

**UNIVERSIDAD DE CANTABRIA/
EUSKAL HERRIKO UNIBERTSITATEA**

**MÁSTER EN
BIOLOGÍA MOLECULAR Y BIOMEDICINA
TRABAJO FIN DE MÁSTER**

**“CONTRIBUCIÓN COOPERATIVA DE LOS
PROMOTORES DE GENES DEL DESARROLLO Y
DE LOS CLÚSTERES DE CTCF AL AISLAMIENTO
DE LOS TAD EN HUMANOS”**

MASTER’S THESIS

**“COOPERATIVE CONTRIBUTION OF
DEVELOPMENTAL GENE PROMOTERS AND
CTCF CLUSTERS TO TAD INSULATION IN
HUMANS”**

Realizada por: ANDREA DINI

Dirigida por: ALVARO RADA IGLESIAS

Laboratorio: regulación transcripcional en enfermedades
congénitas y de desarrollo

SANTANDER 2025

INDEX

ABSTRACT	5
BACKGROUND.....	6
1. Gene regulation in vertebrates.....	6
2. 3D genome organization and Topologically Associating Domains (TADs)	6
3. Cooperative insulation by <i>CTCF</i> clusters and developmental genes promoters.....	8
4. <i>SIX3/SIX2</i> locus.....	9
5. Promoters contribute to insulation through promoter competition	10
6. Frontonasal dysplasia (FND) and recent FND-like pathological mechanisms	11
OBJECTIVES	14
METHODOLOGY	15
1. hiPSC maintenance, splitting and freezing	15
2. DNA extraction and PCR	15
3. CRISPR-Cas9 genome editing.....	17
4. Selection of the clones containing the deletion/insertion after CRISPR editing.....	18
5. NCC differentiation.....	19
6. NPC differentiation	19
7. RNA isolation, cDNA synthesis and RT-qPCR.....	20
8. Western blot.....	20
RESULTS	21
1. <i>SIX3</i> locus re-arrangements using CRISPR technology	21
2. Selection of the clones carrying the deletions of interest through PCR.....	21
3. Evaluation of <i>SIX2</i> expression levels after Fb-NPC differentiation of the established hiPSC lines..	24
4. Engineering of hiPSC lines with deletions identified in FND patients to assess if disrupting the cooperative insulation provided by promoter competition and CTCF cluster can lead to human disease .	26

5. Tagging of <i>SIX3</i> gene with a 3XFLAG	26
6. Testing of the capability of the 3XFLAG to detect the <i>SIX3</i> protein through WB	28
7. Replication of the FND patient deletion by engineering the <i>SIX3</i> ^{3xFLAG/3xFLAG} hiPSC line	28
DISCUSSION	30
CONCLUSIONS	32
BIBLIOGRAPHY	33

ABSTRACT

Chromatin is organized in self-interacting regions called topologically associating domains (TADs), which ensure proper interactions between enhancers and target promoters found within the same domain, thanks to the presence of *CTCF* binding sites at their boundaries that act as insulators.

In a recent study, the Rada-Iglesias lab showed that in mice the promoters of developmental genes can cooperate with *CTCF* clusters to robustly insulate their own regulatory domains. In the current project, using *SIX3/SIX2* as representative developmental locus, we generated human induced pluripotent stem cells (hiPSC) with various genomic re-arrangements and, upon differentiation of these hiPSC into forebrain-like neural progenitors (Fb-NPC), we demonstrate that the recently discovered mechanism is conserved in humans.

To address the medical relevance of this novel regulatory mechanism, we generated hiPSC with deletions identified in frontonasal dysplasia-like phenotypes (FND) patients spanning *SIX2* gene and the *CTCF* cluster separating *SIX2* and *SIX3* TADs. In the future, these hiPSC will be differentiated into neural crest cells (NCC), an embryonic cell type responsible of craniofacial development, to assess whether these deletions lead to ectopic expression of *SIX3* in NCC. If this is confirmed, it would reveal that by disrupting the cooperative insulation provided by *SIX2* promoter competition and *CTCF*-dependent physical insulation, the investigated deletions could cause FND-like phenotypes through ectopic activation of *SIX3* in NCC.

Overall, our results suggest that *CTCF* binding sites and developmental gene promoters cooperate in the insulation of their domains in humans, and this regulatory mechanism may explain the pathological effects of some structural variants identified in patients.

BACKGROUND

1. Gene regulation in vertebrates

Gene regulation is a fundamental process that occurs to activate some lineage-specific genes while repressing others, promoting cell differentiation and cell plasticity. Gene regulation is particularly important during development since it is when the differentiation process ensures that cells acquire specialized functions, starting from common progenitors.

During embryogenesis, the specific and precise expression of developmental genes is largely achieved through the coordinated activity of enhancers and insulators (Pachano *et al*, 2022).

Enhancers are non-coding regions of the DNA that are able to promote gene transcription by interacting with the promoter of their target gene/s (FIG 1). They can exert their function through large linear distance, thanks to the tridimensional organization of the chromatin (Long *et al*, 2016).

Insulators are regions of the DNA that regulate gene expression by acting as barriers or boundaries within the genome (FIG 1). Their focus is to ensure that enhancers specifically induce the expression of their target genes, preventing them from communicating with non-target genes (West *et al*, 2002). In vertebrates, insulators are preferentially bound by the architectural factor *CTCF*, which is the main driver of domain insulation through a cohesin-mediated loop extrusion model (Bell *et al*, 1999).

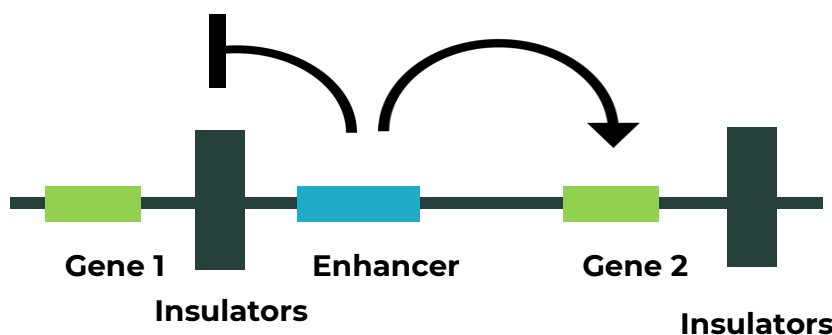


Figure 1. Graphical representation of the activity of enhancers and insulators within the genome.

2. 3D genome organization and Topologically Associating Domains (TADs)

Recent studies revealed that vertebrate genomes are organized in self-interacting regions called topologically associating domains (TADs), where interactions between regulatory elements inside these regions are higher in respect to other regions of the genome (Dixon *et al*, 2012). These results were obtained using novel methodologies, such as Hi-C, which can detect all the DNA-DNA interactions that occur within the genome, therefore resolving the 3D organization of genomes at high resolution (FIG.2A). TADs are fundamental to

ensure proper enhancer-promoter interaction within the same domain, while preventing spurious interactions between enhancers and non-target gene promoters located in different domains.

It has been demonstrated that TADs boundaries often coincide with the presence of *CTCF* clusters (Dixon *et al*, 2012), which can bind *CTCF* proteins, stalling the loop-extruding cohesin complexes and leading to the TAD's insulation (Nora *et al*, 2017; Rao *et al*, 2017).

In the context of developmental loci, enhancers and their target genes tend to co-localize within TADs whose boundaries are evolutionary conserved, suggesting that TADs represent important regulatory domains (Harmston *et al*, 2017).

Recent studies indicate that TAD boundaries provide relatively weak physical insulation, allowing certain enhancers to bypass the boundary and control the expression of genes located in neighboring domains (Chakraborty *et al*, 2023). This suggests that, rather than impenetrable walls, TAD boundaries might act as dynamic and partially permeable barriers, allowing certain level of physical crosstalk across regulatory domains (FIG.2B). This partial permeability of TAD boundaries seems inconsistent with the highly specific expression of developmental genes during embryogenesis, suggesting that additional and unknown mechanisms may have a role in maintaining regulatory domain insulation.

Overall, TAD boundaries seem to play a dual regulatory role, as they can facilitate enhancer-gene communication within TADs while preventing, albeit partly, undesired enhancer-gene contacts across TADs.

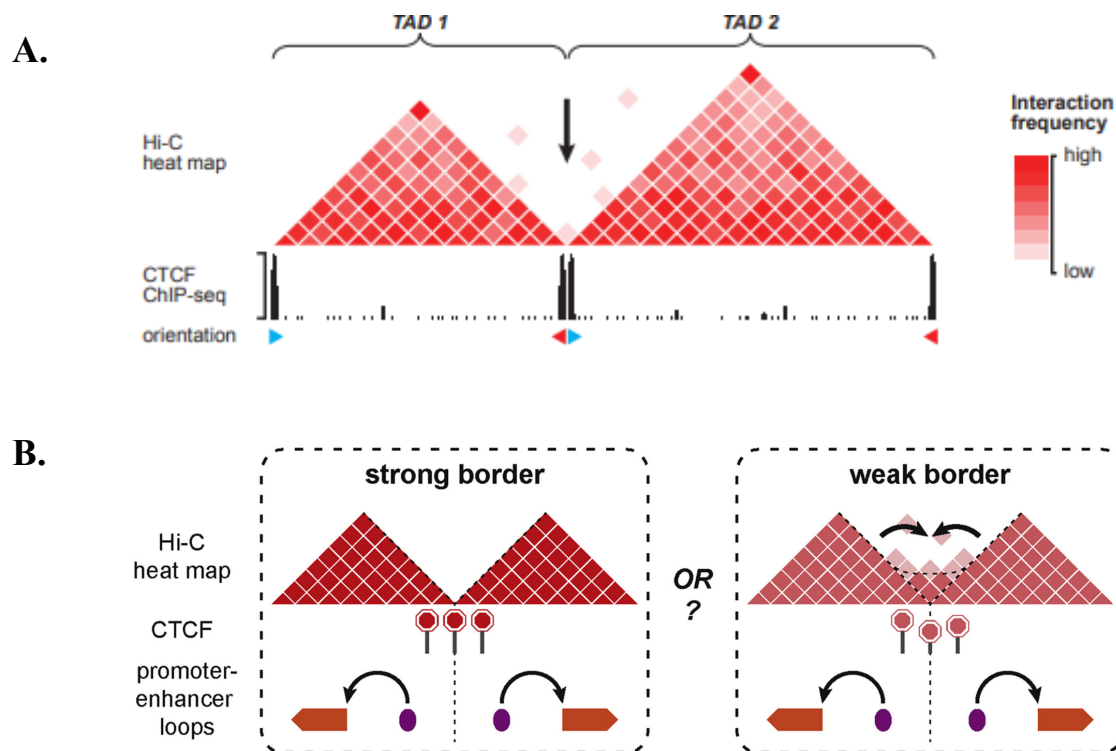


Figure 2.

A. Heat map of two neighbouring TADs obtained through a Hi-C technologies. As well as ChIP-seq representation to show the *CTCF* binding sites at the boundaries of each TAD (Houda Belaghzal *et al*, 2017).

B. Hi-C heat map representation of the interactions between genomic regions belonging to: two neighbouring TADs with a strong boundary vs two neighbouring TADs with a weak boundary (Chang *et al*, 2020).

3. Cooperative insulation by *CTCF* clusters and developmental genes promoters

The partial permeability of TAD boundaries suggests that additional mechanisms might ensure the robust insulation of developmental regulatory domains. In this regard, reporter assays in *Drosophila* and the genetic dissection of the mammalian alpha and beta globin loci suggest that gene promoters can contribute to the insulation of regulatory domains, and they can do it through either enhancer blocking or promoter competition (Ohtsuki *et al*, 1998; Ohtsuki & Levine, 1998; Bozhilov *et al*, 2021). In both cases, the activation of a preferred gene can prevent an enhancer from activating nearby gene/s.

Enhancer blocking usually occurs when the preferred gene is placed in between the enhancer and the other gene/s (Ohtsuki & Levine, 1998), in a similar way to how classical insulators work (FIG.3A). Enhancer blocking might occur through structural mechanisms that involve RNA Pol2 complexes, which, at promoters, can act as weak physical barriers against cohesin-mediated loop extrusion (Banigan *et al*, 2023).

Promoter competition instead can take place regardless of the relative position of the preferred gene with respect to the enhancer and other neighboring gene/s (Ohtsuki *et al*, 1998) (FIG.3B). So, instead of involving physical barriers against cohesin complexes, this mechanism would involve promoters competing for a limited amount of transcription factors or co-activators within shared transcriptional hubs (Oudelaar *et al*, 2019; Sabi & Tuller, 2019).

