JCI insight

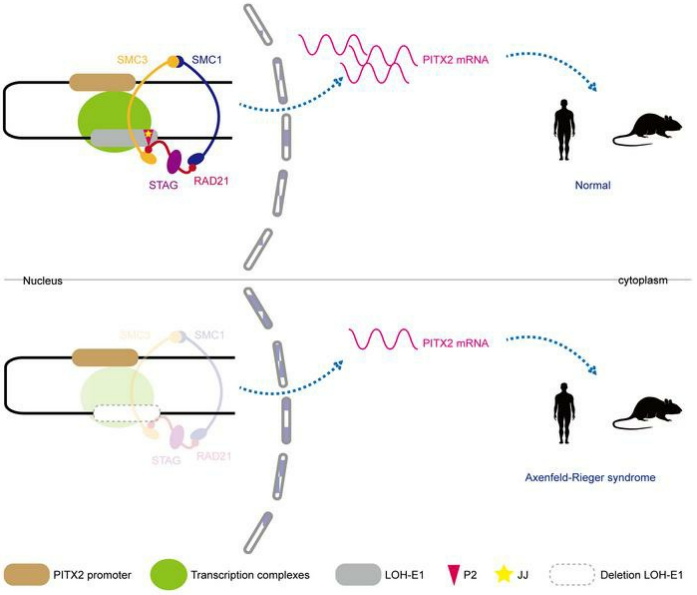
Intergenic sequences harboring potential enhancer elements contribute to Axenfeld-Rieger syndrome by regulating *PITX2*

Yizheng Jiang, … , Liang Zhou, Zhengmao Hu

JCI Insight. 2024. https://doi.org/10.1172/jci.insight.177032.

Research In-Press Preview Genetics Ophthalmology

Graphical abstract



Intergenic sequences harboring potential enhancer elements contribute to Axenfeld-Rieger syndrome by regulating *PITX2*

Yizheng Jiang[1,*], Yu Peng[2,*], Qi Tian[1], Zhe Cheng[1], Bei Feng[1], Junping Hu[1], Lu Xia[1], Hui Guo[1], Kun Xia[1,3], Liang Zhou[4], Zhengmao Hu[1]

[1]MOE Key Lab of Rare Pediatric Diseases & Hunan Key Laboratory of Medical Genetics of the School of Life Sciences, Central South University, Changsha, Hunan, 410078 China

[2]Department of Medical Genetics, The Affiliated children's Hospital Of Xiangya school of Medicine, Central south University, Hunan Children's Hospital, Changsha, Hunan, 410007 China

[3]MOE Key Laboratory of Rare Pediatric Diseases, Hengyang Medical School, University of South China, Hengyang, Hunan, 421001 China

[4]Department of Ophthalmology, The Second Xiangya Hospital, Central South University, Changsha, Hunan, 410011 China

Corresponding authors: Zhengmao Hu, 110 Xiangya Road, Changsha, Hunan, 410078, China, +86 0731-84805357, huzhengmao@sklmg.edu.cn; Liang Zhou, 139 Renmin middle road, Changsha, Hunan, 410011, China, +86 18073137551, zhouliang12@csu.edu.cn; Kun Xia, 110 Xiangya road, Changsha, Hunan, 410078, China, +86 0731-84805357, xiakun@sklmg.edu.cn;

* Yizheng Jiang and Yu Peng are co-first authors.

The authors have declared that no conflict of interest exists.

Abstract

Recent studies have uncovered that non-coding sequence variants may relate to Axenfeld-Rieger syndrome (ARS), a rare developmental anomaly with genetic heterogeneity. However, how these genomic regions are functionally and structurally associated with ARS is still unclear. In this study, we performed genome-wide linkage analysis and whole-genome sequencing in a Chinese ARS family and identified a heterozygous deletion of about 570 kb (termed LOH-1) in the intergenic sequence between *PITX2* and *FAM241A*. Knockout of LOH-1 homologous sequences caused ARS phenotypes in mice. RNA-seq and RT-qPCR revealed a significant reduction in *Pitx2* gene expression in LOH-1$^{-/-}$ mice, while *Foxc1* expression remained unchanged. ChIP-seq and bioinformatics analysis identified a potential enhancer region (LOH-E1) within LOH-1. Deletion of LOH-E1 led to a significant downregulation of the *PITX2* gene. Mechanistically, we found a sequence (hg38 chr4:111,399,594-111,399,691) which is on LOH-E1 could regulate *PITX2* by binding to RAD21, a critical component of the cohesin complex. Knockdown of *RAD21* resulted in reduced *PITX2* expression. Collectively, our findings indicate that a potential enhancer sequence which is within LOH-1 may regulate *PITX2* expression remotely through cohesin-mediated loop domains, leading to ARS when absent.

Introduction

In recent years, with the development of sequencing technology, the rate of genetic diagnosis of Mendelian diseases has been substantially increased. However, the proportion of unsolved exomes was much higher than expected and a significant subset of disorders may relate to non-coding regions of the genome (1). Although the functional relevance of most non-coding variants is not known, a number of diseases have been shown to be associated with non-coding variants (2).

One example of such a Mendelian disorder is Axenfeld-Rieger syndrome (ARS), a rare developmental anomaly with an incidence of approximately 1:50,000-100,000 in newborns (3). The clinical manifestations of ARS include mainly ocular symptoms: the presence of posterior embryotoxon (manifested by a prominent, anteriorly displaced Schwalbe's line near the posterior corneal limbus), iridocorneal adhesions, iris hypoplasia, corectopia and/or polycoria. Typically inherited in an autosomal dominant manner, two causative genes for ARS have been identified: *PITX2* (paired-like homeodomain transcription factor) at 4q25 (4) and *FOXC1* (forkhead box c1) at 6p25 (5). Phillips et al. also linked ARS to 13q14 (RIEG2, MIM: 601499) (6), but so far no pathogenic gene has been clearly identified at this locus. Additionally, mutations in *CYP1B1*, *PRDM5*, *COL4A1* and *PAX6* have also been associated with ARS in a few reports (7-10). Currently, only 40% of ARS cases are linked to mutations in *PITX2* and *FOXC1*, leaving a significant portion of ARS patients with an unknown genetic cause (3).

Recent studies have suggested that non-coding sequence variants upstream of the *PITX2* gene may be associated with disorders. For instance, Volkmann et al. identified a 7600 kb deletion at 106-108 kb upstream of the *PITX2* gene in an ARS patient, which did not affect the integrity of the *PITX2* gene (11). Aguirre et al. discovered an 80 bp segment near the SNP (rs2200733) of the upstream noncoding sequence of the *PITX2* gene with enhancer activity and a topological

linkage between the *PITX2c* and *ENPEP* promoters (12). GWAS studies have also highlighted the SNP (rs6817105) upstream of the *PITX2* gene as a significant locus associated with atrial fibrillation (13). However, the association between upstream non-coding sequences, *PITX2* expression and the development of genetic disorder remains to be fully elucidated.

In this study, we identified a deletion LOH-1 in an intergenic sequence upstream of the *PITX2* gene that co-segregated with the disease phenotype in a Chinese ARS family. We also generated a mouse model with a knockout of the LOH-1 homologous sequence. These mice largely replicated the phenotype of ARS patients, with a notable downregulation in *Pitx2* expression. Subsequent functional experiments revealed that the potential enhancer region P2 which is within LOH-1 may long-rangely regulate the expression of *PITX2* through cohesin-mediated loop domains, and causes ARS when it is lacking.

Results

Clinical characteristics

In this ARS family, household members underwent ocular examination (Figure 1A). Different degrees of corneal embryotoxon, iridocorneal adhesion, iris hypoplasia and iris corectopia were presented in fourteen eyes of all affected patients (Figure 1B-①, ③, ④). Iris polycoria was manifested in patient II:6, III:3, III:7 (Figure 1B-②). Glaucoma was found in patient II:5, II:6, II:8, III:3, III-5 but not obvious in patient III:6 and III:7 who were at the age of 8 and 12 respectively at the time of examination. Teeth and umbilicus were examined by the physicians and no abnormalities were demonstrated in this family (Figure 1B-⑤, ⑥). And the electrocardiographic results in patient II:5 and II:6 did not suggest significant cardiac abnormalities such as atrial fibrillation. The clinical characteristics were summarized in Table S1.

LOH-1 deletion in the intergenic sequence upstream of *PITX2* gene co-segregated with ARS

First, we considered the possibility that variants in the ARS known genes *PITX2* and *FOXC1* might be responsible for the observed clinical manifestations. Exonic regions of *PITX2* and *FOXC1* were examined by Sanger sequencing in proband III:3 from the ARS family and no disease-related variant was found. Then, all collected samples underwent linkage analysis within the ARS family. Parametric multipoint linkage analysis of the ARS family revealed a highly linked locus of approximately 21.2 Mb on 4q22.1- q26, with a maximum logarithm of

odds (LOD) score of 3.258, surrounding the marker rs1354680 (Figure 1C and Table S2). Interestingly, this linkage region contained the known ARS pathogenic gene *PITX2*.

Furthermore, one patient (II:6) was selected for whole genome sequencing (WGS). CNV analysis showed no large deletions/duplications in the *PITX2* gene coding regions and intron regions, while a heterozygous deletion of about 570 kb was detected approximately 200 kb upstream of *PITX2* (Figure S1A), situated between *PITX2* and *FAM241A* (we named the deletion LOH-1).

We used Integrative Genomics Viewer (IGV) software to visualize the WGS reads and to search for sequencing abundance anomalies near the ends of the region from 110,868,844 to 111,438,844 on chromosome 4 (hg38). The sequence information of the reads at the sequencing abundance anomaly was obtained. The precise location of the LOH-1 deletion was inferred to be between 110,869,880 and 111,437,100 (Figure S1B). Sanger sequencing confirmed the putative LOH-1 location and revealed a complete co-segregation of LOH-1 with the ARS phenotype (Figure 1A, D).

