and homeotic phenotypes. *J. Dev. Biol.* **10**, 9 (2022).

75. Moorthy, S. D. et al. Enhancers and super-enhancers have an equivalent regulatory role in embryonic stem cells through regulation of single or multiple genes. *Genome Res.* **27**, 246–258 (2017).

76. Thomas, H. F. et al. Temporal dissection of an enhancer cluster reveals distinct temporal and functional contributions of individual elements. *Mol. Cell* **81**, 969–982 e13 (2021).

77. Hay, D. et al. Genetic dissection of the alpha-globin super-enhancer in vivo. *Nat. Genet.* **48**, 895–903 (2016).

78. Hnisz, D., Shrinivas, K., Young, R. A., Chakraborty, A. K. & Sharp, P. A. A phase separation model for transcriptional control. *Cell* **169**, 13–23 (2017).

79. Frasch, M., Chen, X. & Lufkin, T. Evolutionary-conserved enhancers direct region-specific expression of the murine Hoxa-1 and Hoxa-2 loci in both mice and *Drosophila*. *Development* **121**, 957–974 (1995).

80. Lampe, X., Picard, J. J. & Rezsohazy, R. The Hoxa2 enhancer 2 contains a critical Hoxa2 responsive regulatory element. *Biochem. Biophys. Res. Commun.* **316**, 898–902 (2004).

81. Tumpel, S., Cambronero, F., Sims, C., Krumlauf, R. & Wiedemann, L. M. A regulatory module embedded in the coding region of Hoxa2 controls expression in rhombomere 2. *Proc. Natl Acad. Sci. USA* **105**, 20077–20082 (2008).

82. Sham, M. H. et al. The zinc finger gene Krox20 regulates HoxB2 (Hox2.8) during hindbrain segmentation. *Cell* **72**, 183–196 (1993).

83. Nonchev, S. et al. Segmental expression of Hoxa-2 in the hindbrain is directly regulated by Krox-20. *Development* **122**, 543–554 (1996).

84. Nonchev, S. et al. The conserved role of Krox-20 in directing Hox gene expression during vertebrate hindbrain segmentation. *Proc. Natl Acad. Sci. USA* **93**, 9339–9345 (1996).

85. Tumpel, S. et al. Expression of Hoxa2 in rhombomere 4 is regulated by a conserved cross-regulatory mechanism dependent upon Hoxb1. *Dev. Biol.* **302**, 646–660 (2007).

86. Lampe, X. et al. An ultraconserved Hox-Pbx responsive element resides in the coding sequence of Hoxa2 and is active in rhombomere 4. *Nucleic Acids Res.* **36**, 3214–3225 (2008).

87. Bahr, C. et al. A Myc enhancer cluster regulates normal and leukaemic haematopoietic stem cell hierarchies. *Nature* **553**, 515–520 (2018).

88. Herranz, D. et al. A NOTCH1-driven MYC enhancer promotes T cell development, transformation and acute lymphoblastic leukemia. *Nat. Med.* **20**, 1130–1137 (2014).

89. Lettice, L. A. et al. A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Hum. Mol. Genet.* **12**, 1725–1735 (2003).

90. Cavalheiro, G. R., Pollex, T. & Furlong, E. E. To loop or not to loop: what is the role of TADs in enhancer function and gene regulation? *Curr. Opin. Genet. Dev.* **67**, 119–129 (2021).

91. Despang, A. et al. Functional dissection of the Sox9-Kcnj2 locus identifies nonessential and instructive roles of TAD architecture. *Nat. Genet.* **51**, 1263–1271 (2019).

92. Zufferey, M., Tavernari, D., Oricchio, E. & Ciriello, G. Comparison of computational methods for the identification of topologically associating domains. *Genome Biol.* **19**, 217 (2018).

93. Javierre, B. M. et al. Lineage-specific genome architecture links enhancers and non-coding disease variants to target gene promoters. *Cell* **167**, 1369–1384.e19 (2016).

94. Galupa, R. et al. A conserved noncoding locus regulates random monoallelic Xist expression across a topological boundary. *Mol. Cell* **77**, 352–367.e8 (2020).

95. Finn, E. H. et al. Extensive heterogeneity and intrinsic variation in spatial genome organization. *Cell* **176**, 1502–1515.e10 (2019).

96. Zuin, J. et al. Nonlinear control of transcription through enhancer-promoter interactions. *Nature* **604**, 571–577 (2022).

97. Gentile, C. et al. PRC2-associated chromatin contacts in the developing limb reveal a possible mechanism for the atypical role of PRC2 in HoxA gene expression. *Dev. Cell* **50**, 184–196.e4 (2019).

98. Deng, W. et al. Reactivation of developmentally silenced globin genes by forced chromatin looping. *Cell* **158**, 849–860 (2014).

99. Deng, W. & Blobel, G. A. Manipulating nuclear architecture. *Curr. Opin. Genet. Dev.* **25**, 1–7 (2014).

100. Yokoshi, M., Segawa, K. & Fukaya, T. Visualizing the role of boundary elements in enhancer-promoter communication. *Mol. Cell* **78**, 224–235.e5 (2020).

101. Danielian, P. S., Muccino, D., Rowitch, D. H., Michael, S. K. & McMahon, A. P. Modification of gene activity in mouse embryos in utero by a tamoxifen-inducible form of Cre recombinase. *Curr. Biol.* **8**, 1323–1326 (1998).

102. Luche, H., Weber, O., Nageswara Rao, T., Blum, C. & Fehling, H. J. Faithful activation of an extra-bright red fluorescent protein in "knock-in" Cre-reporter mice ideally suited for lineage tracing studies. *Eur. J. Immunol.* **37**, 43–53 (2007).

103. Concordet, J. P. & Haeussler, M. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Res.* **46**, W242–W245 (2018).

104. Buenrostro, J. D., Wu, B., Chang, H. Y. & Greenleaf, W. J. ATAC-seq: a method for assaying chromatin accessibility genome-wide. *Curr. Protoc. Mol. Biol.* **109**, 21.29.1–21.29.9 (2015).

105. Schoenfelder, S., Javierre, B. M., Furlan-Magaril, M., Wingett, S. W. & Fraser, P. Promoter capture Hi-C: high-resolution, genome-wide profiling of promoter interactions. *J. Vis. Exp.* 57320 (2018).

106. Conlon, R. A. & Rossant, J. Exogenous retinoic acid rapidly induces anterior ectopic expression of murine Hox-2 genes in vivo. *Development* **116**, 357–368 (1992).

107. Duboule, D. & Dolle, P. The structural and functional organization of the murine HOX gene family resembles that of Drosophila homeotic genes. *EMBO J.* **8**, 1497–1505 (1989).

108. Ren, S. Y., Pasqualetti, M., Dierich, A., Le Meur, M. & Rijli, F. M. A Hoxa2 mutant conditional allele generated by Flp- and Cre-mediated recombination. *Genesis* **32**, 105–108 (2002).

109. Gaidatzis, D., Lerch, A., Hahne, F. & Stadler, M. B. QuasR: quantification and annotation of short reads in R. *Bioinformatics* **31**, 1130–1132 (2015).

110. Kent, W. J. et al. The human genome browser at UCSC. *Genome Res.* **12**, 996–1006 (2002).

111. Lopez-Delisle, L. et al. pyGenomeTracks: reproducible plots for multivariate genomic datasets. *Bioinformatics* **37**, 422–423 (2021).

112. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).

113. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).

114. Langmead, B., Wilks, C., Antonescu, V. & Charles, R. Scaling read aligners to hundreds of threads on general-purpose processors. *Bioinformatics* **35**, 421–432 (2019).

115. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008(2021).

116. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

117. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).

118. Wingett, S. et al. HiCUP: pipeline for mapping and processing Hi-C data. *F1000Res* **4**, 1310 (2015).

119. Durand, N. C. et al. Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Syst.* **3**, 95–98 (2016).

120. Abdennur, N. & Mirny, L. A. Cooler: scalable storage for Hi-C data and other genomically labeled arrays. *Bioinformatics* **36**, 311–316 (2020).

121. Wolff, J. et al. Galaxy HiCExplorer 3: a web server for reproducible Hi-C, capture Hi-C and single-cell Hi-C data analysis, quality control and visualization. *Nucleic Acids Res.* **48**, W177–W184 (2020).

122. Wolff, J. et al. Galaxy HiCExplorer: a web server for reproducible Hi-C data analysis, quality control and visualization. *Nucleic Acids Res.* **46**, W11–W16 (2018).

123. Ramirez, F. et al. High-resolution TADs reveal DNA sequences underlying genome organization in flies. *Nat. Commun.* **9**, 189 (2018).

124. Cairns, J. et al. CHiCAGO: robust detection of DNA looping interactions in Capture Hi-C data. *Genome Biol.* **17**, 127 (2016).

125. Freire-Pritchett, P. et al. Detecting chromosomal interactions in Capture Hi-C data with CHiCAGO and companion tools. *Nat. Protoc.* **16**, 4144–4176 (2021).

126. Ben Zouari, Y., Molitor, A. M., Sikorska, N., Pancaldi, V. & Sexton, T. ChiCMaxima: a robust and simple pipeline for detection and visualization of chromatin looping in Capture Hi-C. *Genome Biol.* **20**, 102 (2019).

127. Castro-Mondragon, J. A. et al. JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* **50**, D165–D173 (2022).

128. Grant, C. E., Bailey, T. L. & Noble, W. S. FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**, 1017–1018 (2011).

129. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).

130. Lun, A. T. & Smyth, G. K. csaw: a Bioconductor package for differential binding analysis of ChIP-seq data using sliding windows. *Nucleic Acids Res.* **44**, e45 (2016).

131. Amemiya, H. M., Kundaje, A. & Boyle, A. P. The ENCODE blacklist: identification of problematic regions of the genome. *Sci. Rep.* **9**, 9354 (2019).

Article

132. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016).
133. Khan, A. et al. JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. *Nucleic Acids Res.* **46**, D260–D266 (2018).
134. Blanchette, M. et al. Aligning multiple genomic sequences with the threaded blockset aligner. *Genome Res.* **14**, 708–715 (2004).

Acknowledgements

Author contributions

S.K., M.M., and F.M.R. conceived the study, designed experiments, and analyzed experimental data. S.K. performed most of the experiments. H.K. carried out cell sorting. S.K. and M.B.S. performed motif enrichment analysis. S.K. and J.W. performed alignments and QC of ChIP-seq, RNA-seq, and ATAC-seq data. F.R. performed a super-enhancer analysis. Y.B.Z. performed alignments and QC of Hi-C and promoter-capture Hi-C data and carried out differential gene expression analysis of RNA-seq and analysis of Hi-C and promoter-capture Hi-C data. J.W. analysed Hi-C data from mESC and generated multiple plots and TAD separation scores. O.J. performed promoter-capture Hi-C. S.D. generated CRISPR/Cas9 knockout mice. N.V. carried out in situ hybridizations on sections. A.S. carried out Hi-C in mESCs. M.M., S.K., and F.M.R. analyzed HIRE1 and HIRE2 mutant phenotypes. M.M. and S.K. wrote the first draft, further revised by O.J. and Y.B.Z. F.M.R. revised and wrote the final manuscript. All authors commented on the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-023-38953-0.

Correspondence and requests for materials should be addressed to Filippo M. Rijli.

Peer review information *Nature Communications* thanks Guillaume Andrey, J Crump and the other anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
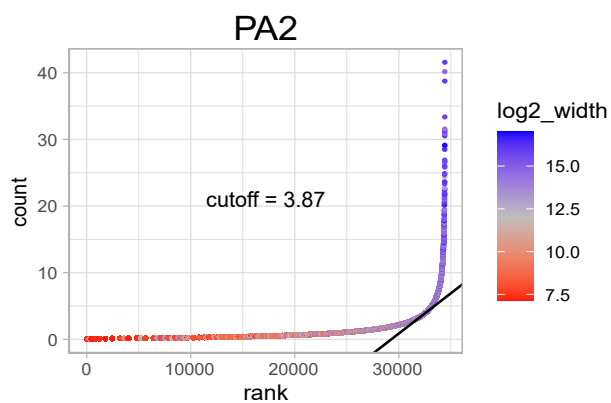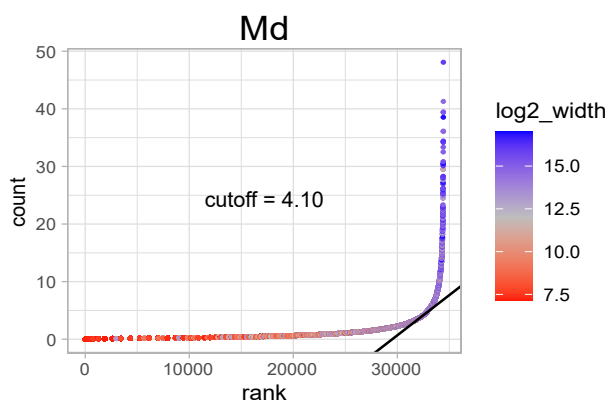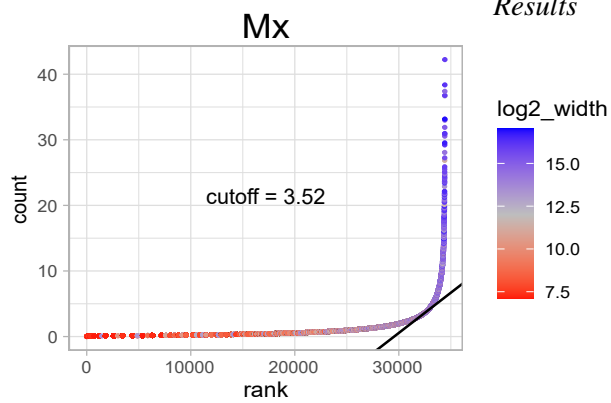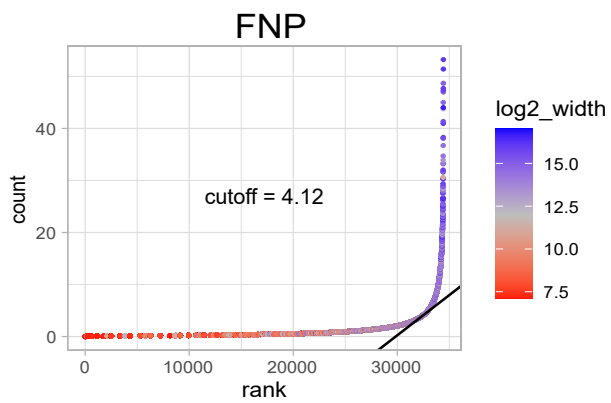
22

57

A multiple super-enhancer region establishes inter-TAD interactions and controls *Hoxa* function in cranial neural crest

Kessler et al.

Supplementary Information

**Supplementary Fig. 1 Identification of craniofacial super-enhancers.**

Hockey stick plots displaying the H3K27 acetylation levels within merged peaks as counts per million on the y-axis and the according rank on the x-axis, separately for each cranial neural crest cell (CNCC) subpopulation. The cutoff value for the SE category was determined as described by the ROSE algorithm using the tangent (shown in black) and is indicated in the figure. The color reflects the log2 transformed width of the merged regions.
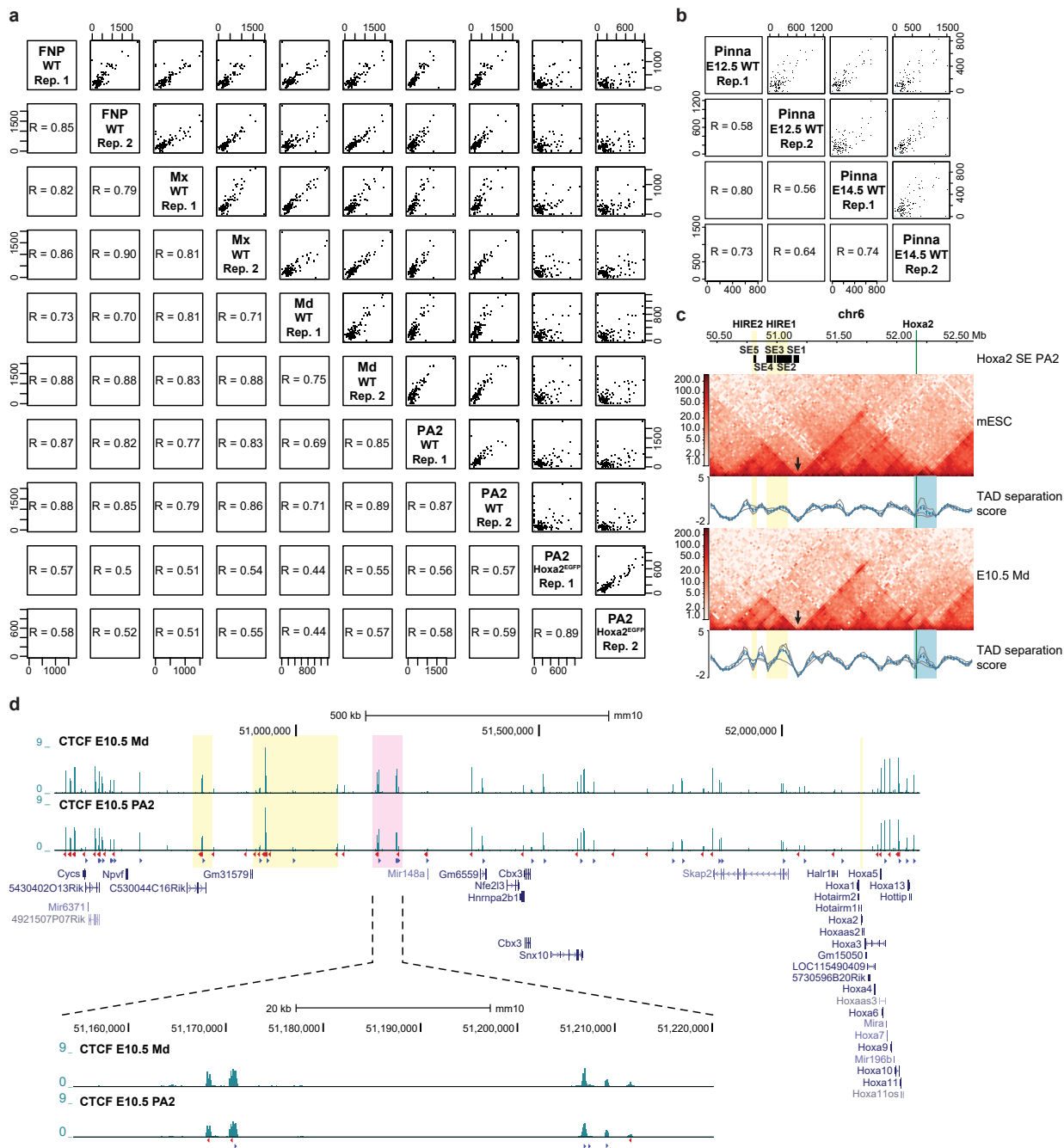
Supplementary Fig. 2 Promoter-capture Hi-C replicates, Hi-C in mESC and Md, and CTCF binding in PA2 and Md at E10.5.

a. Pairwise correlation scatter plots with Pearson's correlation coefficients between promoter capture Hi-C (PCHi-C) replicates from wild type (WT) frontonasal (FNP), maxillary (Mx), mandibular (Md), and second pharyngeal arch (PA2) derived cranial neural crest cells (CNCCs), as well as *Hoxa2*$^{EGFP/EGFP}$ PA2 derived CNCCs at E10.5. **b.** Pairwise correlation scatter plots with Pearson's correlation coefficients between PCHi-C replicates of pinna derived CNCCs at E12.5 and E14.5. **c.** Hi-C interaction heatmaps at 25 kb resolution in a 2.2 Mb region of chromosome 6 including the *Hoxa* cluster (50,442,417-52,636,150 bp) in mouse embryonic stem cells (mESC) and in Md derived CNCCs at E10.5. TAD separation scores are called with HiCExplorer's hicFindTADs and are shown as blue lines. The additional gray lines show the TAD score for different window sizes. The blue highlight marks the polycomb domain encompassing the *Hoxa* cluster and *Evx1* (chr6:52,145,433-52,327,518). Arrows indicate the location of the TAD boundary between *Hoxa2* and HIRE1/HIRE2. **d.** CTCF ChIP-seq tracks in Md and PA2 derived CNCCs at E10.5. CTCF motif orientation at CTCF peaks is indicated with red (minus strand) and blue (plus strand) arrowheads. The pink box highlights the CTCF binding sites at the TAD boundary between *Hoxa2* and HIRE1 / HIRE2. A zoom in on the TAD boundary is shown at the bottom. In panels **c** and d, yellow boxes highlight HIRE1/HIRE2 and a green (**c**) or yellow (**d**) line highlights the *Hoxa2* locus.
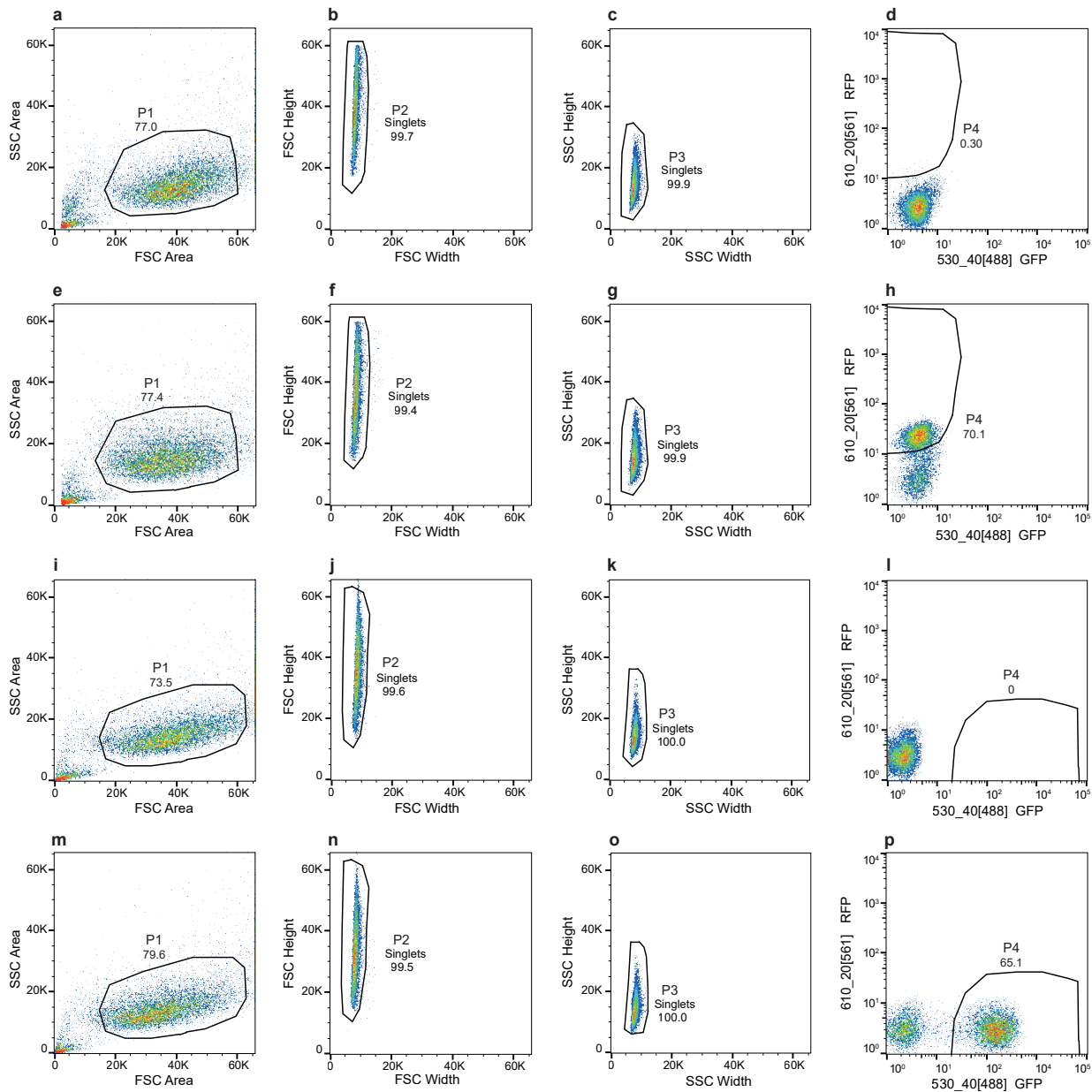
Supplementary Fig. 3 Fluorescence activated cell sorting (FACS) gating strategy.

a-p. FACS of Red Fluorescent Protein (RFP)-positive cranial neural crest cells (CNCCs) from *Wnt1::Cre*;*ROSA^{RFP}*embryos (**a-h**) and FACS of Green Fluorescent Protein (GFP)-positive cells from the second arch of E10.5*Hoxa2^{EGFP/EGFP}* embryos (**i-p**). Living cells are first discriminated from debris by scatter areas P1 (FSC Area versus SSC Area) (**a**, **e**, **i**, **m**). Single cells are discriminated from doublets by firstly assessing forward scatter (FSC Width versus FSC Height; P2) (**b**, **f**, **j**, **n**) and next by assessing side scatter (SSC Width versus SSC Height; P3) (**c**, **g**, **k**, **o**). Subsequently, cells were sorted based on fluorescent markers of interest, GFP (**h**) or RFP (**p**). The baseline control fluorescent signal is from a wild-type (**d**, **l**) littermate (negative control).