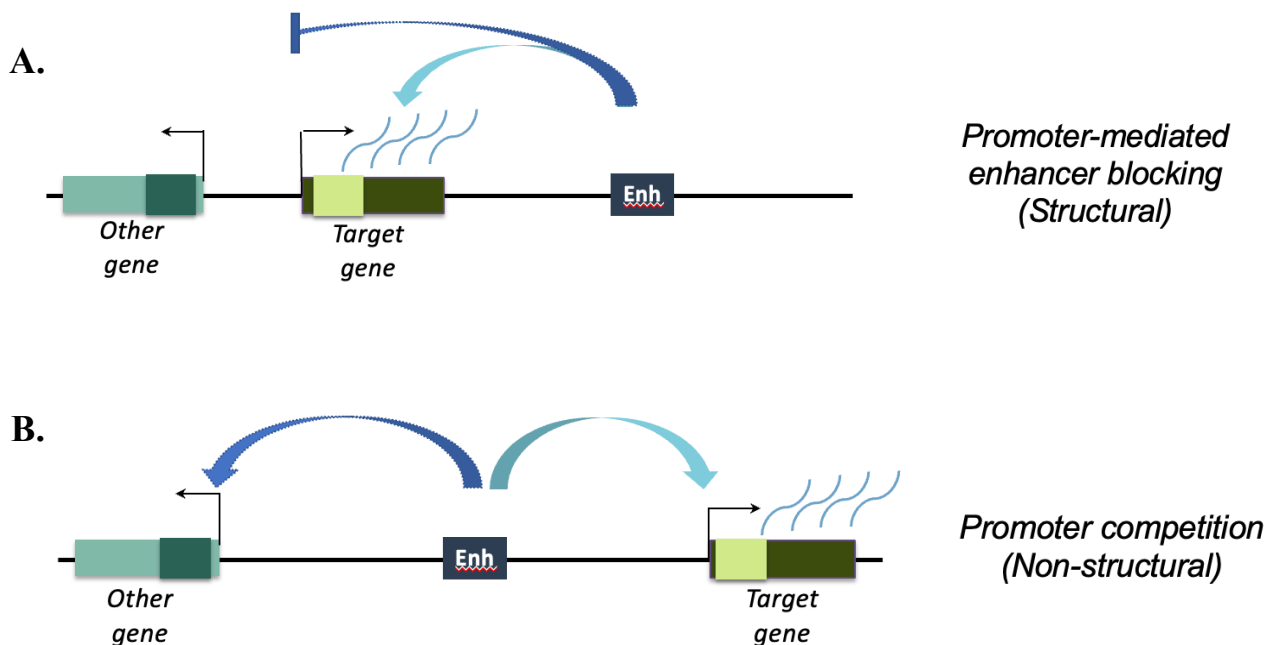


Figure 3.

- A. Schematic representation of the position-dependent promoter-mediated enhancer blocking mechanism.
- B. Schematic representation of the non-position-dependent promoter competition mechanism.

Nevertheless, it is currently unclear whether these promoter-dependent mechanisms significantly contribute, either alone or together with classical insulators, to the insulation of regulatory domains. A recent work from Rada-Iglesias lab has provided novel insights into these open questions (Ealo *et al*, 2024). Briefly, by genetically dissecting a couple of representative developmental loci (*Gbx2/Asb18* and *SIX3/SIX2*) in mouse cells, they showed that developmental gene promoters and nearby *CTCF* clusters cooperatively contribute to the strong insulation of nearby TAD boundaries. Moreover, those data suggests that while the *CTCF* clusters confer physical insulation, the contribution of gene promoters to regulatory insulation preferentially entails promoter competition rather than enhancer blocking.

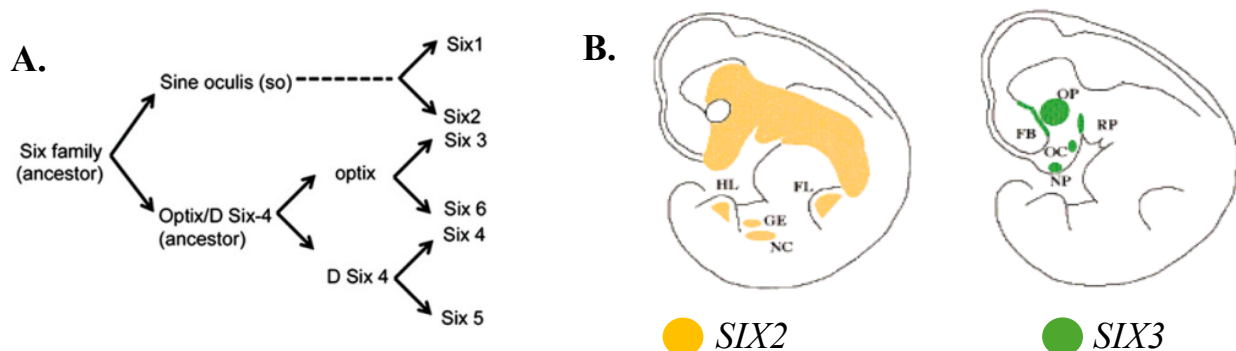
4. *SIX3/SIX2* locus

SIX3 and *SIX2* are developmental genes, situated on chromosome 2, encoding for transcription factors of the Sine Oculis family (Meurer *et al*, 2021; FIG.4A). They have high relevance during development, but they display largely non-overlapping expression patterns, as assessed during mouse embryogenesis (O'Brien *et al*, 2018). In particular *SIX3* is expressed in Neural Progenitor Cells (NPC) and is involved in the formation of the prosencephalon and eyes, while *SIX2* is expressed in kidneys and the craniofacial ectomesenchyme derived from the cranial Neural Crest Cells (NCC) (FIG.4B).

SIX3 and *SIX2* genes are close to each other (~50 Kb) in both mice and humans, but they are found in distinct TADs, separated by a strong and evolutionary conserved boundary containing up to seven *CTCF* sites (FIG.4C).

The non-overlapping expression patterns are also observed *in vitro* and, for example, upon differentiation of mouse ESC (mESC) into neural progenitors (NPC), only *SIX3* gets induced due to the activation of a strong super-enhancer (SE) located within its regulatory domain (Cruz-Molina *et al*, 2017).

Several structural variants associated to defects in this locus are reported in literature. *SIX3* deletions or loss-of-function mutations are reported as one of the main causes of a brain defect known as holoprosencephaly. While, *SIX2* deletions have been reported to be implicated in Frontonasal Dysplasia (FND) (Hufnagel *et al*, 2016; Guan *et al*, 2016; Henn *et al*, 2018).



C.

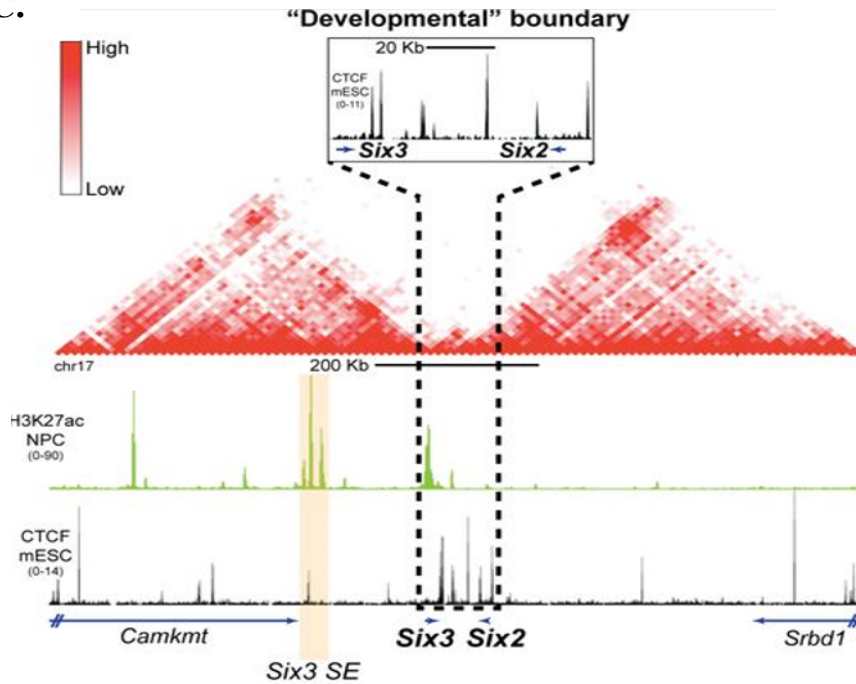


Figure 4.

A. Scheme representing all the genes belonging to the Sine Oculus (SO) family, from which also *SIX3* and *SIX2* derives (Kawakami et al, 2020).

B. Graphical representation of where the *SIX2* and *SIX3* genes are expressed during embryonic development (Kawakami et al, 2020).

C. Hi-C heat map of the *Six3/Six2* locus in mouse cells, showing the relative positions of *Six3* and *Six2* genes in neighbouring TADs, as well as the *CTCF* cluster separating their TADs (Ealo et al, 2024).

5. Promoters contribute to insulation through promoter competition

Recent studies in Rada-Iglesias lab demonstrated that promoters of developmental genes contribute to insulation of their regulatory domain in cooperation with the canonical *CTCF*-mediated physical insulation, and they do it through a promoter competition mechanism.

In particular, they used CRISPR editing to generate mouse Embryonic Stem Cells (mESC) transgenic lines with different genomic re-arrangements to interrogate the relative contribution of *SIX3* and the *CTCF* cluster to the insulation of the *SIX3* regulatory domain (Ealo et al, 2024). First, to assess whether the developmental gene promoter would contribute to the insulation by enhancer blocking, they performed genomic inversions of the region between *SIX3* and its super enhancer (SE), which placed the enhancer close to the TAD boundary and in between *SIX3* and *SIX2*. If the mechanism adopted by developmental gene promoters would have been enhancer blocking, upon deletion of the *CTCF* cluster combined with the inversion of *SIX3*/SE, a strong increase in the expression of *SIX2* would have been observed. But this inversion did not significantly affect *SIX2* expression, especially if compared to the deletion of the *CTCF* cluster without the inversion (FIG.5B,C). Then, to finally demonstrate the cooperative activity of the *CTCF* cluster and the promoter of developmental genes to the insulation of the domain, they performed a deletion spanning both the *CTCF* cluster and the *SIX3* gene and they showed that, upon differentiation into NPC, the expression of *SIX2* was increasing more than observed for the single deletion of the *CTCF* cluster or the single deletion of *SIX3* gene (FIG.5A,C).

Altogether, this work shows that the robust insulation of the *SIX3* regulatory domain depends on the cooperativity between the *CTCF* cluster, which confers physical insulation, and *SIX3*-dependent promoter competition (Ealo *et al*, 2024).

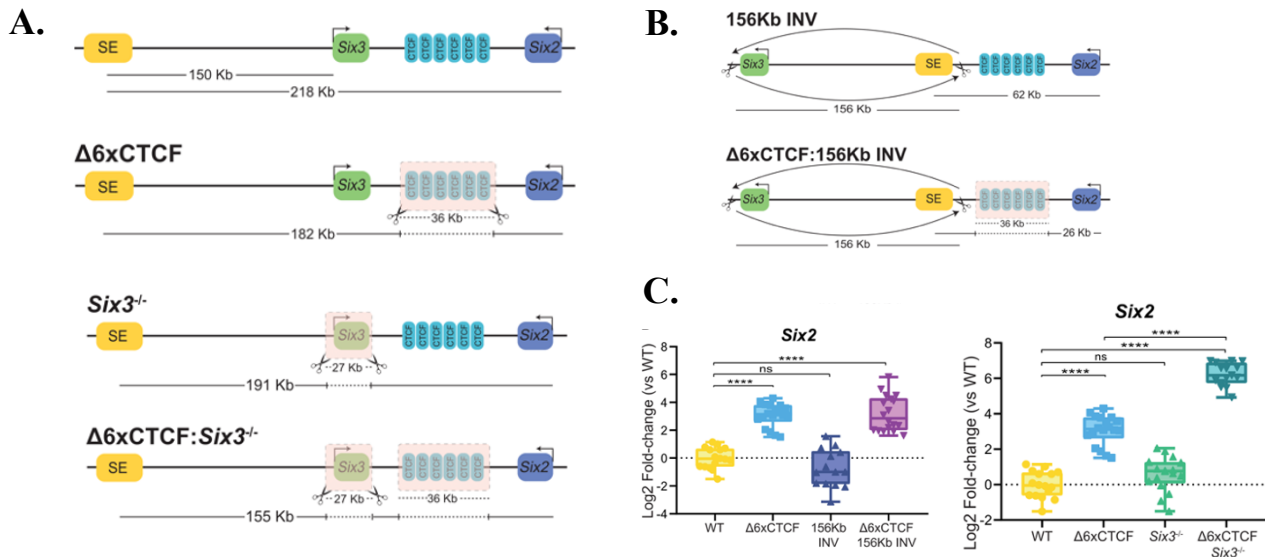


Figure 5. Schematic representation of some of the genomic re-arrangements performed by Ealo *et al* (deletions and inversions) in mESC, together with the relative expression analysis of *SIX2* for each experimental condition.