LOH-1 knockout mice had ARS phenotypes

To investigate the relationship between LOH-1 deletion and ARS, we knocked out the homologous sequence of LOH-1 (mm10 chr3:128,507,082-128,912,082) in mice by CRISPR/Cas9-targeted strategy. Mendelian ratios analysis of heterozygous mating outcomes revealed a near 50% pre-birth lethality rate in LOH-$1^{-/-}$ (KO) mice, while LOH-$1^{+/-}$(HET) mice appeared unaffected. Examination of pre-birth lethal LOH-$1^{-/-}$ embryos at E12.5 revealed the ventral body wall defect and evisceration (Figure S2A), which is consistent with previous

reports (14, 15). Observation of dorsal images of mice at 3 weeks old revealed that LOH-1$^{-/-}$ mice were reduced in size compared to wild-type (WT) mice, while LOH-1$^{+/-}$ mice did not exhibit obvious abnormality (Figure 2A). By examining the body weight of mice from 3 to 8 weeks after birth, we found a marked weight reduction in LOH-1$^{-/-}$ mice compared to WT mice, with no difference observed in LOH-1$^{+/-}$ mice (Figure 2B). Slit lamp examination and anterior segment optical coherence tomography (AS-OCT) showed clear corneas and normal deep anterior chambers in both LOH-1$^{+/-}$ and WT mice. In contrast, LOH-1$^{-/-}$ mice exhibited opacified corneas and disappeared anterior chambers (Figure 2C). Histopathological findings confirmed normal corneas and angle structures in the WT and LOH-1$^{+/-}$ mice. Thickened corneas with disorganized epithelial and stromal cells, disappeared anterior chambers, and closed angle structures were detected in the LOH-1$^{-/-}$ mice (Figure 2D). By measuring the corneal thickness and corneal epithelial thickness with the histopathological slides of the mice, we identified a significant increase in whole corneal thickness but a decrease in corneal epithelial thickness in LOH-1$^{-/-}$ mice compared with WT, while LOH-1$^{+/-}$ mice showed no such changes (Figure 2E). Atomic force microscopy (AFM) showed increased corneal roughness in LOH-1$^{-/-}$ mice compared to WT mice (Figure S2B, C). However, Fundus retinal images and histopathologic results revealed no retinal abnormalities in LOH-1$^{-/-}$, LOH-1$^{+/-}$ and WT mice (Figure S2D, E).

LOH-1 deletion resulted in dramatically decreased expression of *Pitx2* gene

To further explore the molecular mechanism of LOH-1 knockout mice, we analyzed the transcriptional profiles of 8-week-old LOH-1$^{-/-}$ and WT mouse eyes using RNA sequencing

(RNA-seq) analysis. As shown in Figure 2F, 97 genes were down-regulated and 77 genes were up-regulated (Table S3). Among them, the expression of the ARS causative gene *Pitx2* was significantly reduced. However, the expression levels of another causative gene *Foxc1*, and genes located on either sides of *Pitx2* i.e. *Enpep* and *Fam241a* were not significantly changed. Quantitative PCR (RT-qPCR) analysis further corroborated these RNA-seq findings. The expression of the three isoforms of *Pitx2*, *Pitx2a*, *Pitx2b* and *Pitx2c*, showed marked reductions in the eyes of LOH-$1^{-/-}$ mice, with moderate differences observed in LOH-$1^{+/-}$ mice (Figure 2G). In addition, we examined the expression of *Pitx2* by RT-qPCR in the eye, heart, kidney, stomach, and skeletal muscle of WT and LOH-$1^{-/-}$ embryos at E18.5. The results revealed that *Pitx2* was significantly decreased in these tissues of LOH-$1^{-/-}$ embryos (Figure S2F). These findings indicate that the homozygous knockout of LOH-1 in mice leads to a decrease in *Pitx2* gene expression.

Identification of a potential enhancer region in LOH-1

Given the aforementioned findings, we postulated the presence of cis-regulatory elements such as enhancers within LOH-1 that modulate the expression of *Pitx2*. To identify these elements in LOH-1, we analyzed the enhancer-associated histone modification (monomethylation at histone H3 lysine 4 (H3K4Me1) and acetylation of histone H3 at lysine 27 (H3K27Ac)) ChIP-seq data of 15-week-old human embrynic sclera. Numerous prominent peaks of H3K4Me1 and H3K27Ac signals were observed within this region. (Figure 3A, Figure S4A,B, Table S4,5). Meanwhile, we investigated DNase I hypersensitivity in human embryonic eye, retina, and heart tissues, and assessed the H3K4Me1 in human embryonic heart. We also re-analyzed

datasets from the UCSC genome browser (16), sourced from the Encyclopedia of DNA Elements (ENCODE, the ENCODE Project Consortium) (17, 18), to evaluate the enrichment of H3K4Me1, trimethylation at histone H3 lysine 4 (H3K4Me3), H3K27Ac within the LOH-1 locus across 7 cell lines. The 100 vertebrates Basewise Conservation by PhyloP and the public H3K27Ac ChIP-seq data (human retina and heart) from the Cistrome Data Browser (19, 20) were also used for reference. The results from the above data suggest that there may be multiple enhancer regions in LOH-1.

Subsequently, we selected a potential enhancer region named LOH-E1 (hg38 chr4:111,397,892-111,402,926) in LOH-1 which contains the specific epigenetic signals. The pronounced enrichment of H3K4Me1 and H3K27Ac signals in the LOH-E1 indicated potential enhancer activity (Figure 3B). We also examined the epigenetic signals of mouse homologous regions for LOH-1 and LOH-E1 by the UCSC genome browser from the ENCODE datasets (Figure S3A). The results supported our speculation.

Additionally, according to the GeneHancer database (21), LOH-E1 appears to interact with the *PITX2* gene (Figure S3B), suggesting that *PITX2* could be a potential target gene for the LOH-E1 enhancer region.

Deletion of LOH-E1 significantly reduced *PITX2* expression

To investigate the action mechanism of the LOH-E1, we employed the CRISPR/Cas9 system to specifically knock out the LOH-E1 region in the human embryonic kidney cell line (HEK293) (Figure S5A). Sequencing confirmed the generation of homozygous LOH-E1 deletion clones (HEK293-KO). RT-qPCR analysis revealed that the expression of *PITX2A* and *PITX2B* was

strongly downregulated in the HEK293-KO group compared with the wild type cells. However, the expression of *PITX2C* remained unchanged (Figure 4B). Western blot also confirmed that the expression of PITX2B was downregulated in the HEK293-KO group (Figure 4C).

Deletion of LOH-E1 suppressed cell proliferation

Observations on LOH-E1 depleted cells showed no significant morphological changes (Figure 4A). CCK-8 assay indicated a significant reduction in the proliferation rate of HEK293-KO group compared to wild-type cells (Figure 4D). Cell cycle analysis revealed a notable decrease in the G1 phase and an increase in the S phase for LOH-E1$^{-/-}$ cells compared to control cells (Figure 4E). In the scratch assay, 2-well culture slices produced uniform gaps in the confluent monolayer and wound healing was imaged at various time points (Figure S5B). The results showed a slower wound healing rate in the KO group compared to the control group (Figure S5C). In the apoptosis assay, early apoptosis was significantly higher in the HEK293-KO group compared to the control cells, with no significant changes in late apoptosis (Figure S5D).

P2 which is on LOH-E1 could regulate *PITX2* expression by binding to RAD21

We sought to elucidate the molecular mechanism by which LOH-E1 specifically modulates *PITX2* expression. Enhancers can recruit transcription factors (TFs) and coactivators to alter chromatin spatial structure and improve transcription of target genes. To identify potential TFs, we analyzed the LOH-E1 and *PITX2* core promoter region that are bound to TFs, focusing on regions enriched in H3K4Me1, H3K4Me3 and H3K27Ac. This analysis was conducted using the UCSC genome browser which is sourced from ChIP-seq of the ENCODE datasets,

following the method of Deng et al. (22). Among the 340 TFs in the ENCODE dataset, we identified 12 TFs that are associated with LOH-E1 and overlapped with regions enriched in H3K4Me1 and H3K27Ac (Figure S6A). Meanwhile, 26 TFs were found to bind to the core promoter region of *PITX2*, overlapping with regions enriched in both H3K4Me3 and H3K27Ac (Figure S6B). The Veen diagram revealed two overlapping TFs, RAD21 and CREB1(Figure 5A).

To identify the specific TFs recruited by the LOH-E1 enhancer region, we first analyzed DNase I hypersensitivity (human embryonic eye, retina, and heart) and the H3K4Me1 data (HEK293 and human heart) from the ENCODE dataset as well as the human eye- and heart-related single-cell ATAC data from the CATlas database (23). This analysis identified two core regions within LOH-E1 with significant epigenetic signals (Figure S7), located at hg38 chr4:111,398,770-111,400,269 (1500bp) and hg38 chr4:111,400,720-111,401,749 (1030bp), respectively. Based on these regions, we synthesized a double-stranded DNA probe labeled with biotin at the 5' end, which was purified by incubation with nuclear proteins from HEK293 cells for Liquid Chromatograph Mass Spectrometer (LC-MS) detection (Figure S8A). Among the results of the enriched differential proteins, we found that RAD21, which is a critical component of the cohesin complex, may be the main transcription cofactors affected by LOH-E1 (Figure 5B, Table S6).

As LOH-E1 correlates with the expression of *PITX2*, to further investigate the relationship between RAD21 and PITX2, we performed a correlation analysis using GEPIA2 database (24). The results showed a positive correlation between these two genes in heart and muscle (Figure S8B), which are both gene-specific tissue in Genotype-Tissue Expression (GTEx) database

(25). We also visualized Hi-C (high-throughput chromosome conformation capture) and Virtual 4C (Circular chromosomal conformation capture) data in human Adrenal gland online in the 3D Genome Browser (26), and the results suggest a possible long-range interaction between LOH-E1 and *PITX2* (Figure S8C, D).

Considering the potential role of RAD21 in the LOH-E1 mediated regulation of *PITX2*, we hypothesized that downregulation of RAD21 would reduce the expression of the target gene *PITX2*. We designed small interfering RNA (siRNA) targeting *RAD21* (Figure 5C). Subsequently, we selected RAD21_si2 and RAD21_si3 to detect the expression of target genes *PITX2A*, *B*, and *C* in HEK293 cells, and found that the mRNA expression of *PITX2A*, *B*, and *C* was significantly reduced (Figure 5D).

A recent study suggests that RAD21 N-terminal tail binds DNA to guide it through the kleisin gate and finally entry into the cohesin ring (27). To scan the binding sequence of RAD21 in LOH-E1 core enhancer region, we used the TF binding site prediction function in AnimalTFDB 3.0 database (28, 29), and finally selected a 41bp sequence named JJ (hg38 chr4:111,399,619-111,399,659) (Table S7). According to sequence comparison, two regions in JJ sequence can match the RAD21 motif (Figure 5E), so we hypothesized that RAD21 could bind to JJ sequence. In order to verify this speculation, we performed ChIP-qPCR on the JJ sequence-containing region P2 (hg38 chr4:111,399,594-111,399,691) and surrounding regions P1 (hg38 chr4:111,398,648-111,398,756) and P3 (hg38 chr4:111,400,348-111,400,409). After that, we found a stronger enrichment of RAD21 binding to the P2 compared to P1 and P3 in the HEK293 cell line, when the RAD21 antibody was compared to the control antibody immunoglobulinG (IgG) (Figure 5F). This evidence suggested a direct interaction between

RAD21 and P2.