65

Supplementary Fig. 4 Chromatin landscape of HIRE1 and HIRE2.

a. Genome browser view on HIRE1 and HIRE2 (chr6:50,779,104 – 51,097,965) with chromatin accessibility (ATAC-seq, purple) from the frontonasal process (FNP), the maxillary (Mx), the mandibular (Md) and the second pharyngeal arch (PA2) cranial neural crest cells (CNCCs) at E10.5, and from the pinna CNCCs at E12.5 and E14.5. ChIP-seq profiles for H3K27me3 (red), H3K4me2 (green), and H3K27ac (blue) from PA2 CNCCs at E10.5, and from the pinna CNCCs at E12.5 and E14.5 are also shown. At the bottom are binding sites identified by ChIP-seq, for Hoxa2[1], Pbx, and Meis[2] in PA2 at E11.5. Highlighted in yellow are HIRE1 and HIRE2. **b.** Illustration representing HIRE1 [Chr6:50,913,170-51,087,888] and HIRE2 [Chr6:50,789,172-50,828,639] deletions. HIRE1 and HIRE2 regions were deleted in mutant mice using the CRISPR-Cas9 system, with regions preserved after deletion represented in green, and regions deleted represented in black. *HIRE1^{del}* and *HIRE2^{del}* mutations were validated using sanger sequencing.
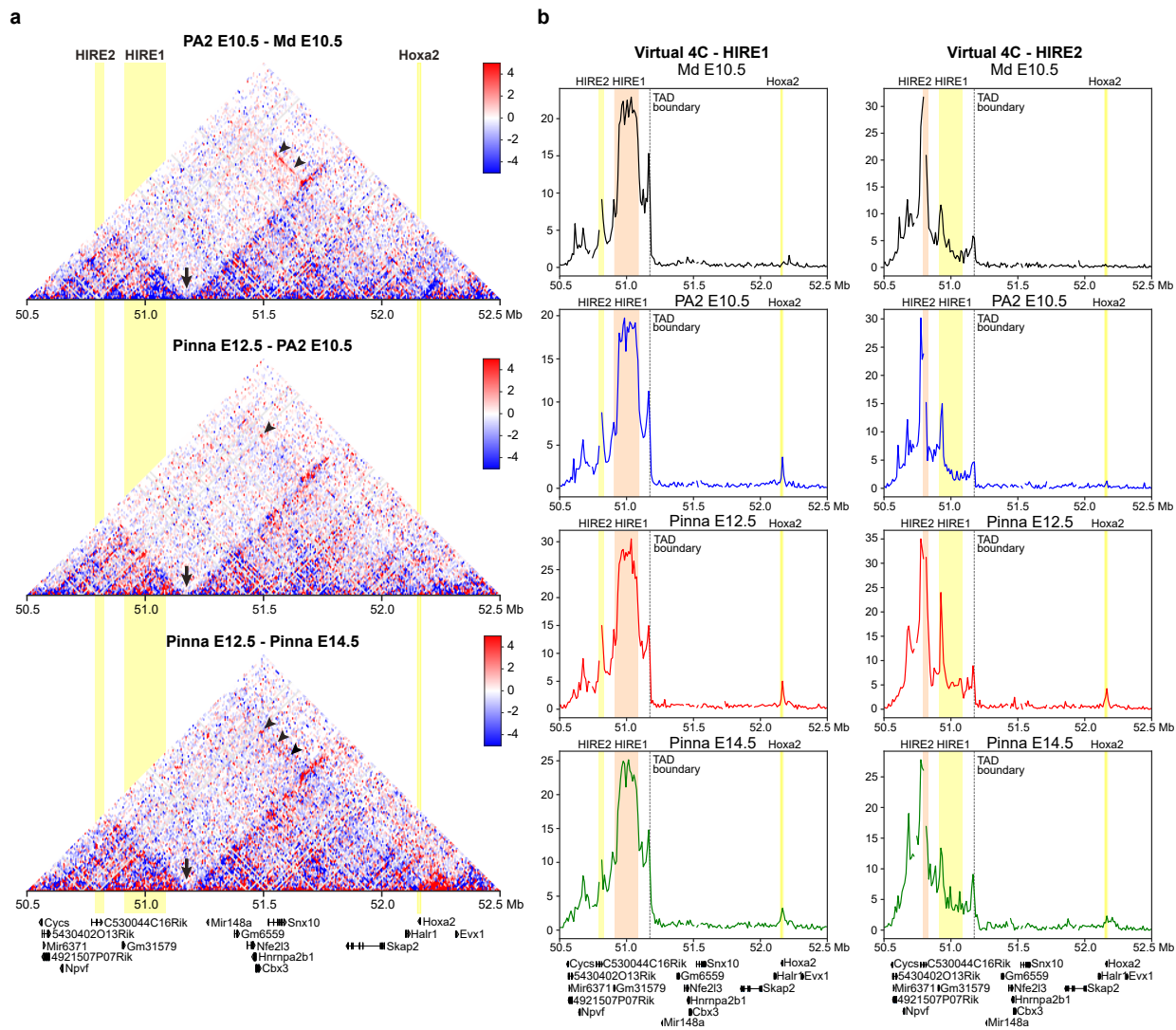
Supplementary Fig. 5 Interactions at HIRE1 and HIRE2

a. Differential HiC maps at 10 kb resolution are shown for a ~2Mb region of chromosome 6 including the *Hoxa* cluster, comparing interactions between E10.5 second pharyngeal arch (PA2) and E10.5 mandibular (Md) cranial neural crest cells (CNCCs) (top), between E12.5 pinna CNCCs and E10.5 PA2 CNCCs (middle), and between E12.5 pinna CNCCs and E14.5 pinna CNCCs (bottom). Shown below are the positions of genes. For better readability, only the genomic position of the *Hoxa2* gene is shown within the *Hoxa* cluster. The yellow rectangles show the HIRE1/HIRE2 regions and *Hoxa2* gene. The arrowheads on the differential map highlight the "stripe" or "spots" of increased interactions between *Hoxa2* and HIRE1/HIRE2 (red) in E10.5 PA2 as compared to Md (top) in E12.5 pinna as compared to E10.5 PA2 (middle) and in E12.5 pinna as compared to E14.5 pinna (bottom). A black arrow between two TADs indicates the presence of the TAD border. **b.** Virtual 4C plots from HIRE1 (left side) and HIRE2 (right side) for E10.5 Md, E10.5 PA2, E12.5 pinna and E14.5 pinna CNCCs, derived from Hi-C data. The orange rectangles highlight the viewpoints whereas the yellow rectangles highlight the *Hoxa2* gene, HIRE2 (left side) and HIRE1 (right side).
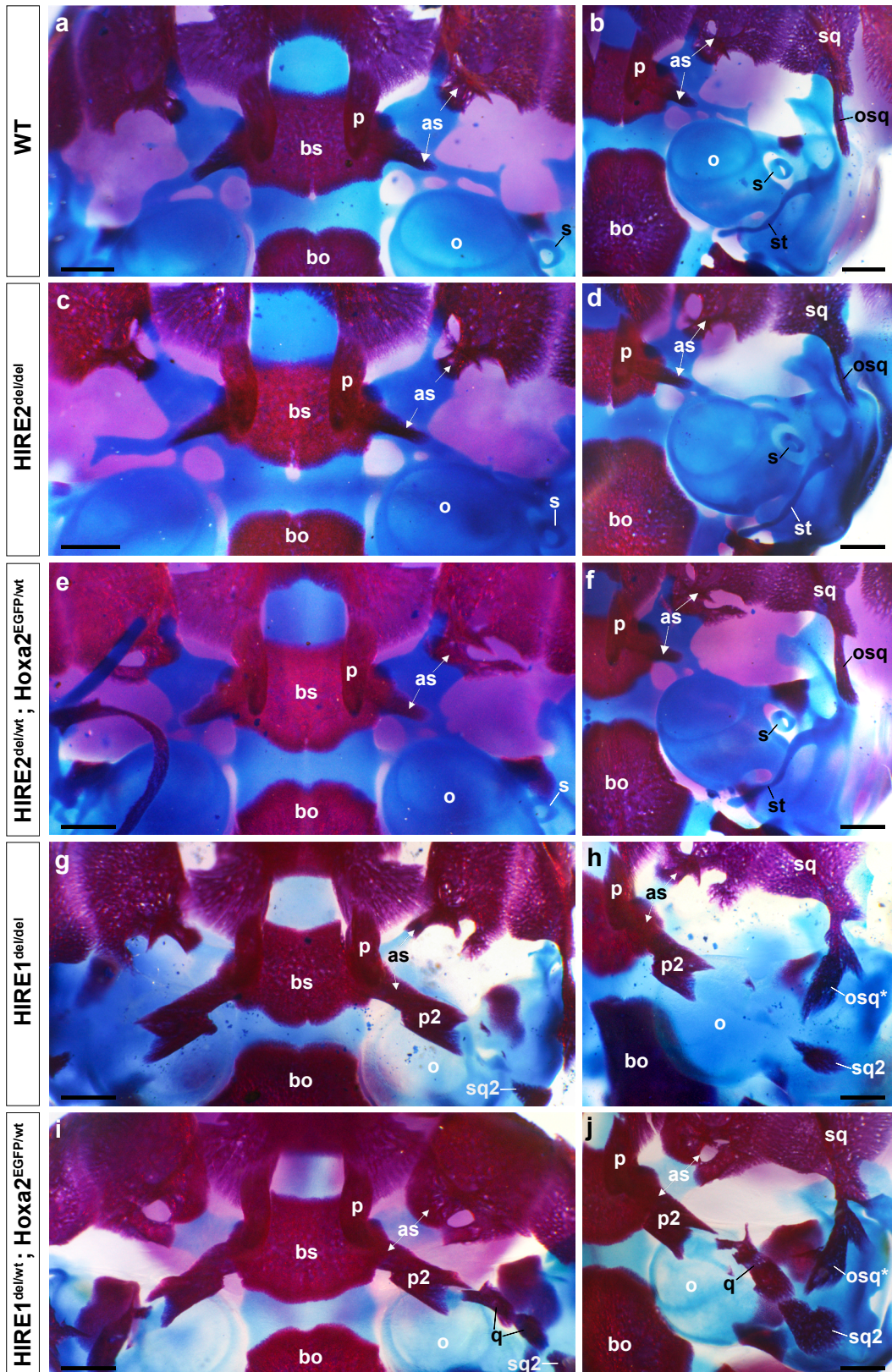
Supplementary Fig. 6 Cranial skeletal changes in HIRE1 and HIRE2 mutant fetuses.

Ventral (**a, c, e, g, i**) and ventrolateral (**b, d, f, h, j**) views of the cranial base of E18.5 wild-type (WT) (**a, b**), *HIRE2$^{del/del}$* homozygous mutant (**c, d**), *HIRE2$^{del/wt}$;Hoxa2$^{EGFP/wt}$* trans-heterozygote mutant (**e, f**), *HIRE1$^{del/del}$* homozygous mutant (**g, h**) and *HIRE1$^{del/wt}$;Hoxa2$^{EGFP/wt}$* trans-heterozygous mutant (**i, j**) fetuses. **a** to **j** are representative images of n=4/4 WT (**a, b**), n=6/6 *HIRE2$^{del/del}$* (**c, d**), n=4/4 *HIRE2$^{del/wt}$;Hoxa2$^{EGFP/wt}$* (**e, f**), n=4/4 *HIRE1$^{del/del}$* (**g, h**) and n=4/4 *HIRE1$^{del/wt}$;Hoxa2$^{EGFP/wt}$* (**i, j**) fetuses. as represents the alisphenoid bone; bo represents the basioccipital bone; bs represents the basisphenoid bone; o represents the cartilaginous otic capsule; p represents the pterygoid bone; p2 represents the ectopic pterygoid bone; osq and osq* represent respectively the normal and modified retrotympanic (otic) process of the squamosal bone; s represents the stapes; sq represents the squamosal bone; sq2 represents the ectopic squamosal bone; st represents the styloid process; q represents the atavistic quadrate bone, ectopically appearing in *Hoxa2$^{-/-}$* mutants[3] and in the *HIRE1$^{del/wt}$;Hoxa2$^{EGFP/wt}$* trans-heterozygous mutants. Scale bars represent 500 µm.
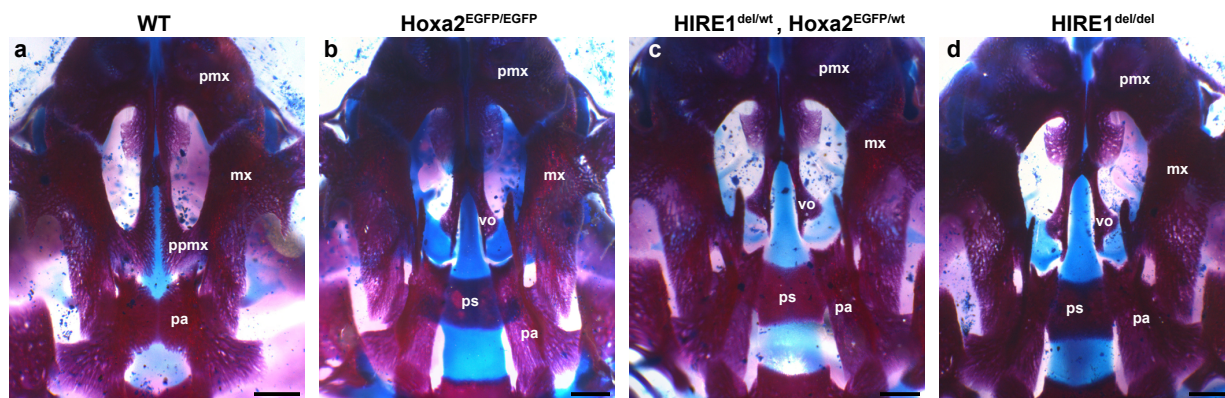
Supplementary Fig. 7 Cleft secondary palate in HIRE1 mutant fetuses

Skeletal preparations of the palate area after the removal of the lower jaw from E18.5 wild-type (WT) (**a**), *Hoxa2$^{EFGP/EGFP}$* (**b**), *HIRE1$^{del/wt}$;Hoxa2$^{EGFP/wt}$* (**c**) and *HIRE1$^{del/del}$* (**d**) fetuses. **a** to **d** are representative images of WT (n=4/4) (**a**), *Hoxa2$^{EFGP/EGFP}$* (n=6/8) (**b**), *HIRE1$^{del/wt}$;Hoxa2$^{EGFP/wt}$* (n=14/24) (**c**) and *HIRE1$^{del/del}$* (n=8/10) (**d**) fetuses. The nose is toward the top. mx represents the maxilla, pa the palatal bone, pmx the premaxilla, ppmx the palatal process of maxilla, ps the presphenoid bone, and vo the vomer. Note that vo and ps in (b, c, d) are not visible in (a) because they are underneath Mx and Pa. Scale bars represent 500 µm.
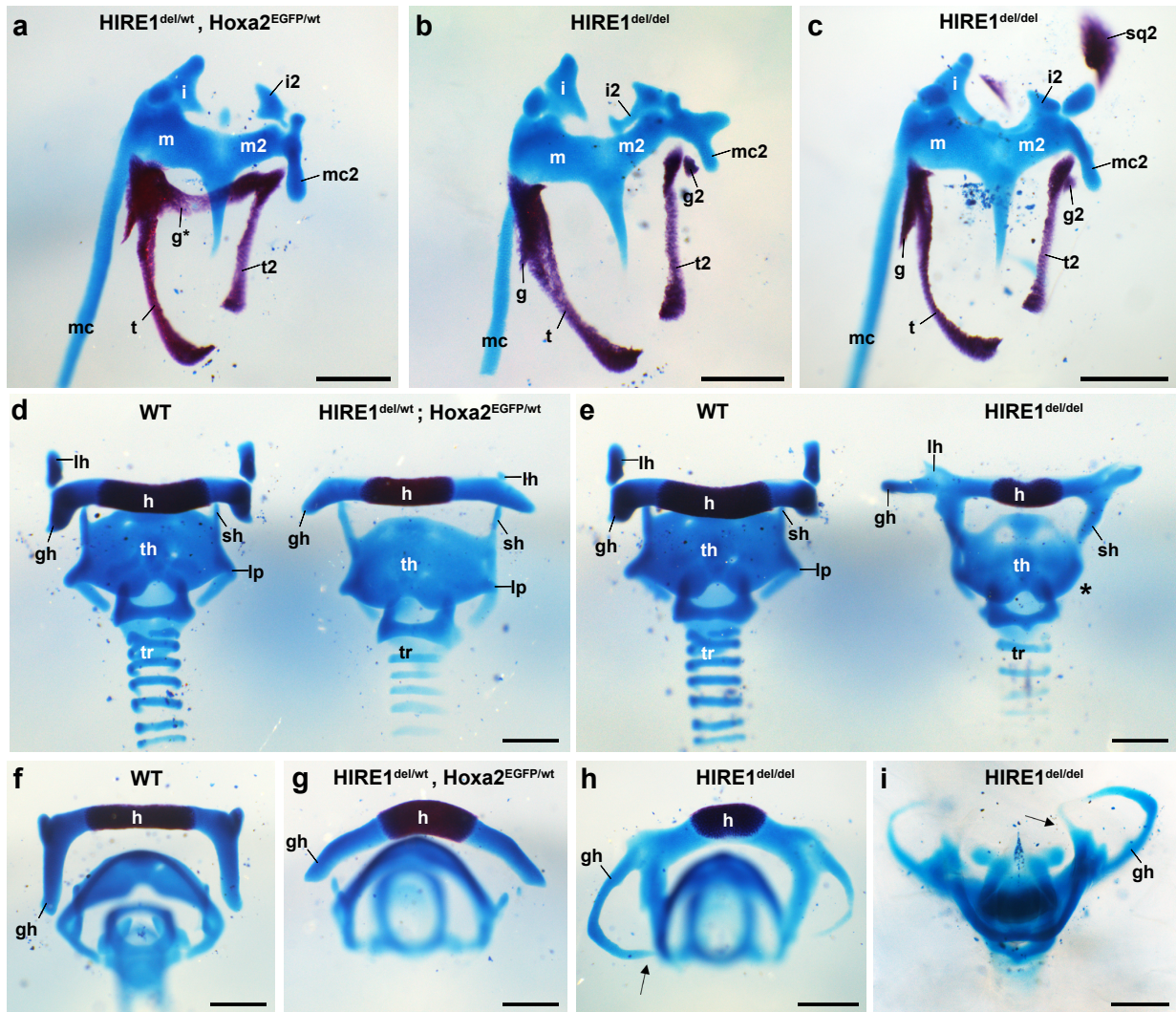
Supplementary Fig. 8 Middle ear and hyoid bone skeletal phenotypes upon deletion of HIRE1
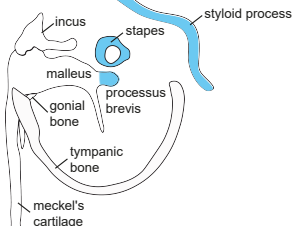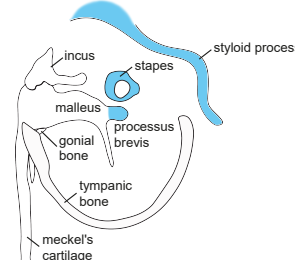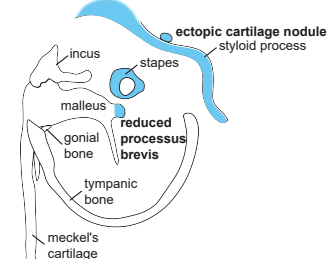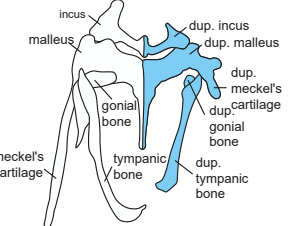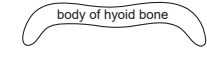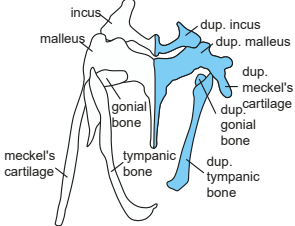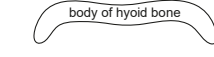
Middle ear (**a-c**) and hyoid (**d-i**) skeletal preparations from E18.5 wild-type (WT) (**d, f**), *HIRE1^{del/del}* homozygous mutant (**b, c, e, h, i**) and *HIRE1^{del/wt};Hoxa2^{EGFP/wt}* trans-heterozygous mutant (**a, d, g**) fetuses. In **a** and **b**, the absence of fusion between the incus and its duplicated counterpart is representative of n=1/8 fetus sides for *HIRE1^{del/wt};Hoxa2^{EGFP/wt}* (**a**) and n=1/8 fetus sides for *HIRE1^{del/del}* (**b**). g* represents the transformed gonial bone, as described in *Hoxa2-/-* mutant fetuses [3]. The transformed gonial bone presented in **a** is representative of n=2/8 *HIRE1^{del/wt};Hoxa2^{EGFP/wt}* fetus sides. **d**, **i** and **j** are representative of WT (n=4/4), *HIRE1^{del/wt};Hoxa2^{EGFP/wt}* (n=4/4) and *HIRE1^{del/del}* (n=4/4) fetuses. In **e**, the asterisk displays the absence of the lateral process (lp) of the laryngeal cartilage in *HIRE1^{del/del}* homozygous mutant fetuses. The black arrows in **h** and **i** display the fusion between the elongated greater horn (gh) of the hyoid bone and the posterior border of the thyroid cartilage (th) observed in *HIRE1^{del/del}* homozygous mutant fetuses. g and g2 represent the WT and duplicated gonial bones, respectively; h represents the body of the hyoid bone; i and i2 represent the WT and duplicated incus respectively; lh represents the lesser horn of the hyoid bone; m and m2 represent the WT and duplicated mallei, respectively; mc and mc2 represent the WT and partially duplicated Meckel's cartilages respectively; sh represents the superior horn of the thyroid cartilage; sq2 represents the ectopic squamosal bone; t and t2 represent the WT and duplicated tympanic bones respectively; tr represents the trachea. Scale bars represent 500 µm.