A. deletions of the *CTCF* cluster, deletions of *SIX3* gene and combined deletion of both *CTCF* cluster and *SIX3* gene.

B. Inversion between *SIX3* gene and the *SIX3* super enhancer (SE)

C. Box plot showing the expression levels of *SIX2* in mNPC derived from the mESC lines presenting the re-arrangements described above

6. Frontonasal dysplasia (FND) and recent FND-like pathological mechanisms

Frontonasal dysplasia (FND) is a rare craniofacial malformation syndrome characterized by facial hypertelorism, broad nasal root, and frontal bossing. It originates from a defect in development during early embryogenesis (Kjaer, 1995). FND has been usually considered a genetically heterogeneous condition, with three major autosomal forms linked to mutations in the ALX family genes: *ALX3* (FND type 1), *ALX4* (FND type 2) and *ALX1* (FND type 3) (Uz *et al.*, 2010; Kayserili *et al.*, 2009; Twigg *et al.*, 2009).

More recently, non-coding and structural genomic alterations have emerged as additional mechanisms contributing to FND-like phenotypes. By reviewing the existing literature, we identified two families in which FND was reported to be caused by deletions spanning *SIX2* and part of the nearby *CTCF* cluster separating the *SIX2* and *SIX3* TADs (Hufnagel *et al*, 2016; Guan *et al*, 2016) (FIG.6). The related pathological mechanism has been associated to *SIX2* haploinsufficiency, since *SIX2* is normally highly expressed in the neural crest-derived ectomesenchyme, an embryonic tissue giving rise to most of the facial bones and cartilage (Fabian & Crump, 2023; Liu *et al*, 2019).

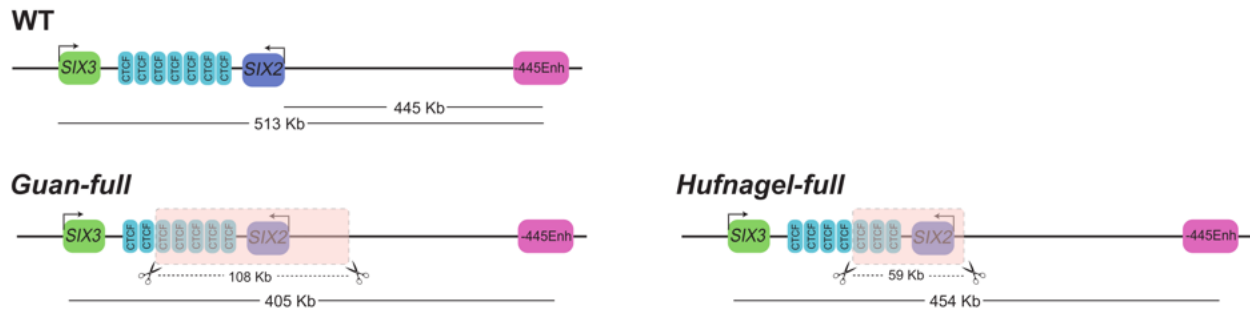
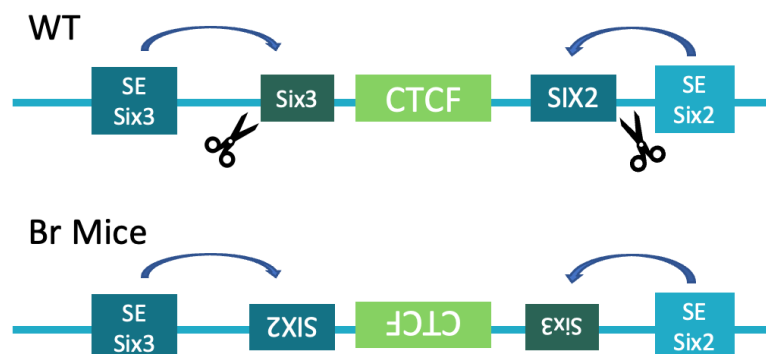


Figure 6. Schematic representation of the two FND patient's deletion reported by the Guan et al (left) and Hufnagel et al (right), within the *SIX3/SIX2* locus.

However, human patients with mutations causing loss of *SIX2* function, as well as *SIX2*^{-/-} mouse models, do not show craniofacial defects compatible with FND but rather kidney abnormalities (Weber *et al*, 2008; Self *et al*, 2006). This could be explained by the redundant activity between *SIX1* and *SIX2* during craniofacial development, the absence of *SIX2* would be compensated by the presence of *SIX1*, as in fact demonstrated by comparing *Six1*^{-/-} or *SIX2*^{-/-} single-KO mice, which show either mild or absent craniofacial defects, respectively, to the *Six1*^{-/-}*SIX2*^{-/-} double-KO mice, which show a severe craniofacial dysmorphism (Liu *et al*, 2019).

The *SIX2*-related FND cases were attributed to *SIX2* haploinsufficiency partly due the FND and renal hypoplasia phenotypes observed in mice heterozygous for the radiation-induced *Brachyrrhine* (*Br*) mutation, an inversion spanning *SIX3*, *SIX2* and the *CTCF* cluster separating their TADs (Fogelgren *et al*, 2008) (FIG.7). *Br* heterozygous mice do not only show reduced *SIX2* expression within the developing kidney and facial mesenchyme, but also ectopic *SIX2* expression in embryonic tissues where *SIX3* is normally expressed, due to the placement of one *SIX2* allele in proximity of *SIX3* enhancers (enhancer adoption) (O'Brien *et al*, 2018). Analogously, the inverted *SIX3* allele gets ectopically expressed within *SIX2* expressing tissues (e.g. nephron progenitors) due to its proximity to *SIX2* enhancers. Therefore, the *Br* inversion might physically disconnect *SIX2* from its cognate enhancers and, thus, reduce *SIX2* expression in both the facial ectomesenchyme and nephron progenitors. However, *SIX2* haploinsufficiency should get manifested in the kidney but not in craniofacial structures due to the redundant function of *SIX1* in the latter (Liu *et al*, 2019).

Figure 7. Schematic representation of the *Br* inversion within the *Six3/Six2* locus, compared to the WT allele



We speculate that the craniofacial defects observed in *Br* heterozygous mice, as well as the *SIX2*-related FND cases described above, might be caused by an enhancer adoption mechanism that leads to the aberrant expression of *SIX3* in the facial ectomesenchyme.

Furthermore, a novel bioinformatic tool implemented to predict the pathological effects of human structural variants called POSTRE (Sánchez-Gaya *et al*, 2023), predicted that the deletions identified in *SIX2*-related FND cases could be pathogenic due to enhancer adoption and gain of *SIX3* expression in human ectomesenchyme cells (hEMC) rather than due to *SIX2* haploinsufficiency.

OBJECTIVES

The main goals of the study are:

1. to assess whether the cooperative insulation mediated by promoter competition and *CTCF* clusters, observed in previous studies within the *SIX3/SIX2* locus in mice, is conserved in humans.
2. to evaluate whether this mechanism could be used to explain developmental defects associated with certain structural variants, using frontonasal dysplasia-like phenotypes (FND) as case examples.

Our hypothesis sustains that, in humans, TADs insulation does not exclusively rely on the activity of *CTCF*-mediated physical insulation. Instead, promoters of developmental genes may give a contribution to the insulation of their own domains, cooperating with the *CTCF* clusters through a promoter competition mechanism. This further layer of regulation could be a step forward in explaining why TAD boundaries sometimes act as partially permeable barriers rather than impenetrable walls.

To address the medical relevance of this novel regulatory mechanism, we hypothesize that it could be used to explain the *SIX2*-related FND cases discussed above, in which the pathological mechanism currently attributed is *SIX2* haploinsufficiency. In particular we propose, instead, that the deletions, spanning *SIX2* along with part of the nearby *CTCF* cluster, might disrupt the cooperative insulation provided by the latter two, which in turn might lead to the ectopic interaction between *SIX3* and enhancers that normally control the expression of *SIX2* in neural crest cells (NCC). Ultimately, this could result in the ectopic and detrimental expression of *SIX3* in NCC and/or the ectomesenchymal mesenchyme, which might alter the transcriptional landscape of these cells during development, causing the typical craniofacial defects associated with FND.

METHODOLOGY

1. hiPSC maintenance, splitting and freezing

UKKi011-A hiPSC were cultured on Geltrex-coated plates (Geltrex™ LDEV-Free, hESC-Qualified, Thermo Fisher Scientific), a laminin-rich extracellular matrix commonly used to maintain pluripotency. Cells were maintained in mTeSR medium (STEMCELL Technologies), a chemically defined, feeder-free culture medium optimized for the long-term maintenance of human pluripotent stem cells. The medium was eventually supplemented with ROCK inhibitor (Y-27632, R&D Systems) when needed, an inhibitor of the Rho-associated protein kinase involved in the cytoskeleton regulation and cell adhesion, in order to prevent apoptosis and to promote cell attachment when cells are dissociated. Cells were stored in the incubator at 37°C with 5% CO₂.

When cells were reaching high levels of confluency they were split into different plate wells, or they were frozen at -80°C. Briefly, PBS was added to wash the cells and Accutase was added to detach them. The cells were centrifuged at 1000 rpm for 3 min to collect the cell pellet, which is then resuspended in mTeSR. Then, the resuspended pellet could be either frozen, adding 500 µl of freezing media (a mix of 40% mTeSR, 40% of KOSR and 20% of DMSO) or split to a different plate, using mTeSR supplemented with 10 µM ROCK inhibitor.

2. DNA extraction and PCR

The DNA was extracted using the NZY tissue DNA isolation kit (NZYTech), following the manufacturer's instructions. The PCR was performed using the NZY Taq II 2x Green Master Mix (NZYTech). Briefly, the PCR mix was prepared with 9,5 µl of water, 0,5 µl of forward primer (1:10), 0,5 µl of reverse primer (1:10), 12,5 µl of NZY Taq II 2x Green Master Mix (NZYTech), and 2 µl of extracted DNA, for each sample (25 µl of total volume). The PCR program was as follows:

- I. 3 min at 95°C
- II. 37xcycles: 30 sec denaturation at 94 °C, 30 sec annealing at 60 °C and 30 sec extension at 72 °C
- III. 5 min at 72 °C for a final extension

All the primers used for the PCRs are reported in TAB.1, TAB.2 and TAB.3.