Discussion

In this study, we identified a heterozygous deletion LOH-1 in the upstream intergenic region of the *PITX2* gene by genome-wide linkage analysis and whole-genome sequencing, which co-segregates with ARS in a Chinese ARS family. Knockout of LOH-1 homologous sequence in mice revealed that LOH-$1^{-/-}$ mice developed ARS-associated phenotypes and that *Pitx2* gene expression level was significantly decreased. These results suggest that deletion of non-coding intergenic sequence LOH-1 can induce ARS in both humans and mice.

Several studies have previously highlighted the presence of functional elements in the intergenic non-coding sequences upstream of the *PITX2* gene, which are linked to genetic diseases. However, the pathogenic mechanisms involved are still unclear. Combined with our genetic results, we hypothesized that intergenic sequences upstream of *PITX2* should play a regulatory role in *PITX2* gene expression. We then performed a series of bioinformatics analyses of epigenetic data in public databases, along with ChIP-seq to target a potential enhancer region LOH-E1 in LOH-1. In the HEK293 cell line, cell proliferation, cell cycle, apoptosis and migration were all affected by LOH-E1 deletion, and the expression levels of *PITX2* were downregulated, verifying that LOH-E1 may be a candidate enhancer region that regulates *PITX2* expression. We summarize the deletions for the 4q25 locus that have been reported and the deletions that we found (Figure S9). Among them ARS_Volkman, Aniridia_Ansari and ARS_Protas are explicitly shown to overlap with LOH-E1. Through data analysis by relevant experiments and public database, RAD21 emerged as a potential protein binding to P2 sequence which is on LOH-E1. As the core subunit of cohesin, RAD21 is the only physical linkage between the SMC1/SMC3 heterodimer and the STAG subunit, which regulates the binding or dissociation of cohesin from chromatin and is involved in regulating gene transcriptional expression (30, 31). And whether DNA loading is successful or results in loop extrusion might be determined by a conserved RAD21 N-terminal tail that guides the

DNA through the kleisin gate (27). The above evidence demonstrated that P2 which is on LOH-E1 binds to RAD21 to regulate the expression of *PITX2*.

However, we found that the expression of the three isoforms of *PITX2* was not completely consistent in LOH-E1 knockout cells and LOH-1 knockout mice, as reflected in *PITX2C*. Among the three isoforms of *Pitx2*, the *Pitx2a, b* splice variants share the same promoter and are expressed bilaterally in some tissues, whereas *Pitx2c* is transcribed from a separate promoter and is expressed asymmetrically (32). We hypothesized that LOH-E1 may act only on the *PITX2A, B* promoter and not the *PITX2C* promoter. And LOH-E1 knockout mice will be constructed subsequently to further clarify the mechanism of LOH-E1 with *Pitx2*. Epigenetic data analysis further suggested that LOH-E1 might not be the only *PITX2* enhancer within LOH-1, hinting at the existence of other enhancer active regions that warrant further exploration.

PITX2 plays a crucial role early in embryonic development to regulate the left-right asymmetric development of internal organs such as the gut, heart, liver and stomach (e. g. gut rotation) (33). Human and mouse embryonic development have different requirements for the dose of *PITX2*/*Pitx2*, with human development being more sensitive to the appropriate dose of *PITX2* (11). This may explain the fact that LOH-1 heterozygous deletion causes ARS in humans. In previous studies, *Pitx2* knockout homozygous mice would die prematurely due to various developmental defects, while heterozygous mice were usually described as normal (14, 15, 34, 35). In our study, the LOH-$1^{+/-}$ (HET) mice exhibited diminished expression of *Pitx2* compared to wild type, but were able to reproduce and survive normally and did not have a significant ARS phenotype. Moreover, the LOH-$1^{-/-}$ (KO) mice are not completely lethal, with about 50% surviving to adulthood. Surviving KO mice displayed pronounced reduced expression of *Pitx2*

and a distinct ARS phenotype. Therefore, the LOH-1 knockout mice present an ideal model for delving deeper into ARS mechanisms.

In summary, our study unveiled an enhancer region P2 which is on LOH-1, regulating the expression of *PITX2* by binding to RAD21, and elucidated its significance in ARS. In the absence of LOH-1, RAD21 fails to successfully guide DNA into the cohesin ring, leading to the pathogenicity of ARS. This work helps to improve the understanding of intergenic sequence variants, enhance the diagnosis of related genetic diseases, and offer potential avenues for the prevention and treatment of ARS.

Methods

Sex as a biological variant. This study examined male and female mice and similar results were obtained for both sexes.

Subjects and clinical examination. This study recruited seventeen individuals from a Chinese ARS family. Blood samples were collected from all seventeen family members and the gDNA was extracted from peripheral blood. 7 out of 9 patients underwent detailed clinical examinations in the Second Xiangya hospital to further confirm the clinical manifestations of ARS. The other two were not able to come to the hospital, including II:2 who was blind for years and hence refused to leave his hometown and IV:1 who was too young to cooperate with the examinations. In addition to 7 patients, 1 normal individual (III:4) also received clinical examination. The average age of eight individuals was 30.1, ranging from 8 to 49. And the gender ratio was 1:1 (4 male:4 female). All participants were checked in the Second Xiangya Hospital, Central South University by ophthalmologists for ocular manifestations and internal physicians for extra-ocular manifestations. Best-corrected visual acuity (BCVA) was determined using a logMAR chart and an auto refractometer (Topcon KR-800, Topcon Optical Company, Tokyo, Japan). Intraocular pressure was measured using a Goldmann applanation tonometer. Anterior segment and angles were checked with a slit lamp (Huvitz Slit Lamp HS-5000, Coburn Technologies, Singapore) and a gonioscope (Suzhou Liuliu Vision Technology Co.). Fundus examination was performed with an ophthalmoscope.

Genome-wide linkage analysis. Seventeen samples from the ARS family were genotyped

using Illumina iScan system (Illumina, USA) and Illumina HumanCytoSNP-12 V.2.1 BeadChIP kit. The called Genotypes were quality controlled by Illumina GenomeStudio V2011.1, after excluding SNPs with low quality and check PC_error as well as PPC_error, a data file (ARS.dat), a pedigree file (ARS.ped) and a map file (ARS.map) were extracted. Genome-wide linkage disequilibrium of the ARS family was tested by merlin V.1.1.2 (36) under multiple-parameter analysis with model file settings (para.model) 'VERY_RARE_DISEASE 0 0,0.99,0.99 Dominant_Model'. The 'merlin' prompt was used to analyze autosomal linkage disequilibrium, while the 'minx' prompt was used for X-linked linkage disequilibrium analysis.

Whole genome sequencing. The gDNA of ARS-II:6 was analysed through WGS. After passing the DNA quality check, the DNA was broken into fragments of 200~300bp using a Biorupter ultrasonic fragmenter. Then, after repairing the sticky ends, a phosphate group was added to the 5' end and A to the 3' end, and the splice sequence with index was added to both ends of the DNA fragment by TA ligation. Finally, the DNA library was amplified by PCR. The libraries were quantified using the Qubit instrument, the qualified libraries were placed into the cBot for bridge amplification, and the clusters were generated and sequenced using the Illumina Hiseq sequencing system. Next the reads were aligned to the human genome assembly GRCh38/hg38 using iSAAC-01.15.04.01 (37). Variants were called with Isaac Variant Caller v1.0.6 (37) and Control-freeC v9.1 (38) annotated using ANNOVAR (39). Integrative Genomics Viewer (40) was used to import WGS data in .bam format for visual analysis.

Primer design. All primers were designed using the online software Primer3 (https://primer3.ut.ee/) based on the human GRCh38/hg38 or mouse GRCm38/mm10 assemblies. The sequence information of all primers is listed in Table S8. PCR and Sanger sequencing were conducted in the seventeen samples from the ARS family to confirm the co-segregation status of LOH-1 deletion.

Mice. The LOH-1 knockout mice were generated by zygote injection of CRISPR/Cas9 mRNA and a pair of gRNAs. The mice were genotyped by PCR and the PCR products were confirmed by Sanger sequencing. Mice live in an environment with relatively stable temperature (typically 22-24°C) and humidity (typically 70%), in both light and darkness for 12 hours, and were allowed to eat and drink freely. The genetic background of all mice used in this project was C56BL/6J. All animal experiments were complied with all relevant ethical regulations and were approved by the Committee for Experimental Animals at Central South University.

Ophthalmic examination of mice. The ocular manifestations of the mice were examined using a slit lamp (Huvitz Slit Lamp HS-5000, Coburn Technologies, Singapore) and a gonioscope (Suzhou Liuliu Vision Technology Co.). Anterior segment and angles were checked using optical coherence tomography (OCT) (Visante OCT, Zeiss, Germany). Fundus retinal images were captured by a small animal retinal microscopic imaging system (Micron IV Retinal Imaging Microscope, Phoenix Research Labs, USA). Mice were anesthetized by intraperitoneal injection of a mixture with ketamine and xylazine prior to the testing.

H&E staining. Mouse eyeballs were fixed in 4% paraformaldehyde/PBS at 4°C overnight. Following dehydration in graded alcohol, the samples were embedded in paraffin and sectioned sagittally at 5 μm using a paraffin microtome (RM2235, Leica Biosystems, Germany). The sections were then dewaxed, rehydrated, and stained with hematoxylin and eosin using the H&E staining kit (Cat: ab245880, Abcam). The slices are digitally scanned and imported into CaseViewer 2.4 (3DHISTECH, Hungary) for processing.