E10.5 PA2
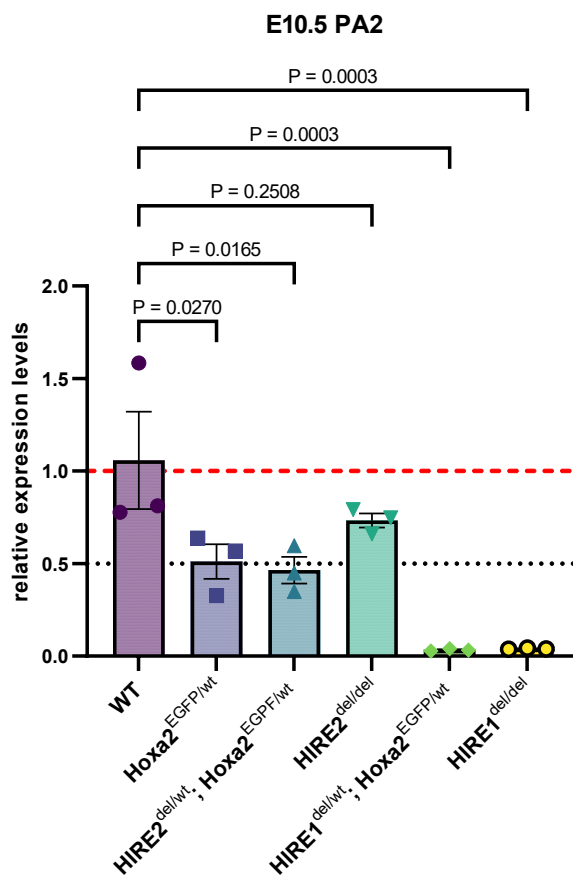
Supplementary Fig. 9 *Hoxa2* down-regulation in E10.5 second pharyngeal arch upon HIRE1 or HIRE2 deletion.

Relative expression levels of *Hoxa2,* detected by qRT-PCR analysis, in E10.5 second pharyngeal arch (PA2) of WT and *Hoxa2$^{EGFP/wt}$*, *HIRE2$^{del/wt}$;Hoxa2$^{EGFP/wt}$,* *HIRE2$^{del/del}$,* *HIRE1$^{del/wt}$;Hoxa2$^{EGFP/wt}$,* and *HIRE1$^{del/del}$* mutant embryos (n = 3 per genotype). Data are represented as means ± SEM. Significance is based on a one-way ANOVA followed by Dunnett's multiple comparisons test with WT as the control sample.

PA2 CNCC-derived skeletal and cartilaginous structures

Control

Hoxa2 transcript levels
in PA2 at E10.5:
WT : 100 %
Hoxa2$^{EGFP/wt}$: about 50 %
(See supplementary figure 9)

Middle ear

incus
stapes
styloid process
malleus
processus brevis
gonial bone
tympanic bone
meckel's cartilage

(See figure 6a, 6b)

Hyoid

lesser horn of the hyoid bone
body of hyoid bone

(See figure 6c)

Pinna
(See figure 5b, 5e, 5f)

■ PA2 derived

HIRE2$^{del/del}$

Hoxa2 transcript levels
in PA2 at E10.5:
about 75% compared to WT
(See supplementary figure 9)

Middle ear

incus
stapes
styloid process
malleus
processus brevis
gonial bone
tympanic bone
meckel's cartilage

No visible malformations
(See figure 6d, 6e)

Hyoid

lesser horn of the hyoid bone
body of hyoid bone

No visible malformation
(See figure 6f)

Pinna

No visible malformations
at birth or in adult animals
(See figure 5c)

■ PA2 derived

HIRE2$^{del/wt}$;Hoxa2$^{EGFP/wt}$

Hoxa2 transcript levels
in PA2 at E10.5:
about 46% compared to WT
(See supplementary figure 9)

Middle ear

incus
stapes
malleus
ectopic cartilage nodule styloid process
reduced processus brevis
gonial bone
tympanic bone
meckel's cartilage

(See figure 6g, 6h)

Hyoid

lesser horn of the hyoid bone
body of hyoid bone

No visible malformations
(See figure 6i)

Pinna

No visible malformations at birth
Microtia in adult animals
(See figure 5d, 5e, 5g)

■ PA2 derived

HIRE1$^{del/del}$

Hoxa2 transcript levels
in PA2 at E10.5:
about 0-4% compared to WT
(See supplementary figure 9)

Middle ear

incus
malleus
dup. incus
dup. malleus
dup. meckel's cartilage
dup. gonial bone
gonial bone
tympanic bone
dup. tympanic bone
meckel's cartilage

Homeotic transformation of PA2
skeletal elements to PA1-like morphology
(See figure 6j, 6k)

Hyoid

body of hyoid bone

Absence of the lesser horns
(See figure 6l)

Pinna
Absent
(See figure 5h)

■ PA2 derived

HIRE1$^{del/wt}$;Hoxa2$^{EGFP/wt}$

Hoxa2 transcript levels
in PA2 at E10.5:
about 0-3% compared to WT
(See supplementary figure 9)

Middle ear

incus
malleus
dup. incus
dup. malleus
dup. meckel's cartilage
dup. gonial bone
gonial bone
tympanic bone
dup. tympanic bone
meckel's cartilage

Homeotic transformation of PA2
skeletal elements to PA1-like morphology
(See figure 6m, 6n)

Hyoid

body of hyoid bone

Absence of the lesser horns
(See figure 6o)

Pinna
Absent
(See figure 5i)

■ PA2 derived

77

Supplementary Fig. 10 Morphological changes of second arch-derived skeletal and cartilaginous structures correlate to reduction of *Hoxa2* transcript levels.

(Left) *Hoxa2* transcript levels in E10.5 second pharyngeal arch (PA2) of *Hoxa2$^{EGFP/wt}$*, *HIRE2$^{del/del}$*, *HIRE2$^{del/wt}$;Hoxa2$^{EGFP/wt}$* , *HIRE1$^{del/del}$* and *HIRE1$^{del/wt}$;Hoxa2$^{EGFP/wt}$* mutant embryos, as compared to wild-type (WT), and obtained by qRT-PCR analysis (see suppl. fig. 9). (Right) For each genotype, the corresponding phenotype at PA2 cranial neural crest cells (CNCC)-derived skeletal and cartilaginous structures (blue) is represented. Note that no visible malformations are observed in *Hoxa2$^{EGFP/wt}$* fetuses.
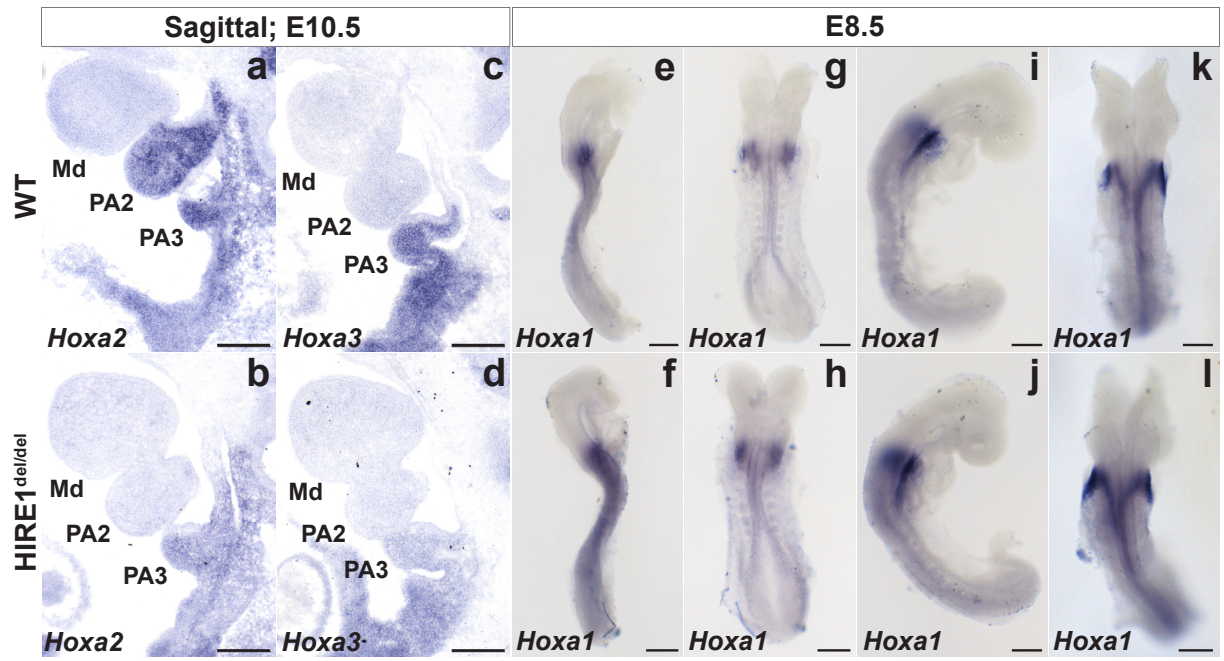
Supplementary Fig. 11 Effect of HIRE1 deletion on anterior *Hoxa* gene expression.

In-situ hybridization on sagittal sections of wild-type (WT) (**a**, **c**) and *HIRE1*[^del/del] (**b**, **d**) embryos at E10.5 using *Hoxa2* (**a**, **b**) and *Hoxa3* (**c**, **d**) antisense probes. **a-d** are representative images of serial sections from n=4 WT and n=4 *HIRE1*[^del/del] embryos. *Hoxa2* expression is undetectable in the second pharyngeal arch (PA2) and strongly reduced in the third pharyngeal arch (PA3) of *HIRE1*[^del/del] embryos (**b**) compared to WT (**a**). *Hoxa3* expression is strongly reduced in PA3 of *HIRE1*[^del/del] embryos (**d**) compared to WT (**c**). Whole mount *in-situ* hybridization on WT (**e, g, i, k**; n=5 embryos) and *HIRE1*[^del/del] (**f, h, j, l**; n=2 embryos) embryos at E8.5 using *Hoxa1* antisense probes. No differences in *Hoxa1* expression could be observed between WT and *HIRE1*[^del/del] embryos. Md represents the mandibular process. Scale bars represent 200 µm.
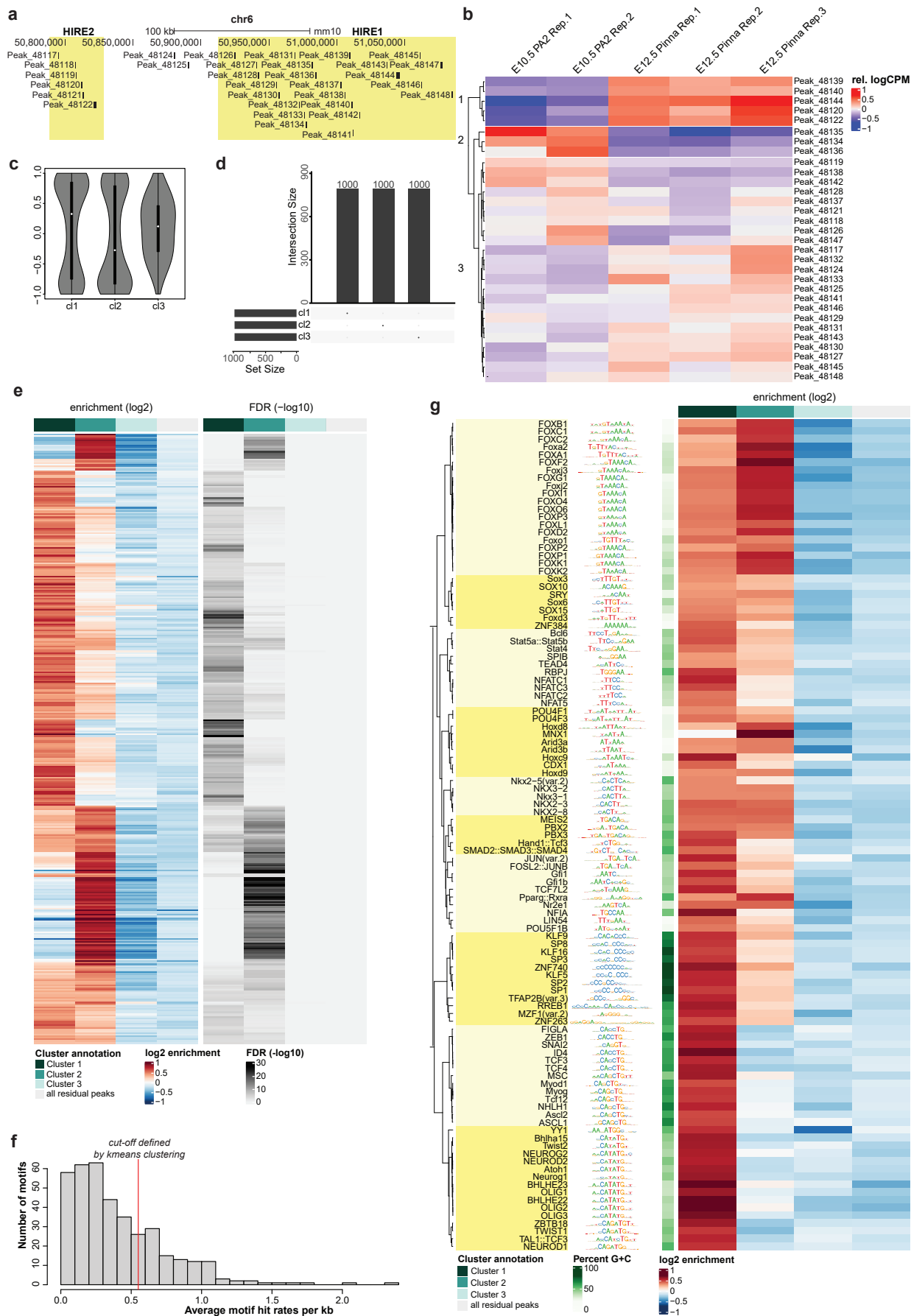
Supplementary Fig.12 Transcription factor motif enrichment analysis.

a. Schematic representation of a genomic window, depicting the locations and names of 31 merged ATAC-seq peaks overlapping HIRE1 and HIRE2 from the second arch (PA2) derived CNCCs at E10.5 and from pinna derived CNCCs at E12.5. **b.** Clustering of ATAC-seq peaks overlapping HIRE1 and HIRE2 (chr6:50789172-51087888, mouse GRCm38/mm10) from PA2 derived CNCCs at E10.5 and from pinna derived CNCCs at E12.5 according to their relative accessibility levels at E10.5 and E12.5. Peaks were classified into three groups using hierarchical clustering. **c.** Violin plot depicting the distribution of Pearson's correlation coefficients of the ATAC-seq signal at all ATAC-seq peaks in E10.5 and E12.5 (n = 133466 ATAC-seq peaks) to the average accessibility profile of each of the three clusters. For clusters 1 and 2, the distribution of correlation coefficients was bimodal, corresponding to peaks that were highly correlated or anti-correlated with the profiles identified from HIRE2 to HIRE1 peaks. Plots extend from the data minima to the maxima; the white dot indicates the median, the box shows the interquartile range and whiskers extend to the most extreme data points.**d.** UpSet plot showing the top 1000 ATAC-seq peaks (ranked by their correlation coefficient) for each of the three clusters. There is no overlap between the top 1000 peaks for each cluster. **e.** Transcription factor motif enrichment heatmap. Shown are the log2-fold enrichments and -log10 (FDR) values for all transcription factor motifs with an FDR less than 0.001 in any of the three clusters (n = 382). Each row depicts one transcription factor motif. While there are many significantly enriched motifs for clusters 1 and 2, no specific motifs are enriched in cluster 3 or all residual ATAC-seq peaks. **f.** Histogram of the average motif hit rates per kilobase (kb). The 31 ATAC-seq peaks overlapping HIRE1 and HIRE2 were scanned for motif hits using all transcription factor motifs with an FDR less than 0.001 in any of the three clusters as input (see Methods). To select motifs with a high frequency of predicted binding sites, the average motif hit rates per kb were calculated and clustered into two sets with k-means clustering (k=2). All motifs from the set with higher hit-rates (n = 107) were used in a motif similarity analysis. **g.** Transcription factor motif enrichment heatmap. Shown are the log2-fold enrichments for all 107 transcription factor motifs with high hit-rates from **f** and an FDR less than 0.001 in any of the three clusters. The rows are ordered according to the motif similarity analysis shown in Supplementary Fig. 13. and yellow highlights indicate the different motif similarity clusters defined in Supplementary Fig. 13.
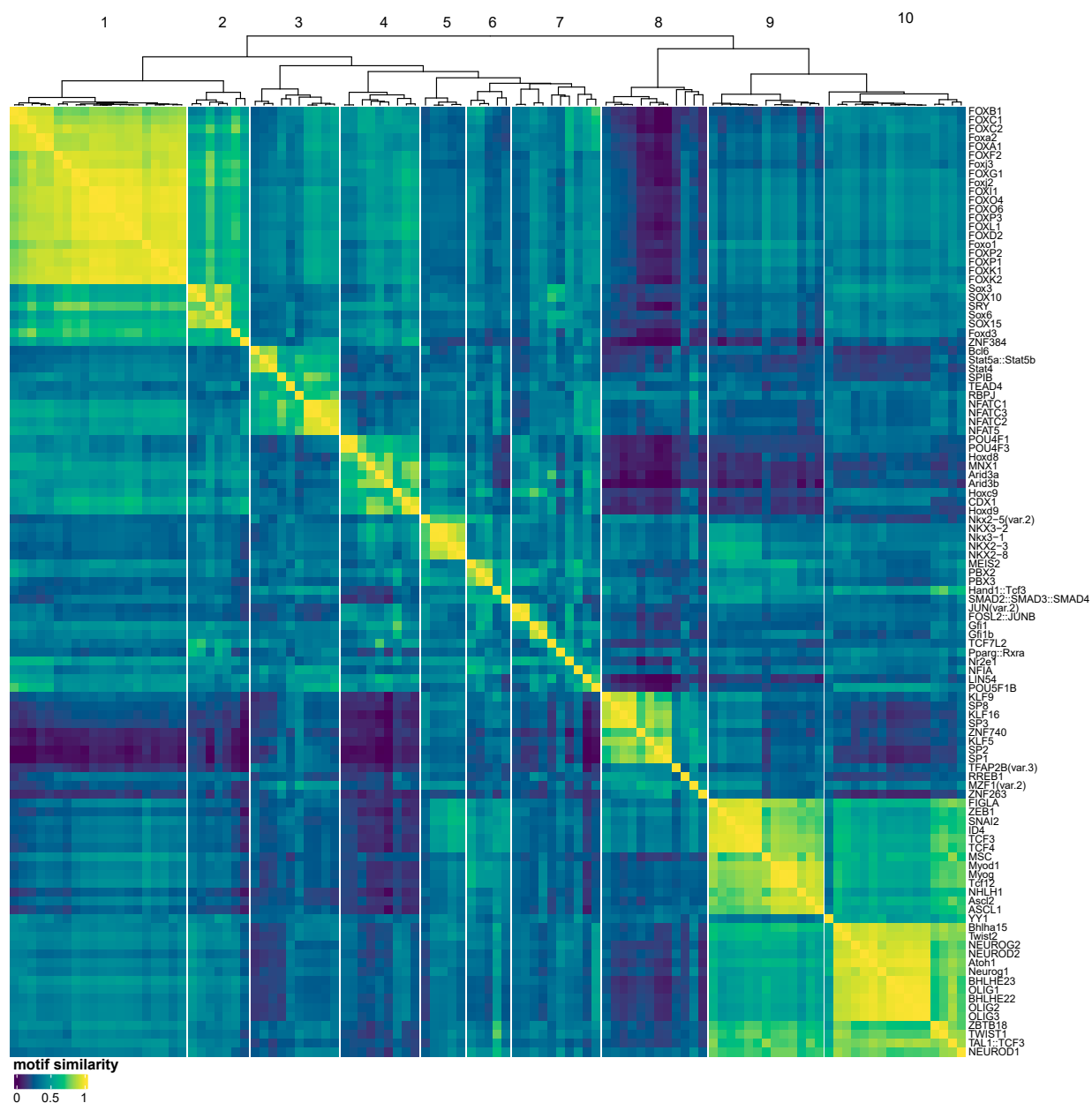
motif similarity
0   0.5   1

Supplementary Fig. 13 Motif similarity analysis. All 107 transcription factor motifs with high hit-rates defined in Supplementary Fig. 12f were used to perform a motif similarity analysis using *monaLisa*[4]. See Methods for details.

Supplementary Table 1 Genomic locations of the 2,232 identified CNCC super-enhancers

The table contains the genomic locations (mm10) of all super-enhancers identified in CNCCs at E10.5 based on their H3K27ac signal.

Supplementary Table 2 SEs interacting with genes encoding for transcription factor genes differentially expressed in CNCCs subpopulations at E10.5.

The table contains the genomic coordinates (mm10) of the 147 super-enhancers (SEs) with at least one significant interaction to an expressed positional transcription factor (TF) gene as depicted in figure 1, including the information to which TF each SE is linked. In addition, the mean strengths of all interactions from the positional TF promoter to restriction fragment(s) that overlap with a SE are given for each population. The last column indicates the number of the k-means cluster as shown on the left side of the heatmap. The rows follow the same order as the heatmap of figure 1. Note that the table contains 148 unique SE-promoter pairs (rows) because one SE is shared by 2 different promoters (*Twist2* and *Hes6*), and thus appears twice.

Supplementary Table 3 Differential gene expression analysis. The table provides on five separate sheets the differential gene expression analysis between second pharyngeal arch (PA2) derived CNCCs at E10.5 and pinna derived CNCCs at E12.5, between PA2 derived CNCCs at E10.5 and pinna derived CNCCs at E14.5, between pinna derived CNCCs at E12.5 and pinna derived CNCCs at E14.5, between PA2 cells from *HIRE1*[del/del] and wild-type embryos at E10.5, and between PA3 cells from *HIRE1*[del/del] and wild-type embryos at E10.5. Each sheet contains the following information in the columns: Entrez gene ID, gene width, ENSEMBL gene ID, gene symbol, log2 fold change (column name logFC), mean expression level (log2 counts per million, column name logCPM), quasi-likelihood F-test score (F), raw P value (PValue), and false discovery rate (FDR). P-values and FDR were calculated by *edgeR*[5]. See Methods for details on statistical analysis.

Supplementary Table 4 Transcription factor motif enrichment analysis. The table contains the results of the transcription factor motif analysis. Each row stands for one of the motifs of the JASPAR2018 database. The columns contain information about the motif ID, motif name, motif family, and motif class, as well as the log2-fold enrichments and -log10(FDR) values for each of the ATAC-seq peak clusters (as defined in Supplementary Fig. 12b-d) and all residual peaks. P-values and FDR were calculated by *monaLisa*[4]. See Methods for details.

Supplementary Table 5 Annotation of merged ATAC-seq peaks from PA2 derived CNCCs at E10.5 and pinna derived CNCCs at E12.5 overlapping the region spanning from HIRE2 to HIRE1 with predicted motifs. The number of occurrences of all transcription factor motifs with an FDR less than 0.001 in any of the three cluster profiles (Suppl. Fig. 8D) was predicted within each of the ATAC-seq peaks from second pharyngeal arch (PA2) derived cranial neural crest cells (CNCCs) at E10.5 and pinna derived CNCCs at E12.5, which overlap the region spanning from HIRE2 to HIRE1 (chr6:50789172-51087888, mouse GRCm38/mm10). See Methods for details.

Supplementary Table 6 List of selected representative motifs clustered by motif similarity analysis. Each row in the table stands for one of the motifs identified in the motif enrichment analysis (see Methods). The columns contain information about the motif similarity cluster that a motif was assigned to in Supplementary Fig. 13, the motif ID, motif name, motif family and motif class.

Supplementary Table 7 Sample information. The type of experiments performed on the different cell types, the protocols used for the libraries, the sample names, and for previously published datasets the GEO accession number are indicated.

Supplementary References

1.   Donaldson IJ, *et al.* Genome-wide occupancy links Hoxa2 to Wnt-beta-catenin signaling in mouse embryonic development. *Nucleic Acids Res* **40**, 3990-4001 (2012).

2.   Amin S, *et al.* Hoxa2 selectively enhances Meis binding to change a branchial arch ground state. *Dev Cell* **32**, 265-277 (2015).

3.   Rijli FM, Mark M, Lakkaraju S, Dierich A, Dolle P, Chambon P. A homeotic transformation is generated in the rostral branchial region of the head by disruption of Hoxa-2, which acts as a selector gene. *Cell* **75**, 1333-1349 (1993).

4.   Machlab D, Burger L, Soneson C, Rijli FM, Schubeler D, Stadler MB. monaLisa: an R/Bioconductor package for identifying regulatory motifs. *Bioinformatics*, (2022).

5.   Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140 (2010).

Discussion and Outlook

The results presented in this dissertation give fundamental insights into the importance of long-range SEs in controlling key developmental TFs in craniofacial morphogenesis through genetic perturbation of two newly identified extremely long-range SE regions (HIRE1 and HIRE2) that regulate *Hoxa2* and *Hoxa3* gene expression in CNCCs and hindbrain rhombomeres. Furthermore, the list of predicted SEs that specifically target genes establishing CNCC positional identity during facial development (Minoux et al., 2017) is a valuable resource for future research on regulatory elements involved in craniofacial development.