Primer name	Forward	Reverse
<i>SIX3_KO</i>	GGTCACGAGCTGCTTTCAAA	AACCTGTCAGCTCTACTCGG
<i>CTCF_del</i>	GAGTGCGCCCTACTTAGAA	TCAGGGAAGGAGGGAAATCG
<i>SIX3_CTCF_del</i>	GATGCAGTTTCGGGGTCAC	TCAGGGAAGGAGGGAAATCG
<i>SIX3_inv_L</i>	GGTCACGAGCTGCTTTCAAA	GACCCTGTCGCTAGCAAAAG
<i>SIX3_inv_R</i>	GGGGTAGCAGGTCTTCAACA	AACCTGTCAGCTCTACTCGG
<i>CTCF_inv_L</i>	GTCTCTCGAGCCCTAAGAC	TGACCGGAATTCCTCTTGGG
<i>CTCF_inv_R</i>	CTGCTGTGGAGTCTGATGGA	CCCCAGACTTAAGCTCCAGA
<i>SIX3_CTCF_inv_L</i>	GGTCACGAGCTGCTTTCAAA	GAAACACAGAGCAGTTCCCG
<i>SIX3_CTCF_inv_R</i>	TCACTTAGCCAGAGACAGCC	CCCCAGACTTAAGCTCCAGA
<i>SIX3_wt_L</i>	GGTCACGAGCTGCTTTCAAA	GGGGTAGCAGGTCTTCAACA
<i>SIX3_wt_R</i>	GACCCTGTCGCTAGCAAAAG	ACTCTGAAGAACTGGCGGT
<i>CTCF_wt_L</i>	GTCTCTCGAGCCCTAAGAC	ACCTGCTCCTTGATGTCCTC
<i>CTCF_wt_R</i>	TGACCGGAATTCCTCTTGGG	CCCCAGACTTAAGCTCCAGA
<i>SIX3_CTCF_wt_L</i>	GGTCACGAGCTGCTTTCAAA	GGGGTAGCAGGTCTTCAACA
<i>SIX3_CTCF_wt_R</i>	TGACCGGAATTCCTCTTGGG	CCCCAGACTTAAGCTCCAGA

Table 1. Primers used in the PCRs after *SIX3* locus CRISPR-mediated deletions

Primer name	Forward	Reverse
<i>FLAG_out</i>	TCCGACTTCTGCTTCTCCAG	CTTTCTAGGACAAGCACGGC
<i>FLAG_dup_inv</i>	CGGCGGATCTGACTACAAAG	CGGCGGATCTGACTACAAAG
<i>FLAG_dup</i>	GCGGATCTGACTACAAAGACC	CGGCGGATCTGACTACAAAG
<i>FLAG_in_L</i>	AATGTGATGTAGGTGGCGGTGG	CTTTCTAGGACAAGCACGGC
<i>FLAG_in_R</i>	TCCGACTTCTGCTTCTCCAG	GGGCGGCCTTGCTATCACTGTGCATC

Table 2. Primers used in the PCSs after *FLAG* CRISPR-mediated insertion

Primer name	Forward	Reverse
<i>SIX2_CTCF_full_del</i>	GTGTCAGGGAGAAATGCAGG	CCTCGTAGGTACCTGCATGT
<i>SIX2_KO</i>	GAAACACAGAGCAGTTCCCG	GGGGACTGAGTGTTGAAGGA
<i>CTCF_del</i>	GTGTCAGGGAGAAATGCAGG	CCCCAGACTTAAGCTCCAGA

Table 3. Primers used in the PCRs after *SIX2* locus CRISPR-mediated deletions

3. CRISPR-Cas9 genome editing

The CRISPR-Cas9 gRNAs used were designed using the CRISPR Benchling software tool (<https://www.benchling.com/crispr/>). A Guanine was added to the first position of the sequence, in order to increase the cutting efficiency, and a restriction site for the BbsI restriction enzyme was added at the beginning of the gRNA for cloning purposes into the vector. For each sgRNA, two complementary oligonucleotides were synthesized. The complementary oligos were annealed by incubation at 95°C for 5 min and subsequent cooling at 25°C at a cooling rate of 5°C/min (FIG 8A, step 1).

The annealed oligos were cloned into a CRISPR-Cas9 expression vector (pX330A_hCas9_long_chimeric_gRNA_G2P) (FIG.8B): 1 µl of vector was digested using 1 µl of BbsI restriction enzyme, then 1 µl of previously annealed oligos (diluted 1:200) were ligated with 50 ng of digested vector, using 1 µl of T4 ligase (New England Biolabs), with 2 µl of relative T4 ligase buffer and 15,35 µl of water (20 µl of total volume). The ligation reaction was incubated for 3 h at room temperature and then overnight at 16°C (FIG 8A, step 2).

A.

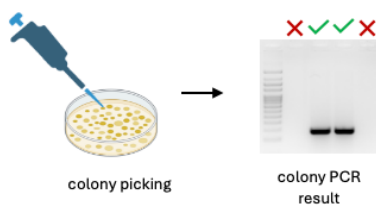
1 Annealing of the gRNA

5' caccGGGAGGGTTGGCTTAAATGT 3'
3' CCCTCCCAACCGAATTACaaaa 5'

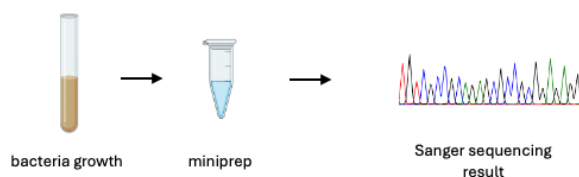
2 Ligation of the annealed oligos with px330A vector



3 Bacteria transformation & positive colony selection



4 Miniprep preparation & Sanger sequencing



B.

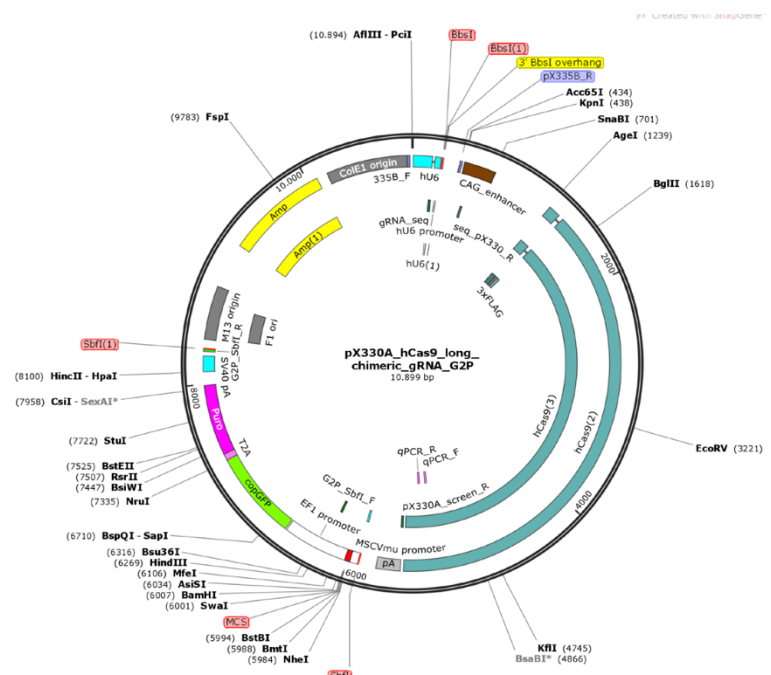


Figure 8.

A. representative scheme of the different steps for the preparation of the sgRNA.

B. Graphical representation of the vector used for gRNAs cloning. The vector includes different BbsI restriction sites, which we used for the gRNA cloning, the CAS9 gene, GFP gene for transfected bacteria detection and Amp resistance gene to select the transfected bacteria.

The ligated plasmids were then used to transform competent cells of *E. Coli*: 30 µl of competent cells were mixed with 2,5 µl of ligated plasmid, and the transfection performed through heat shock (1 min on ice, 1 min at 37°C, 1 min on ice). The transformed bacteria were added to a tube with 1 ml of LB medium and incubated for 1 h at 37°C while shaking. Then, they were plated on LB agar plates with ampicillin and incubated overnight at 37°C (FIG 8A, step 3).

The positive colonies were identified by colony PCR, using the NZY Master Mix protocol described above. The forward primer was the respective gRNA used as oligo at the beginning, the reverse was the pX330_seq_R primer, complementary to the relative sequence in the vector (sequence of pX330_seq_R “GGAAAGTCCCTATTGGCGTT”) (FIG 8A, step 3).

Each positive colony was added to a tube with 4 ml LB medium tube with 8 µl of ampicillin and incubated overnight at 37°C with shaking, let bacteria grow. The plasmid DNA was then extracted from the bacterial culture, using the NZY Miniprep Kit (NZYTech) and a Sanger sequencing was performed to confirm the correct insertion of the gRNA (FIG 8A, step 4).

The transfection was performed when the hIPS cell culture was at around 85% confluency in a six-well plate, we add 1 ml of Accutase and wait 3-5 min at 37°C, in order to have small clusters of 3-5 cells, to increase the transfection efficiency. Then we add 1 ml of mTeSR and we centrifuged the cells 3 min at 1000 rpm to collect the pellet, we resuspend it in mTeSR medium and we count and dilute the cells in order to have a final amount of 75.000 cells per well. We plate the cells in a 24 well plate and incubate them at 37°C with 5% CO₂, overnight.

The transfection was performed using the Lipofectamine reagent, a mix of cationic lipids that bind to DNA, forming liposomal complexes and promoting their endocytosis to the cell. For each transfection we prepared one tube containing both 25 µl of Opti-MEM I Reduced Serum Medium (Thermo Fisher Scientific) and 1 µl of Lipofectamine Stem Transfection Reagent (Thermo Fisher Scientific) and one tube containing both 25 µl of Opti-MEM medium and 500 ng of the plasmid DNA previously prepared. We add the tube with the DNA to the tube with the Lipofectamine, mixed well and waited 10 min. Then we added the solution to the cultures in the 24 well plates and we incubate the cells overnight at 37°C with 5% CO₂.

The following day we check the presence of correctly transfected cells with a fluorescence microscope, since the plasmid contains the GFP gene, and we add puromycin 1000 to select them, since the plasmid also contains the puromycin resistance gene and we incubate overnight. The next day we remove the puromycin and we let the cells recover for some days in mTeSR, at 37°C with 5% CO₂.

4. Selection of the clones containing the deletion/insertion after CRISPR editing

To select the populations containing the correct deletion/insertion, after the gRNA transfection, we performed NZY Master Mix PCRs, using specific primer pairs. For deletions, primers were designed to anneal outside

the targeted boundaries, while for insertions, one primer was located within the inserted sequence and the other in the flanking genomic region.

After identifying a population carrying the deletion, we isolated single clones. Starting from a confluent 24-well plate, cells were detached with Accutase, diluted to 1 cell/100 μ l, and plated into 96-well plates. For the first 4 days, cells were cultured in mTeSR with 10 μ M ROCK inhibitor, switching to unsupplemented mTeSR from day 6, with medium changes every 2 days. Clones were screened via NZY Master Mix PCR using the same primers as previously. Positive clones were further analyzed with additional primer combinations to assess for inversions, duplications, or WT alleles, depending on whether homozygous or heterozygous deletions were desired. The same analysis was done on the initial population as a positive control.

For insertions, due to lower CRISPR knock-in efficiency, we first plated subpopulations of $\sim$ 10 cells per well (10 cells/100 μ l). The rest of the protocol mirrored that used for deletions. Subpopulations were screened via PCR, and once a positive one was found, single clones were plated and analyzed as described above.

All the primers used for each experiment are indicated within the relative experiment description in the result section and are listed in the primer tables (TAB.1, TAB.2, TAB.3).

5. NCC differentiation

For the differentiation of hiPSCs into NCC we used previously reported protocols (Bajpai *et al*, 2010; Prescott *et al*, 2015; Rada-Iglesias *et al*, 2012).

Briefly, confluent hiPSCs colonies were detached by 2mg/mL collagenase treatment, washed with PBS and plated in Petri dishes in a human NCC differentiation medium (Neurobasal and DMEM F12 media in 1:1 ratio, 0.5x B27 with Vitamin A and 0.5x N2 supplements, 20ng/mL epidermal fibroblast growth factor, 20ng/mL basic fibroblast growth factor and 5ug/mL Insulin). Embryoid body (EB) formation was induced already 24h post splitting and medium was changed every 2-3 days. Typically, at Day 7 EBs attached to the Petri dishes and gave rise to NCC outgrowths. On Day 11, the NCC were either harvested for downstream analyses or dissociated by accutase treatment and seeded (50.000 cells per cm^2) on cell culture dishes coated with 5mg/mL fibronectin in human NCC maintenance medium.

6. NPC differentiation

For the NPC differentiation, human iPSCs were plated at 250,000–300,000 cells/ cm^2 on Geltrex-coated plates (ThermoFisher). Cells were seeded in mTeSR medium supplemented with 10 μ M ROCK inhibitor (Y-27632) and cultured overnight to reach a high-density monolayer. Differentiation for the first 3 days was carried out only in KSR-based medium (KnockOut DMEM) supplemented with 15% KnockOut Serum Replacement,

2 mM L-glutamine, 1% MEM Non-Essential Amino Acids (NEAA), 0.1 mM 2-mercaptoethanol and 1% Pen Strep. Then, from day 4, followed a gradual transition to N2 medium (1:1 DMEM/F12 and Neurobasal medium supplemented with 1× N2, 1× B27, 2 mM L-glutamine, and 0.1 mM 2-mercaptoethanol), in particular a media transition from 100% KSR to 75% KSR + 25% N2.

During the differentiation, the medium was supplemented with the following molecules to promote anterior neuroectoderm differentiation: LDN193189 (Sigma-Aldrich) 500 nM and SB431542 (R&D Systems) 10 µM from day 0 to day 6, XAV939 (Sigma-Aldrich) 5 µM from day 0 to day 4. Cells were fed daily, and media composition was adjusted according to differentiation stage as described.