RNA-sequencing and data analysis. Total RNA was extracted from mouse eyes using TRIzol (Cat: AM9738, Invitrogen) depending on the manufacturer's instructions. RNA quality was assessed using the RNA Nano 6000 Assay kit on the Agilent 2100 bioanalyzer system (Agilent Technologies, CA, USA). The RNA-seq library was performed using the NEB Next Ultra RNA Library Prep Kit for Illumina (Cat: E7530, NEB). In short, mRNA was purified from total RNA by using Oligo (dT) magnetic beads. Fragmentation was conducted in First Strand Synthesis Reaction Buffer using divalent cations at high temperatures. First-strand cDNA was synthesized using fragmentation mRNA and random oligonucleotides primers in the M-MuLV Reverse Transcriptase System, followed by degradation of the RNA strand with RNaseH. Afterwards, the second strand cDNA was synthesized with DNA polymerase I and dNTPs. Library fragments were sorted and purified with the AMPure XP system (Beckman Coulter, Beverly, USA). After PCR amplification, the PCR products were purified again using AMPure XP beads to get the final library. After established, the library was initially quantified with a Qubit 2.0 Fluorometer, and then the insert size of the library was checked using an Agilent

2100 bioanalyzer. After the insert size meets expectations, the effective library concentration was accurately quantified by RT-qPCR (effective library concentration above 2nM). Libraries were qualified and sequenced by Illumina NovaSeq 6000. The sequencing data were filtered using SOAPnuke v1.5.2 (41), then paired-end clean reads were aligned to the reference genome with Hisat2 v2.0.5 (42). Differential expression analysis of two groups was carried out with the DESeq2 R package v1.20.0 (43). The resulting P-values were adjusted using the Benjamini and Hochberg's method to control for false discovery rate. Padj<=0.05 and |log2(foldchange)| >= 1 were chosen as the threshold for significantly differential expression. The RNA-sequencing data reported in this paper have been deposited in the Genome Sequence Archive (44) in National Genomics Data Center (45), China National Center for Bioinformation / Beijing Institute of Genomics, Chinese Academy of Sciences.

RT-qPCR. The cDNA was synthesized with RevertAid First Strand cDNA Synthesis Kit (Cat: K1622, Thermo Fisher Scientific). RT-qPCR was performed using the Maxima SYBR Green qPCR Master Mix (Cat: K0251, Thermo Fisher Scientific) on CFX96 Real-Time PCR Detection System (Bio-Rad, USA). Beta-Actin (β-Actin) and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) were used as internal controls to normalize the mRNA levels of candidate genes. Data were imported into CFX manager 3.1 (Bio-Rad, USA) and then analyzed.

Western blotting. Samples were lysed in 2×SDS lysis buffer. Proteins were separated by 10% SDS-PAGE and transferred to a PVDF membrane. The membranes were incubated with a primary antibody overnight at 4°C. The primary antibodies used were PITX2 antibody (Cat:

ab221142, Abcam) and α-tubulin antibody (Cat: 5335, CST). Membranes were incubated with HRP-conjugated secondary antibody. The proteins were visualized using the SuperSignal West Femto (Cat: 34096, Thermo Fisher Scientific). Band intensities were quantified by Quantity One (Bio-Rad, USA).

Chromatin immunoprecipitation (ChIP). ChIP assays were conducted on HEK293 cells and sclera from human embryos at 15 weeks. The kit used for the experiment was SimpleChIP Plus Sonication Chromatin IP Kit (Cat: 56383, CST). The tissue samples were minced into 1-2 mm cubes with a clean scalpel, fixed in 16% formaldehyde (Cat: 12606, CST) and incubated for 10 min at room temperature in order to crosslink the proteins to the DNA, then quenched in glycine solution on ice for 5 min. The tissue was resuspended in ChIP Sonication Cell Lysis Buffer, and the tissue suspension was transferred into a Dounce homogenizer using a cut pipet tip, and the tissue pieces were broken up using a tight fitting pestle (type A) until no large pieces of tissue were observed. Resuspend tissue suspension in ice-cold ChIP Sonication Nuclear Lysis Buffer and incubate on ice for 10 min. After transfer of tissue suspensions into Covaris microTUBEs for sonication using Covaris S2 (Covaris, USA), the chromatin was incubated with H3K4Me1 antibody (10 μl per reaction, Cat: 5326, CST), H3K27Ac antibody (5 μl per reaction, Cat: 8173, CST), RAD21 antibody (10 μl per reaction, Cat: ab217678, Abcam) or IgG for 4 h at 4°C with rotation. Immediately add Protein G Magnetic Beads to each IP reaction and incubate for 2 h at 4°C with rotation. The chromatin was eluted from the antibody/Protein G Magnetic Beads and gently vortexed at 65°C for 30 min (1200 rpm). Next, reverse cross-links by adding 5M NaCl and Proteinase K, and incubate 2 h at 65°C. Finally, DNA was

purified with DNA purification spin columns. Immunoprecipitated and input DNA was tested to downstream assay using high throughput sequencing or qPCR.

ChIP-seq data analysis. The ChIP-seq sequencing project was completed on the MGISEQ-T7 sequencing platform, and PE libraries (~300bp) were constructed for sequencing. The raw data was filtered using fastp v0.23.0 (46) to obtain high quality sequencing clean data. The filtered clean reads were mapped to the Human Genome Overview GRCh38/hg38 assembly using bowtie2 v2.5.1 (47). MACS2 v2.2.7.1 (48) was used to call peaks using the broad peak settings, with input used as the control.

Cell culture. HEK293 cells (Cat: CRL-1573, American Type Culture Collection) were cultured in Dulbecco's modified Eagle's medium (DMEM) (Cat: C11995500BT, Gibco) with 10% fetal bovine serum (FBS) (Cat: 10099141, Gibco) and 1% penicillin and streptomycin (Cat: 15140122, Thermo Fisher Scientific). All cells were cultured at 37°C in a humidified incubator containing 5% CO2 and passaged every 3 days.

LOH-E1 region knockout cell line construction. The target-specific sgRNAs were designed using CRISPOR (49), and then were inserted into the plasmid VB105n at the AscI and KpnI restriction site. The constructs were confirmed by Sanger sequencing. Plasmids were transfected into HEK293 cells by electroporator, and the parameter setting voltage is 1200V, pulse width is 10ms, the number of electric shock is 3 times. After 24 hours of transfection, the cells were selected with 600 μg/ml hygromycin for 3 days. Next, the surviving cells were

trypsinized and diluted to 96-well plates for monoclonal cell screening. Genomic DNA of different monoclonal cells was extracted, and the CRISPR/Cas9-edited sites were subjected to PCR amplification and Sanger sequencing to obtain monoclonal cell lines with knockdown of the target fragments. The sequence information of all sgRNAs is listed in Table S8.

CCK8 cell proliferation assay. Cells were trypsinized and diluted to 96-well plates for cell proliferation assay. Using the Countess II automated cell counter (Cat: AMQAX1000, Thermo Fisher Scientific) to seed each well with 2000 cells. CCK-8 solution (Cat: A311-01, Vazyme) was then added to each well at 24, 48, 72 and 96 hours after seeding, and three replicate experiments for each time point. After incubation at 37 °C for 2 h, the absorbance values of individual wells at 450 nm were measured separately using the Synergy H4 multifunctional microplate detector (BioTek, USA).

Cell cycle analysis. The Cell Cycle Analysis Kit (Cat: C1052, Beyotime) was used to perform cell cycle analysis depending on the manufacturer's instructions. Briefly, first carefully collect the cell culture fluid, set aside and treat the cells with trypsin, then add the previously collected cell culture fluid, blow down all the cells, centrifuge at 1000 × g for 5 minutes, and precipitate the cells. Carefully aspirate off the supernatant, add pre-chilled PBS to resuspend the cells, then add pre-chilled 70% ethanol, mix with gentle blowing, and fix for 24 hours at 4°C. Prepare fresh propidium iodide staining solution to resuspend the cell precipitate at 37°C for 30 minutes protected from light. Finally, the cell cycle assay was conducted on the DXP Athena Flow Cytometer (Cytek, USA), followed by analysis in FlowJo v10.

Cell apoptosis assay. Cell apoptosis was assessed using the Annexin V-FITC/PI Apoptosis Detection Kit (Cat: A211-01, Vazyme) depending on the manufacturer's instructions. Briefly, cells were treated via trypsinization without EDTA, centrifuge at 300 × g for 5 min at 4°C and discard supernatant. Cells were washed twice with pre-cooled PBS and centrifuged at 300 × g for 5 min at 4°C each time. Then add 1 × Binding Buffer and gently blow well to a single cell suspension. Finally, Annexin V-FITC and PI Staining Solution were added, gently blown well, incubated for 10 min at room temperature (20 ~ 25°C) avoiding light and then 1 × Binding Buffer was added and gently mixed. The stained samples were detected by the DXP Athena Flow Cytometer (Cytek, USA) within 1 h, followed by analysis in FlowJo v10.

Wound healing capacity assay. The Culture-Insert 2 Well in μ-Dish 35 mm (Cat: 81176, Ibidi) was used for wound healing capacity test depending on the manufacturer's instructions. Briefly, cells were trypsinized and 70 μL of cell suspension was applied to each well of the Culture-Insert 2 Well. After the cells had been cultured for 24 hours to develop an optically confluent monolayer, the Culture-Insert was removed to establish the wound gap. Wash the cell layer with PBS to remove cell debris and non-adherent cells. Next the cells were cultured with 2 ml serum-free medium. Wound gap photography was observed using Eclipse inverted microscope (Leica, Germany) at 0h, 3h, 6h, 12h, 24h, 48h and 72h.

Cell nuclear protein extraction. Nuclear and Cytoplasmic Protein Extraction Kit (Cat: P0027, Beyotime) was employed to extract nuclear proteins depending on the manufacturer's

instructions. Briefly, cells were washed with PBS and then scraped off with a cell scraper. Cell Cytoplasmic Protein Extraction Reagent A was added to the cell precipitate then vigorously Vortexed at highest speed for 5 seconds to completely suspend and disperse the cell precipitate. After 15 min of ice bath, add Cell Cytoplasmic Protein Extraction Reagent B and centrifuge at maximum speed for 5 sec, then centrifuge at 4°C for 5 min at 16000 × g. Next, remove the supernatant, add the nucleoprotein extraction reagent and centrifuge at the highest speed after vigorous Vortex. Finally, immediately aspirate the supernatant into a pre-cooled plastic tube to obtain the nuclear protein extracted from the cells.

DNA pulldown. The DNA sequence sense strand probes and antisense strand probes were generated by gene synthesis, and then biotin labels the 5' end of the sense strand probe and antisense strand probe to form a double-stranded DNA probe by annealing. Add DNA probes to cell nuclear protein and incubate overnight at room temperature (18~25°C). Afterwards, the mixture was incubated with streptomycin magnetic beads for 4 h in a shaker at room temperature, and the proteins were eluted after centrifugation for mass spectrometry identification. The sequence information of all DNA probes is listed in Table S8.