Here we used the ROSE algorithm (Loven et al., 2013; Whyte et al., 2013) on H3K27ac ChIP-seq data from four CNCCs subpopulations, namely FNP, Mx, Md, and PA2, to predict putative CNCC subpopulation-specific SEs genome-wide. By using PCHi-C profiles from these four subpopulations, we were able to link SEs to previously identified differentially expressed positional TFs involved in craniofacial morphogenesis (Minoux et al., 2017), including five SEs that exclusively target *Hoxa2* in PA2 derived CNCCs over an extremely large genomic distance (>1 Mb). SE1 was not further pursued as it overlaps with the TAD boundary. Remarkably, these SEs are able to interact with *Hoxa2* across a strong TAD boundary that was identified based on both TAD separation scores and visual inspection of Hi-C contact maps. Due to their linear proximity and the distribution of H3K27ac, we grouped SE2-4 into HIRE1, a 175 kb long gene-poor region, while SE5 encompassed a 39kb region, termed HIRE2. Both of these large regulatory elements are highly conserved within eutherian mammals and certain constituent enhancers of the SEs even showed some sequence conservation in fish. Upon deletion of either HIRE1 or HIRE2 in mice, we observed a range of structural abnormalities. These abnormalities partially or fully mimic the skeletal phenotypes caused by the loss of *Hoxa2* (Gendron-Maguire et al., 1993; Rijli et al., 1993) and also resemble some aspects of the *Hoxa3* mutant skeletal phenotypes (Chojnowski et al., 2016; Condie & Capecchi, 1994; Manley & Capecchi, 1995). In addition, we show that HIRE1 and HIRE2 are selectively required for the expression of *Hoxa2* and *Hoxa3* in CNCCs and hindbrain rhombomeres, and provide evidence that *Hoxa2* is partially involved in the maintenance of its own expression likely in concert with co-factors, such as Pbx and Meis.

SEs have been defined as large clusters of enhancers that exhibit an exceptionally high

level of H3K27ac, possess an unusually high density of TF binding sites, demonstrate a greater capacity to initiate transcription when compared to typical enhancers, and are highly enriched in disease-associated genetic variation (Hnisz et al., 2013; Loven et al., 2013; Whyte et al., 2013). Furthermore, they are oftentimes located close to genes that play crucial roles in determining cell identity (Hnisz et al., 2013; Siersbæk et al., 2014; Whyte et al., 2013). However, the SE concept has been put into question by multiple researchers (Blobel et al., 2021), as their prediction via the ROSE algorithm is based on thresholds that have been defined arbitrarily by the Young lab, such as the maximum distance of 12.5 kb between constituent enhancers (Hnisz et al., 2013; Loven et al., 2013; Whyte et al., 2013). Similar to other enhancer predictions based on chromatin marks or chromatin accessibility, the ROSE algorithm does not give any information on the target gene or enhancer function. Hence, there is a high risk of calling false positives and missing enhancers that are non-canonically marked or exceed the arbitrary inter-enhancer distance cut-off (Blobel et al., 2021). While the risk of missing relevant regulatory elements is certainly a limitation of the ROSE algorithm, we show that coupling SE calling with spatial data from 3C-based techniques, such as PCHi-C, may reduce the risk of calling false positive *cis*-regulatory elements, as it specifically assigns SE-gene pairs. In addition, it has been debated, if SEs are indeed a specific type of regulatory domain that is functionally different from typical enhancers, or if they are just clusters of functionally independent enhancers targeting the same gene. There is a wide range of studies that support the notion that SEs are functionally different from typical enhancers showing a functional interdependence between individual constituent enhancers (Dowen et al., 2014; Huang et al., 2018; Jiang et al., 2016; Loven et al., 2013; Proudhon et al., 2016; Shin et al., 2016; Siersbæk et al., 2014; Thomas et al., 2021; Whyte et al., 2013), while at other putative SEs the constituent enhancers appear to act independently in an entirely additive fashion (Hay et al., 2016). Here we investigated the functional role of two putative SEs of *Hoxa2*, HIRE1 and HIRE2, which exhibit SE features, such as strong enrichment of H3K27ac and a high density of TF binding sites. Our functional analysis of these two regions by CRISPR-Cas9 mediated deletion indicates that HIRE1 is indeed a regulatory region essential for the expression of anterior *Hoxa* genes in CNCCs and thereby craniofacial morphogenesis, whereas HIRE2 appears to be dispensable for the morphogenesis of craniofacial structures under unperturbed conditions. Thus, HIRE2 may not be a SE by function, but may be required to buffer *Hoxa2* expression levels in order to mediate robustness against environmental or genetic perturbations similar to other redundant enhancers (Kvon et al., 2021; Long et al., 2020; Osterwalder et al., 2018). While the SE terminology may be questionable, one should not undermine the fact that multiple studies, including ours, were able to functionally

validate the importance of SEs for controlling gene expression of key cell identity-defining genes (Hnisz et al., 2016; Jiang et al., 2016; Proudhon et al., 2016; Shin et al., 2016; Siersbæk et al., 2014; Thomas et al., 2021). Taken together, our findings show that the SE prediction algorithm is able to identify functionally relevant *cis*-regulatory regions, although not all of them might show SE activity. Hence, if a predicted region functions as a SE should be validated on a case-by-case basis.

A number of *cis*-regulatory elements located near or within the *Hoxa2* gene have been discovered in the past, which drive the expression of reporter genes in CNCCs (M. Maconochie et al., 1999; McEllin et al., 2016), R3/R5 (M. K. Maconochie et al., 2001; Nonchev, Maconochie, et al., 1996; Nonchev, Vesque, et al., 1996; Tümpel et al., 2006), R4 (Lampe et al., 2008; Tumpel et al., 2007), and R2 (Frasch et al., 1995; Tumpel et al., 2008). However, none of these enhancers have been functionally validated *in vivo* through e.g. deletion experiments. We show that *Hoxa2* expression in both CNCC and most of the hindbrain rhombomeres is dependent upon the long-range inter-TAD interaction between *Hoxa2* and HIRE1, which puts into question if all of the proximal enhancers are relevant for *Hoxa2* expression. Consequently, our results significantly shift the perspective on how *Hoxa2* is regulated from a model that solely involves multiple proximal enhancers to a model that is based on multiple long-range inter-TAD SEs.

Although we were able to identify two novel regions that regulate *Hoxa2* and *Hoxa3* expression in CNCCs and hindbrain, we have little understanding of how the interactions between HIRE1/HIRE2 and *Hoxa2* are mediated across such a large distance, how these contacts are occurring across a TAD boundary, and why HIRE1 and HIRE2 activate *Hox* genes only in PA2 and PA3, despite being enriched for H3K37ac in all CNCCs subpopulations. Different types of tethering elements have been implicated with mediating long-range enhancer-promoter contacts, of which some are depending on cohesin-mediated loop extrusion (Pachano et al., 2022; Schoenfelder & Fraser, 2019). One such element are CTCF-binding sites which are found at many mouse lineage-specific genes. CTCF binding to the promoter of such genes has recently been shown to be necessary for establishing their enhancer-promoter and promoter-promoter contacts (Kubo et al., 2021; Oh et al., 2021). Secondly, CGIs have been proposed to act as tethering elements between CpG-rich enhancers and promoters likely mediated by PcG complexes (Crispatzu et al., 2021; Cruz-Molina et al., 2017; Pachano et al., 2021). Thirdly, several TFs, such as YY1 (Beagrie et al., 2017; Weintraub et al., 2017) and LDB1 (Deng et al., 2012; Song et al., 2007), have been shown to contribute to the establishment of enhancer-promoter contacts. Firstly, the CTCF binding patterns around HIRE1, HIRE2, and the *Hoxa* cluster are the same in PA2 and Md, it seems

unlikely that CTCF plays a significant role in the specific interaction between HIRE1/HIRE2 and *Hoxa2* in PA2 CNCCs. Secondly, while *Hoxa* gene promoters do contain CGIs, neither HIRE1 nor HIRE2 contains any CGIs. Hence, CGIs are unlikely to mediate the interaction between HIRE1 and HIRE2. Thirdly, one can speculate that the cooperative binding of TFs to HIRE1 and HIRE2 may be responsible for bringing *Hoxa2* in contact with HIRE1 and HIRE2. Such a model could also explain, why HIRE1 and HIRE2 only interact with *Hoxa2* in PA2, as certain TFs are differentially expressed between the CNCCs subpopulations (Minoux et al., 2017). Using published ChIP-seq data we observe a high density of Meis, Pbx and Hoxa2 (Amin et al., 2015; Donaldson et al., 2012) binding at HIRE1 and HIRE2 in PA2 CNCCs. Furthermore, *Hoxa2* is involved in the regulation of *Meis1* and *Meis2* expression in PA2 CNCCs (Amin et al., 2015), both *Meis1* and *Meis2* are specifically highly expressed in PA2 CNCCs (Minoux et al., 2017) and we identified two PA2-specific SEs interacting with *Meis1*. Additionally, our TF motif enrichment analysis gives insights into which type of TFs potentially bind at HIRE1 and HIRE2 constituent enhancers. TALE TFs, such as Meis and Pbx, are known co-factors of Hox proteins and have been shown to form dimeric and trimeric complexes with Hox conferring unique sequence specificity to the Hox proteins *in vitro* (Bobola & Sagerström, 2022). ChIP-seq analysis of Meis, Pbx, Hoxa2 and Hoxa3 in PA2 and posterior PAs confirmed that Hoxa2 and Hoxa3 binding almost completely coincides with regions occupied by Meis and Pbx and have been found to be enriched at active enhancers (Amin et al., 2015; Bridoux et al., 2020; Donaldson et al., 2012). Moreover, it has been shown that in CNCCs TALE-Hox binding specificity can be further modulated by direct or indirect cooperativity with other tissue-specific TFs (Bridoux et al., 2020). However, TALE do not only form complexes with Hox proteins, but also with other TFs (Bobola & Sagerström, 2022). In fact, TALE TFs appear to be able to create a chromatin environment that allows for the binding of other tissue-specific TFs, such as Hox. Hence, acting as pioneer factors (Bobola & Sagerström, 2022; Iwafuchi-Doi & Zaret, 2016). This leads to the conversion of a general pool of low-affinity TALE binding sites into tissue-restricted, high-confidence binding events associated with transcriptional activation. This results in TALE:TF complexes being linked to active chromatin and specific gene expression within certain tissue domains and lineages (Bobola & Sagerström, 2022). Considering the dense binding of Meis and Pbx at HIRE1/HIRE2 and *Hoxa2* promoter and the high expression of *Meis1* and *Meis2* specifically in PA2 CNCCs, it is plausible to assume that TALE TFs may be of particular importance in mediating the interaction between HIRE1/HIRE2 and *Hoxa2*. Such a model could potentially explain why HIRE1/HIRE2 and *Hoxa2* only interact in PA2 CNCCs and why we see only minor changes in the interaction frequency between HIRE1/HIRE2 and *Hoxa2* in

Hoxa2[EGFP/EGFP] mice. It could be of broader interest to functionally evaluate this model in the future. This could be done by transient enhancer reporter assay, such as enSERT (Kvon et al., 2020) since the reporter should only show activity if the required TFs are expressed. In addition, conditional knock-out of *Meis1* or *Meis2* in CNCCs using Meis1[fl/fl] (Kocabas et al., 2012) or Meis2[fl/fl] (Machon et al., 2015) and Wnt1::Cre mice could be useful to investigate the potential role of these two TALE TFs in facilitating the interactions between HIRE1/HIRE2 and *Hoxa2*.

SEs are very potent transcriptional activators and thus need to be structurally isolated from the surrounding genomic environment. Indeed, insulated neighborhoods, formed by cohesin-mediated loop extrusion between two convergent CTCF sites, have been suggested to insulate SEs and their target gene from interaction with other genomic regions (Dowen et al., 2014). In support of this model, we found that HIRE1 and HIRE2 are enclosed in a TAD flanked by convergent CTCF sites, that contains only one coding gene (*Npvf*), which is not expressed in any of CNCC subpopulations, and few non-coding genes. Considering, that HIRE1 and HIRE2 are strongly enriched for H3K27ac in all CNCCs, the spatial insulation of those two regions is likely necessary to prevent the activation of nearby genes. Multiple studies have shown that the perturbation of the boundaries of insulated neighborhoods containing SEs is associated with aberrant transcriptional activation of nearby genes (Dowen et al., 2014; Hnisz et al., 2016; Ji et al., 2016). The spatial insulation of the HIRE1 and HIRE2 may further promote the formation of structures similar to highly interconnected enhancers (HICE), which have been shown to act as integration hubs that coordinate differentiation by the formation of 3D enhancer communities during adipocyte differentiation (Madsen et al., 2020), or phase-separated condensates (Hnisz et al., 2017). It should be noted though that TAD boundary insulation is not absolute (Bonev et al., 2017; Javierre et al., 2016; Paulsen et al., 2019; Szabó & Mayor, 2018). Although interactions within a TAD are favored by loop extrusion (Dekker & Mirny, 2016), TADs and CTCF loops dynamically fold and unfold (Gabriele et al., 2022; Mach et al., 2022) and inter-TAD contacts may occur, depending on boundary strength and chromatin composition (Bonev et al., 2017; Paulsen et al., 2019; Szabó & Mayor, 2018). Nevertheless, TADs can functionally insulate regulatory sequences (Zuin et al., 2022).

The novel *Hoxa2* regulatory landscape, including HIRE1 and HIRE2, that we have discovered can be a useful model system to study regulatory mechanisms *in vivo*. For instance, deletion of subparts, single constituent enhancers, or CTCF binding sites within HIRE1/HIRE2 could determine if there are certain elements that mediate the interaction between HIRE1/HIRE2 and *Hoxa2* and if the constituent enhancers act in an additive or synergistic manner. Apart from

deletions, enhancer reporter assay, such as enSERT (Kvon et al., 2020), could provide information on the activity of the single constituent enhancers. Furthermore, it could be of interest to investigate if the deletion of the TAD boundary between HIRE1/HIRE2 and *Hoxa2* would result in the ectopic expression of *Hoxa* genes in FNP, Mx, and Md, suggesting an insulating function of the boundary, and/or a reduction of *Hoxa2* expression in PA2, indicating a role of the boundary in mediating the contact between HIRE1/HIRE2 and *Hoxa2*. In order to identify so called "hotspot" (Siersbæk et al., 2014) or "hub" (Huang et al., 2018) enhancers within the regulatory landscape, 5C-seq or enhancer-capture Hi-C could be performed. In addition, the study of constituent enhancers and their regulatory environment is relevant from a clinical and evolutionary perspective. Changes in enhancer sequences, such as single nucleotide polymorphisms (SNPs), and structural variation of the 3D genome are major drivers of regulatory innovation and morphological variation and have high relevance for cancer genetics and rare congenital diseases (Claringbould & Zaugg, 2021; Long et al., 2016; Prescott et al., 2015; Spielmann et al., 2018). For instance, *Hoxa2* haploinsufficiency has been shown to cause microtia in humans (Alasti & Van Camp, 2009; Brown et al., 2013; Piceci et al., 2017). Similar to microtia in humans, we have shown here that the reduction of *Hoxa2* expression in mice below a certain threshold is sufficient to induce morphological changes of the external ear, namely a small ear, and minimal changes in the middle ear skeleton. Considering, the high sequence conservation of HIRE1 and HIRE2 between mice and humans, databases of variant-to-phenotype links such as ClinVar (Landrum et al., 2014) could be checked to look for known pathogenic variants in the region corresponding to HIRE1 and HIRE2 in humans. Similarly, this could be done for all of the here identified SEs, which could help to determine which SEs are of functional relevance in craniofacial morphogenesis. Each of these experiments has the potential to increase our general understanding of the mechanisms involved in promoter-enhancer communication throughout development and could lead to the identification of further functional SE potentially implicated in congenital disorders.

In conclusion, there are still many open questions regarding the mechanisms that regulate the interaction between HIRE1 and HIRE2 and *Hoxa2*. Most of these are however difficult to address due to the limitations of our *in vivo* model system. While studying the role of enhancers *in vivo* has many advantages, it also sets major constraints on the depth of functional analyses that can be performed, due to the time it takes to generate new mouse lines and the limited number of cells that can be obtained per embryo. An *in vitro* system to generate mouse NCC in culture from mESC, similar to the human NCC culture system (Bajpai et al., 2010; Rada-Iglesias et al., 2011), would enable large-scale functional analyses of the mechanisms involved in mediating such long-

range enhancer-promoter contacts. Especially, proteomics studies, which are largely unfeasible *in vivo* due to the number of cells required, could be employed to identify which TFs bind to specific enhancers.

References

Adam, R. C., Yang, H., Rockowitz, S., Larsen, S. B., Nikolova, M., Oristian, D. S., Polak, L., Kadaja, M., Asare, A., Zheng, D., & Fuchs, E. (2015). Pioneer factors govern super-enhancer dynamics in stem cell plasticity and lineage choice (2015/03/25). *Nature*, *521*(7552), 366–370. https://doi.org/10.1038/nature14289

Alasti, F., & Van Camp, G. (2009). Genetics of microtia and associated syndromes (20090316th ed.). *J Med Genet*, *46*(6), 361–369. https://doi.org/10.1136/jmg.2008.062158

Alexander, J. M., Guan, J., Li, B., Maliskova, L., Song, M., Shen, Y., Huang, B., Lomvardas, S., & Weiner, O. D. (2019). Live-cell imaging reveals enhancer-dependent Sox2 transcription in the absence of enhancer proximity (R. H. Singer, K. Struhl, & Z. Liu, Eds.). *eLife*, *8*, e41769. https://doi.org/10.7554/eLife.41769

Alexander, T., Nolte, C., & Krumlauf, R. (2009). Hox Genes and Segmentation of the Hindbrain and Axial Skeleton. *Annual Review of Cell and Developmental Biology*, *25*(1), 431–456. https://doi.org/10.1146/annurev.cellbio.042308.113423

Allahyar, A., Vermeulen, C., Bouwman, B. A. M., Krijger, P. H. L., Verstegen, M. J. A. M., Geeven, G., van Kranenburg, M., Pieterse, M., Straver, R., Haarhuis, J. H. I., Jalink, K., Teunissen, H., Renkens, I. J., Kloosterman, W. P., Rowland, B. D., de Wit, E., de Ridder, J., & de Laat, W. (2018). Enhancer hubs and loop collisions identified from single-allele topologies. *Nature Genetics*, *50*(8), 1151–1160. https://doi.org/10.1038/s41588-018-0161-5

Amin, S., Donaldson, I. J., Zannino, D. A., Hensman, J., Rattray, M., Losa, M., Spitz, F., Ladam, F., Sagerstrom, C., & Bobola, N. (2015). Hoxa2 selectively enhances Meis binding to change a branchial arch ground state (20150129th ed.). *Dev Cell*, *32*(3), 265–277. https://doi.org/10.1016/j.devcel.2014.12.024

Anderson, E., Devenney, P. S., Hill, R. E., & Lettice, L. A. (2014). Mapping the Shh long-range regulatory domain. *Development*, *141*(20), 3934–3943. https://doi.org/10.1242/dev.108480

Andrey, G., Schopflin, R., Jerkovic, I., Heinrich, V., Ibrahim, D. M., Paliou, C., Hochradel, M., Timmermann, B., Haas, S., Vingron, M., & Mundlos, S. (2017). Characterization of hundreds of regulatory landscapes in developing limbs reveals two regimes of chromatin folding (20161206th ed.). *Genome Res*, *27*(2), 223–233. https://doi.org/10.1101/gr.213066.116

Andrey, G., & Mundlos, S. (2017). The three-dimensional genome: regulating gene expression during pluripotency and development. *Development*, *144*(20), 3646–3658. https://doi.org/10.1242/dev.148304

Antosova, B., Smolikova, J., Klimova, L., Lachova, J., Bendova, M., Kozmikova, I., Machon, O., & Kozmik, Z. (2016). The Gene Regulatory Network of Lens Induction Is Wired through Meis-Dependent

Shadow Enhancers of Pax6. *PLOS Genetics*, *12*(12), e1006441–. https://doi.org/10.1371/journal.pgen.1006441

Attanasio, C., Nord, A. S., Zhu, Y., Blow, M. J., Li, Z., Liberton, D. K., Morrison, H., Plajzer-Frick, I., Holt, A., Hosseini, R., Phouanenavong, S., Akiyama, J. A., Shoukry, M., Afzal, V., Rubin, E. M., FitzPatrick, D. R., Ren, B., Hallgrimsson, B., Pennacchio, L. A., & Visel, A. (2013). Fine tuning of craniofacial morphology by distant-acting enhancers. *Science*, *342*(6157), 1241006. https://doi.org/10.1126/science.1241006

Baggiolini, A., Varum, S., Mateos, J. M., Bettosini, D., John, N., Bonalli, M., Ziegler, U., Dimou, L., Clevers, H., Furrer, R., & Sommer, L. (2015). Premigratory and Migratory Neural Crest Cells Are Multipotent In Vivo. *Cell Stem Cell*, *16*(3), 314–322. https://doi.org/10.1016/J.STEM.2015.02.017

Bahr, C., von Paleske, L., Uslu, V. V., Remeseiro, S., Takayama, N., Ng, S. W., Murison, A., Langenfeld, K., Petretich, M., Scognamiglio, R., Zeisberger, P., Benk, A. S., Amit, I., Zandstra, P. W., Lupien, M., Dick, J. E., Trumpp, A., & Spitz, F. (2018). A Myc enhancer cluster regulates normal and leukaemic haematopoietic stem cell hierarchies (2018/01/18). *Nature*, *553*(7689), 515–520. https://doi.org/10.1038/nature25193

Bajpai, R., Chen, D. A., Rada-Iglesias, A., Zhang, J., Xiong, Y., Helms, J., Chang, C.-P., Zhao, Y., Swigut, T., & Wysocka, J. (2010). CHD7 cooperates with PBAF to control multipotent neural crest formation. *Nature*, *463*(7283), 958–962. https://doi.org/10.1038/nature08733

Balsalobre, A., & Drouin, J. (2022). Pioneer factors as master regulators of the epigenome and cell fate. *Nature Reviews Molecular Cell Biology*, *23*(7), 449–464. https://doi.org/10.1038/s41580-022-00464-z

Baltzinger, M., Ori, M., Pasqualetti, M., Nardi, I., & Rijli, F. M. (2005). Hoxa2 knockdown in Xenopus results in hyoid to mandibular homeosis (2005/10/14). *Dev Dyn*, *234*(4), 858–867. https://doi.org/10.1002/dvdy.20567

Bartman, C. R., Hsu, S. C., Hsiung, C. C., Raj, A., & Blobel, G. A. (2016). Enhancer Regulation of Transcriptional Bursting Parameters Revealed by Forced Chromatin Looping. *Molecular Cell*, *62*(2), 237–247. https://doi.org/10.1016/J.MOLCEL.2016.03.007

Basch, M. L., Bronner-Fraser, M., & García-Castro, M. I. (2006). Specification of the neural crest occurs during gastrulation and requires Pax7. *Nature*, *441*(7090), 218–222. https://doi.org/10.1038/nature04684

Batut, P. J., Bing, X. Y., Sisco, Z., Raimundo, J., Levo, M., & Levine, M. S. (2022). Genome organization controls transcriptional dynamics during development. *Science*, *375*(6580), 566–570. https://doi.org/10.1126/science.abi7178

Beagrie, R. A., Scialdone, A., Schueler, M., Kraemer, D. C. A., Chotalia, M., Xie, S. Q., Barbieri, M., de Santiago, I., Lavitas, L.-M., Branco, M. R., Fraser, J., Dostie, J., Game, L., Dillon, N., Edwards, P. A. W., Nicodemi, M., & Pombo, A. (2017). Complex multi-enhancer contacts captured by genome architecture mapping. *Nature*, *543*(7646), 519–524. https://doi.org/10.1038/nature21411

