PRMT1-SFPQ regulates intron retention to control matrix gene expression during craniofacial development

Julia Raulino Lima[1,#], Nicha Ungvijanpunya[1,2,#], Qing Chen[3,#], Greg Park[1], Mohammadreza Vatankhah[1], Tal Rosen[1], Yang Chai[1], Amy Merrill-Brugger[1], Weiqun Peng[3,*], Jian Xu[1,4,*]

[1]Center for Craniofacial Molecular Biology, Herman Ostrow School of Dentistry, University of Southern California, Los Angeles, CA, USA; [2]Department of Orthodontics, Faculty of Dentistry, Chulalongkorn University, Bangkok, Thailand; [3]Department of Physics, The George Washington University, USA; [4]Department of Biochemistry and Molecular Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.

# These authors contributed equally to this work.

[*] To whom correspondence should be addressed: Jian Xu: Center for Craniofacial Molecular Biology, University of Southern California, Los Angeles, CA 90033; xujian@usc.edu. Weiqun Peng: Department of Physics, The George Washington University, Washington, DC 20052; wpeng@gwu.edu

RUNNING TITLE: **PRMT1-SFPQ REGULATES INTRON RETENTION IN CRANIOFACIAL DEVELOPMENT.**

ABSTRACT

Spliceosomopathies, which are a group of disorders caused by defects in the splicing machinery, frequently affect the craniofacial skeleton and limb, but the molecular mechanism underlying this tissue-specific sensitivity remains unclear. Splicing factors and small nuclear ribonucleoproteins (snRNPs) are core components of splicing machinery, and splicing factors are further controlled by post-translational modifications, among which arginine methylation is one of the most frequent modifications. To determine the splicing mechanisms in cranial neural crest cells (CNCCs), which give rise to the majority of the craniofacial skeleton, we focused on upstream regulators for splicing proteins responsible for arginine methylation, protein arginine methyltransferases (PRMT). These enzymes catalyze arginine methylation of splicing factors to modify splicing factor expression and activity, influencing the splicing product. PRMT1 is the highest expressing enzyme of the PRMT family in CNCCs and its role in craniofacial development is evident based on our earlier investigation, where CNCC-specific *Prmt1* deletion caused cleft palate and mandibular hypoplasia. In the present study, we uncover the roles of PRMT1 in CNCCs in the regulation of intron retention, a type of alternative splicing where introns are retained in the mature mRNA sequence. Mandibular primordium of *Prmt1*-deficient embryos demonstrated an increase in the percentage of intron-retaining mRNA of matrix genes, which triggered NMD, causing a reduction in matrix transcript expression. We further identified SFPQ as a substrate of PRMT1 that depends on PRMT1 for arginine methylation and protein expression in the developing craniofacial structures. Depletion of SFPQ in CNCCs phenocopied PRMT1 deletion in that matrix, Wnt signaling components and neuronal gene transcripts contained higher IR and exhibited lower expression. We further recognized gene length as a common feature among SFPQ-regulated genes in CNCCs. Altogether, these findings demonstrate that the PRMT1-SFPQ pathway modulates matrix Wnt signaling components and neuronal gene expression via intron retention in CNCCs during craniofacial development.