7. RNA isolation, cDNA synthesis and RT-qPCR

Total RNA was isolated using the NZY Total RNA Isolation kit (NZYTech) following the manufacturer's instructions. RNA was reverse transcribed into cDNA using the NZY First-Strand cDNA Synthesis Kit (NZYTech). In particular, 1 µg of RNA was incubated with 10 µl of NZYRT 2x Master Mix, 2 µl of NZYRT Enzyme Mix and nuclease-free water to a total volume of 20 µl, for 10 min at 25°C, followed by 25 min at 50 °C. The enzyme was then heat inactivated at 85°C for 5 min. To digest the remaining RNA, 1 µl of NZY RnaseH was added to the reaction and incubated at 37 °C for 20 min.

RT-qPCRs were performed using the CFX 384 detection system (Bio-Rad) using NZYSpeedy qPCR Green Master Mix 2x (NZYtech) and the respective primers. For each sample, RT-qPCRs were performed as technical triplicates.

8. Western blot

For western blot analysis, cells were lysed on ice using RIPA buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS) supplemented with protease and phosphatase inhibitors (Roche). Lysates were incubated for 30 min with agitation at 4 °C. After centrifugation at 16.000 g for 20 min at 4 °C, supernatants were collected and protein concentration was quantified. Samples of 20 µg were mixed with 2X Laemmli buffer, boiled at 95 °C for 5 min, and separated through SDS-PAGE. Proteins were transferred to a PVDF membrane in a transfer buffer (25 mM Tris, 190 mM glycine, 20% methanol) overnight at 10 mA at 4 °C. Membranes were blocked with 3% BSA in TBST and incubated overnight at 4 °C with primary antibodies diluted in blocking buffer. The membrane was washed and incubated with HRP-conjugated secondary antibodies for 1 h at room temperature. Detection was performed using chemiluminescent substrate and visualized with a CCD-based imaging system.

RESULTS

1. *SIX3* locus re-arrangements using CRISPR technology

To test whether the cooperative insulation of the *SIX3* regulatory domain provided by *SIX3*-dependent promoter competition and the nearby *CTCF* cluster is conserved in humans, we used CRISPR editing to generate hiPSC lines carrying deletions analogous to those previously obtained by Rada-Iglesias lab in mouse Embryonic Stem Cells (mESC) (Ealo *et al*, 2024). Since our aim was to evaluate how these deletions might affect human forebrain development using an *in vitro* differentiation system (Fb-NPC differentiation), the deletions were all generated in heterozygosity in order to resemble the human *SIX3*-related pathologic conditions (i.e. HPE). More specifically, we used CRISPR genome editing to establish hiPSC lines carrying three different deletions, each of them in heterozygosity (FIG.9):

- *SIX3* KO, a 10Kb deletion spanning the *SIX3* gene, to generate *SIX3*^{-/+} hiPSC line
- *CTCF* deletion, a 57Kb deletion spanning the *CTCF* cluster separating *SIX2* and *SIX3* domains, to generate *CTCF*^{Δ/+} hiPSC line
- *SIX3* KO + *CTCF* deletion, a 67Kb deletion spanning both the *SIX3* gene and the *CTCF* cluster separating *SIX2* and *SIX3* domains, to generate *CTCF*^{Δ/+}:*SIX3*^{-/+} hiPSC line

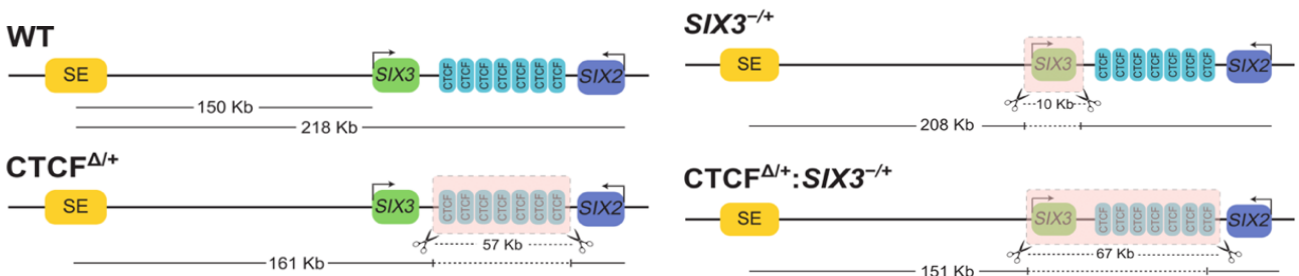


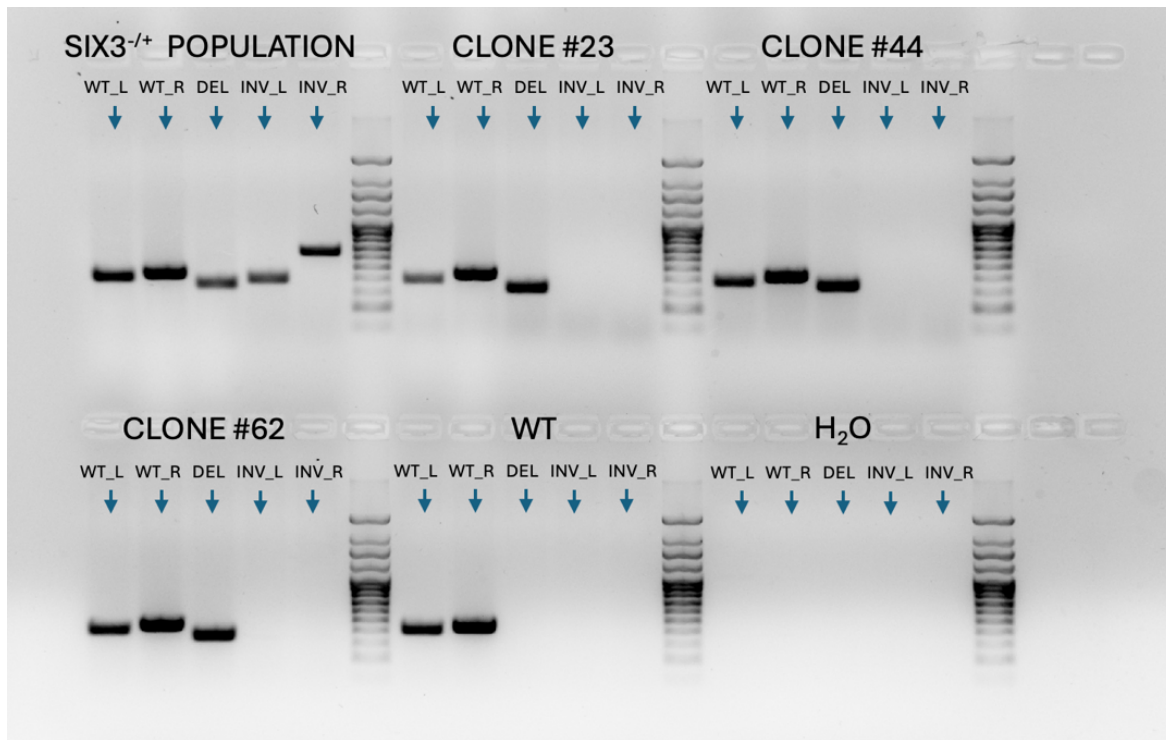
Figure 9. Schematic representation of the different deletions performed within the *SIX3/SIX2* locus on hiPSC, in order to obtain an hiPSC line for each one of them.

2. Selection of the clones carrying the deletions of interest through PCR

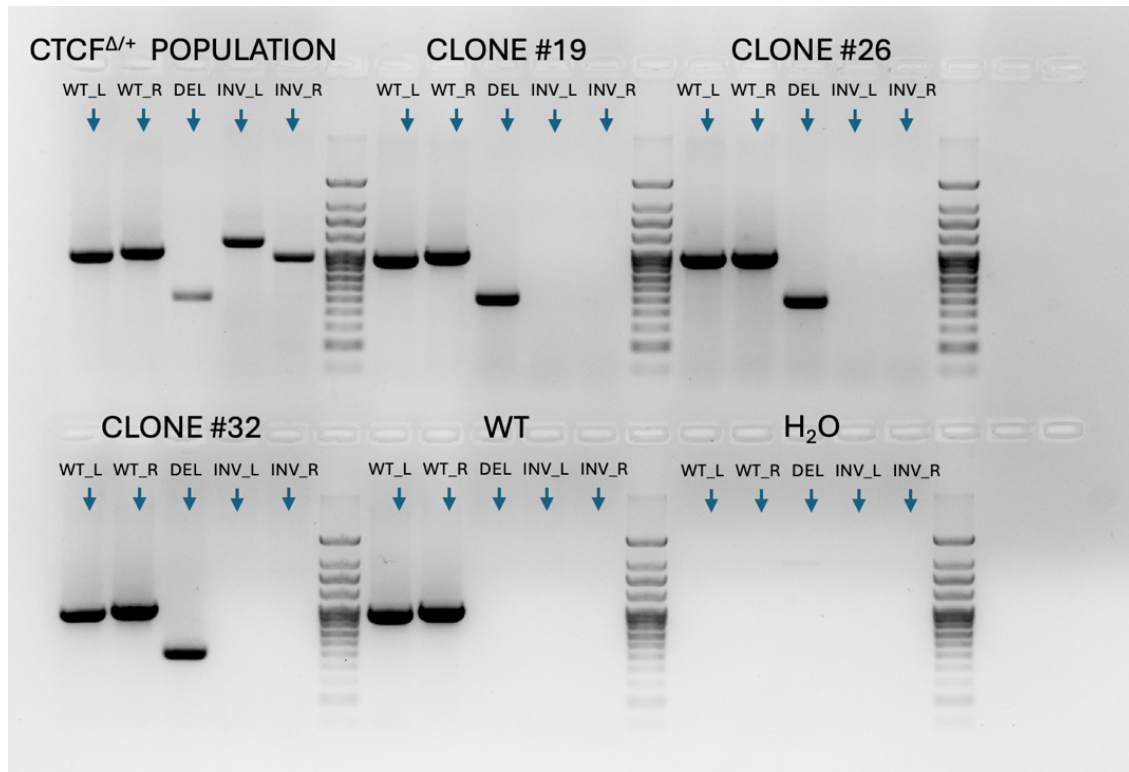
To select the populations and then the clones containing the deletions of interest, we performed several PCRs. First, we performed a PCR on the initial populations, directly deriving from the original gRNA transfection. We used primers *SIX3*_KO, *CTCF*_del and *SIX3*_CTCF_del (primers TAB.1) to check for the presence of the *SIX3* KO, *CTCF* deletion and *SIX3* KO+*CTCF* deletions, respectively. After we found the populations carrying

the deletions of interest, we proceeded on seeding single clones from these populations, and we repeated the PCRs using the same primers. Once some clones presenting the desired deletions were identified, we performed another PCR to select the clones without inversions and having the deletions in heterozygosis. We used primers SIX3_inv_L and SIX3_inv_R (primers TAB.1) to check the inversions in the *SIX3*^{-/+} clones, primers CTCF_inv_L and CTCF_inv_R (primers TAB.1) for *CTCF*^{Δ/+} clones and primers SIX3_CTCF_inv_L and SIX3_CTCF_inv_R (primers TAB.1) for *CTCF*^{Δ/+}:*SIX3*^{-/+} clones. While, to check the WT allele presence we used primers SIX3_wt_L and SIX3_wt_R (primers TAB.1) for the *SIX3*^{-/+} clones, primers CTCF_wt_L and CTCF_wt_R (primers TAB.1) for *CTCF*^{Δ/+} clones and primers SIX3_CTCF_wt_L and SIX3_CTCF_wt_R (primers TAB.1) for *CTCF*^{Δ/+}:*SIX3*^{-/+} clones. We found 3 clones (clones #23, #44, #62) for the *SIX3*^{-/+} population (FIG.10A), 3 clones (clones #19, #26, #32) for the *CTCF*^{Δ/+} population (FIG.10B) and 2 clones (clones #42, #50) for the *CTCF*^{Δ/+}:*SIX3*^{-/+} population (FIG.10C). All these clones presented the targeted deletions respecting the requirements settled: absence of inversions and presence of the WT allele (deletion in heterozygosis). For each one of the initial populations of transfected cells, we detected the presence of the deletion and of inversions as positive controls, while in the WT we confirmed the absence of deletions or inversions as negative controls (FIG.10).

A.



B.



C.