REFERENCES

Benabdallah, N. S., Williamson, I., Illingworth, R. S., Kane, L., Boyle, S., Sengupta, D., Grimes, G. R., Therizols, P., & Bickmore, W. A. (2019). Decreased Enhancer-Promoter Proximity Accompanying Enhancer Activation. *Molecular Cell*, *76*(3), 473–484. https://doi.org/10.1016/j.molcel.2019.07.038

Benko, S., Fantes, J. A., Amiel, J., Kleinjan, D.-J., Thomas, S., Ramsay, J., Jamshidi, N., Essafi, A., Heaney, S., Gordon, C. T., McBride, D., Golzio, C., Fisher, M., Perry, P., Abadie, V., Ayuso, C., Holder-Espinasse, M., Kilpatrick, N., Lees, M. M., . . . Lyonnet, S. (2009). Highly conserved non-coding elements on either side of SOX9 associated with Pierre Robin sequence. *Nature Genetics*, *41*(3), 359–364. https://doi.org/10.1038/ng.329

Bernstein, B. E., Meissner, A., & Lander, E. S. (2007). The Mammalian Epigenome. *Cell*, *128*(4), 669–681. https://doi.org/10.1016/j.cell.2007.01.033

Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., Jaenisch, R., Wagschal, A., Feil, R., Schreiber, S. L., & Lander, E. S. (2006). A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. *Cell*, *125*(2), 315–326. https://doi.org/10.1016/J.CELL.2006.02.041

Bintu, B., Mateo, L. J., Su, J. H., Sinnott-Armstrong, N. A., Parker, M., Kinrot, S., Yamaya, K., Boettiger, A. N., & Zhuang, X. (2018). Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. *Science*, *362*(6413). https://doi.org/10.1126/science.aau1783

Bird, A., Taggart, M., Frommer, M., Miller, O. J., & Macleod, D. (1985). A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. *Cell*, *40*(1), 91–99. https://doi.org/https://doi.org/10.1016/0092-8674(85)90312-5

Birgbauer, E., Sechrist, J., Bronner-Fraser, M., & Fraser, S. (1995). Rhombomeric origin and rostrocaudal reassortment of neural crest cells revealed by intravital microscopy. *Development*, *121*(4), 935–945. https://doi.org/10.1242/dev.121.4.935

Blanco, E., González-Ramírez, M., Alcaine-Colet, A., Aranda, S., & Di Croce, L. (2020). The Bivalent Genome: Characterization, Structure, and Regulation. *Trends in Genetics*, *36*(2), 118–131. https://doi.org/10.1016/j.tig.2019.11.004

Blobel, G. A., Higgs, D. R., Mitchell, J. A., Notani, D., & Young, R. A. (2021). Testing the super-enhancer concept (20210903rd ed.). *Nat Rev Genet*, *22*(12), 749–755. https://doi.org/10.1038/s41576-021-00398-w

Bobola, N., Carapuco, M., Ohnemus, S., Kanzler, B., Leibbrandt, A., Neubuser, A., Drouin, J., & Mallo, M. (2003). Mesenchymal patterning by Hoxa2 requires blocking Fgf-dependent activation of Ptx1. *Development*, *130*(15), 3403–3414. https://doi.org/10.1242/dev.00554

Bobola, N., & Sagerström, C. G. (2022). TALE transcription factors: Cofactors no more. *Seminars in Cell & Developmental Biology*. https://doi.org/10.1016/J.SEMCDB.2022.11.015

Boettiger, A. N., Bintu, B., Moffitt, J. R., Wang, S., Beliveau, B. J., Fudenberg, G., Imakaev, M., Mirny, L. A., Wu, C.-t., & Zhuang, X. (2016). Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature*, *529*(7586), 418–422. https://doi.org/10.1038/nature16496

Bolande, R. P. (1974). The neurocristopathies: A unifying concept of disease arising in neural crest malde-velopment. *Human Pathology*, *5*(4), 409–429. https://doi.org/10.1016/S0046-8177(74)80021-3

Boltsis, I., Grosveld, F., Giraud, G., & Kolovos, P. (2021). Chromatin Conformation in Development and Disease. *Frontiers in Cell and Developmental Biology*, *9*. https://www.frontiersin.org/articles/10.3389/fcell.2021.723859

Bonev, B., Mendelson Cohen, N., Szabo, Q., Fritsch, L., Papadopoulos, G. L., Lubling, Y., Xu, X., Lv, X., Hugnot, J. P., Tanay, A., & Cavalli, G. (2017). Multiscale 3D Genome Rewiring during Mouse Neural Development. *Cell*, *171*(3), 557–572. https://doi.org/10.1016/J.CELL.2017.09.043

Bothma, J. P., Garcia, H. G., Ng, S., Perry, M. W., Gregor, T., & Levine, M. (2015). Enhancer additivity and non-additivity are determined by enhancer strength in the Drosophila embryo (R. Krumlauf, Ed.). *eLife*, *4*, e07956. https://doi.org/10.7554/eLife.07956

Boulet, A. M., & Capecchi, M. R. (1996). Targeted Disruption ofhoxc-4Causes Esophageal Defects and Vertebral Transformations. *Developmental Biology*, *177*(1), 232–249. https://doi.org/10.1006/DBIO.1996.0159

Bridoux, L., Zarrineh, P., Mallen, J., Phuycharoen, M., Latorre, V., Ladam, F., Losa, M., Baker, S. M., Sagerstrom, C., Mace, K. A., Rattray, M., & Bobola, N. (2020). HOX paralogs selectively convert binding of ubiquitous transcription factors into tissue-specific patterns of enhancer activation (20201214th ed.). *PLoS Genet*, *16*(12), e1009162. https://doi.org/10.1371/journal.pgen.1009162

Bronner, M. E., & LeDouarin, N. M. (2012). Development and evolution of the neural crest: An overview. https://doi.org/10.1016/j.ydbio.2011.12.042

Brown, K. K., Viana, L. M., Helwig, C. C., Artunduaga, M. A., Quintanilla-Dieck, L., Jarrin, P., Osorno, G., McDonough, B., DePalma, S. R., Eavey, R. D., Seidman, J. G., & Seidman, C. E. (2013). HOXA2 haploinsufficiency in dominant bilateral microtia and hearing loss (20130711th ed.). *Hum Mutat*, *34*(10), 1347–1351. https://doi.org/10.1002/humu.22367

Buenrostro, J. D., Wu, B., Chang, H. Y., & Greenleaf, W. J. (2015). ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide (20150105th ed.). *Curr Protoc Mol Biol*, *109*, 1–21. https://doi.org/10.1002/0471142727.mb2129s109

Cannavò, E., Khoueiry, P., Garfield, D. A., Geeleher, P., Zichner, T., Gustafson, E. H., Ciglar, L., Korbel, J. O., & Furlong, E. E. (2016). Shadow Enhancers Are Pervasive Features of Developmental Regulatory Networks. *Current Biology*, *26*(1), 38–51. https://doi.org/10.1016/J.CUB.2015.11.034

Chapuy, B., McKeown, M. R., Lin, C. Y., Monti, S., Roemer, M. G., Qi, J., Rahl, P. B., Sun, H. H., Yeda, K. T., Doench, J. G., Reichert, E., Kung, A. L., Rodig, S. J., Young, R. A., Shipp, M. A., & Bradner, J. E. (2013). Discovery and Characterization of Super-Enhancer-Associated Dependencies in Diffuse Large B Cell Lymphoma. *Cancer Cell*, *24*(6), 777–790. https://doi.org/10.1016/J.CCR.2013.11.003

Chen, H., Levo, M., Barinov, L., Fujioka, M., Jaynes, J. B., & Gregor, T. (2018). Dynamic interplay between enhancer–promoter topology and gene activity. *Nature Genetics*, *50*(9), 1296–1303. https://doi.org/10.1038/s41588-018-0175-z

REFERENCES

Chisaka, O., & Capecchi, M. R. (1991). Regionally restricted developmental defects resulting from targeted disruption of the mouse homeobox gene hox-1.5. *Nature*, *350*(6318), 473–479. https://doi.org/10.1038/350473a0

Cho, W.-K., Spille, J.-H., Hecht, M., Lee, C., Li, C., Grube, V., & Cisse, I. I. (2018). Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science*, *361*(6400), 412–415. https://doi.org/10.1126/science.aar4199

Chojnowski, J. L., Masuda, K., Trau, H. A., Thomas, K., Capecchi, M., & Manley, N. R. (2014). Multiple roles for HOXA3 in regulating thymus and parathyroid differentiation and morphogenesis in mouse. *Development*, *141*(19), 3697–3708. https://doi.org/10.1242/dev.110833

Chojnowski, J. L., Trau, H. A., Masuda, K., & Manley, N. R. (2016). Temporal and spatial requirements for Hoxa3 in mouse embryonic development. *Developmental Biology*, *415*(1), 33–45. https://doi.org/10.1016/J.YDBIO.2016.05.010

Claringbould, A., & Zaugg, J. B. (2021). Enhancers in disease: molecular basis and emerging treatment strategies (20210820th ed.). *Trends Mol Med*, *27*(11), 1060–1073. https://doi.org/10.1016/j.molmed.2021.07.012

Condie, B. G., & Capecchi, M. R. (1994). Mice with targeted disruptions in the paralogous genes hoxa-3 and hoxd-3 reveal synergistic interactions (1994/07/28). *Nature*, *370*(6487), 304–307. https://doi.org/10.1038/370304a0

Couly, G., Grapin-Botton, A., Coltey, P., & Le Douarin, N. M. (1996). The regeneration of the cephalic neural crest, a problem revisited: the regenerating cells originate from the contralateral or from the anterior and posterior neural fold. *Development*, *122*(11), 3393–3407. https://doi.org/10.1242/dev.122.11.3393

Couly, G., Grapin-Botton, A., Coltey, P., Ruhin, B., & Le Douarin, N. M. (1998). Determination of the identity of the derivatives of the cephalic neural crest: incompatibility between Hox gene expression and lower jaw development. *Development*, *125*(17), 3445–3459. https://doi.org/10.1242/dev.125.17.3445

Couly, G., Creuzet, S., Bennaceur, S., Vincent, C., & Le Douarin, N. M. (2002). Interactions between Hox-negative cephalic neural crest cells and the foregut endoderm in patterning the facial skeleton in the vertebrate head. *Development*, *129*(4), 1061–1073. https://doi.org/10.1242/dev.129.4.1061

Crane, J. F., & Trainor, P. A. (2006). Neural Crest Stem and Progenitor Cells. *Annual Review of Cell and Developmental Biology*, *22*(1), 267–286. https://doi.org/10.1146/annurev.cellbio.22.010305.103814

Creuzet, S., Couly, G., Vincent, C., & Le Douarin, N. M. (2002). Negative effect of Hox gene expression on the development of the neural crest-derived facial skeleton. *Development*, *129*(18), 4301–4313. https://doi.org/10.1242/dev.129.18.4301

Creyghton, M. P., Cheng, A. W., Welstead, G. G., Kooistra, T., Carey, B. W., Steine, E. J., Hanna, J., Lodato, M. A., Frampton, G. M., Sharp, P. A., Boyer, L. A., Young, R. A., & Jaenisch, R. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proceedings*

of the National Academy of Sciences, *107*(50), 21931–21936. https://doi.org/10.1073/pnas.1016071107

Crispatzu, G., Rehimi, R., Pachano, T., Bleckwehl, T., Cruz-Molina, S., Xiao, C., Mahabir, E., Bazzi, H., & Rada-Iglesias, A. (2021). The chromatin, topological and regulatory properties of pluripotency-associated poised enhancers are conserved in vivo. *Nature Communications*, *12*(1), 4344. https://doi.org/10.1038/s41467-021-24641-4

Cruz-Molina, S., Respuela, P., Tebartz, C., Kolovos, P., Nikolic, M., Fueyo, R., van Ijcken, W. F., Grosveld, F., Frommolt, P., Bazzi, H., & Rada-Iglesias, A. (2017). PRC2 Facilitates the Regulatory Topology Required for Poised Enhancer Function during Pluripotent Stem Cell Differentiation. *Cell Stem Cell*, *20*(5), 689–705. https://doi.org/10.1016/J.STEM.2017.02.004

Davidson, E. H. (2010). Emerging properties of animal gene regulatory networks. *Nature*, *468*(7326), 911–920. https://doi.org/10.1038/nature09645

Davidson, I. F., Bauer, B., Goetz, D., Tang, W., Wutz, G., & Peters, J. M. (2019). DNA loop extrusion by human cohesin (20191121st ed.). *Science*, *366*(6471), 1338–1345. https://doi.org/10.1126/science.aaz3418

Davidson, I. F., Barth, R., Zaczek, M., van der Torre, J., Tang, W., Nagasaka, K., Janissen, R., Kerssemakers, J., Wutz, G., Dekker, C., & Peters, J.-M. (2022). CTCF is a DNA-tension-dependent barrier to cohesin-mediated DNA loop extrusion. *bioRxiv*, 2022.09.08.507093. https://doi.org/10.1101/2022.09.08.507093

Dekker, J., & Mirny, L. (2016). The 3D Genome as Moderator of Chromosomal Communication. *Cell*, *164*(6), 1110–1121. https://doi.org/10.1016/j.cell.2016.02.007

Deng, W., Lee, J., Wang, H., Miller, J., Reik, A., Gregory, P. D., Dean, A., & Blobel, G. A. (2012). Controlling Long-Range Genomic Interactions at a Native Locus by Targeted Tethering of a Looping Factor. *Cell*, *149*(6), 1233–1244. https://doi.org/10.1016/J.CELL.2012.03.051

Deng, W., Rupon, J. W., Krivega, I., Breda, L., Motta, I., Jahn, K. S., Reik, A., Gregory, P. D., Rivella, S., Dean, A., & Blobel, G. A. (2014). Reactivation of Developmentally Silenced Globin Genes by Forced Chromatin Looping. *Cell*, *158*(4), 849–860. https://doi.org/10.1016/J.CELL.2014.05.050

Denholtz, M., Bonora, G., Chronis, C., Splinter, E., de Laat, W., Ernst, J., Pellegrini, M., & Plath, K. (2013). Long-Range Chromatin Contacts in Embryonic Stem Cells Reveal a Role for Pluripotency Factors and Polycomb Proteins in Genome Organization. *Cell Stem Cell*, *13*(5), 602–616. https://doi.org/10.1016/J.STEM.2013.08.013

Deschamps, J., & Duboule, D. (2017). Embryonic timing, axial stem cells, chromatin dynamics, and the Hox clock. *Genes & Development*, *31*(14), 1406–1416. https://doi.org/10.1101/gad.303123.117

Despang, A., Schopflin, R., Franke, M., Ali, S., Jerkovic, I., Paliou, C., Chan, W. L., Timmermann, B., Wittler, L., Vingron, M., Mundlos, S., & Ibrahim, D. M. (2019). Functional dissection of the Sox9-Kcnj2 locus identifies nonessential and instructive roles of TAD architecture (2019/07/31). *Nat Genet*, *51*(8), 1263–1271. https://doi.org/10.1038/s41588-019-0466-z

REFERENCES

de Wit, E., Vos, E. S., Holwerda, S. J., Valdes-Quezada, C., Verstegen, M. J., Teunissen, H., Splinter, E., Wijchers, P. J., Krijger, P. H., & de Laat, W. (2015). CTCF Binding Polarity Determines Chromatin Looping. *Molecular Cell*, *60*(4), 676–684. https://doi.org/10.1016/J.MOLCEL.2015.09.023

Diao, Y., Fang, R., Li, B., Meng, Z., Yu, J., Qiu, Y., Lin, K. C., Huang, H., Liu, T., Marina, R. J., Jung, I., Shen, Y., Guan, K.-L., & Ren, B. (2017). A tiling-deletion-based genetic screen for cis-regulatory element identification in mammalian cells. *Nature Methods*, *14*(6), 629–635. https://doi.org/10.1038/nmeth.4264

Dickel, D. E., Ypsilanti, A. R., Pla, R., Zhu, Y., Barozzi, I., Mannion, B. J., Khin, Y. S., Fukuda-Yuzawa, Y., Plajzer-Frick, I., Pickle, C. S., Lee, E. A., Harrington, A. N., Pham, Q. T., Garvin, T. H., Kato, M., Osterwalder, M., Akiyama, J. A., Afzal, V., Rubenstein, J. L. R., . . . Visel, A. (2018). Ultraconserved Enhancers Are Required for Normal Development (20180118th ed.). *Cell*, *172*(3), 491–499. https://doi.org/10.1016/j.cell.2017.12.017

Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J. S., & Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions (20120411th ed.). *Nature*, *485*(7398), 376–380. https://doi.org/10.1038/nature11082

Donaldson, I. J., Amin, S., Hensman, J. J., Kutejova, E., Rattray, M., Lawrence, N., Hayes, A., Ward, C. M., & Bobola, N. (2012). Genome-wide occupancy links Hoxa2 to Wnt-beta-catenin signaling in mouse embryonic development (20120105th ed.). *Nucleic Acids Res*, *40*(9), 3990–4001. https://doi.org/10.1093/nar/gkr1240

Dowen, J. M., Fan, Z. P., Hnisz, D., Ren, G., Abraham, B. J., Zhang, L. N., Weintraub, A. S., Schuijers, J., Lee, T. I., Zhao, K., & Young, R. A. (2014). Control of Cell Identity Genes Occurs in Insulated Neighborhoods in Mammalian Chromosomes. *Cell*, *159*(2), 374–387. https://doi.org/https://doi.org/10.1016/j.cell.2014.09.030

Dunipace, L., Ozdemir, A., & Stathopoulos, A. (2011). Complex interactions between cis-regulatory modules in native conformation are critical for Drosophila snail expression. *Development*, *138*(18), 4075–4084. https://doi.org/10.1242/dev.069146

Fabik, J., Psutkova, V., & Machon, O. (2021). The mandibular and hyoid arches—from molecular patterning to shaping bone and cartilage. *International Journal of Molecular Sciences*, *22*(14). https://doi.org/10.3390/ijms22147529

Franke, M., Ibrahim, D. M., Andrey, G., Schwarzer, W., Heinrich, V., Schöpflin, R., Kraft, K., Kempfer, R., Jerković, I., Chan, W.-L., Spielmann, M., Timmermann, B., Wittler, L., Kurth, I., Cambiaso, P., Zuffardi, O., Houge, G., Lambie, L., Brancati, F., . . . Mundlos, S. (2016). Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature*, *538*(7624), 265–269. https://doi.org/10.1038/nature19800

Frankel, N., Davis, G. K., Vargas, D., Wang, S., Payre, F., & Stern, D. L. (2010). Phenotypic robustness conferred by apparently redundant transcriptional enhancers. *Nature*, *466*(7305), 490–493. https://doi.org/10.1038/nature09158

Frasch, M., Chen, X., & Lufkin, T. (1995). Evolutionary-conserved enhancers direct region-specific expression of the murine Hoxa-1 and Hoxa-2 loci in both mice and Drosophila (1995/04/01). *Development*, *121*(4), 957–974. https://doi.org/10.1242/dev.121.4.957

Freire-Pritchett, P., Schoenfelder, S., Várnai, C., Wingett, S. W., Cairns, J., Collier, A. J., García-Vílchez, R., Furlan-Magaril, M., Osborne, C. S., Fraser, P., Rugg-Gunn, P. J., & Spivakov, M. (2017). Global reorganisation of cis-regulatory units upon lineage commitment of human embryonic stem cells (J. Dekker, Ed.). *eLife*, *6*, e21926. https://doi.org/10.7554/eLife.21926

Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., & Mirny, L. A. (2016). Formation of Chromosomal Domains by Loop Extrusion (20160519th ed.). *Cell Rep*, *15*(9), 2038–2049. https://doi.org/10.1016/j.celrep.2016.04.085

Fukaya, T., Lim, B., & Levine, M. (2016). Enhancer Control of Transcriptional Bursting. *Cell*, *166*(2), 358–368. https://doi.org/10.1016/J.CELL.2016.05.025

Fulco, C. P., Munschauer, M., Anyoha, R., Munson, G., Grossman, S. R., Perez, E. M., Kane, M., Cleary, B., Lander, E. S., & Engreitz, J. M. (2016). Systematic mapping of functional enhancer–promoter connections with CRISPR interference. *Science*, *354*(6313), 769–773. https://doi.org/10.1126/science.aag2445

Furlong, E. E. M., & Levine, M. (2018). Developmental enhancers and chromosome topology. *Science*, *361*(6409), 1341–1345. https://doi.org/10.1126/science.aau0320

Gabriele, M., Brandão, H. B., Grosse-Holz, S., Jha, A., Dailey, G. M., Cattoglio, C., Hsieh, T.-H. S., Mirny, L., Zechner, C., & Hansen, A. S. (2022). Dynamics of CTCF- and cohesin-mediated chromatin looping revealed by live-cell imaging. *Science*, *376*(6592), 496–501. https://doi.org/10.1126/science.abn6583

Gammill, L. S., & Bronner-Fraser, M. (2003). Neural crest specification: migrating into genomics. *Nat Rev Neurosci*, *4*(10), 795–805. https://doi.org/10.1038/nrn1219

Gammill, L. S., & Bronner-Fraser, M. (2002). Genomic analysis of neural crest induction. *Development*, *129*(24), 5731–5741. https://doi.org/10.1242/dev.00175

Garcıa-Castro, M. I., Marcelle, C., & Bronner-Fraser, M. (2002). Ectodermal Wnt Function as a Neural Crest Inducer. *Science*, *297*(5582), 848–851. https://doi.org/10.1126/science.1070824

Garnett, A. T., Square, T. A., & Medeiros, D. M. (2012). BMP, Wnt and FGF signals are integrated through evolutionarily conserved enhancers to achieve robust expression of Pax3 and Zic genes at the zebrafish neural plate border. *Development*, *139*(22), 4220–4231. https://doi.org/10.1242/dev.081497

Gendron-Maguire, M., Mallo, M., Zhang, M., & Gridley, T. (1993). Hoxa-2 mutant mice exhibit homeotic transformation of skeletal elements derived from cranial neural crest. *Cell*, *75*(7), 1317–1331. https://doi.org/10.1016/0092-8674(93)90619-2

Ghavi-Helm, Y., Jankowski, A., Meiers, S., Viales, R. R., Korbel, J. O., & Furlong, E. E. M. (2019). Highly rearranged chromosomes reveal uncoupling between genome topology and gene expression (20190715th ed.). *Nat Genet*, *51*(8), 1272–1282. https://doi.org/10.1038/s41588-019-0462-3

REFERENCES

Gorkin, D. U., Barozzi, I., Zhao, Y., Zhang, Y., Huang, H., Lee, A. Y., Li, B., Chiou, J., Wildberg, A., Ding, B., Zhang, B., Wang, M., Strattan, J. S., Davidson, J. M., Qiu, Y., Afzal, V., Akiyama, J. A., Plajzer-Frick, I., Novak, C. S., . . . Ren, B. (2020). An atlas of dynamic chromatin landscapes in mouse fetal development. *Nature*, *583*(7818), 744–751. https://doi.org/10.1038/s41586-020-2093-3

Gould, A., Morrison, A., Sproat, G., White, R. A., & Krumlauf, R. (1997). Positive cross-regulation and enhancer sharing: two mechanisms for specifying overlapping Hox expression patterns. *Genes & Development*, *11*(7), 900–913. https://doi.org/10.1101/gad.11.7.900

Graham, A., & Richardson, J. (2012). Developmental and evolutionary origins of the pharyngeal apparatus. *EvoDevo*, *3*(1), 24. https://doi.org/10.1186/2041-9139-3-24

Grammatopoulos, G. A., Bell, E., Toole, L., Lumsden, A., & Tucker, A. S. (2000). Homeotic transformation of branchial arch identity after Hoxa2 overexpression (2000/11/15). *Development*, *127*(24), 5355–5365. https://doi.org/10.1242/dev.127.24.5355

Green, S. A., Simoes-Costa, M., & Bronner, M. E. (2015). Evolution of vertebrates as viewed from the crest. *Nature*, *520*(7548), 474–482. https://doi.org/10.1038/nature14436

Groves, A. K., & LaBonne, C. (2014). Setting appropriate boundaries: Fate, patterning and competence at the neural plate border. *Developmental Biology*, *389*(1), 2–12. https://doi.org/10.1016/J.YDBIO.2013.11.027