INTRODUCTION

Craniofacial abnormalities affecting bone formation in the skull and face are the most common birth defects in infants. Proper formation of these structures involves coordinated processes including proliferation, migration, and differentiation of cranial neural crest cells (CNCCs) (Martik and Bronner 2017; Plein, Fantin, and Ruhrberg 2015). CNCCs are a transient population of progenitor cells that populate the first and second pharyngeal arches and give rise to the maxilla, mandible, palates, and other structures of the developing head (Chai et al. 2000; Jiang et al. 2002; Martik and Bronner 2017). Dysregulation of transcription factors, chromatin remodelers, RNA regulatory proteins, and signaling molecules have been implicated in impaired neural crest development that results in craniofacial defects (Bélanger et al. 2018; Martik and Bronner 2017; Plein, Fantin, and Ruhrberg 2015; Strobl-Mazzulla, Marini, and Buzzi 2012). Spliceosomopathies, which are a group of disorders caused by defects in the splicing machinery, frequently affect the craniofacial skeleton and limb. However, why these tissues are more sensitive to spliceosomal defects remain incompletely understood. Splicing factors and small nuclear ribonucleoproteins (snRNPs) are core components of splicing machinery, which control pre-mRNA splicing and mRNA maturation through the process of alternative splicing (AS). AS includes seven types of events, exon skipping, mutually exclusive exons, alternative 5' and 3'-splice sites, alternative promoters, alternative polyadenylation, and intron retention (Hooper et al. 2020). Intron retention (IR) is a type of AS abundant in plants and fungi, where introns are retained within the mature mRNA sequence (E.T. Wang et al. 2008). Although initially overlooked in mammals due to low occurrence and challenges in detection, recent technical advances have unveiled its significant roles in various aspects, including cell differentiation and stress response, especially in neuronal, immune, and cancer cells (Marquez et al. 2012; Monteuuis et al. 2019; Braunschweig et al. 2014; Wong and Schmitz 2022).

In exploring the splicing mechanisms in CNCCs, we focused on upstream regulators for splicing factors. The activity of splicing factors is tightly controlled by post-translational modifications, among which methylation is the most abundant, surpassing phosphorylation (Thandapani et al. 2013; Ruta, Pagliarini, and Sette 2021). Splicing regulator methylation mostly occurs on arginine motifs, catalyzed by the protein

arginine methyltransferase (PRMT) family of enzymes. PRMTs-mediated methylation of splicing factors determines their subcellular localization, expression, and activity, influencing the splicing product (Blanc and Richard 2017; Liu and Dreyfuss 1995; Snijders et al. 2015). PRMT1 is the most abundant PRMT in most cell types and specifically catalyzes methylation of arginine (R) residues within RG/RGG/GAR repeats, which are motifs enriched in RNA-binding proteins, particularly splicing regulators, resulting in asymmetric dimethylation of arginine (ADMA) (Thandapani et al. 2013; Smith et al. 2004). The importance of PRMT1 in craniofacial development is evident from our earlier investigation in which genetic deletion of *Prmt1* in CNCCs leads to cleft palate and mandibular hypoplasia (Gou, Li, Wu, et al. 2018; Gou, Li, Jackson-Weaver, et al. 2018). In the present study, we uncovered a previously unrecognized role of PRMT1 in regulating intron retention (IR) in CNCCs. The mandibular primordium of *Prmt1*-deficient embryos demonstrated augmented retention of introns in mRNAs that encode bone and cartilage matrix components. We further identified SFPQ as a downstream mediator of PRMT1 in the developing craniofacial structures, that depends on PRMT1 for arginine methylation and protein expression. Depletion of SFPQ in CNCCs partially phenocopied PRMT1 deletion, causing elevated retention of introns in mRNA transcripts of matrix, neuronal, and Wnt signaling pathways genes. We further recognized long gene/intron splicing as a common feature among SFPQ-regulated introns and demonstrated these retained introns contain premature termination codons (PTCs) that trigger nonsense-mediate decay (NMD) to reduce mRNA expression levels. Together, these findings demonstrate that the PRMT1-SFPQ pathway modulates CNCC gene expression via control of splicing. Deficiency of this pathway leads to aberrant intron retention, which induces mRNA decay that downregulates matrix expression in CNCCs during craniofacial development.

RESULTS

CNCCs from the embryonic mandibular process exhibit abundant intron retention, which was further elevated by loss of *Prmt1*.

PRMT1 is the most abundant enzyme of the arginine methyltransferase family in cranial neural crest cells (CNCCs), with an expression level 10~100 folds higher than other PRMTs (Figure-1A). We previously demonstrated critical roles for PRMT1 in CNCCs during craniofacial skeleton formation using genetic