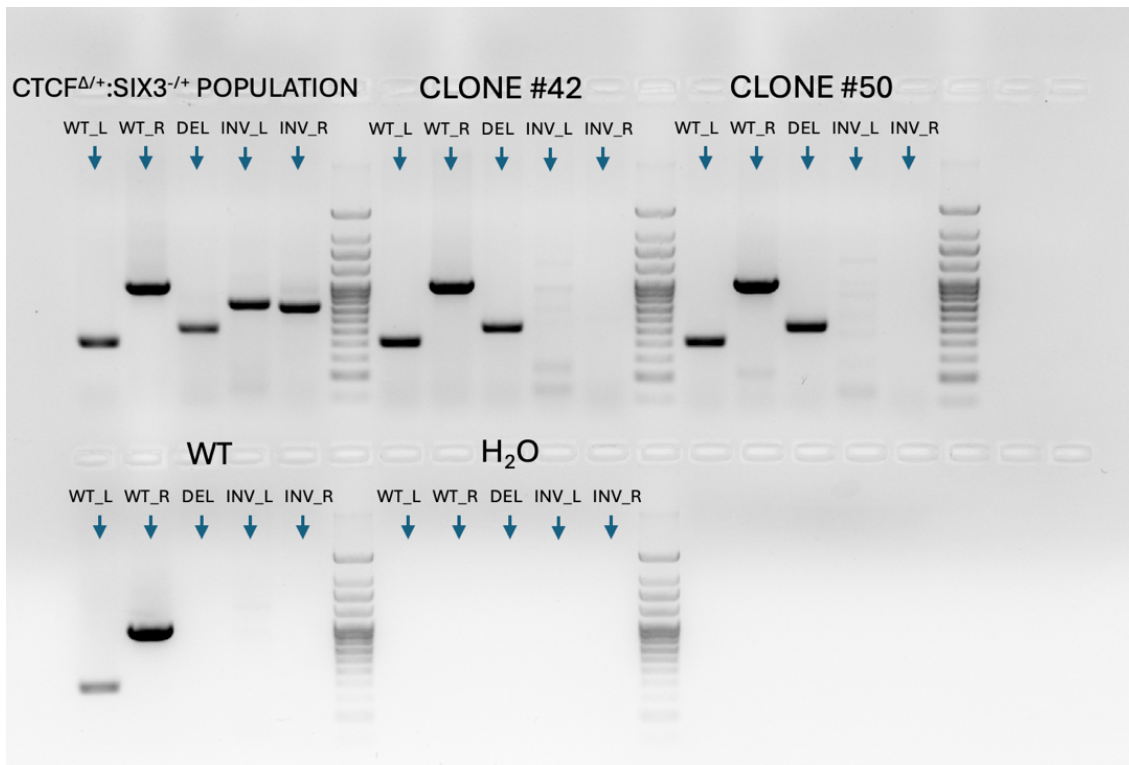


Figure 10. PCRs of the transfected populations and derived clones for the *SIX3* KO deletion (A), the *CTCF* deletion (B) and the *SIX3* KO + *CTCF* deletion (C). The first two bands for each sample correspond to the presence of the WT allele, using primers across the left boundary of the deletion (WT_L) and across the right boundary of the deletion (WT_R). The third band for each sample corresponds to the presence of the deletion relative to that population. The last two bands for each sample correspond to the presence of the inversion, using primers across the left boundary of the supposed inversion (INV_L) and across the right boundary of the supposed inversion (INV_R).

3. Evaluation of *SIX2* expression levels after Fb-NPC differentiation of the established hiPSC lines

Once established, the hiPSC lines containing the three deletions of interest, they were differentiated into Forebrain-like Neural Progenitor Cells (Fb-NPC) with an eight days differentiation protocol. RNA extraction and RT-qPCR were performed in all the differentiated Fb-NPC clones, in order to check the expression levels of both *SIX3* and *SIX2*. If our hypothesis was correct we should observe a higher expression of *SIX2* in the *CTCF*^{Δ/+}:*SIX3*^{-/+} NPC respect to the *CTCF*^{Δ/+} NPC. With the eight days differentiation protocol we did not find any relevant difference in the expression levels of *SIX2* between the *CTCF*^{Δ/+} Fb-NPC (*CTCF*) and the *CTCF*^{Δ/+}:*SIX3*^{-/+} Fb-NPC (KO+*CTCF*) conditions (FIG.11A). Nevertheless, the *SIX3* expression levels were in line with the expected difference between the two conditions (FIG.11B).

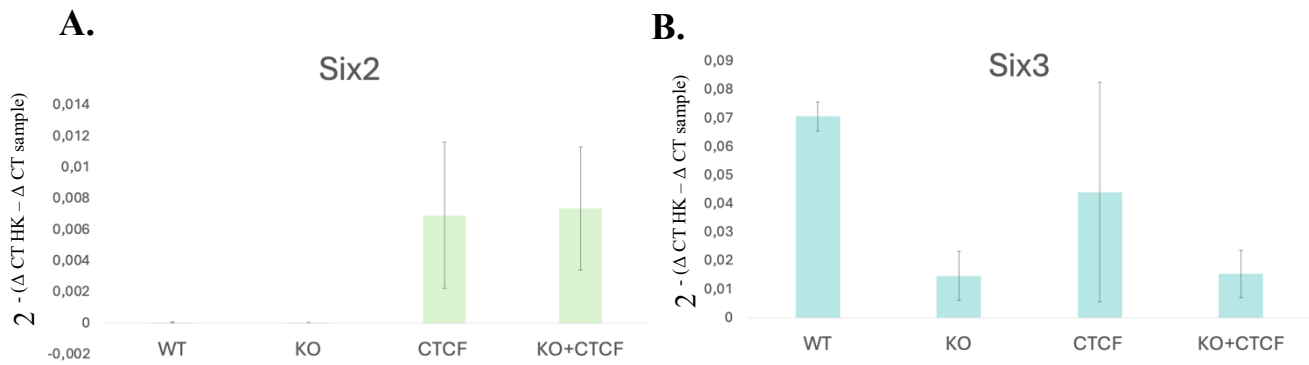


Figure 11. Bar plots of the results obtained with RT-qPCR on the Fb-NPC presenting the previously described genomic rearrangements. **A.** Bar plots of the expression levels of *SIX2* in Fb-NPC across all the conditions indicated. **B.** Bar plots of the expression levels of *SIX3* in Fb-NPC across all the conditions indicated. The error bars correspond to the standard deviation.

Next, we repeated the Fb-NPC differentiation starting from the same hiPSC lines containing the deletions of interest, but this time *SIX3* and *SIX2* expression was measured after four days. The results, represented as an average between the technical triplicates for each Fb-NPC clone for each deletion, showed an increase in *SIX2* expression in the *CTCF*^{Δ/+}:*SIX3*^{-/+} (KO+*CTCF*) cells in comparison to the *CTCF*^{Δ/+} (*CTCF*) cells (FIG.12A,C). Moreover, *SIX3* expression was also lower in the *CTCF*^{Δ/+} (*CTCF*) cells compared to the WT control (FIG.12B,C). Overall, these results confirm our hypothesis that *SIX3* promoter competition and *CTCF* physical insulation cooperate in the insulation of the *SIX3* regulatory domain also in humans. However, the contribution of *SIX3* promoter competition seems to be transient (Day 4) and as differentiation progresses (Day 8), the *CTCF*-dependent insulation becomes dominant.

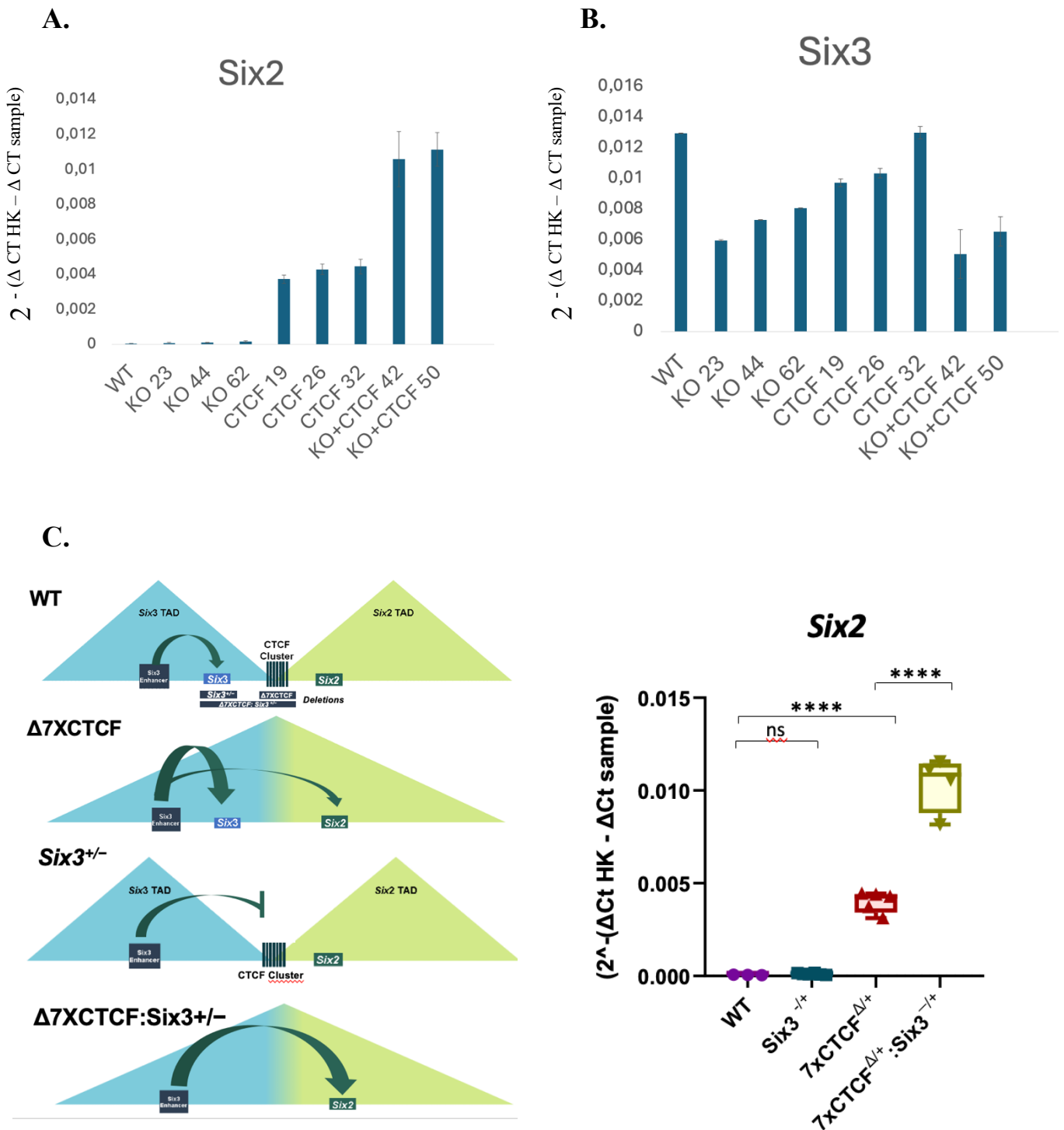


Figure 12. **A.** Box plots of the expression levels of *SIX2* in Fb-NPC across all the previously established cell lines presenting the indicated deletions. **B.** Box plots of the expression levels of *SIX3* in Fb-NPC across all the previously established cell lines presenting the deletions indicated. **C.** Box plots of *SIX3* and *SIX2* expression levels across the indicated conditions in Fb-NPC, obtained by the average between the different Fb-NPC clones within each condition, respectively *SIX3* KO, 7XCTCF del (p-value: 0,000025) and *SIX3* KO + CTCF del (p-value: 0,000087). And schematic representation of the interaction occurring between *SLX3* enhancer and *SLX3/SIX2* within each condition. Statistical analysis carried out using t-test. * represents 0,01 < p-value < 0,05. ** represents 0,001 p-value < 0,01. *** represents 0,0001 < p-value < 0,001. **** represents 0,00001 < 0,0001.

4. Engineering of hiPSC lines with deletions identified in FND patients to assess if disrupting the cooperative insulation provided by promoter competition and CTCF cluster can lead to human disease

To address the medical relevance of this novel regulatory mechanism, we wanted to understand if this novel pathological mechanism could be used to explain the *SIX2*-related FND cases described above. In particular we focused on creating hiPSCs carrying the previously described patient deletion (Hufnagel *et al*, 2016), and differentiating these hiPSCs into NCC to assess the expression levels of *SIX3*, to test whether *SIX3* expression gets induced by the disruption of the TAD boundary combined with the *SIX2* KO, therefore confirming our hypothesis. There are no commercial antibodies against *SIX3* with the required specificity, so before generating the intended deletion we had to tag the *SIX3* protein by inserting a 3XFLAG at its C-terminus (FIG.13). This would enable us to detect the expression of *SIX3* by western blot and immunofluorescence using the anti-3XFLAG antibodies. Once we created a stable cell line expressing the *SIX3*-3XFLAG in homozygosis, we used it for all the further experiments.