Gu, B., Swigut, T., Spencley, A., Bauer, M. R., Chung, M., Meyer, T., & Wysocka, J. (2018). Transcription-coupled changes in nuclear mobility of mammalian cis-regulatory elements. *Science*, *359*(6379), 1050–1055. https://doi.org/10.1126/science.aao3136

Guo, Y., Xu, Q., Canzio, D., Shou, J., Li, J., Gorkin, D. U., Jung, I., Wu, H., Zhai, Y., Tang, Y., Lu, Y., Wu, Y., Jia, Z., Li, W., Zhang, M. Q., Ren, B., Krainer, A. R., Maniatis, T., & Wu, Q. (2015). CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function. *Cell*, *162*(4), 900–910. https://doi.org/10.1016/J.CELL.2015.07.038

Hay, D., Hughes, J. R., Babbs, C., Davies, J. O. J., Graham, B. J., Hanssen, L., Kassouf, M. T., Marieke Oudelaar, A. M., Sharpe, J. A., Suciu, M. C., Telenius, J., Williams, R., Rode, C., Li, P. S., Pennacchio, L. A., Sloane-Stanley, J. A., Ayyub, H., Butler, S., Sauka-Spengler, T., . . . Higgs, D. R. (2016). Genetic dissection of the alpha-globin super-enhancer in vivo (2016/07/05). *Nat Genet*, *48*(8), 895–903. https://doi.org/10.1038/ng.3605

Herranz, D., Ambesi-Impiombato, A., Palomero, T., Schnell, S. A., Belver, L., Wendorff, A. A., Xu, L., Castillo-Martin, M., Llobet-Navas, D., Cordon-Cardo, C., Clappier, E., Soulier, J., & Ferrando, A. A. (2014). A NOTCH1-driven MYC enhancer promotes T cell development, transformation and acute lymphoblastic leukemia (2014/09/10). *Nat Med*, *20*(10), 1130–1137. https://doi.org/10.1038/nm.3665

His, W. (1868). *Untersuchungen über die erste Anlage des Wirbelthierleibes : die erste Entwickelung des Hühnchens im Ei / von Wilhelm His.* F.C.W. Vogel, https://doi.org/10.5962/bhl.title.15288

Hnisz, D., Abraham, B. J., Lee, T. I., Lau, A., Saint-Andre, V., Sigova, A. A., Hoke, H. A., & Young, R. A. (2013). Super-enhancers in the control of cell identity and disease (20131010th ed.). *Cell*, *155*(4), 934–947. https://doi.org/10.1016/j.cell.2013.09.053

Hnisz, D., Shrinivas, K., Young, R. A., Chakraborty, A. K., & Sharp, P. A. (2017). A Phase Separation Model for Transcriptional Control. *Cell*, *169*(1), 13–23. https://doi.org/10.1016/j.cell.2017.02.007

Hnisz, D., Weintraub, A. S., Day, D. S., Valton, A.-L., Bak, R. O., Li, C. H., Goldmann, J., Lajoie, B. R., Fan, Z. P., Sigova, A. A., Reddy, J., Borges-Rivera, D., Lee, T. I., Jaenisch, R., Porteus, M. H., Dekker, J., & Young, R. A. (2016). Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science*, *351*(6280), 1454–1458. https://doi.org/10.1126/science.aad9024

Hu, D., Marcucio, R. S., & Helms, J. A. (2003). A zone of frontonasal ectoderm regulates patterning and growth in the face. *Development*, *130*(9), 1749–1758. https://doi.org/10.1242/dev.00397

Hu, N., Strobl-Mazzulla, P. H., & Bronner, M. E. (2014). Epigenetic regulation in neural crest development. https://doi.org/10.1016/j.ydbio.2014.09.034

Huang, J., Li, K., Cai, W., Liu, X., Zhang, Y., Orkin, S. H., Xu, J., & Yuan, G. C. (2018). Dissecting super-enhancer hierarchy based on chromatin interactions. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-03279-9

Hughes, J. R., Roberts, N., McGowan, S., Hay, D., Giannoulatou, E., Lynch, M., De Gobbi, M., Taylor, S., Gibbons, R., & Higgs, D. R. (2014). Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-throughput experiment (20140112th ed.). *Nat Genet*, *46*(2), 205–212. https://doi.org/10.1038/ng.2871

Hunt, P., Ferretti, P., Krumlauf, R., & Thorogood, P. (1995). Restoration of normal Hox code and branchial arch morphogenesis after extensive deletion of hindbrain neural crest. *Developmental biology*, *168*(2), 584–97. https://doi.org/10.1006/dbio.1995.1104

Hunt, P., Gulisano, M., Cook, M., Sham, M. H., Faiella, A., Wilkinson, D., Boncinelli, E., & Krumlauf, R. (1991). A distinct Hox code for the branchial region of the vertebrate head (1991/10/31). *Nature*, *353*(6347), 861–864. https://doi.org/10.1038/353861a0

Hunt, P., Whiting, J., Nonchev, S., Sham, M. H., Marshall, H., Graham, A., Cook, M., Allemann, R., Rigby, P. W., Gulisano, M., & et al. (1991). The branchial Hox code and its implications for gene regulation, patterning of the nervous system and head evolution (1991/01/01). *Dev Suppl*, *Suppl 2*, 63–77. https://www.ncbi.nlm.nih.gov/pubmed/1688181

Hunt, P., & Krumlauf, R. (1991). Deciphering the Hox code: Clues to patterning branchial regions of the head. *Cell*, *66*(6), 1075–1078. https://doi.org/10.1016/0092-8674(91)90029-X

Hunter, M. P., & Prince, V. E. (2002). Zebrafish hox paralogue group 2 genes function redundantly as selector genes to pattern the second pharyngeal arch (2002/06/28). *Dev Biol*, *247*(2), 367–389. https://doi.org/10.1006/dbio.2002.0701

Irving, C., & Mason, I. (2000). Signalling by FGF8 from the isthmus patterns anterior hindbrain and establishes the anterior limit of Hox gene expression. *Development*, *127*(1), 177–186. https://doi.org/10.1242/dev.127.1.177

REFERENCES

Isbel, L., Grand, R. S., & Schübeler, D. (2022). Generating specificity in genome regulation through transcription factor sensitivity to chromatin. *Nature Reviews Genetics*, *23*(12), 728–740. https://doi.org/10.1038/s41576-022-00512-6

Iwafuchi-Doi, M., & Zaret, K. S. (2016). Cell fate control by pioneer transcription factors. *Development*, *143*(11), 1833–1837. https://doi.org/10.1242/dev.133900

Javierre, B. M., Burren, O. S., Wilder, S. P., Kreuzhuber, R., Hill, S. M., Sewitz, S., Cairns, J., Wingett, S. W., Varnai, C., Thiecke, M. J., Burden, F., Farrow, S., Cutler, A. J., Rehnstrom, K., Downes, K., Grassi, L., Kostadima, M., Freire-Pritchett, P., Wang, F., ... Fraser, P. (2016). Lineage-Specific Genome Architecture Links Enhancers and Non-coding Disease Variants to Target Gene Promoters. *Cell*, *167*(5), 1369–1384. https://doi.org/10.1016/j.cell.2016.09.037

Ji, X., Dadon, D. B., Powell, B. E., Fan, Z. P., Borges-Rivera, D., Shachar, S., Weintraub, A. S., Hnisz, D., Pegoraro, G., Lee, T. I., Misteli, T., Jaenisch, R., & Young, R. A. (2016). 3D Chromosome Regulatory Landscape of Human Pluripotent Cells. *Cell Stem Cell*, *18*(2), 262–275. https://doi.org/10.1016/J.STEM.2015.11.007

Jiang, T., Raviram, R., Snetkova, V., Rocha, P. P., Proudhon, C., Badri, S., Bonneau, R., Skok, J. A., & Kluger, Y. (2016). Identification of multi-loci hubs from 4C-seq demonstrates the functional importance of simultaneous interactions. *Nucleic Acids Research*, *44*(18), 8714–8725. https://doi.org/10.1093/nar/gkw568

Joshi, O., Wang, S. Y., Kuznetsova, T., Atlasi, Y., Peng, T., Fabre, P. J., Habibi, E., Shaik, J., Saeed, S., Handoko, L., Richmond, T., Spivakov, M., Burgess, D., & Stunnenberg, H. G. (2015). Dynamic Reorganization of Extremely Long-Range Promoter-Promoter Interactions between Two States of Pluripotency. *Cell Stem Cell*, *17*(6), 748–757. https://doi.org/10.1016/J.STEM.2015.11.010

Kane, L., Williamson, I., Flyamer, I. M., Kumar, Y., Hill, R. E., Lettice, L. A., & Bickmore, W. A. (2022). Cohesin is required for long-range enhancer action at the Shh locus. *Nature Structural & Molecular Biology*, *29*(9), 891–897. https://doi.org/10.1038/s41594-022-00821-8

Kanzler, B., Kuschert, S. J., Liu, Y. H., & Mallo, M. (1998). Hoxa-2 restricts the chondrogenic domain and inhibits bone formation during development of the branchial area. *Development*, *125*(14), 2587–2597. https://doi.org/10.1242/dev.125.14.2587

Khudyakov, J., & Bronner-Fraser, M. (2009). Comprehensive spatiotemporal analysis of early chick neural crest network genes. *Developmental Dynamics*, *238*(3), 716–723. https://doi.org/https://doi.org/10.1002/dvdy.21881

Kim, J. H., Rege, M., Valeri, J., Dunagin, M. C., Metzger, A., Titus, K. R., Gilgenast, T. G., Gong, W., Beagan, J. A., Raj, A., & Phillips-Cremins, J. E. (2019). LADL: light-activated dynamic looping for endogenous gene expression control. *Nature Methods*, *16*(7), 633–639. https://doi.org/10.1038/s41592-019-0436-5

Kirilenko, P., He, G., Mankoo, B. S., Mallo, M., Jones, R., & Bobola, N. (2011). Transient activation of meox1 is an early component of the gene regulatory network downstream of hoxa2 (20110118th ed.). *Mol Cell Biol*, *31*(6), 1301–1308. https://doi.org/10.1128/MCB.00705-10

Kitazawa, T., Fujisawa, K., Narboux-Neme, N., Arima, Y., Kawamura, Y., Inoue, T., Wada, Y., Kohro, T., Aburatani, H., Kodama, T., Kim, K. S., Sato, T., Uchijima, Y., Maeda, K., Miyagawa-Tomita, S., Minoux, M., Rijli, F. M., Levi, G., Kurihara, Y., & Kurihara, H. (2015). Distinct effects of Hoxa2 overexpression in cranial neural crest populations reveal that the mammalian hyomandibular-ceratohyal boundary maps within the styloid process (2015/04/19). *Dev Biol*, *402*(2), 162–174. https://doi.org/10.1016/j.ydbio.2015.04.007

Kitazawa, T., Minoux, M., Ducret, S., & Rijli, F. M. (2022). Different Ectopic Hoxa2 Expression Levels in Mouse Cranial Neural Crest Cells Result in Distinct Craniofacial Anomalies and Homeotic Phenotypes (20220131st ed.). *J Dev Biol*, *10*(1). https://doi.org/10.3390/jdb10010009

Klann, T. S., Black, J. B., Chellappan, M., Safi, A., Song, L., Hilton, I. B., Crawford, G. E., Reddy, T. E., & Gersbach, C. A. (2017). CRISPR–Cas9 epigenome editing enables high-throughput screening for functional regulatory elements in the human genome. *Nature Biotechnology*, *35*(6), 561–568. https://doi.org/10.1038/nbt.3853

Klemm, S. L., Shipony, Z., & Greenleaf, W. J. (2019). Chromatin accessibility and the regulatory epigenome. *Nature Reviews Genetics*, *20*(4), 207–220. https://doi.org/10.1038/s41576-018-0089-8

Kmita, M., & Duboule, D. (2003). Organizing Axes in Time and Space; 25 Years of Colinear Tinkering. *Science*, *301*(5631), 331–333. https://doi.org/10.1126/science.1085753

Kocabas, F., Zheng, J., Thet, S., Copeland, N. G., Jenkins, N. A., DeBerardinis, R. J., Zhang, C., & Sadek, H. A. (2012). Meis1 regulates the metabolic phenotype and oxidant defense of hematopoietic stem cells. *Blood*, *120*(25), 4963–4972. https://doi.org/10.1182/blood-2012-05-432260

Kontges, G., & Lumsden, A. (1996). Rhombencephalic neural crest segmentation is preserved throughout craniofacial ontogeny. *Development*, *122*(10), 3229–3242. https://doi.org/10.1242/dev.122.10.3229

Kragesteen, B. K., Spielmann, M., Paliou, C., Heinrich, V., Schopflin, R., Esposito, A., Annunziatella, C., Bianco, S., Chiariello, A. M., Jerkovic, I., Harabula, I., Guckelberger, P., Pechstein, M., Wittler, L., Chan, W. L., Franke, M., Lupianez, D. G., Kraft, K., Timmermann, B., . . . Andrey, G. (2018). Dynamic 3D chromatin architecture contributes to enhancer specificity and limb morphogenesis (20180927th ed.). *Nat Genet*, *50*(10), 1463–1473. https://doi.org/10.1038/s41588-018-0221-x

Kubo, N., Ishii, H., Xiong, X., Bianco, S., Meitinger, F., Hu, R., Hocker, J. D., Conte, M., Gorkin, D., Yu, M., Li, B., Dixon, J. R., Hu, M., Nicodemi, M., Zhao, H., & Ren, B. (2021). Promoter-proximal CTCF binding promotes distal enhancer-dependent gene activation. *Nature Structural & Molecular Biology*, *28*(2), 152–161. https://doi.org/10.1038/s41594-020-00539-5

Kulesa, P. M., & Fraser, S. E. (2002). Segmentation of the vertebrate hindbrain: a time-lapse analysis. *International Journal of Developmental Biology*, *42*(3), 385–392. https://doi.org/10.1387/IJDB.9654023

Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., Ziller, M. J., Amin, V., Whitaker, J. W., Schultz, M. D., Ward, L. D., Sarkar, A., Quon, G., Sandstrom, R. S., Eaton, M. L., Wu, Y.-C., . . . investigators, P. (2015). Integrative analysis

REFERENCES

of 111 reference human epigenomes. *Nature*, *518*(7539), 317–330. https://doi.org/10.1038/nature14248

Kundu, S., Ji, F., Sunwoo, H., Jain, G., Lee, J. T., Sadreyev, R. I., Dekker, J., & Kingston, R. E. (2017). Polycomb Repressive Complex 1 Generates Discrete Compacted Domains that Change during Differentiation. *Molecular Cell*, *65*(3), 432–446. https://doi.org/10.1016/J.MOLCEL.2017.01.009

Kutejova, E., Engist, B., Mallo, M., Kanzler, B., & Bobola, N. (2005). Hoxa2 downregulates Six2 in the neural crest-derived mesenchyme. *Development*, *132*(3), 469–478. https://doi.org/10.1242/dev.01536

Kutejova, E., Engist, B., Self, M., Oliver, G., Kirilenko, P., & Bobola, N. (2008). Six2 functions redundantly immediately downstream of Hoxa2. *Development*, *135*(8), 1463–1470. https://doi.org/10.1242/dev.017624

Kvon, E. Z. (2015). Using transgenic reporter assays to functionally characterize enhancers in animals. *Genomics*, *106*(3), 185–192. https://doi.org/10.1016/J.YGENO.2015.06.007

Kvon, E. Z., Waymack, R., Gad, M., & Wunderlich, Z. (2021). Enhancer redundancy in development and disease. https://doi.org/10.1038/s41576-020-00311-x

Kvon, E. Z., Zhu, Y., Kelman, G., Novak, C. S., Plajzer-Frick, I., Kato, M., Garvin, T. H., Pham, Q., Harrington, A. N., Hunter, R. D., Godoy, J., Meky, E. M., Akiyama, J. A., Afzal, V., Tran, S., Escande, F., Gilbert-Dussardier, B., Jean-Marçais, N., Hudaiberdiev, S., . . . Pennacchio, L. A. (2020). Comprehensive In Vivo Interrogation Reveals Phenotypic Impact of Human Enhancer Variants. *Cell*, *180*(6), 1262–1271. https://doi.org/10.1016/J.CELL.2020.02.031

La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M. E., Lönnerberg, P., Furlan, A., Fan, J., Borm, L. E., Liu, Z., van Bruggen, D., Guo, J., He, X., Barker, R., Sundström, E., Castelo-Branco, G., . . . Kharchenko, P. V. (2018). RNA velocity of single cells. *Nature*, *560*(7719), 494–498. https://doi.org/10.1038/s41586-018-0414-6

Lampe, X., Samad, O. A., Guiguen, A., Matis, C., Remacle, S., Picard, J. J., Rijli, F. M., & Rezsohazy, R. (2008). An ultraconserved Hox-Pbx responsive element resides in the coding sequence of Hoxa2 and is active in rhombomere 4 (2008/04/18). *Nucleic Acids Res*, *36*(10), 3214–3225. https://doi.org/10.1093/nar/gkn148

Landrum, M. J., Lee, J. M., Riley, G. R., Jang, W., Rubinstein, W. S., Church, D. M., & Maglott, D. R. (2014). ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Research*, *42*(D1), D980–D985. https://doi.org/10.1093/nar/gkt1113

Le Douarin, N., & Kalcheim, C. (1999). *The Neural Crest* (2nd ed.). Cambridge University Press. https://doi.org/DOI:10.1017/CBO9780511897948

Le Douarin, N. M., Creuzet, S., Couly, G., & Dupin, E. (2004). Neural crest cell plasticity and its limits. *Development*, *131*(19), 4637–4650. https://doi.org/10.1242/dev.01350

Le Douarin, N. M., & Teillet, M.-A. (1973). The migration of neural crest cells to the wall of the digestive tract in avian embryo. *Development*, *30*(1), 31–48. https://doi.org/10.1242/dev.30.1.31

Lettice, L. A., Heaney, S. J., Purdie, L. A., Li, L., de Beer, P., Oostra, B. A., Goode, D., Elgar, G., Hill, R. E., & de Graaff, E. (2003). A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly (2003/07/03). *Hum Mol Genet*, *12*(14), 1725–1735. https://doi.org/10.1093/hmg/ddg180

Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B. R., Sabo, P. J., Dorschner, M. O., Sandstrom, R., Bernstein, B., Bender, M. A., Groudine, M., Gnirke, A., Stamatoyannopoulos, J., Mirny, L. A., Lander, E. S., & Dekker, J. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, *326*(5950), 289–293. https://doi.org/10.1126/science.1181369

Lignell, A., Kerosuo, L., Streichan, S. J., Cai, L., & Bronner, M. E. (2017). Identification of a neural crest stem cell niche by Spatial Genomic Analysis. *Nature Communications*, *8*(1), 1830. https://doi.org/10.1038/s41467-017-01561-w

Long, H. K., Osterwalder, M., Welsh, I. C., Hansen, K., Davies, J. O. J., Liu, Y. E., Koska, M., Adams, A. T., Aho, R., Arora, N., Ikeda, K., Williams, R. M., Sauka-Spengler, T., Porteus, M. H., Mohun, T., Dickel, D. E., Swigut, T., Hughes, J. R., Higgs, D. R., ... Wysocka, J. (2020). Loss of Extreme Long-Range Enhancers in Human Neural Crest Drives a Craniofacial Disorder (20200928th ed.). *Cell Stem Cell*, *27*(5), 765–783. https://doi.org/10.1016/j.stem.2020.09.001

Long, H. K., Prescott, S. L., & Wysocka, J. (2016). Ever-Changing Landscapes: Transcriptional Enhancers in Development and Evolution. *Cell*, *167*(5), 1170–1187. https://doi.org/10.1016/j.cell.2016.09.018

Loven, J., Hoke, H. A., Lin, C. Y., Lau, A., Orlando, D. A., Vakoc, C. R., Bradner, J. E., Lee, T. I., & Young, R. A. (2013). Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell*, *153*(2), 320–334. https://doi.org/10.1016/j.cell.2013.03.036

Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F., & Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8Å resolution. *Nature*, *389*(6648), 251–260. https://doi.org/10.1038/38444

Lukoseviciute, M., Gavriouchkina, D., Williams, R. M., Hochgreb-Hagele, T., Senanayake, U., Chong-Morrison, V., Thongjuea, S., Repapi, E., Mead, A., & Sauka-Spengler, T. (2018). From Pioneer to Repressor: Bimodal foxd3 Activity Dynamically Remodels Neural Crest Regulatory Landscape In Vivo. *Developmental Cell*, *47*(5), 608–628. https://doi.org/10.1016/J.DEVCEL.2018.11.009

Lumsden, A., Sprawson, N., & Graham, A. (1991). *Segmental origin and migration of neural crest cells in the hindbrain region of the chick embryo* (tech. rep.).