Mass spectrometric identification and analysis. Proteins enriched by DNA Pulldown were subjected to SDS-PAGE gel electrophoresis. The entire stacking gel was rinsed with water for several times and the bands of interest were excised and cut into cubes (~ 1×1 mm). Next, the digested peptides were extracted from the gel pieces and lyophilized for the following step. The sample was analyzed by on-line nanospray LC-MS/MS on Q Exactive HF mass

spectrometer (Thermo Fisher Scientific, MA, USA) coupled to an EASY-nanoLC 1000system (Thermo Fisher Scientific, MA, USA). Tandem mass spectra were processed by PEAKS Studio version 10.6 (Bioinformatics Solutions Inc., Waterloo, Canada). PEAKS DB was set up to search the database of uniprot-homo sapiens (version201907, 20428 entries) assuming trypsin as the digestion enzyme. The peptides with -10lgP≥20 and the proteins with -10lgP≥20 and containing at least 1 unique peptide were filtered.

siRNA. siRNAs targeting RAD21 were designed using siCatch and purchased from RiboBio (RiboBio, Guangzhou, China). Cells were seeded in 6-well plates 1 day in advance to reach 60-70% confluence. The siRNA against RAD21 and the control siRNA were then transfected with Lipofectamine 3000 reagent (Cat: 2399133, Thermo Fisher Scientific) on the following day according to the manufacturer's instructions. After incubation for 48h, cells were harvested for RNA extraction. The sequence information of all siRNAs is listed in Table S8.

ENCODE data. The DNase-seq data from the ENCODE database used in this study included ENCSR820ICX, ENCSR474GZQ, ENCSR621ENC, ENCSR127PWK, ENCSR782XFY, ENCSR154ZNQ, ENCSR000CNV and ENCSR409LVQ. The ChIP-seq data from the ENCODE database used in this study included ENCSR676ZKW, ENCSR000FCG, ENCSR000CDL and ENCSR000CDK.

Statistical analysis. All experiments were conducted in at least three independent replicates. Significances were assessed by Student’s t test (parametric), one-way or two-way analysis of variance (ANOVA). All data are represented as mean ± SEM and p-value < 0.05 was

considered to indicate a statistically significant difference. *p<0.05, **p<0.01, ***p<0.001, ****P<0.0001, ns, not significant.

Study approval. This study was approved by the Institutional Review Board (IRB) of the School of Life Sciences, Central South University (IRB #2022-2-22). All study participants had thoroughly read and signed the informed consent form before blood samples were collected for further analysis, and that the record of informed consent has been retained.

Data availability. Values for all data points in graphs are reported in the Supporting Data Values file. The RNA-seq data have been deposited in the Genome Sequence Archive (GSA: CRA011549).

Author contributions

Yizheng Jiang and Yu Peng share first-author position. The order of first authorship was determined by the volume of work each author contributed to the study. Yizheng Jiang, Yu Peng, Bei Feng, Junping Hu and Zhe Cheng carried out the experiments and analyzed the data. Yizheng Jiang, Liang Zhou, Yu Peng and Zhengmao Hu participated in the sample collection and clinical examination. Yizheng Jiang, Yu Peng, Qi Tian, Lu Xia, Hui Guo, Kun Xia, and Zhengmao Hu designed and supervised the study. Yizheng Jiang, Liang Zhou, and Zhengmao Hu wrote the manuscript. All authors read and approved the final manuscript.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (grant number 82271908,82130043), the National Key R&D Program of China (grant number 2021YFA0805202), the Key R&D Program of Hunan Province (grant number 2019SK2051, 2021SK1010). The funders had no role in the study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

REFERENCES

1. Valente EM, and Bhatia KP. Solving Mendelian Mysteries: The Non-coding Genome May Hold the Key. *Cell.* 2018;172(5):889-91.
2. Eichler EE. Genetic Variation, Comparative Genomics, and the Diagnosis of Disease. *The New England journal of medicine.* 2019;381(1):64-74.
3. Seifi M, and Walter MA. Axenfeld-Rieger syndrome. *Clin Genet.* 2018;93(6):1123-30.
4. Semina EV, Reiter R, Leysens NJ, Alward WL, Small KW, Datson NA, et al. Cloning and characterization of a novel bicoid-related homeobox transcription factor gene, RIEG, involved in Rieger syndrome. *Nature genetics.* 1996;14(4):392-9.
5. Mears AJ, Jordan T, Mirzayans F, Dubois S, Kume T, Parlee M, et al. Mutations of the forkhead/winged-helix gene, FKHL7, in patients with Axenfeld-Rieger anomaly. *American journal of human genetics.* 1998;63(5):1316-28.
6. Phillips JC, del Bono EA, Haines JL, Pralea AM, Cohen JS, Greff LJ, et al. A second locus for Rieger syndrome maps to chromosome 13q14. *American journal of human genetics.* 1996;59(3):613-9.
7. Micheal S, Siddiqui SN, Zafar SN, Venselaar H, Qamar R, Khan MI, et al. Whole exome sequencing identifies a heterozygous missense variant in the PRDM5 gene in a family with Axenfeld-Rieger syndrome. *Neurogenetics.* 2016;17(1):17-23.
8. Riise R, Storhaug K, and Brøndum-Nielsen K. Rieger syndrome is associated with PAX6 deletion. *Acta ophthalmologica Scandinavica.* 2001;79(2):201-3.
9. Sibon I, Coupry I, Menegon P, Bouchet JP, Gorry P, Burgelin I, et al. COL4A1 mutation in Axenfeld-Rieger anomaly with leukoencephalopathy and stroke. *Annals of neurology.* 2007;62(2):177-84.
10. Tanwar M, Dada T, and Dada R. Axenfeld-Rieger Syndrome Associated with Congenital Glaucoma and Cytochrome P4501B1 Gene Mutations. *Case reports in medicine.* 2010;2010.
11. Volkmann BA, Zinkevich NS, Mustonen A, Schilter KF, Bosenko DV, Reis LM, et al. Potential novel mechanism for Axenfeld-Rieger syndrome: deletion of a distant region containing regulatory elements of PITX2. *Investigative ophthalmology & visual science.* 2011;52(3):1450-9.
12. Aguirre LA, Alonso ME, Badía-Careaga C, Rollán I, Arias C, Fernández-Miñán A, et al. Long-range regulatory interactions at the 4q25 atrial fibrillation risk locus involve PITX2c and ENPEP. *BMC biology.* 2015;13:26.
13. Ellinor PT, Lunetta KL, Albert CM, Glazer NL, Ritchie MD, Smith AV, et al. Meta-analysis identifies six new susceptibility loci for atrial fibrillation. *Nature genetics.* 2012;44(6):670-5.

14. Lin CR, Kioussi C, O'Connell S, Briata P, Szeto D, Liu F, et al. Pitx2 regulates lung asymmetry, cardiac positioning and pituitary and tooth morphogenesis. *Nature.* 1999;401(6750):279-82.
15. Lu MF, Pressman C, Dyer R, Johnson RL, and Martin JF. Function of Rieger syndrome gene in left-right asymmetry and craniofacial development. *Nature.* 1999;401(6750):276-8.
16. Lee BT, Barber GP, Benet-Pagès A, Casper J, Clawson H, Diekhans M, et al. The UCSC Genome Browser database: 2022 update. *Nucleic acids research.* 2022;50(D1):D1115-d22.
17. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. *Nature.* 2012;489(7414):57-74.
18. Davis CA, Hitz BC, Sloan CA, Chan ET, Davidson JM, Gabdank I, et al. The Encyclopedia of DNA elements (ENCODE): data portal update. *Nucleic acids research.* 2018;46(D1):D794-d801.
19. Mei S, Qin Q, Wu Q, Sun H, Zheng R, Zang C, et al. Cistrome Data Browser: a data portal for ChIP-Seq and chromatin accessibility data in human and mouse. *Nucleic acids research.* 2017;45(D1):D658-d62.
20. Zheng R, Wan C, Mei S, Qin Q, Wu Q, Sun H, et al. Cistrome Data Browser: expanded datasets and new tools for gene regulatory analysis. *Nucleic acids research.* 2019;47(D1):D729-d35.
21. Fishilevich S, Nudel R, Rappaport N, Hadar R, Plaschkes I, Iny Stein T, et al. GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. *Database : the journal of biological databases and curation.* 2017;2017.
22. Deng R, Huang JH, Wang Y, Zhou LH, Wang ZF, Hu BX, et al. Disruption of super-enhancer-driven tumor suppressor gene RCAN1.4 expression promotes the malignancy of breast carcinoma. *Molecular cancer.* 2020;19(1):122.
23. Zhang K, Hocker JD, Miller M, Hou X, Chiou J, Poirion OB, et al. A single-cell atlas of chromatin accessibility in the human genome. *Cell.* 2021;184(24):5985-6001.e19.
24. Tang Z, Kang B, Li C, Chen T, and Zhang Z. GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic acids research.* 2019;47(W1):W556-w60.
25. Consortium G. The Genotype-Tissue Expression (GTEx) project. *Nature genetics.* 2013;45(6):580-5.
26. Wang Y, Song F, Zhang B, Zhang L, Xu J, Kuang D, et al. The 3D Genome Browser: a web-based browser for visualizing 3D genome organization and long-range chromatin interactions. *Genome biology.* 2018;19(1):151.
27. Higashi TL, Eickhoff P, Sousa JS, Locke J, Nans A, Flynn HR, et al. A Structure-Based Mechanism for DNA Entry into the Cohesin Ring. *Molecular cell.* 2020;79(6):917-33.e9.
28. Hu H, Miao YR, Jia LH, Yu QY, Zhang Q, and Guo AY. AnimalTFDB 3.0: a comprehensive resource for annotation and prediction of animal transcription factors. *Nucleic acids research.* 2019;47(D1):D33-d8.
29. Zhang Q, Liu W, Zhang HM, Xie GY, Miao YR, Xia M, et al. hTFtarget: A Comprehensive Database for Regulations of Human Transcription Factors and Their Targets. *Genomics, proteomics & bioinformatics.* 2020;18(2):120-8.
30. Deardorff MA, Wilde JJ, Albrecht M, Dickinson E, Tennstedt S, Braunholz D, et al. RAD21 mutations cause a human cohesinopathy. *American journal of human genetics.* 2012;90(6):1014-27.
31. Yun J, Song SH, Kang JY, Park J, Kim HP, Han SW, et al. Reduced cohesin destabilizes high-level gene amplification by disrupting pre-replication complex bindings in human cancers with chromosomal instability. *Nucleic acids research.* 2016;44(2):558-72.
32. Hu S, Mahadevan A, Elysee IF, Choi J, Souchet NR, Bae GH, et al. The asymmetric Pitx2 gene

regulates gut muscular-lacteal development and protects against fatty liver disease. *Cell reports*. 2021;37(8):110030.