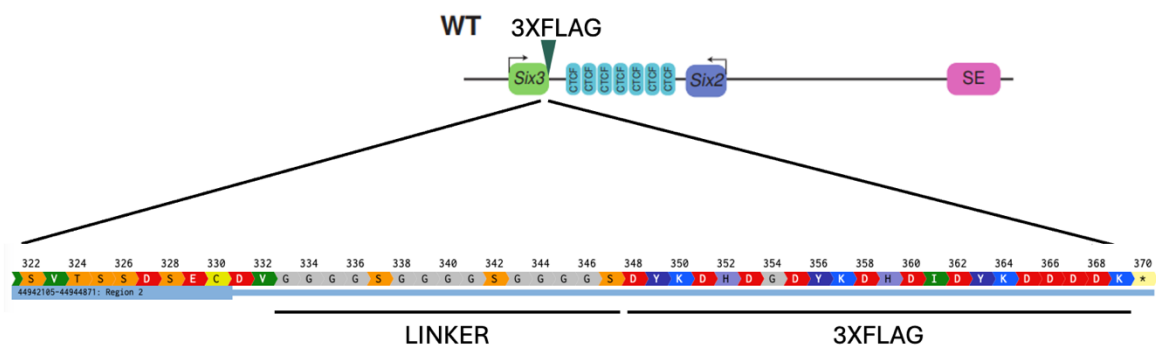
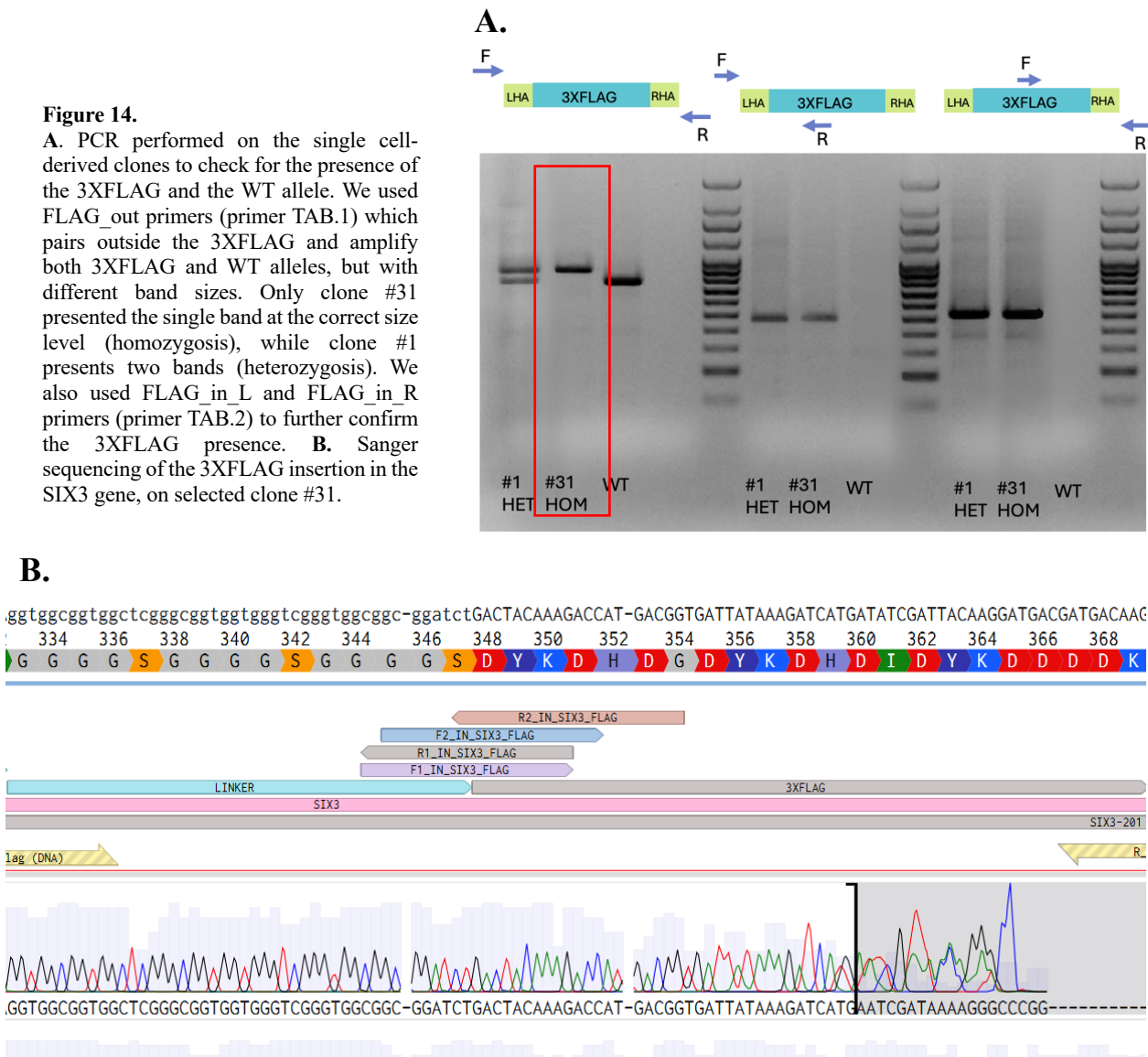


Figure 13. Representation of the 3XFLAG positioning inside the *SIX3* gene and sequence of the 3XFLAG as well as of the LINKER used to prevent the 3XFLAG from disturbing *SIX3* function.

5. Tagging of *SIX3* gene with a 3XFLAG

To knock-in the 3XFLAG we used CRISPR-Cas9 genome editing. In particular, we inserted the 3XFLAG in the region of the *SIX3* gene corresponding to the C-terminus of the *SIX3* protein, into hiPSCs. DNA extraction and PCR were performed to select firstly the transfected population and then the single clones containing the 3XFLAG, using primers FLAG_out (TAB.2 primers). Then, we performed a PCR to select the clones that did not present duplications or duplications with inversions of the inserted DNA fragment, using respectively FLAG_dup and FLAG_dup_inv primers (TAB.2 primers). We found two clones respecting all these parameters: clone #1 and clone #31. Finally, to select the clone/s that were containing the 3XFLAG in

homozygosis, we performed additional PCRs using FLAG_out (primers TAB.2) to check the FLAG allele and/or WT allele, as well as FLAG_in_L and FLAG_in_R (primers TAB.2) to confirm only the FLAG presence respectively from the left and right boundaries of the insertion. We found only clone #31 to respect all the parameters and to express the 3XFLAG in homozygosis (FIG.14A). After we found the candidate clone, we performed Sanger sequencing and alignment of the sequence with the original one to finally check if the 3XFLAG was correctly placed (FIG.14B)



6. Testing of the capability of the 3XFLAG to detect the *SIX3* protein through WB

Finally, we needed to test whether the 3XFLAG tagging of *SIX3* was or not affecting (i) the capacity of the hiPSC to express *SIX3*, (ii) the *SIX3* protein synthesis and (iii) to understand if anti-FLAG antibodies could be used to detect the *SIX3* protein. So, we performed a Fb-like NPC differentiation (Fb-NPC) on the *SIX3*^{3xFLAG/3xFLAG} hiPSC line just created, obtained nuclear extracts and performed a western blot using the anti-FLAG antibodies. Importantly, the WB was able to detect the *SIX3* protein (FIG.15), meaning that (i) the *SIX3* gene was correctly transcribed, (ii) the *SIX3* protein was normally expressed and (iii) the anti-FLAG antibodies were able to detect *SIX3*.

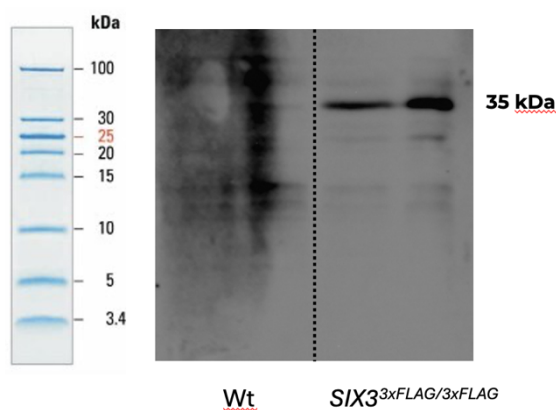


Figure 15.

Western blot gel using the anti-FLAG antibodies on both a WT hiPSC and *SIX3*^{3xFLAG/3xFLAG} hiPSC. The first two columns represent the WT protein extract respectively with a lower amount and a higher amount of proteins. The second two columns represent the *SIX3*^{3xFLAG/3xFLAG} hiPSC protein extract respectively with a lower amount and a higher amount of extract charged on the gel.

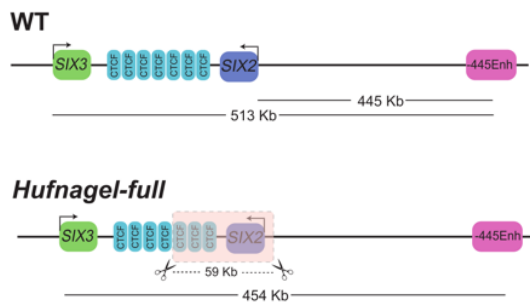
7. Replication of the FND patient deletion by engineering the *SIX3*^{3xFLAG/3xFLAG} hiPSC line

Once obtained and tested, we used the *SIX3*^{3xFLAG/3xFLAG} hiPSC line to engineer a deletion identified in one FND-like family found in literature (Hufnagel *et al*, 2016). These patients were carrying a heterozygous deletion spanning both *SIX2* and part of the nearby *CTCF* cluster separating the *SIX3* and *SIX2* TADs (FIG.16A). To study the effects of this deletion as well as the effects of partial deletions spanning only either the *CTCF* cluster or the *SIX2* gene, we used CRISPR-Cas9 genome editing in *SIX3*^{3xFLAG/3xFLAG} hiPSC to generate hiPSC lines with the whole patient deletion (FIG.16B), as well as hiPSC lines carrying only the left part of the deletion, spanning 3 out of 7 *CTCF* binding sites (FIG.16C), and a cell line carrying only the right part of the deletion, spanning the *SIX2* gene alone (FIG.16C). All the hiPS cell lines established were carrying the deletions in heterozygosis, in order to replicate as faithfully as possible the clinical condition. Our goal at the end was to assess if NCC-derived ectomesenchymal cells carrying these deletions ectopically expressed *SIX3*.

A.



B.



C.

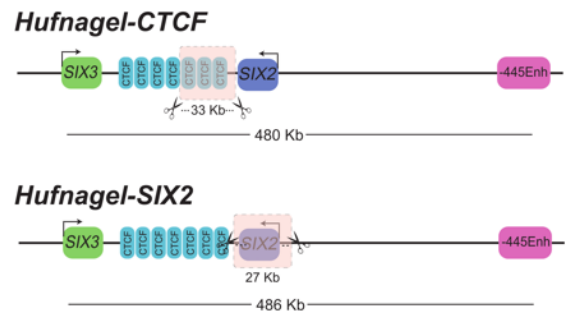


Figure 16. A. representation of the deletion carried by the patient reported by Hufnagel *et al.*, 2016. B. Schematic representation of the three types of deletion we performed using CRISPR genome editing on hiPSC.

In particular, for what concern the patient reported by Hufnagel *et al.* we used CRISPR-Cas9 to generate:

- I. the full deletion, a 59 kb deletion, replicating the patient one, containing both the *SIX2* gene and part of the *CTCF* cluster (3 out of 7 of the *CTCF* sites) that separates *SIX2* and *SIX3* domains (FIG.16B)
- II. the *CTCF* deletion, a 33 kb deletion corresponding to the left part of the patient's deletion, only spanning 3 out of 7 sites of the *CTCF* cluster (FIG.16C)
- III. the *SIX2* KO, a 27 kb deletion corresponding to the right part of the patient's deletion, only spanning the *SIX2* gene (FIG.16C)

To select the hiPSC populations that contained the intended deletions, we performed different PCRs using primers specific for *SIX2*_whole_del, *SIX2*_KO and *SIX2*_CTCF_del (primers TAB.3). After we will find clones containing the deletions of interest, we will perform other NZY Master Mix PCRs to check for the presence of inversions and WT alleles, since we wanted the deletion to be in heterozygosis.