Lupiáñez, D. G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J. M., Laxova, R., Santos-Simarro, F., Gilbert-Dussardier, B., Wittler, L., Borschiwer, M., Haas, S. A., Osterwalder, M., Franke, M., Timmermann, B., Hecht, J., ... Mundlos, S. (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell*, *161*(5), 1012–1025. https://doi.org/10.1016/j.cell.2015.04.004

REFERENCES

Mach, P., Kos, P. I., Zhan, Y., Cramard, J., Gaudin, S., Tünnermann, J., Marchi, E., Eglinger, J., Zuin, J., Kryzhanovska, M., Smallwood, S., Gelman, L., Roth, G., Nora, E. P., Tiana, G., & Giorgetti, L. (2022). Cohesin and CTCF control the dynamics of chromosome folding. *Nature Genetics*, *54*(12), 1907–1918. https://doi.org/10.1038/s41588-022-01232-7

Machon, O., Masek, J., Machonova, O., Krauss, S., & Kozmik, Z. (2015). Meis2 is essential for cranial and cardiac neural crest development. *BMC Developmental Biology*, *15*(1), 40. https://doi.org/10.1186/s12861-015-0093-6

Maconochie, M., Krishnamurthy, R., Nonchev, S., Meier, P., Manzanares, M., Mitchell, P. J., & Krumlauf, R. (1999). Regulation of Hoxa2 in cranial neural crest cells involves members of the AP-2 family. *Development*, *126*(7), 1483–1494. https://doi.org/10.1242/dev.126.7.1483

Maconochie, M. K., Nonchev, S., Studer, M., Chan, S. K., Pöpperl, H., Sham, M. H., Mann, R. S., & Krumlauf, R. (1997). Cross-regulation in the mouse HoxB complex: the expression of Hoxb2 in rhombomere 4 is regulated by Hoxb1. *Genes & Development*, *11*(14), 1885–1895. https://doi.org/10.1101/gad.11.14.1885

Maconochie, M. K., Nonchev, S., Manzanares, M., Marshall, H., & Krumlauf, R. (2001). Differences in Krox20-Dependent Regulation of Hoxa2 and Hoxb2 during Hindbrain Development. *Developmental Biology*, *233*(2), 468–481. https://doi.org/10.1006/DBIO.2001.0197

Madsen, J. G., Madsen, M. S., Rauch, A., Traynor, S., Van Hauwaert, E. L., Haakonsson, A. K., Javierre, B. M., Hyldahl, M., Fraser, P., & Mandrup, S. (2020). Highly interconnected enhancer communities control lineage-determining genes in human mesenchymal stem cells. *Nature Genetics*, *52*(11), 1227–1238. https://doi.org/10.1038/s41588-020-0709-z

Makki, N., & Capecchi, M. R. (2010). Hoxa1 lineage tracing indicates a direct role for Hoxa1 in the development of the inner ear, the heart, and the third rhombomere. *Developmental Biology*, *341*(2), 499–509. https://doi.org/10.1016/J.YDBIO.2010.02.014

Mallen, J., Kalsan, M., Zarrineh, P., Bridoux, L., Ahmad, S., & Bobola, N. (2021). Molecular Characterization of HOXA2 and HOXA3 Binding Properties (20211203rd ed.). *J Dev Biol*, *9*(4). https://doi.org/10.3390/jdb9040055

Mallo, M., Wellik, D. M., & Deschamps, J. (2010). Hox genes and regional patterning of the vertebrate body plan. *Developmental Biology*, *344*(1), 7–15. https://doi.org/10.1016/J.YDBIO.2010.04.024

Manley, N. R., & Capecchi, M. R. (1995). The role of Hoxa-3 in mouse thymus and thyroid development. *Development*, *121*(7), 1989–2003. https://doi.org/10.1242/dev.121.7.1989

Manley, N. R., & Capecchi, M. R. (1997). Hox group 3 paralogous genes act synergistically in the formation of somitic and neural crest-derived structures (1998/01/27). *Dev Biol*, *192*(2), 274–288. https://doi.org/10.1006/dbio.1997.8765

Mansour, M. R., Abraham, B. J., Anders, L., Berezovskaya, A., Gutierrez, A., Durbin, A. D., Etchin, J., Lawton, L., Sallan, S. E., Silverman, L. B., Loh, M. L., Hunger, S. P., Sanda, T., Young, R. A., & Look, A. T. (2014). An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element. *Science*, *346*(6215), 1373–1377. https://doi.org/10.1126/science.1259037

Manzanares, M., Cordes, S., Ariza-McNaughton, L., Sadl, V., Maruthainar, K., Barsh, G., & Krumlauf, R. (1999). Conserved and distinct roles of kreisler in regulation of the paralogous Hoxa3 and Hoxb3 genes. *Development*, *126*(4), 759–769. https://doi.org/10.1242/dev.126.4.759

Manzanares, M., Cordes, S., Kwan, C.-T., Sham, M. H., Barsh, G. S., & Krumlauf, R. (1997). Segmental regulation of Hoxb-3 by kreisler. *Nature*, *387*(6629), 191–195. https://doi.org/10.1038/387191a0

Manzanares, M., Nardelli, J., Gilardi-Hebenstreit, P., Marshall, H., Giudicelli, F., Martínez-Pastor, M. T., Krumlauf, R., & Charnay, P. (2002). Krox20 and kreisler co-operate in the transcriptional control of segmental expression of Hoxb3 in the developing hindbrain. *The EMBO Journal*, *21*(3), 365–376. https://doi.org/https://doi.org/10.1093/emboj/21.3.365

Martik, M. L., & Bronner, M. E. (2017). Regulatory Logic Underlying Diversification of the Neural Crest. https://doi.org/10.1016/j.tig.2017.07.015

Martik, M. L., & Bronner, M. E. (2021). Riding the crest to get a head: neural crest evolution in vertebrates. https://doi.org/10.1038/s41583-021-00503-2

McCord, R. P., Kaplan, N., & Giorgetti, L. (2020). Chromosome Conformation Capture and Beyond: Toward an Integrative View of Chromosome Structure and Function. *Molecular Cell*, *77*(4), 688–708. https://doi.org/10.1016/j.molcel.2019.12.021

McEllin, J. A., Alexander, T. B., Tümpel, S., Wiedemann, L. M., & Krumlauf, R. (2016). Analyses of fugu hoxa2 genes provide evidence for subfunctionalization of neural crest cell and rhombomere cis-regulatory modules during vertebrate evolution. *Developmental Biology*, *409*(2), 530–542. https://doi.org/10.1016/j.ydbio.2015.11.006

Mifsud, B., Tavares-Cadete, F., Young, A. N., Sugar, R., Schoenfelder, S., Ferreira, L., Wingett, S. W., Andrews, S., Grey, W., Ewels, P. A., Herman, B., Happe, S., Higgs, A., LeProust, E., Follows, G. A., Fraser, P., Luscombe, N. M., & Osborne, C. S. (2015). Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C (20150504th ed.). *Nat Genet*, *47*(6), 598–606. https://doi.org/10.1038/ng.3286

Minoux, M., Antonarakis, G. S., Kmita, M., Duboule, D., & Rijli, F. M. (2009). Rostral and caudal pharyngeal arches share a common neural crest ground pattern. *Development*, *136*(4), 637–645. https://doi.org/10.1242/dev.028621

Minoux, M., Holwerda, S., Vitobello, A., Kitazawa, T., Kohler, H., Stadler, M. B., & Rijli, F. M. (2017). Gene bivalency at Polycomb domains regulates cranial neural crest positional identity. *Science*, *355*(6332). https://doi.org/10.1126/science.aal2913

Minoux, M., Kratochwil, C. F., Ducret, S., Amin, S., Kitazawa, T., Kurihara, H., Bobola, N., Vilain, N., & Rijli, F. M. (2013). Mouse Hoxa2 mutations provide a model for microtia and auricle duplication (20130925th ed.). *Development*, *140*(21), 4386–4397. https://doi.org/10.1242/dev.098046

Minoux, M., & Rijli, F. M. (2010). Molecular mechanisms of cranial neural crest cell migration and patterning in craniofacial development. *Development*, *137*(16), 2605–2621. https://doi.org/10.1242/dev.040048

REFERENCES

Moens, C. B., & Selleri, L. (2006). Hox cofactors in vertebrate development (20060303rd ed.). *Dev Biol*, *291*(2), 193–206. https://doi.org/10.1016/j.ydbio.2005.10.032

Monsoro-Burq, A. H., Wang, E., & Harland, R. (2005). Msx1 and Pax3 Cooperate to Mediate FGF8 and WNT Signals during Xenopus Neural Crest Induction. *Developmental Cell*, *8*(2), 167–178. https://doi.org/10.1016/J.DEVCEL.2004.12.017

Montavon, T., Soshnikova, N., Mascrez, B., Joye, E., Thevenet, L., Splinter, E., De Laat, W., Spitz, F., & Duboule, D. (2011). A Regulatory Archipelago Controls Hox Genes Transcription in Digits. *Cell*, *147*(5), 1132–1145. https://doi.org/10.1016/J.CELL.2011.10.023

Moorthy, S. D., Davidson, S., Shchuka, V. M., Singh, G., Malek-Gilani, N., Langroudi, L., Martchenko, A., So, V., Macpherson, N. N., & Mitchell, J. A. (2017). Enhancers and super-enhancers have an equivalent regulatory role in embryonic stem cells through regulation of single or multiple genes (2016/11/30). *Genome Res*, *27*(2), 246–258. https://doi.org/10.1101/gr.210930.116

Morrison, A., Moroni, M. C., Ariza-McNaughton, L., Krumlauf, R., & Mavilio, F. (1996). In vitro and transgenic analysis of a human HOXD4 retinoid-responsive enhancer. *Development*, *122*(6), 1895–1907. https://doi.org/10.1242/dev.122.6.1895

Murillo-Rincón, A. P., & Kaucka, M. (2020). Insights Into the Complexity of Craniofacial Development From a Cellular Perspective. *Frontiers in Cell and Developmental Biology*, *8*. https://www.frontiersin.org/articles/10.3389/fcell.2020.620735

Nikitina, N., Sauka-Spengler, T., & Bronner-Fraser, M. (2008). Dissecting early regulatory relationships in the lamprey neural crest gene network. *Proceedings of the National Academy of Sciences*, *105*(51), 20083–20088. https://doi.org/10.1073/pnas.0806009105

Noden, D. M. (1983). The role of the neural crest in patterning of avian cranial skeletal, connective, and muscle tissues. *Developmental Biology*, *96*(1), 144–165. https://doi.org/https://doi.org/10.1016/0012-1606(83)90318-4

Nolte, C., Rastegar, M., Amores, A., Bouchard, M., Grote, D., Maas, R., Kovacs, E. N., Postlethwait, J., Rambaldi, I., Rowan, S., Yan, Y. L., Zhang, F., & Featherstone, M. (2006). Stereospecificity and PAX6 function direct Hoxd4 neural enhancer activity along the antero-posterior axis. *Developmental Biology*, *299*(2), 582–593. https://doi.org/10.1016/J.YDBIO.2006.08.061

Nonchev, S., Maconochie, M., Vesque, C., Aparicio, S., Ariza-McNaughton, L., Manzanares, M., Maruthainar, K., Kuroiwa, A., Brenner, S., Charnay, P., & Krumlauf, R. (1996). The conserved role of Krox-20 in directing Hox gene expression during vertebrate hindbrain segmentation (1996/09/03). *Proc Natl Acad Sci U S A*, *93*(18), 9339–9345. https://doi.org/10.1073/pnas.93.18.9339

Nonchev, S., Vesque, C., Maconochie, M., Seitanidou, T., Ariza-McNaughton, L., Frain, M., Marshall, H., Sham, M. H., Krumlauf, R., & Charnay, P. (1996). Segmental expression of Hoxa-2 in the hindbrain is directly regulated by Krox-20. *Development*, *122*(2), 543–554. https://doi.org/10.1242/dev.122.2.543

Nora, E. P., Goloborodko, A., Valton, A. L., Gibcus, J. H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L. A., & Bruneau, B. G. (2017). Targeted Degradation of CTCF Decouples Local Insulation of

Chromosome Domains from Genomic Compartmentalization. *Cell*, *169*(5), 930–944. https://doi.org/10.1016/j.cell.2017.05.004

Nora, E. P., Lajoie, B. R., Schulz, E. G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N. L., Meisig, J., Sedat, J., Gribnau, J., Barillot, E., Bluthgen, N., Dekker, J., & Heard, E. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre (20120411th ed.). *Nature*, *485*(7398), 381–385. https://doi.org/10.1038/nature11049

Novo, C. L., Javierre, B. M., Cairns, J., Segonds-Pichon, A., Wingett, S. W., Freire-Pritchett, P., Furlan-Magaril, M., Schoenfelder, S., Fraser, P., & Rugg-Gunn, P. J. (2018). Long-Range Enhancer Interactions Are Prevalent in Mouse Embryonic Stem Cells and Are Reorganized upon Pluripotent State Transition. *Cell Reports*, *22*(10), 2615–2627. https://doi.org/10.1016/J.CELREP.2018.02.040

Oh, S., Shao, J., Mitra, J., Xiong, F., D'Antonio, M., Wang, R., Garcia-Bassets, I., Ma, Q., Zhu, X., Lee, J.-H., Nair, S. J., Yang, F., Ohgi, K., Frazer, K. A., Zhang, Z. D., Li, W., & Rosenfeld, M. G. (2021). Enhancer release and retargeting activates disease-susceptibility genes. *Nature*, *595*(7869), 735–740. https://doi.org/10.1038/s41586-021-03577-1

Osterwalder, M., Barozzi, I., Tissieres, V., Fukuda-Yuzawa, Y., Mannion, B. J., Afzal, S. Y., Lee, E. A., Zhu, Y., Plajzer-Frick, I., Pickle, C. S., Kato, M., Garvin, T. H., Pham, Q. T., Harrington, A. N., Akiyama, J. A., Afzal, V., Lopez-Rios, J., Dickel, D. E., Visel, A., & Pennacchio, L. A. (2018). Enhancer redundancy provides phenotypic robustness in mammalian development (20180131st ed.). *Nature*, *554*(7691), 239–243. https://doi.org/10.1038/nature25461

Pachano, T., Sánchez-Gaya, V., Ealo, T., Mariner-Faulí, M., Bleckwehl, T., Asenjo, H. G., Respuela, P., Cruz-Molina, S., Muñoz-San Martín, M., Haro, E., van IJcken, W. F. J., Landeira, D., & Rada-Iglesias, A. (2021). Orphan CpG islands amplify poised enhancer regulatory activity and determine target gene responsiveness. *Nature Genetics*, *53*(7), 1036–1049. https://doi.org/10.1038/s41588-021-00888-x

Pachano, T., Haro, E., & Rada-Iglesias, A. (2022). Enhancer-gene specificity in development and disease. *Development*, *149*(11), dev186536. https://doi.org/10.1242/dev.186536

Paliou, C., Guckelberger, P., Schöpflin, R., Heinrich, V., Esposito, A., Chiariello, A. M., Bianco, S., Annunziatella, C., Helmuth, J., Haas, S., Jerković, I., Brieske, N., Wittler, L., Timmermann, B., Nicodemi, M., Vingron, M., Mundlos, S., & Andrey, G. (2019). Preformed chromatin topology assists transcriptional robustness of Shh during limb development. *Proceedings of the National Academy of Sciences*, *116*(25), 12390–12399. https://doi.org/10.1073/pnas.1900672116

Parker, H. J., Bronner, M. E., & Krumlauf, R. (2016). The vertebrate Hox gene regulatory network for hindbrain segmentation: Evolution and diversification. *BioEssays*, *38*(6), 526–538. https://doi.org/https://doi.org/10.1002/bies.201600010

Parker, H. J., De Kumar, B., Green, S. A., Prummel, K. D., Hess, C., Kaufman, C. K., Mosimann, C., Wiedemann, L. M., Bronner, M. E., & Krumlauf, R. (2019). A Hox-TALE regulatory circuit for neural crest patterning is conserved across vertebrates. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-09197-8

REFERENCES

Parker, H. J., Pushel, I., & Krumlauf, R. (2018). Coupling the roles of Hox genes to regulatory networks patterning cranial neural crest. https://doi.org/10.1016/j.ydbio.2018.03.016

Pasqualetti, M., Ori, M., Nardi, I., & Rijli, F. M. (2000). Ectopic Hoxa2 induction after neural crest migration results in homeosis of jaw elements in Xenopus (2000/11/15). *Development*, *127*(24), 5367–5378. https://doi.org/10.1242/dev.127.24.5367

Paulsen, J., Liyakat Ali, T. M., Nekrasov, M., Delbarre, E., Baudement, M.-O., Kurscheid, S., Tremethick, D., & Collas, P. (2019). Long-range interactions between topologically associating domains shape the four-dimensional genome during differentiation. *Nature Genetics*, *51*(5), 835–843. https://doi.org/10.1038/s41588-019-0392-0

Pennacchio, L. A., Ahituv, N., Moses, A. M., Prabhakar, S., Nobrega, M. A., Shoukry, M., Minovitsky, S., Dubchak, I., Holt, A., Lewis, K. D., Plajzer-Frick, I., Akiyama, J., De Val, S., Afzal, V., Black, B. L., Couronne, O., Eisen, M. B., Visel, A., & Rubin, E. M. (2006). In vivo enhancer analysis of human conserved non-coding sequences. *Nature*, *444*(7118), 499–502. https://doi.org/10.1038/nature05295

Perry, M. W., Boettiger, A. N., & Levine, M. (2011). Multiple enhancers ensure precision of gap gene-expression patterns in the Drosophila embryo. *Proceedings of the National Academy of Sciences*, *108*(33), 13570–13575. https://doi.org/10.1073/pnas.1109873108

Piceci, F., Morlino, S., Castori, M., Buffone, E., De Luca, A., Grammatico, P., & Guida, V. (2017). Identification of a second HOXA2 nonsense mutation in a family with autosomal dominant non-syndromic microtia and distinctive ear morphology (20160913th ed.). *Clin Genet*, *91*(5), 774–779. https://doi.org/10.1111/cge.12845

Plouhinec, J. L., Roche, D. D., Pegoraro, C., Figueiredo, A. L., Maczkowiak, F., Brunet, L. J., Milet, C., Vert, J. P., Pollet, N., Harland, R. M., & Monsoro-Burq, A. H. (2014). Pax3 and Zic1 trigger the early neural crest gene regulatory network by the direct activation of multiple key neural crest specifiers. *Developmental Biology*, *386*(2), 461–472. https://doi.org/10.1016/J.YDBIO.2013.12.010

Prescott, S. L., Srinivasan, R., Marchetto, M. C., Grishina, I., Narvaiza, I., Selleri, L., Gage, F. H., Swigut, T., & Wysocka, J. (2015). Enhancer divergence and cis-regulatory evolution in the human and chimp neural crest (20150910th ed.). *Cell*, *163*(1), 68–83. https://doi.org/10.1016/j.cell.2015.08.036

Prince, V., & Lumsden, A. (1994). Hoxa-2 expression in normal and transposed rhombomeres: independent regulation in the neural tube and neural crest. *Development*, *120*(4), 911–923. https://doi.org/10.1242/dev.120.4.911

Proudhon, C., Snetkova, V., Raviram, R., Lobry, C., Badri, S., Jiang, T., Hao, B., Trimarchi, T., Kluger, Y., Aifantis, I., Bonneau, R., & Skok, J. A. (2016). Active and Inactive Enhancers Cooperate to Exert Localized and Long-Range Control of Gene Regulation. *Cell Reports*, *15*(10), 2159–2169. https://doi.org/10.1016/J.CELREP.2016.04.087

Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S. A., Flynn, R. A., & Wysocka, J. (2011). A unique chromatin signature uncovers early developmental enhancers in humans. *Nature*, *470*(7333), 279–283. https://doi.org/10.1038/nature09692

Rao, S. S. P., Huang, S. C., Glenn St Hilaire, B., Engreitz, J. M., Perez, E. M., Kieffer-Kwon, K. R., Sanborn, A. L., Johnstone, S. E., Bascom, G. D., Bochkov, I. D., Huang, X., Shamim, M. S., Shin, J., Turner, D., Ye, Z., Omer, A. D., Robinson, J. T., Schlick, T., Bernstein, B. E., . . . Aiden, E. L. (2017). Cohesin Loss Eliminates All Loop Domains. *Cell*, *171*(2), 305–320. https://doi.org/10.1016/j.cell.2017.09.026

Rao, S. S., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., Robinson, J. T., Sanborn, A. L., Machol, I., Omer, A. D., Lander, E. S., & Aiden, E. L. (2014). A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell*, *159*(7), 1665–1680. https://doi.org/10.1016/J.CELL.2014.11.021

Rijli, F. M., Mark, M., Lakkaraju, S., Dierich, A., Dolle, P., & Chambon, P. (1993). A homeotic transformation is generated in the rostral branchial region of the head by disruption of Hoxa-2, which acts as a selector gene. *Cell*, *75*(7), 1333–1349. https://doi.org/10.1016/0092-8674(93)90620-6

Rinzema, N. J., Sofiadis, K., Tjalsma, S. J. D., Verstegen, M. J. A. M., Oz, Y., Valdes-Quezada, C., Felder, A.-K., Filipovska, T., van der Elst, S., de Andrade dos Ramos, Z., Han, R., Krijger, P. H. L., & de Laat, W. (2022). Building regulatory landscapes reveals that an enhancer can recruit cohesin to create contact domains, engage CTCF sites and activate distant genes. *Nature Structural & Molecular Biology*, *29*(6), 563–574. https://doi.org/10.1038/s41594-022-00787-7

Robson, M. I., Ringel, A. R., & Mundlos, S. (2019). Regulatory Landscaping: How Enhancer-Promoter Communication Is Sculpted in 3D. https://doi.org/10.1016/j.molcel.2019.05.032

Rosin, J. M., Li, W., Cox, L. L., Rolfe, S. M., Latorre, V., Akiyama, J. A., Visel, A., Kuramoto, T., Bobola, N., Turner, E. E., & Cox, T. C. (2016). A distal 594bp ECR specifies Hmx1 expression in pinna and lateral facial morphogenesis and is regulated by the Hox-Pbx-Meis complex. *Development*, *143*(14), 2582–2592. https://doi.org/10.1242/dev.133736

Rowley, M. J., & Corces, V. G. (2018). Organizational principles of 3D genome architecture. *Nature Reviews Genetics*, *19*(12), 789–800. https://doi.org/10.1038/s41576-018-0060-8

Rubin, A. J., Barajas, B. C., Furlan-Magaril, M., Lopez-Pajares, V., Mumbach, M. R., Howard, I., Kim, D. S., Boxer, L. D., Cairns, J., Spivakov, M., Wingett, S. W., Shi, M., Zhao, Z., Greenleaf, W. J., Kundaje, A., Snyder, M., Chang, H. Y., Fraser, P., & Khavari, P. A. (2017). Lineage-specific dynamic and pre-established enhancer-promoter contacts cooperate in terminal differentiation (20170814th ed.). *Nat Genet*, *49*(10), 1522–1528. https://doi.org/10.1038/ng.3935

Sabari, B. R., Dall'Agnese, A., Boija, A., Klein, I. A., Coffey, E. L., Shrinivas, K., Abraham, B. J., Hannett, N. M., Zamudio, A. V., Manteiga, J. C., Li, C. H., Guo, Y. E., Day, D. S., Schuijers, J., Vasile, E., Malik, S., Hnisz, D., Lee, T. I., Cisse, I. I., . . . Young, R. A. (2018). Coactivator condensation at super-enhancers links phase separation and gene control. *Science*, *361*(6400), eaar3958. https://doi.org/10.1126/science.aar3958

Sagai, T., Hosoya, M., Mizushina, Y., Tamura, M., & Shiroishi, T. (2005). Elimination of a long-range cis-regulatory module causes complete loss of limb-specific Shh expression and truncation of the mouse limb. *Development*, *132*(4), 797–803. https://doi.org/10.1242/dev.01613

REFERENCES

Sanborn, A. L., Rao, S. S., Huang, S. C., Durand, N. C., Huntley, M. H., Jewett, A. I., Bochkov, I. D., Chinnappan, D., Cutkosky, A., Li, J., Geeting, K. P., Gnirke, A., Melnikov, A., McKenna, D., Stamenova, E. K., Lander, E. S., & Aiden, E. L. (2015). Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes (20151023rd ed.). *Proc Natl Acad Sci U S A*, *112*(47), 6456–65. https://doi.org/10.1073/pnas.1518552112

Sanjana, N. E., Wright, J., Zheng, K., Shalem, O., Fontanillas, P., Joung, J., Cheng, C., Regev, A., & Zhang, F. (2016). High-resolution interrogation of functional elements in the noncoding genome. *Science*, *353*(6307), 1545–1549. https://doi.org/10.1126/science.aaf7613

Santagati, F., & Rijli, F. M. (2003). Cranial neural crest and the building of the vertebrate head. *Nat Rev Neurosci*, *4*(10), 806–818. https://doi.org/10.1038/nrn1221

Santagati, F., Minoux, M., Ren, S. Y., & Rijli, F. M. (2005). Temporal requirement of Hoxa2 in cranial neural crest skeletal morphogenesis. *Development*, *132*(22), 4927–4936. https://doi.org/10.1242/dev.02078

Sato, T., Sasai, N., & Sasai, Y. (2005). Neural crest determination by co-activation of Pax3 and Zic1 genes in Xenopus ectoderm. *Development*, *132*(10), 2355–2363. https://doi.org/10.1242/dev.01823

Sauka-Spengler, T., & Bronner-Fraser, M. (2008). A gene regulatory network orchestrates neural crest formation. *Nature Reviews Molecular Cell Biology*, *9*(7), 557–568. https://doi.org/10.1038/nrm2428

Sauka-Spengler, T., Meulemans, D., Jones, M., & Bronner-Fraser, M. (2007). Ancient Evolutionary Origin of the Neural Crest Gene Regulatory Network. *Developmental Cell*, *13*(3), 405–420. https://doi.org/https://doi.org/10.1016/j.devcel.2007.08.005

Schilling, T. F., Prince, V., & Ingham, P. W. (2001). Plasticity in Zebrafish hox Expression in the Hindbrain and Cranial Neural Crest. *Developmental Biology*, *231*(1), 201–216. https://doi.org/10.1006/DBIO.2000.9997

Schoenfelder, S., & Fraser, P. (2019). Long-range enhancer-promoter contacts in gene expression control. *Nat Rev Genet*, *20*(8), 437–455. https://doi.org/10.1038/s41576-019-0128-0

Schoenfelder, S., Furlan-Magaril, M., Mifsud, B., Tavares-Cadete, F., Sugar, R., Javierre, B. M., Nagano, T., Katsman, Y., Sakthidevi, M., Wingett, S. W., Dimitrova, E., Dimond, A., Edelman, L. B., Elderkin, S., Tabbada, K., Darbo, E., Andrews, S., Herman, B., Higgs, A., . . . Fraser, P. (2015). The pluripotent regulatory circuitry connecting promoters to their long-range interacting elements (20150309th ed.). *Genome Res*, *25*(4), 582–597. https://doi.org/10.1101/gr.185272.114

Schoenfelder, S., Javierre, B. M., Furlan-Magaril, M., Wingett, S. W., & Fraser, P. (2018). Promoter Capture Hi-C: High-resolution, Genome-wide Profiling of Promoter Interactions (20180628th ed.). *J Vis Exp*, (136). https://doi.org/10.3791/57320