33. Sanketi BD, Zuela-Sopilniak N, Bundschuh E, Gopal S, Hu S, Long J, et al. Pitx2 patterns an accelerator-brake mechanical feedback through latent TGFβ to rotate the gut. *Science (New York, NY)*. 2022;377(6613):eabl3921.
34. Campione M, Steinbeisser H, Schweickert A, Deissler K, van Bebber F, Lowe LA, et al. The homeobox gene Pitx2: mediator of asymmetric left-right signaling in vertebrate heart and gut looping. *Development (Cambridge, England)*. 1999;126(6):1225-34.
35. Gage PJ, Suh H, and Camper SA. Dosage requirement of Pitx2 for development of multiple organs. *Development (Cambridge, England)*. 1999;126(20):4643-51.
36. Abecasis GR, Cherny SS, Cookson WO, and Cardon LR. Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. *Nature genetics*. 2002;30(1):97-101.
37. Raczy C, Petrovski R, Saunders CT, Chorny I, Kruglyak S, Margulies EH, et al. Isaac: ultra-fast whole-genome secondary analysis on Illumina sequencing platforms. *Bioinformatics (Oxford, England)*. 2013;29(16):2041-3.
38. Boeva V, Popova T, Bleakley K, Chiche P, Cappo J, Schleiermacher G, et al. Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data. *Bioinformatics (Oxford, England)*. 2012;28(3):423-5.
39. Wang K, Li M, and Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic acids research*. 2010;38(16):e164.
40. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nature biotechnology*. 2011;29(1):24-6.
41. Chen Y, Chen Y, Shi C, Huang Z, Zhang Y, Li S, et al. SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. *GigaScience*. 2018;7(1):1-6.
42. Mortazavi A, Williams BA, McCue K, Schaeffer L, and Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature methods*. 2008;5(7):621-8.
43. Love MI, Huber W, and Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*. 2014;15(12):550.
44. Chen T, Chen X, Zhang S, Zhu J, Tang B, Wang A, et al. The Genome Sequence Archive Family: Toward Explosive Data Growth and Diverse Data Types. *Genomics, proteomics & bioinformatics*. 2021;19(4):578-83.
45. Partners C-NMa. Database Resources of the National Genomics Data Center, China National Center for Bioinformation in 2023. *Nucleic acids research*. 2023;51(D1):D18-d28.
46. Chen S, Zhou Y, Chen Y, and Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics (Oxford, England)*. 2018;34(17):i884-i90.
47. Langmead B, and Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nature methods*. 2012;9(4):357-9.
48. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). *Genome biology*. 2008;9(9):R137.
49. Concordet JP, and Haeussler M. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic acids research*. 2018;46(W1):W242-w5.

Figures and figure legends

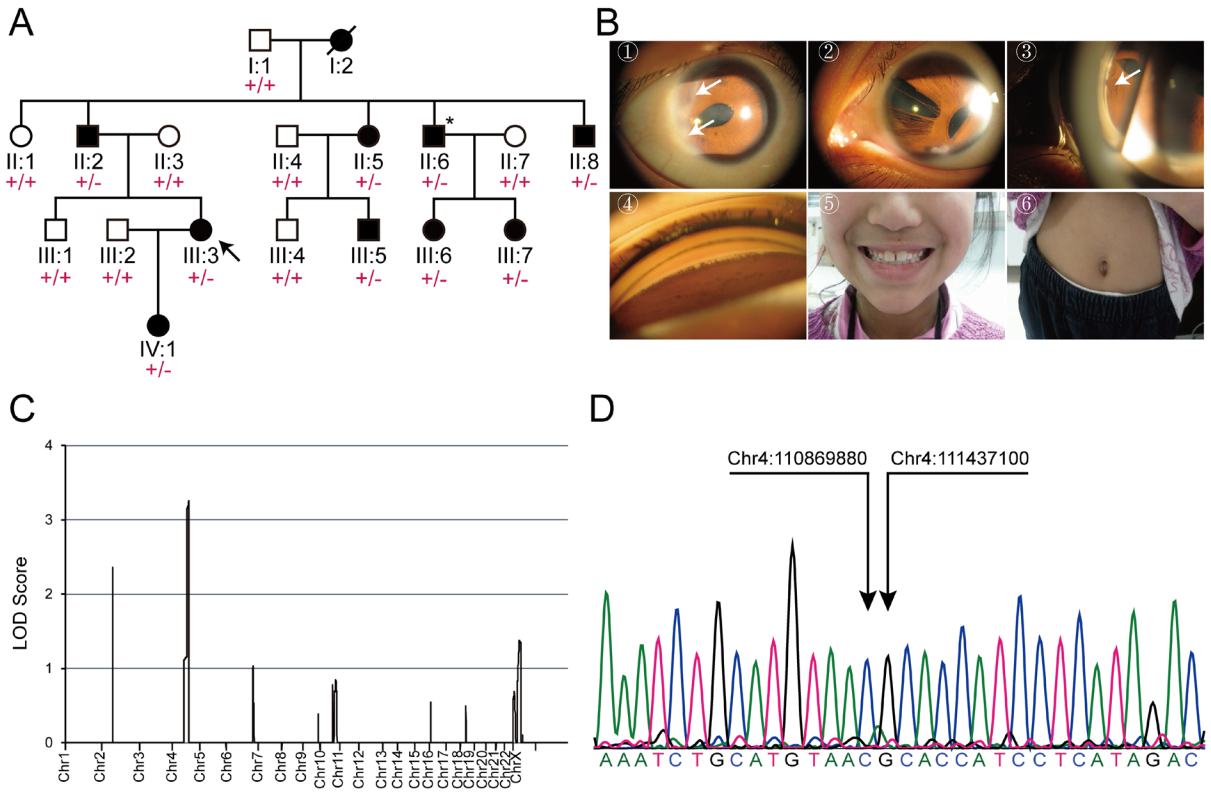


Figure 1. Clinical features and genetic studies of the Axenfeld-Rieger syndrome (ARS) family. (**A**) Pedigree of an ARS family. The arrow in the pedigree indicates the proband. "□" and "○" symbols represent normal male and female subjects, "■" and "●" elements stand for male and female patients, respectively. Samples selected for whole genome sequencing (WGS) were marked with '*'. '+' stands for wild-type allele and '-' refers to LOH-1 deletion allele. (**B**) Clinical manifestations of patients III:3, III:6 and III:7 from the ARS family. ①. An anterior segment photograph of the right eye (III:3) shows posterior embryotoxon (arrow) at the temporal corneal limbus and corectopia towards the temporal side. ②. Pseudopolycoria and atrophic iris in the left eye (III:3). ③. Gonioscopic photographs demonstrating iridocorneal attachment (arrow) and corectopia in the left eye (III:6). ④. Extensions of the peripheral iris to Schwalbe's line are seen on gonioscopy in the right eye (III:6). ⑤. Normal teeth are shown for affected individual III:7. ⑥. Normal umbilicus are shown for affected individual III:7. (**C**) Genome-wide multipoint linkage analysis of ARS family. (**D**) Sanger sequencing identifies the precise location of the LOH-1 deletion as hg38 chr4:110,869,880-111,437,100.

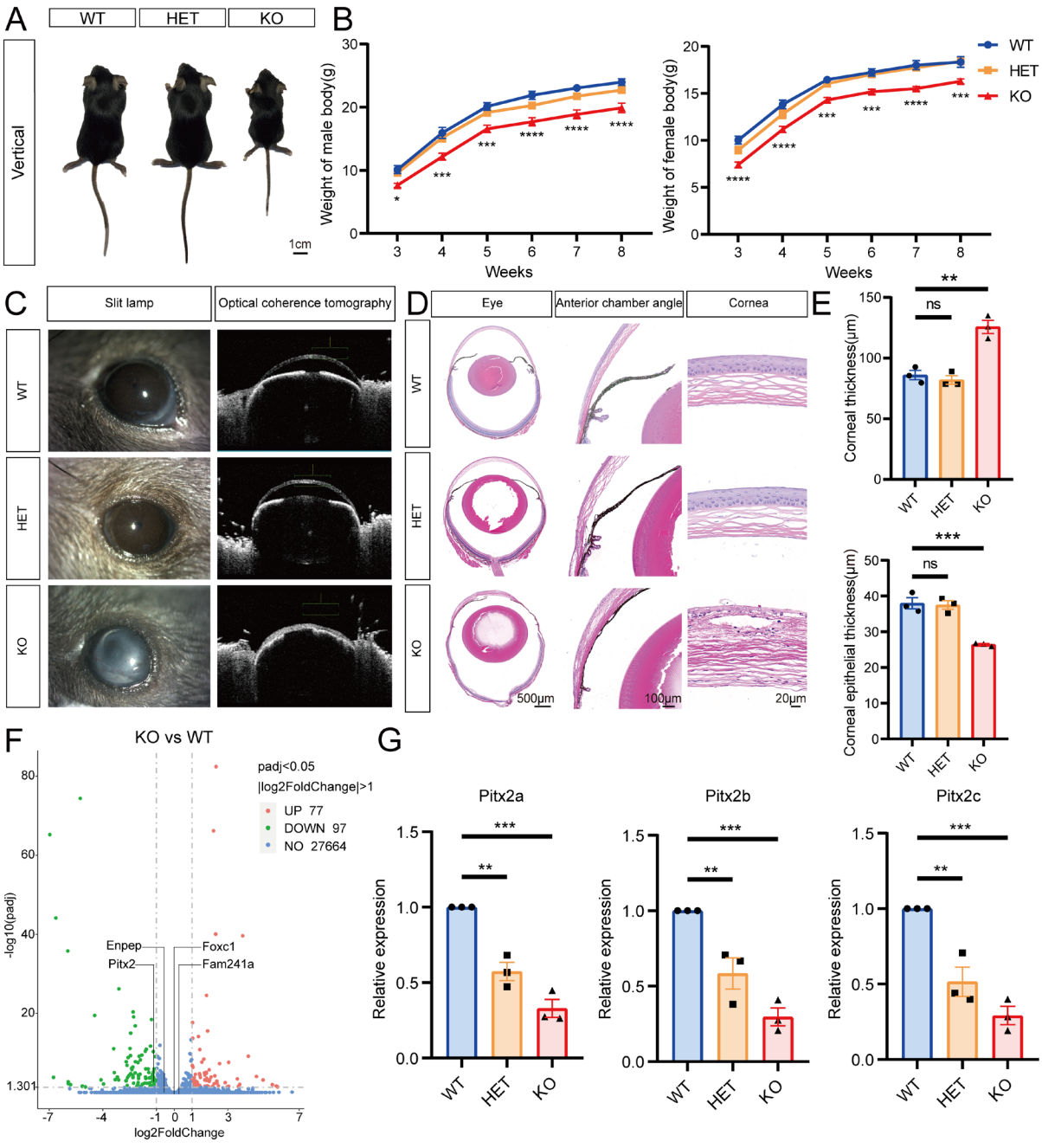

A
WT
HET
KO
Vertical
1cm
B
Weight of male body(g)
Weeks
Weight of female body(g)
Weeks
WT
HET
KO
C
Slit lamp
Optical coherence tomography
WT
HET
KO
D
Eye
Anterior chamber angle
Cornea
500μm
100μm
20μm
E
Corneal thickness(μm)
Corneal epithelial thickness(μm)
ns
F
KO vs WT
padj<0.05
|log2FoldChange|>1
UP 77
DOWN 97
NO 27664
Enpep
Pitx2
Foxc1
Fam241a
-log10(padj)
log2FoldChange
G
Pitx2a
Pitx2b
Pitx2c
Relative expression