In the future, after we successfully obtain at least one clone for each deletion respecting all the requirements, we will differentiate them into NCC and ectomesenchymal cells. After that, we will perform a total RNA extraction and RT-qPCR to check the expression levels of *SIX2* and *SIX3*, as well as a western blot and an immunofluorescence, using anti-3XFLAG antibodies, to confirm the ectopic expression of the *SIX3* protein.

DISCUSSION

Since Topologically Associating Domains (TADs) were uncovered, they have always been considered insulated regulatory regions of the genome that, through the insulator activity of the *CTCF* clusters located at their boundaries, were able to restrict enhancer-promoter interactions within each domain, inhibiting the interactions of enhancers with promoters found in neighboring TADs. However, recent studies demonstrated that TAD boundaries do not always work as strong physical barriers, and in some cases enhancers can bypass these boundaries and activate genes in nearby domains (Chakraborty *et al.*, 2023). This suggests that TAD boundaries do not work as impenetrable walls, but rather as partially permeable barriers that still allow some leaky interaction between different regulatory regions. However, this mechanism seems inconsistent with the very precise and specific control of developmental gene expression during embryonic development, suggesting that additional and unknown mechanisms might be involved.

Here, we demonstrated that in humans, TAD insulation does not only rely on *CTCF*-dependent physical insulation but also on promoter competition, in agreement with the recent findings of the Rada-Iglesias lab in mice (Ealo *et al.*, 2024). In particular, we engineered hiPSC with deletions spanning the *CTCF* cluster separating the *SIX3/SIX2* domains and the *SIX3* gene (*CTCF*^{Δ/+};*SIX3*^{-/+} hiPSC line), as well as hiPSC with the deletions spanning only the *CTCF* cluster (the *CTCF*^{Δ/+} hiPSC line) or the *SIX3* gene (*SIX3*^{-/+} hiPSC line). We showed that, upon differentiation into Fb-NPC for four days, the cells with the deletion spanning both the *CTCF* cluster and *SIX3* transiently expressed a higher level of *SIX2* compared to cells with only the *CTCF* deletion. These results are in line with our hypothesis that the insulation of the *SIX3* TAD is due to a cooperative activity between the local *CTCF* cluster and *SIX3* promoter competition. By disrupting the TAD boundary through the *CTCF* cluster deletion, the *SIX3* Super Enhancer (SE) is free to interact with *SIX2* in the neighboring TAD, promoting its expression. But when the TAD boundary disruption through *CTCF* deletion is combined with *SIX3* KO, *SIX2* expression increases even further. This is likely due to the absence of the *SIX3* promoter, which would otherwise compete for the SE activity.

Notably, after the eight days NPC differentiation we didn't find any relevant difference in *SIX3* expression between the *CTCF*^{Δ/+} Fb-NPC and the *CTCF*^{Δ/+};*SIX3*^{-/+} Fb-NPC. Therefore, the insulator effect of promoter competition seems to be transient (i.e. observed on Day4 but not on Day8), which was not previously observed in mouse cells (Ealo *et al.*, 2024). One potential explanation for these inconsistent results between mice and humans is that the *CTCF* deletion generated in both species is not exactly the same. Namely, in humans we eliminated all the *CTCF* sites located between the *SIX3* and *SIX2* genes, while in mice one of those sites was left intact (Ealo *et al.*, 2024). Therefore, we are already working to further investigate this topic by generating hiPSC with deletions that better resemble the re-arrangements previously generated in mouse cells.

To assess the medical relevance of this novel regulatory mechanism, we are going to investigate whether it can be used to explain the pathomechanism causing FND in several patient families (). We started by generating hiPSCs carrying the deletion found in one of this FND families and that span *SIX2* along with part of the nearby

CTCF cluster, as well as hiPSCs carrying deletions that span either *SIX2* or the *CTCF* cluster. In the future we will differentiate these hiPSCs lines into NCC and ectomesenchymal cells and we will analyse the expression levels of *SIX2* and *SIX3*. If, as we expect, *SIX3* gets more induced in NCC and/or ectomesenchymal cells with the whole patient deletion compared to WT cells or cells carrying only the *CTCF* cluster deletion, this would support the involvement of the proposed regulatory mechanism in the observed phenotype. The increased expression of *SIX3* would result from the disruption of the TAD boundary between the *SIX3* and *SIX2* domains, in combination with the absence of *SIX2*, allowing enhancers from the neighbouring domain to ectopically activate *SIX3* in NCCs.

Although, even if our data would show that the FND patient-specific deletion results in increased expression of *SIX3*, we cannot fully conclude that this overexpression is the primary cause of the observed FND phenotype. Additional studies are required to determine the pathogenic relevance of *SIX3* upregulation. For instance, ChIP-seq on the *SIX2* and *SIX3* proteins could help to identify the cis-regulatory elements (CREs) and genes directly regulated by *SIX3* when ectopically expressed in patients hEMC. Furthermore, RNA-seq could suggest whether changes in differentially expressed genes in hEMC are likely to result in craniofacial defects.

CONCLUSIONS

In conclusion, we demonstrated that the promoters of developmental genes and *CTCF* clusters cooperate in the robust insulation of TADs in humans.

Furthermore, we are working to demonstrate that, when disrupted, this cooperative insulating activity might be the causative pathomechanism of some FND phenotypes observed in the literature.

This project represents an important step forward in the understanding of TADs insulation, which can be used to explain pathological phenotypes associated to some structural variants.

BIBLIOGRAPHY

1. Banigan, E. J. *et al.* Transcription shapes 3D chromatin organization by interacting with loop extrusion. *Proc. Natl Acad. Sci. USA*, 120, e2210480120. (2023)
2. Bell, A. C., West, A. G. & Felsenfeld, G. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell*, 98, 387–396. (1999)
3. Bozhilov, Y. K. *et al.* A gain-of-function single nucleotide variant creates a new promoter which acts as an orientation-dependent enhancer-blocker. *Nat. Commun.*, 12, 3806. (2021)
4. Chakraborty, S. *et al.* Enhancer–promoter interactions can bypass CTCF-mediated boundaries and contribute to phenotypic robustness. *Nat. Genet.*, 55, 280–290. (2023)
5. Cruz-Molina, S. *et al.* PRC2 facilitates the regulatory topology required for poised enhancer function during pluripotent stem cell differentiation. *Cell Stem Cell*, 20, 689–705. (2017)
6. Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, 485, 376–380. (2012)
7. Ealo, T., Sanchez-Gaya, V., Respuela, P. *et al.* Cooperative insulation of regulatory domains by CTCF-dependent physical insulation and promoter competition. *Nat Commun* **15**, 7258 (2024).
8. Fabian, P., & Crump, J. G. Reassessing the embryonic origin and potential of craniofacial ectomesenchyme. *Seminars in Cell & Developmental Biology*, 138, 45–53. (2023)
9. Fogelgren, B., Kuroyama, M. C., McBratney Owen, B., Spence, A. A., Malahn, L. E., Anawati, M. K., Cabatbat, C., Alarcon, V. B., Marikawa, Y., & Lozanoff, S. Misexpression of *Six2* is associated with heritable frontonasal dysplasia and renal hypoplasia in 3H1 Br mice. *Developmental Dynamics*, 237(7), 1767–1779. (2008)
10. Guan, J., Wang, D., Cao, W. *et al.* *SIX2* haploinsufficiency causes conductive hearing loss with ptosis in humans. *J Hum Genet* **61**, 917–922 (2016).
11. Harmston, N. *et al.* Topologically associating domains are ancient features that coincide with Metazoan clusters of extreme non-coding conservation. *Nat. Commun.*, 8, 441. (2017)
12. Henn, A., Weng, H., Novak, S., Rettenberger, G., Gerhardinger, A., Rossier, E., & Zirn, B. *Six2* gene haploinsufficiency leads to a recognizable phenotype with ptosis, frontonasal dysplasia, and conductive hearing loss. *Clinical Dysmorphology*, 27(2), 27–30. (2018)
13. Hufnagel, R. B. *et al.* A New Frontonasal Dysplasia Syndrome Associated with Deletion of the *SIX2* Gene. *Am. J. Med. Genet. A*, 170, 2760–2765. (2016)

14. Kayserili, H., Uz, E., Niessen, C., Vargel, İ., Alanay, Y., Tuncbilek, G., Yigit, G., Uyguner, O., Candan, S., Okur, H., Kaygın, S., Balci, S., Mavili, E., Alikasifoglu, M., Haase, I., Wollnik, B., & Akarsu, N. A. ALX4 dysfunction disrupts craniofacial and epidermal development. *Human Molecular Genetics*, 18(22), 4357–4366. (2009)
15. Kjaer, I. Morphological characteristics of dentitions developing excessive root resorption during orthodontic treatment. *European Journal of Orthodontics*, 17(1), 25–34. (1995)
16. Liu, Z. *et al.* Six1 and Six2 maintain murine nephron progenitors and synergistically regulate nephrogenesis. *Dev. Biol.*, 445, 151–161. (2019)
17. Long, H. K., Prescott, S. L. & Wysocka, J. Ever-changing landscapes: transcriptional enhancers in development and evolution. *Cell*, 167, 1170–1187. (2016)
18. Meurer, L., Ferdman, L., Belcher, B., & Camarata, T. The SIX Family of Transcription Factors: Common Themes Integrating Developmental and Cancer Biology. *Frontiers in Cell and Developmental Biology*, 9, 707854. (2021)
19. Nora, E. P. *et al.* Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell*, 169, 930–944. (2017)
20. O'Brien, L. L. *et al.* Transcriptional regulatory control of mammalian nephron progenitors revealed by multi-factor cistromic analysis and genetic studies. *PLOS Genet.*, 14, e1007181. (2018)
21. Ohtsuki, S. & Levine, M. GAGA mediates the enhancer blocking activity of the eve promoter in the Drosophila embryo. *Genes Dev.*, 12, 3325–3330. (1998)
22. Oudelaar, A. M. *et al.* A revised model for promoter competition based on multi-way chromatin interactions at the α -globin locus. *Nat. Commun.*, 10, 5412. (2019)
23. Pachano, T., Haro, E. & Rada-Iglesias, Á. Enhancer-gene specificity in development and disease. *Development*, 149(11), 1–12. (2022)
24. Rao, S. S. P. *et al.* Cohesin loss eliminates all loop domains. *Cell*, 171, 305–320. (2017)
25. Sabi, R. & Tuller, T. Modelling and measuring intracellular competition for finite resources during gene expression. *J. R. Soc. Interface*, 16, 20180887. (2019)
26. Sánchez Gaya, V., & Rada Iglesias, Á. POSTRE: a tool to predict the pathological effects of human structural variants. *Nucleic Acids Research*, 51(9), e54. (2023)
27. Self, M., Lagutin, O. V., Bowling, B. A., Hendrix, J., Cai, Y., Dressler, G. R., & Oliver, G. Six2 is required for suppression of nephrogenesis and progenitor renewal in the developing kidney. *The EMBO Journal*, 25(21), 5214–5228. (2006)
28. Twigg, S. R. F., Versnel, S. L., Nürnberg, G., Lees, M. M., Bhat, M., Hammond, P., Hennekam, R. C. M., Hooogbeem, A. J. M., Hurst, J. A., Johnson, D., Robinson, A. A., Scambler, P. J., Gerrelli, D., Nürnberg, P.,

Mathijssen, I. M. J., & Wilkie, A. O. M. Frontorhiny, a Distinctive Presentation of Frontonasal Dysplasia Caused by Recessive Mutations in the ALX3 Homeobox Gene. *American Journal of Human Genetics*, 84(5), 698–705. (2009)

29. Uz, E., Alanay, Y., Aktas, D., Vargel, I., Gucer, S., Tuncbilek, G., von Eggeling, F., Yilmaz, E., Deren, O., Posorski, N., Ozdag, H., Liehr, T., Balci, S., Alikasifoglu, M., Wollnik, B., & Akarsu, N. A. Disruption of ALX1 causes extreme microphthalmia and severe facial clefting: expanding the spectrum of autosomal recessive ALX related frontonasal dysplasia. *American Journal of Human Genetics*, 86(5), 789–796. (2010)
30. Weber, S., Taylor, J. C., Winyard, P., Baker, K. F., Sullivan Brown, J., Schild, R., Knüppel, T., Zurowska, A. M., Caldas Alfonso, A., Litwin, M., Emre, S., Ghiggeri, G. M., Bakkaloglu, A., Mehls, O., Antignac, C., Schaefer, F., & Burdine, R. D. SIX2 and BMP4 mutations associate with anomalous kidney development. *Journal of the American Society of Nephrology*, 19(5), 891–903. (2008)