Schoenfelder, S., Sugar, R., Dimond, A., Javierre, B. M., Armstrong, H., Mifsud, B., Dimitrova, E., Matheson, L., Tavares-Cadete, F., Furlan-Magaril, M., Segonds-Pichon, A., Jurkowski, W., Wingett, S. W., Tabbada, K., Andrews, S., Herman, B., LeProust, E., Osborne, C. S., Koseki, H., . . . Elderkin, S. (2015). Polycomb repressive complex PRC1 spatially constrains the mouse embryonic

stem cell genome (20150831st ed.). *Nat Genet*, *47*(10), 1179–1186. https://doi.org/10.1038/ng.3393

Scholes, C., Biette, K. M., Harden, T. T., & DePace, A. H. (2019). Signal Integration by Shadow Enhancers and Enhancer Duplications Varies across the Drosophila Embryo. *Cell Reports*, *26*(9), 2407–2418. https://doi.org/10.1016/J.CELREP.2019.01.115

Schorle, H., Meier, P., Buchert, M., Jaenisch, R., & Mitchell, P. J. (1996). Transcription factor AP-2 essential for cranial closure and craniofacial development. *Nature*, *381*(6579), 235–238. https://doi.org/10.1038/381235a0

Schwarz, D., Varum, S., Zemke, M., Schöler, A., Baggiolini, A., Draganova, K., Koseki, H., Schübeler, D., & Sommer, L. (2014). Ezh2 is required for neural crest-derived cartilage and bone formation. *Development*, *141*(4), 867–877. https://doi.org/10.1242/dev.094342

Schwarzer, W., Abdennur, N., Goloborodko, A., Pekowska, A., Fudenberg, G., Loe-Mie, Y., Fonseca, N. A., Huber, W., Haering, C. H., Mirny, L., & Spitz, F. (2017). Two independent modes of chromatin organization revealed by cohesin removal (20170927th ed.). *Nature*, *551*(7678), 51–56. https://doi.org/10.1038/nature24281

Sechrist, J., Serbedzija, G. N., Scherson, T., Fraser, S. E., & Bronner-Fraser, M. (1993). Segmental migration of the hindbrain neural crest does not arise from its segmental generation. *Development*, *118*(3), 691–703. https://doi.org/10.1242/dev.118.3.691

Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., & Cavalli, G. (2012). Three-dimensional folding and functional organization principles of the Drosophila genome (20120119th ed.). *Cell*, *148*(3), 458–472. https://doi.org/10.1016/j.cell.2012.01.010

Shen, Y., Yue, F., McCleary, D. F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Lobanenkov, V. V., & Ren, B. (2012). A map of the cis-regulatory sequences in the mouse genome. *Nature*, *488*(7409), 116–120. https://doi.org/10.1038/nature11243

Shin, H. Y., Willi, M., Yoo, K. H., Zeng, X., Wang, C., Metser, G., & Hennighausen, L. (2016). Hierarchy within the mammary STAT5-driven Wap super-enhancer. *Nature Genetics*, *48*(8), 904–911. https://doi.org/10.1038/ng.3606

Shlyueva, D., Stampfel, G., & Stark, A. (2014). Transcriptional enhancers: From properties to genome-wide predictions. https://doi.org/10.1038/nrg3682

Siersbæk, R., Rabiee, A., Nielsen, R., Sidoli, S., Traynor, S., Loft, A., Poulsen, L. L. C., Rogowska-Wrzesinska, A., Jensen, O. N., & Mandrup, S. (2014). Transcription Factor Cooperativity in Early Adipogenic Hotspots and Super-Enhancers. *Cell Reports*, *7*(5), 1443–1455. https://doi.org/10.1016/J.CELREP.2014.04.042

Siersbaek, R., Madsen, J. G. S., Javierre, B. M., Nielsen, R., Bagge, E. K., Cairns, J., Wingett, S. W., Traynor, S., Spivakov, M., Fraser, P., & Mandrup, S. (2017). Dynamic Rewiring of Promoter-Anchored Chromatin Loops during Adipocyte Differentiation. *Mol Cell*, *66*(3), 420–435. https://doi.org/10.1016/j.molcel.2017.04.010

REFERENCES

Siismets, E. M., & Hatch, N. E. (2020). Cranial Neural Crest Cells and Their Role in the Pathogenesis of Craniofacial Anomalies and Coronal Craniosynostosis. *Journal of Developmental Biology*, *8*(3), 18. https://doi.org/10.3390/jdb8030018

Simeonov, D. R., Gowen, B. G., Boontanrart, M., Roth, T. L., Gagnon, J. D., Mumbach, M. R., Satpathy, A. T., Lee, Y., Bray, N. L., Chan, A. Y., Lituiev, D. S., Nguyen, M. L., Gate, R. E., Subramaniam, M., Li, Z., Woo, J. M., Mitros, T., Ray, G. J., Curie, G. L., . . . Marson, A. (2017). Discovery of stimulation-responsive immune enhancers with CRISPR activation. *Nature*, *549*(7670), 111–115. https://doi.org/10.1038/nature23875

Simões-Costa, M., & Bronner, M. E. (2015). Establishing neural crest identity: a gene regulatory recipe. https://doi.org/10.1242/dev.105445

Soldatov, R., Kaucka, M., Kastriti, M. E., Petersen, J., Chontorotzea, T., Englmaier, L., Akkuratova, N., Yang, Y., Häring, M., Dyachuk, V., Bock, C., Farlik, M., Piacentino, M. L., Boismoreau, F., Hilscher, M. M., Yokota, C., Qian, X., Nilsson, M., Bronner, M. E., . . . Adameyko, I. (2019). Spatiotemporal structure of cell fate decisions in murine neural crest. *Science*, *364*(6444), eaas9536. https://doi.org/10.1126/science.aas9536

Song, S. H., Hou, C., & Dean, A. (2007). A Positive Role for NLI/Ldb1 in Long-Range β-Globin Locus Control Region Function. *Molecular Cell*, *28*(5), 810–822. https://doi.org/10.1016/J.MOLCEL.2007.09.025

Spielmann, M., Lupianez, D. G., & Mundlos, S. (2018). Structural variation in the 3D genome. *Nat Rev Genet*, *19*(7), 453–467. https://doi.org/10.1038/s41576-018-0007-0

Studer, M., Pöpperl, H., Marshall, H., Kuroiwa, A., & Krumlauf, R. (1994). Role of a Conserved Retinoic Acid Response Element in Rhombomere Restriction of Hoxb-1. *Science*, *265*(5179), 1728–1732. https://doi.org/10.1126/science.7916164

Stuhlmiller, T. J., & García-Castro, M. I. (2012a). Current perspectives of the signaling pathways directing neural crest induction. *Cellular and Molecular Life Sciences*, *69*(22), 3715–3737. https://doi.org/10.1007/s00018-012-0991-8

Stuhlmiller, T. J., & García-Castro, M. I. (2012b). FGF/MAPK signaling is required in the gastrula epiblast for avian neural crest induction. *Development*, *139*(2), 289–300. https://doi.org/10.1242/dev.070276

Symmons, O., Uslu, V. V., Tsujimura, T., Ruf, S., Nassari, S., Schwarzer, W., Ettwiller, L., & Spitz, F. (2014). Functional and topological characteristics of mammalian regulatory domains. *Genome Research*, *24*(3), 390–400. https://doi.org/10.1101/gr.163519.113

Szabó, A., & Mayor, R. (2018). Mechanisms of Neural Crest Migration. *Annual Review of Genetics*, *52*(1), 43–63. https://doi.org/10.1146/annurev-genet-120417-031559

Takahashi, Y., Sipp, D., & Enomoto, H. (2013). Tissue Interactions in Neural Crest Cell Development and Disease. *Science*, *341*(6148), 860–863. https://doi.org/10.1126/science.1230717

Tamaru, H. (2010). Confining euchromatin/heterochromatin territory: jumonji crosses the line. *Genes & Development*, *24*(14), 1465–1478. https://doi.org/10.1101/gad.1941010

Thiecke, M. J., Wutz, G., Muhar, M., Tang, W., Bevan, S., Malysheva, V., Stocsits, R., Neumann, T., Zuber, J., Fraser, P., Schoenfelder, S., Peters, J. M., & Spivakov, M. (2020). Cohesin-Dependent and -Independent Mechanisms Mediate Chromosomal Contacts between Promoters and Enhancers. *Cell Reports*, *32*(3), 107929. https://doi.org/10.1016/J.CELREP.2020.107929

Thomas, H. F., Kotova, E., Jayaram, S., Pilz, A., Romeike, M., Lackner, A., Penz, T., Bock, C., Leeb, M., Halbritter, F., Wysocka, J., & Buecker, C. (2021). Temporal dissection of an enhancer cluster reveals distinct temporal and functional contributions of individual elements (2021/01/23). *Mol Cell*, *81*(5), 969–982. https://doi.org/10.1016/j.molcel.2020.12.047

Trainor, P., & Krumlauf, R. (2000). Plasticity in mouse neural crest cells reveals a new patterning role for cranial mesoderm. *Nature Cell Biology*, *2*(2), 96–102. https://doi.org/10.1038/35000051

Trainor, P. A., Ariza-McNaughton, L., & Krumlauf, R. (2002). Role of the Isthmus and FGFs in Resolving the Paradox of Neural Crest Plasticity and Prepatterning. *Science*, *295*(5558), 1288–1291. https://doi.org/10.1126/science.1064540

Trainor, P. A., & Krumlauf, R. (2000). Patterning the cranial neural crest: Hinbrain segmentation and hox gene plasticity. *Nature Reviews Neuroscience*, *1*(2), 116–124. https://doi.org/10.1038/35039056

Trainor, P. A., & Krumlauf, R. (2001). Hox genes, neural crest cells and branchial arch patterning. *Current Opinion in Cell Biology*, *13*(6), 698–705. https://doi.org/10.1016/S0955-0674(00)00273-8

Tsai, A., Alves, M. R. P., & Crocker, J. (2019). Multi-enhancer transcriptional hubs confer phenotypic robustness (D. N. Arnosti, J. K. Tyler, A. H. DePace, & H. Garcia, Eds.). *eLife*, *8*, e45325. https://doi.org/10.7554/eLife.45325

Tumpel, S., Cambronero, F., Ferretti, E., Blasi, F., Wiedemann, L. M., & Krumlauf, R. (2007). Expression of Hoxa2 in rhombomere 4 is regulated by a conserved cross-regulatory mechanism dependent upon Hoxb1 (20061025th ed.). *Dev Biol*, *302*(2), 646–660. https://doi.org/10.1016/j.ydbio.2006.10.029

Tumpel, S., Cambronero, F., Sims, C., Krumlauf, R., & Wiedemann, L. M. (2008). A regulatory module embedded in the coding region of Hoxa2 controls expression in rhombomere 2 (20081222nd ed.). *Proc Natl Acad Sci U S A*, *105*(51), 20077–20082. https://doi.org/10.1073/pnas.0806360105

Tümpel, S., Cambronero, F., Wiedemann, L. M., & Krumlauf, R. (2006). Evolution of cis elements in the differential expression of two Hoxa2 coparalogous genes in pufferfish (Takifugu rubripes). *Proceedings of the National Academy of Sciences*, *103*(14), 5419–5424. https://doi.org/10.1073/pnas.0600993103

Tümpel, S., Wiedemann, L. M., & Krumlauf, R. (2009). Chapter 8 Hox Genes and Segmentation of the Vertebrate Hindbrain. In *Current topics in developmental biology* (pp. 103–137). Academic Press. https://doi.org/10.1016/S0070-2153(09)88004-6

Vieux-Rochas, M., Mascrez, B., Krumlauf, R., & Duboule, D. (2013). Combined function of HoxA and HoxB clusters in neural crest cells. *Developmental Biology*, *382*(1), 293–301. https://doi.org/10.1016/J.YDBIO.2013.06.027

REFERENCES

Visel, A., Akiyama, J. A., Shoukry, M., Afzal, V., Rubin, E. M., & Pennacchio, L. A. (2009). Functional autonomy of distant-acting human enhancers. *Genomics*, *93*(6), 509–513. https://doi.org/10.1016/J.YGENO.2009.02.002

Visel, A., Minovitsky, S., Dubchak, I., & Pennacchio, L. A. (2007). VISTA Enhancer Browser—a database of tissue-specific human enhancers. *Nucleic Acids Research*, *35*(suppl_1), D88–D92. https://doi.org/10.1093/nar/gkl822

Wang, X., Cairns, M. J., & Yan, J. (2019). Super-enhancers in transcriptional regulation and genome organization. *Nucleic Acids Research*, *47*(22), 11481–11496. https://doi.org/10.1093/nar/gkz1038

Weintraub, A. S., Li, C. H., Zamudio, A. V., Sigova, A. A., Hannett, N. M., Day, D. S., Abraham, B. J., Cohen, M. A., Nabet, B., Buckley, D. L., Guo, Y. E., Hnisz, D., Jaenisch, R., Bradner, J. E., Gray, N. S., & Young, R. A. (2017). YY1 Is a Structural Regulator of Enhancer-Promoter Loops. *Cell*, *171*(7), 1573–1588. https://doi.org/10.1016/J.CELL.2017.11.008

Whyte, W. A., Orlando, D. A., Hnisz, D., Abraham, B. J., Lin, C. Y., Kagey, M. H., Rahl, P. B., Lee, T. I., & Young, R. A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell*, *153*(2), 307–319. https://doi.org/10.1016/j.cell.2013.03.035

Wilderman, A., VanOudenhove, J., Kron, J., Noonan, J. P., & Cotney, J. (2018). High-Resolution Epigenomic Atlas of Human Embryonic Craniofacial Development. *Cell Reports*, *23*(5), 1581–1597. https://doi.org/10.1016/J.CELREP.2018.03.129

Wilkinson, D. G., Bhatt, S., Cook, M., Boncinelli, E., & Krumlauf, R. (1989). Segmental expression of Hox-2 homoeobox-containing genes in the developing mouse hindbrain. *Nature*, *341*(6241), 405–409. https://doi.org/10.1038/341405a0

Will, A. J., Cova, G., Osterwalder, M., Chan, W. L., Wittler, L., Brieske, N., Heinrich, V., De Villartay, J. P., Vingron, M., Klopocki, E., Visel, A., Lupianez, D. G., & Mundlos, S. (2017). Composition and dosage of a multipartite enhancer cluster control developmental expression of Ihh (Indian hedgehog). *Nature Genetics*, *49*(10), 1539–1545. https://doi.org/10.1038/ng.3939

Williams, R. M., Candido-Ferreira, I., Repapi, E., Gavriouchkina, D., Senanayake, U., Ling, I. T., Telenius, J., Taylor, S., Hughes, J., & Sauka-Spengler, T. (2019). Reconstruction of the Global Neural Crest Gene Regulatory Network In Vivo. *Developmental Cell*, *51*(2), 255–276. https://doi.org/10.1016/J.DEVCEL.2019.10.003

Williamson, I., Kane, L., Devenney, P. S., Flyamer, I. M., Anderson, E., Kilanowski, F., Hill, R. E., Bickmore, W. A., & Lettice, L. A. (2019). Developmentally regulated Shh expression is robust to TAD perturbations (20190930th ed.). *Development*, *146*(19). https://doi.org/10.1242/dev.179523

Wilson, N. K., Schoenfelder, S., Hannah, R., Sanchez Castillo, M., Schutte, J., Ladopoulos, V., Mitchelmore, J., Goode, D. K., Calero-Nieto, F. J., Moignard, V., Wilkinson, A. C., Jimenez-Madrid, I., Kinston, S., Spivakov, M., Fraser, P., & Gottgens, B. (2016). Integrated genome-scale analysis of the transcriptional regulatory landscape in a blood stem/progenitor cell model (20160125th ed.). *Blood*, *127*(13), 12–23. https://doi.org/10.1182/blood-2015-10-677393

Wunderlich, Z., Bragdon, M. D., Vincent, B. J., White, J. A., Estrada, J., & DePace, A. H. (2015). Krüppel Expression Levels Are Maintained through Compensatory Evolution of Shadow Enhancers. *Cell Reports*, *12*(11), 1740–1747. https://doi.org/https://doi.org/10.1016/j.celrep.2015.08.021

Yanagisawa, H., Clouthier, D. E., Richardson, J. A., Charité, J., & Olson, E. N. (2003). Targeted deletion of a branchial arch-specific enhancer reveals a role ofdHAND in craniofacial development. *Development*, *130*(6), 1069–1078. https://doi.org/10.1242/dev.00337

Zabidi, M. A., & Stark, A. (2016). Regulatory Enhancer–Core-Promoter Communication via Transcription Factors and Cofactors. *Trends in Genetics*, *32*(12), 801–814. https://doi.org/10.1016/j.tig.2016.10.003

Zalc, A., Sinha, R., Gulati, G. S., Wesche, D. J., Daszczuk, P., Swigut, T., Weissman, I. L., & Wysocka, J. (2021). Reactivation of the pluripotency program precedes formation of the cranial neural crest. *Science*, *371*(6529), eabb4776. https://doi.org/10.1126/science.abb4776

Zaugg, J. B., Sahlén, P., Andersson, R., Alberich-Jorda, M., de Laat, W., Deplancke, B., Ferrer, J., Mandrup, S., Natoli, G., Plewczynski, D., Rada-Iglesias, A., & Spicuglia, S. (2022). Current challenges in understanding the role of enhancers in disease. *Nature Structural & Molecular Biology*, *29*(12), 1148–1158. https://doi.org/10.1038/s41594-022-00896-3

Zentner, G. E., Tesar, P. J., & Scacheri, P. C. (2011). Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. *Genome Research*, *21*(8), 1273–1283. https://doi.org/10.1101/gr.122382.111

Zhan, Y., Mariani, L., Barozzi, I., Schulz, E. G., Blüthgen, N., Stadler, M., Tiana, G., & Giorgetti, L. (2017). Reciprocal insulation analysis of Hi-C data shows that TADs represent a functionally but not structurally privileged scale in the hierarchical folding of chromosomes. *Genome Research*, *27*(3), 479–490. https://doi.org/10.1101/gr.212803.116

Zhou, V. W., Goren, A., & Bernstein, B. E. (2011). Charting histone modifications and the functional organization of mammalian genomes. *Nature Reviews Genetics*, *12*(1), 7–18. https://doi.org/10.1038/nrg2905

Zuin, J., Roth, G., Zhan, Y., Cramard, J., Redolfi, J., Piskadlo, E., Mach, P., Kryzhanovska, M., Tihanyi, G., Kohler, H., Eder, M., Leemans, C., van Steensel, B., Meister, P., Smallwood, S., & Giorgetti, L. (2022). Nonlinear control of transcription through enhancer-promoter interactions (2022/04/15). *Nature*, *604*(7906), 571–577. https://doi.org/10.1038/s41586-022-04570-y

Sandra Kessler PhD

Contact

Address: Basel, Schweiz   Mobile: +41 76 814 05 22

E-Mail: sandy.kessler22@gmail.com

Skills & Expertise

Bioinformatics

- NGS data analysis (read mapping, peak calling, data QC, differential gene expression, TF motif analysis, etc.) and visualization using R (Rstudio, Bioconductor) and UNIX command line interface

High-throughput methods

- ChIP-seq, bulk RNA-seq, single-cell RNA-seq, ATAC-seq, Hi-C, capture Hi-C on dissected tissue of mouse embryos and cultured cells

Standard techniques

- Isolation of DNA/RNA/proteins, cell culture, molecular cloning, Western blot, RT-qPCR, CRISPR/Cas9 genome editing in mouse ESCs, dissection of mouse embryos, whole mount in situ hybridization

Languages

- German: native
- English: Professional Proficiency (C1)

Experience

Mar 2023 – present	**Postdoc**, *Friedrich Miescher Institute for Biomedical Research, Switzerland* Group: Filippo M. Rijli

Sep 2017 – Mar 2023   **PhD Student,** *Friedrich Miescher Institute for Biomedical Research, Switzerland*
Supervisor: Filippo M. Rijli

- Identification of enhancers required for facial morphogenesis using a multi-omics approach resulting in a publication in Nature Communications.
- Extensive hands-on experience in molecular biology techniques (e.g. bulk RNA-seq, single-cell RNA-seq, Hi-C, capture Hi-C, ChIP-seq, ATAC-seq).
- Analysis of RNA-seq, ChIP-seq and ATAC-seq data using R and command line (e.g., read mapping, differential gene expression analysis, peak calling, data visualization, TF motif analysis).

Feb 2017 – Aug 2017   **Intern**, *Oncology, Novartis Institutes for Biomedical Research, Switzerland*
Group: Deborah Castelletti

- Gained valuable experience in pharmaceutical research and an understanding of drug discovery processes.
- Regularly conducted cancer cell culture, TaqMan Real-Time PCR and Western Blot.

Sep 2014 – Sep 2016   **M.Sc. thesis project**, *Biotech Research & Innovation Centre, Denmark*
Group: Kristian Helin

- Goal was to elucidate the mechanism involved in the regulation of TET2 binding to DNA and how hydroxymethylation and DNA methylation patterns are regulated.
- Used CRISPR/Cas9 to tag endogenous Tet2 with V5 and 2xFLAG in mESC.
- Optimized ChIP-seq conditions to map TET2 chromatin occupancy.

Mar 2014 – Aug 2014   **Intern**, *Institute of Biosciences and Medical Technology, Finland*
Group: Dan Hultmark

- Research on the role of serine proteases in the Drosophila cellular immune reaction.

- Gained experience in working with Drosophila as well as an understanding of fly genetics and the immune system of Drosophila larva.

Feb 2013 – June 2013   **Intern**, *University of Salford, UK*
Group: Rhoderick Elder
Project: Expression of humanNEIL1 and mouseNEIL3 in *Escherichia coli*

Education

Sep 2017 – Mar 2023   **PhD in Neuroscience**
University of Basel (Basel, Switzerland)
Overall grade: 6.0 (summa cum laude)

Sep 2014 – Sep 2016   **M.Sc. Molecular Biomedicine**
University of Copenhagen (Copenhagen, Denmark)

Oct 2010 – Feb 2014   **B.Sc. Industrial and Environmental Biology**
University of Applied Sciences Bremen (Bremen, Germany)

Conferences & Workshops

Swiss Institute of Bioinformatics course: Single Cell Transcriptomics, Bern, Switzerland (July 2023)

Gulbenkian Training Program in Bioinformatics Course: 3C-based data analysis and 3D reconstruction of chromatin folding, Oeiras, Portugal (October 2021)

Presented at "7th annual Tri-Regional Stem Cell and Developmental Biology Meeting", Tri-National DevStemCell Network, Virtual (October 2020)

Wellcome Genome Campus Advanced Course: Single Cell Technologies and Analysis, Wellcome Genome Campus, Hinxton, Cambridge, UK (July 2019)

Publications

Kessler S., Minoux M., Joshi O., Ben Zouari Y., Ducret S., Ross F., Vilain N., Salvi A., Wolff J., Kohler H., Stadler M.B., Rijli F.M. A multiple super-enhancer region establishes inter-TAD interaction and controls *Hoxa* gene function in cranial neural crest cells. 2023. ***Nat Commun*** 14: 3242 (2023).

Kitazawa T., Machlab D. Joshi O., Maiorano N., Kohler H., Ducret S., **Kessler S**., Gezelius H., Soneson C., Papasaikas P., López-Bendito G., Stadler M.B., Rijli F.M. A unique bipartite Polycomb signature regulates stimulus-response transcription during development. 2021. ***Nat Genet*** 53(3): 379-391.

Rasmussen K.D., Berest I., **Keβler S.**, Nishimura K., Simón-Carrasco L., Vassiliou G.S., Pedersen M.T., Christensen J., Zaugg J.B., Helin K. TET2 binding to enhancers facilitates transcription factor recruitment in hematopoietic cells. 2019. ***Genome Res*** 29(4): 564-575.

Other activities

- Zoom@Novartis Mentorship program participating as mentee with a Novartis mentor. I led the organization of a 3-hour workshop by the mentees for the mentees.
- Mentoring first-year PhD students to help them connect to the FMI and their peers.
- Co-organizer of the Friedrich Miescher Institute's running group, setting up regular training sessions and planning and organizing participations at competitive races.
- Organized a data science career talk and scientific seminar.