Figure 2. Phenotypic Observations and Expression Alterations of ARS-Related Genes in LOH-1 Knockout Mice. (**A**) Size of male LOH-$1^{-/-}$(KO), LOH-$1^{+/-}$(HET) and wild-type (WT) mice at 3 weeks of age. The scale bar represents 1 cm. (**B**) Body weight of male (left) and female (right) KO mice was compared to HET and WT mice from 3 to 8 weeks of age. n=23 for each group. Statistic data were analyzed using two-way ANOVA. (**C**) Slit lamp and optical coherence tomography examinations of the eyes from different groups showed normal anterior segment structure in WT and HET groups, while opacified cornea and disappeared anterior chamber was seen in the KO group. (**D**) HE staining and histopathologic findings demonstrated normal eyeball structures, wide angles and regularly arranged corneal layers in WT and HET groups. However, a disorganized anterior segment, closed angle as well as thick corneal stromal layer were seen in the KO group. Scale bars represent 500μm in eye, 100μm in anterior chamber angle and 20μm in cornea, respectively. (**E**)The whole corneal thickness and corneal epithelial thickness were compared among the groups. n=3 for each group. Statistic data were analyzed using one-way ANOVA. (**F**) Volcano plot displaying the differentially expressed genes (DEGs) (KO VS WT) with 97 downregulated genes and 77 upregulated genes (|log2(FoldChange)| >= 1 and padj<= 0.05). Among them, the expression of *Pitx2* was significantly reduced. However, the expression levels of *Foxc1*, *Enpep* and *Fam241a* were not significantly changed. (**G**) RT-qPCR detection of relative *Pitx2a*, *Pitx2b* and *Pitx2c* mRNA expression in the WT, HET and KO mice. n=3 for each group. Statistic data were analyzed using one-way ANOVA. All data are represented as mean ± SEM. *p<0.05, **p<0.01, ***p<0.001, ****P<0.0001, ns, not significant.

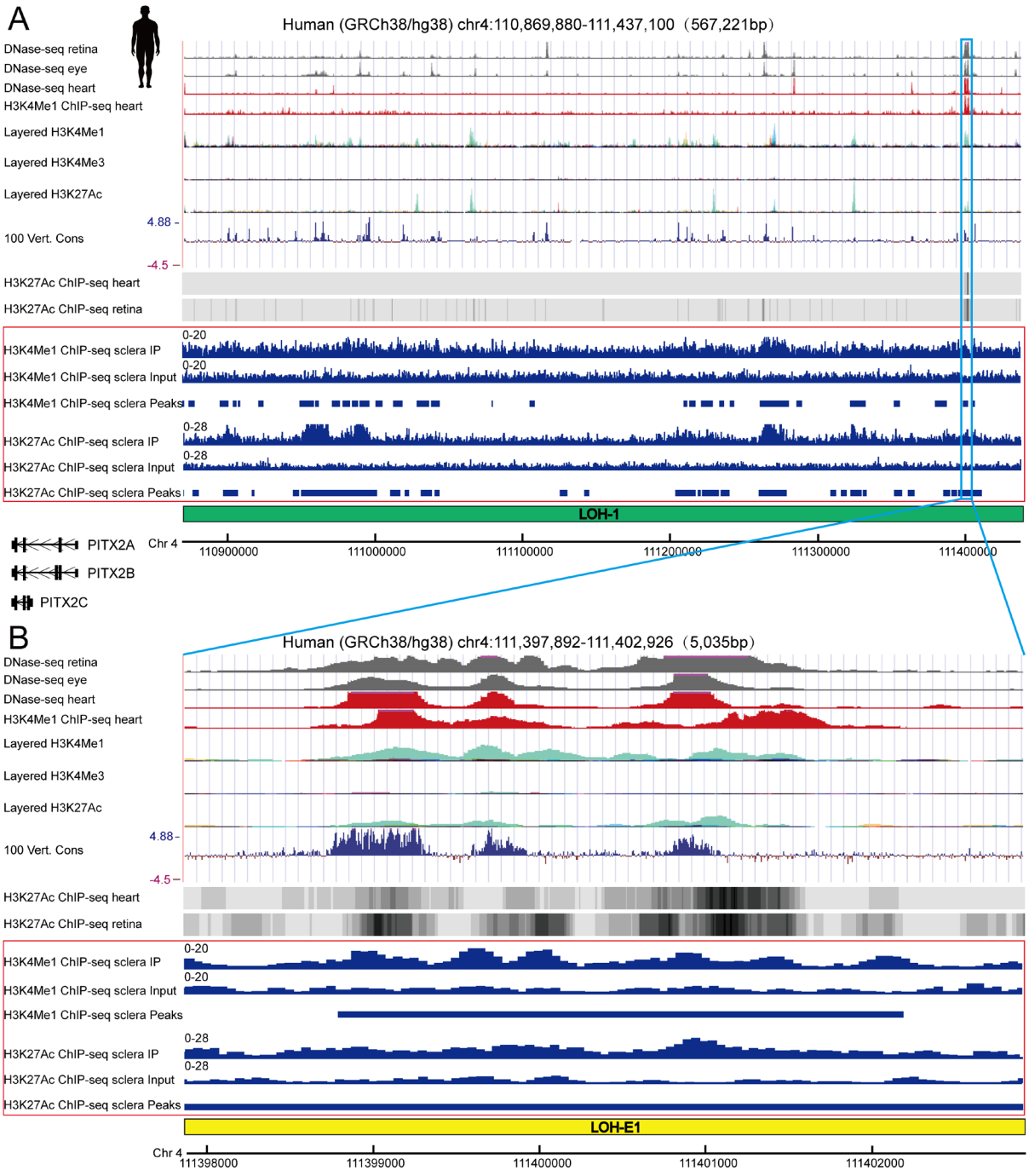

A
Human (GRCh38/hg38) chr4:110,869,880-111,437,100 (567,221bp)
DNase-seq retina
DNase-seq eye
DNase-seq heart
H3K4Me1 ChIP-seq heart
Layered H3K4Me1
Layered H3K4Me3
Layered H3K27Ac
4.88
100 Vert. Cons
-4.5
H3K27Ac ChIP-seq heart
H3K27Ac ChIP-seq retina
0-20
H3K4Me1 ChIP-seq sclera IP
0-20
H3K4Me1 ChIP-seq sclera Input
H3K4Me1 ChIP-seq sclera Peaks
0-28
H3K27Ac ChIP-seq sclera IP
0-28
H3K27Ac ChIP-seq sclera Input
H3K27Ac ChIP-seq sclera Peaks
LOH-1
PITX2A
PITX2B
PITX2C
Chr 4
110900000
111000000
111100000
111200000
111300000
111400000
B
Human (GRCh38/hg38) chr4:111,397,892-111,402,926 (5,035bp)
DNase-seq retina
DNase-seq eye
DNase-seq heart
H3K4Me1 ChIP-seq heart
Layered H3K4Me1
Layered H3K4Me3
Layered H3K27Ac
4.88
100 Vert. Cons
-4.5
H3K27Ac ChIP-seq heart
H3K27Ac ChIP-seq retina
0-20
H3K4Me1 ChIP-seq sclera IP
0-20
H3K4Me1 ChIP-seq sclera Input
H3K4Me1 ChIP-seq sclera Peaks
0-28
H3K27Ac ChIP-seq sclera IP
0-28
H3K27Ac ChIP-seq sclera Input
H3K27Ac ChIP-seq sclera Peaks
LOH-E1
Chr 4
111398000
111399000
111400000
111401000
111402000

Figure 3. Identification of the enhancer LOH-E1 within LOH-1 region. (**A**) Overview of the DNase-seq of retina, eye and heart, the H3K4Me1 ChIP-seq of heart, the H3K4Me1, H3K4Me3 and H3K27Ac marks on 7 cell lines, the 100 vertebrates Basewise Conservation, the H3K27Ac ChIP-seq of heart and retina in LOH-1 region of human supported from the UCSC genome browser. H3K4Me1 and H3K27Ac ChIP-seq results from human embryonic 15 weeks sclera are indicated by the red box. LOH-1 is indicated by the green rectangle. *PITX2* gene is shown to be downstream of the LOH-1. (**B**) Determination of the LOH-E1 enhancer location based on markers of DNase I, H3K4Me1 and H3K27Ac. LOH-E1 is indicated by the yellow rectangle. H3K4Me1 and H3K27Ac ChIP-seq results from human embryonic 15 weeks sclera are indicated by the red box.

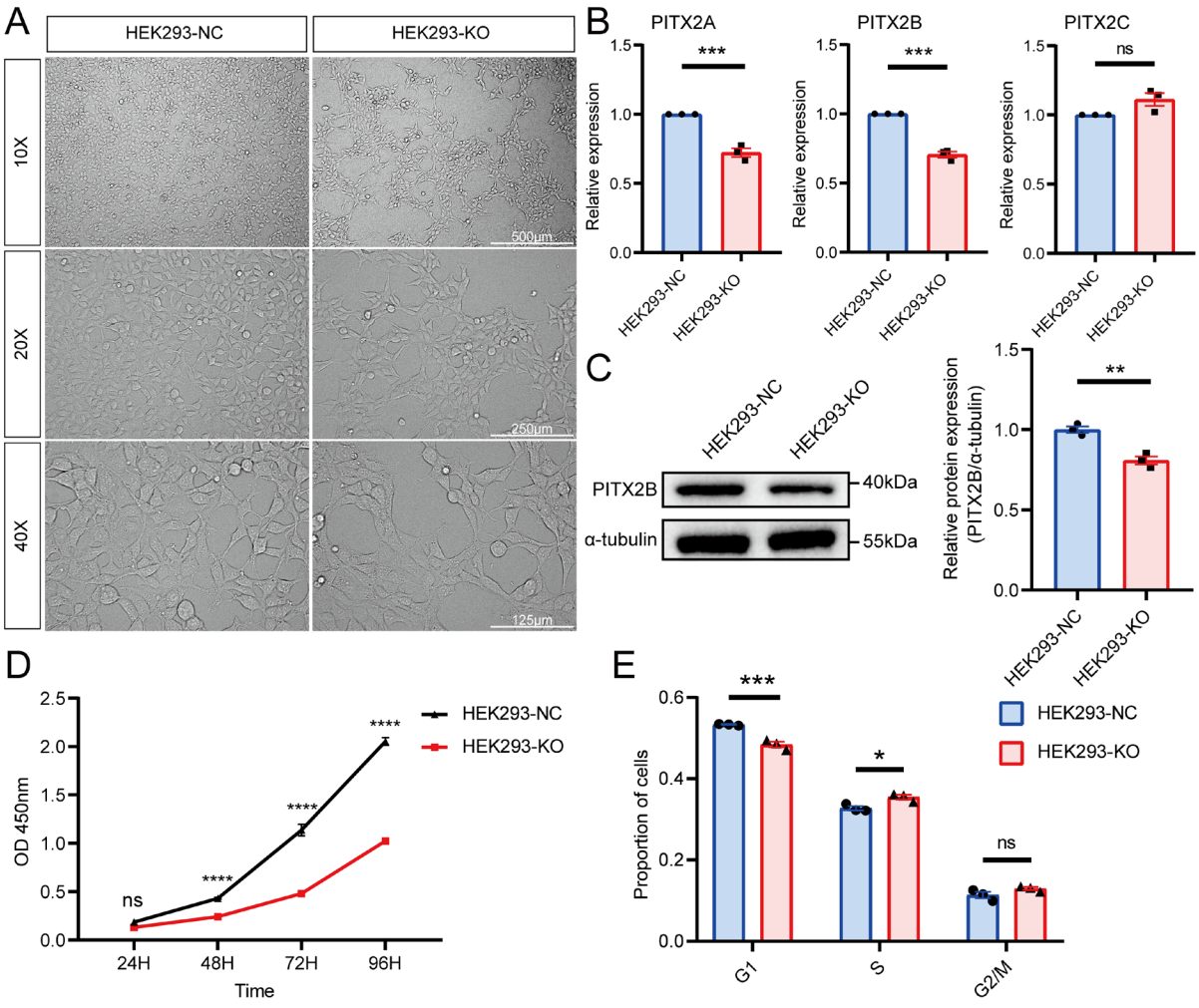

A
HEK293-NC
HEK293-KO
10X
20X
40X
500μm
250μm
125μm
B
PITX2A
PITX2B
PITX2C
***
***
ns
Relative expression
HEK293-NC
HEK293-KO
C
PITX2B
α-tubulin
40kDa
55kDa
Relative protein expression (PITX2B/α-tubulin)
**
D
OD 450nm
Time
24H
48H
72H
96H
ns
****
****
****
HEK293-NC
HEK293-KO
E
Proportion of cells
***
*
ns
G1
S
G2/M
HEK293-NC
HEK293-KO

Figure 4. CRISPR/Cas9-mediated deletion of LOH-E1 in HEK293 cells. (**A**) The typical morphology of the HEK293-NC (negative control) and HEK293-KO (deletion of LOH-E1) under 10×, 20× or 40× magnification, respectively. Scale bars represent 500μm in 10×, 250μm in 20× and 125μm in 40×. (**B**) RT-qPCR detection of relative *PITX2A*, *PITX2B* and *PITX2C* mRNA expression levels in the HEK293-NC and HEK293-KO cells. n=3 for each group. Statistic data were analyzed using two-tailed Student's T-test. (**C**) Western blot analysis of PITX2B in HEK293-KO cells and the control. And relative protein quantification of grey scale value for PITX2B and α-tubulin. n=3 for each group. Statistic data were analyzed using two-tailed Student's T-test. (**D**) CCK-8 assay in HEK293-KO cells and the control. n=3 for each group. Statistic data were analyzed using two-way ANOVA. (**E**) Cell cycle analysis of HEK293-KO cells and the control. n=3 for each group. Statistic data were analyzed using two-way ANOVA. All data are represented as mean ± SEM. *p<0.05, **p<0.01, ***p<0.001, ****P<0.0001, ns, not significant.

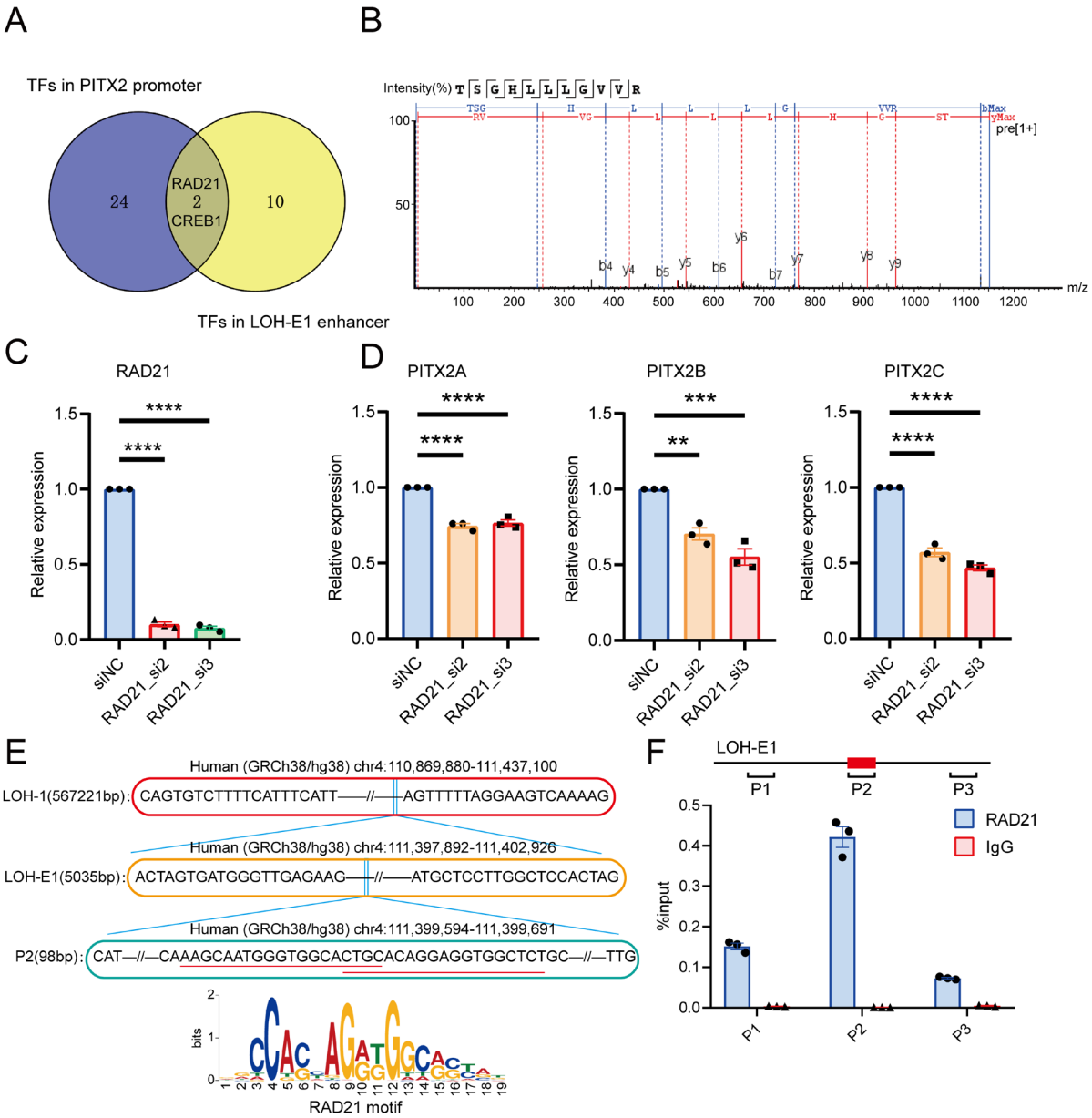

A
TFs in PITX2 promoter
24
RAD21
2
CREB1
10
TFs in LOH-E1 enhancer
B
Intensity(%)
pre[1+]
m/z
C
RAD21
Relative expression
siNC
RAD21_si2
RAD21_si3
D
PITX2A
PITX2B
PITX2C
E
Human (GRCh38/hg38) chr4:110,869,880-111,437,100
LOH-1(567221bp):
CAGTGTCTTTTCATTTCATT——//——AGTTTTTAGGAAGTCAAAAG
Human (GRCh38/hg38) chr4:111,397,892-111,402,926
LOH-E1(5035bp):
ACTAGTGATGGGTTGAGAAG——//——ATGCTCCTTGGCTCCACTAG
Human (GRCh38/hg38) chr4:111,399,594-111,399,691
P2(98bp):
CAT—//—CAAAGCAATGGGTGGCACTGCACAGGAGGTGGCTCTGC—//—TTG
bits
RAD21 motif
F
LOH-E1
P1
P2
P3
RAD21
IgG
%input

Figure 5. P2 which is within LOH-E1 could bind RAD21 to regulate the expression of *PITX2*. (**A**) Veen diagram presents transcription factors that bind to both the LOH-E1 enhancer region and the PITX2 core promoter region. (**B**) Peptide Spectrum Match (PSM) of peptide from RAD21 protein by Liquid Chromatograph Mass Spectrometer (LC-MS) in the DNA pulldown. (**C**) RT-qPCR detection of *RAD21* expression following its downregulation in HEK293 cells through siRNAs. For detection of *RAD21* expression level, *RAD21*_small interfering RNA (RAD21_si)2 or RAD21_si3 compared to small interfering RNA negative control (siNC), respectively. n=3 for each group. Statistic data were analyzed using one-way ANOVA. (**D**) RT-qPCR detection of *PITX2A*, *PITX2B* and *PITX2C* expressions after the downregulation of *RAD21* in HEK293 cells through RAD21_si2 and RAD21_si3. n=3 for each group. Statistic data were analyzed using one-way ANOVA. (**E**) Identification of the potential RAD21 binding sequence within the LOH-E1 region. And the position of the red line can match the RAD21 motif. (**F**) ChIP-qPCR assay for the binding of RAD21 to the P2 and surrounding regions. qPCR detection for the enrichment of the P1, P2 and P3 regions within LOH-E1 upon anti-RAD21 ChIP and IgG in the HEK293 cell line. P1 and P3 are the upstream and downstream random regions of P2. JJ is indicated by the red rectangle. n=3 for each group. All data are represented as mean ± SEM.