deletion, whereby neural crest-specific deletion of *Prmt1* resulted in cleft palate and shorter mandibles (Gou, Li, Wu, et al. 2018; Gou, Li, Jackson-Weaver, et al. 2018). Shorter mandibles, or mandibular hypoplasia is a craniofacial phenotype shared by most spliceosomopathies. To investigate the molecular mechanisms underlying PRMT1's role in mandible development, we conducted a genome-wide analysis of *Prmt1* deletion-induced transcriptional and splicing alterations within the embryonic mandibles. For this purpose, CNCCs, the cell lineage that gives rise to the bone and cartilages in the mandible were isolated from mandibular processes of control (*Wnt1-Cre; Rosa*^{Td}) or *Prmt1 CKO* mutant (*Wnt1-Cre; Prmt1*$^{fl/fl}$*; Rosa*^{Td}) mice at embryonic day 13.5 (E13.5) for mRNA extraction and sequencing (Figure-1B). The ensuing comparative analysis using rMATS (Y. Wang et al. 2024) unveiled changes across five types of alternative splicing events (Figure-1C), echoing our earlier finding where PRMT1 regulates exon usage during cardiac morphogenesis (Jackson-Weaver et al. 2020). In contrast to embryonic epicardial cells, we noted a higher abundance of altered intron retention (IR) in mandibular CNCCs (Figure-1C), with a predominance of IR increase when compared to control CNCCs (Figure-1Ca'). The IR changes were further exemplified by the elevation of intronic expression in *Pex12, Nkx3-2, Nkiras1, 1810026B05Rik,* and the reduction of intronic expression in *Tbx1* (Figure-1D-F). To analyze the intronic expression comprehensively at a gene-specific level, we employed IRI (intron reads index) analysis, which is an algorithm that quantifies the ratio of normalized read counts between intronic and exonic regions (Ni et al. 2016; Sun et al. 2023; Tian et al. 2020). Subsequent analysis for retained introns demonstrated that *Prmt1* deficiency-induced change in intron retention altered an array of biology pathways. We noted protein binding, metal ion binding, and cartilage development among genes that present increased intronic expression, while mitochondrion and metabolic pathways were identified among genes with decreased intronic expression (Supplemental Table S1). These findings identify intron retention as a prevalent phenomenon in CNCCs during craniofacial development, and the deletion of *Prmt1* further enhanced intron retention in genes involved in diverse biological processes.

Neural crest-specific deletion of *Prmt1* caused a significant reduction of matrix gene expression in the developing mandibles.

The main documented roles of retained intron in mature mRNA is to reduce gene expression via non-sense mediated decay (NMD) (Schmitz et al. 2017). When intron-retaining transcripts are incorporated into the translation machinery, introns that contain premature termination codons (PTCs) within their open reading frames trigger NMD, leading to transcript degradation and reduced expression of genes harboring these introns (Wong et al. 2016). We observed that intron-retaining transcripts illustrated in Figure-1, *Pex12,* and *Nkiras1* demonstrated increased intronic expression and decreased exonic expression, suggesting reduced gene expression. To systematically analyze genes that are differentially downregulated, we compared the transcriptional profile between control and *Prmt1* CKO mutant of mandibular CNCCs and revealed downregulation of 303 genes and concurrent upregulation of 160 genes by *Prmt1* deletion (Figure- 2A, 2B). Upon pathway analysis, glycosaminoglycan (GAG) degradation (35 out of 303 genes) and extracellular matrix (ECM) (33 out of 303 genes) emerged as the top pathways among downregulated genes (Figure-2C). Findings from GAG degradation and ECM genes align with our understanding that these downregulated genes encode matrix proteins pivotal to osteogenic and chondrogenic differentiation and the formation of bone and cartilage matrix. This notion was supported by bioinformatic analysis using Ingenuity Pathway Analysis (IPA) based on downregulated genes, which suggested a disruption in bone, cartilage, and connective tissue development within *Prmt1* CKO mutant mandibles (Figure-2D). The data collected from transcriptional analysis underscore the importance of PRMT1 in orchestrating matrix gene expression critical to bone, cartilage, and connective tissue formation during mandible morphogenesis.

Deletion of *Prmt1* in CNCCs elevated intron retention in mRNAs that encode ECM and GAG degradation enzymes.

To determine whether the downregulation of matrix genes is regulated by intron retention, we first examined the intronic expression profile of ECM genes. A prominent elevation of intron retention within ECM genes was revealed by scatter plot where most ECM data points landed above the red line that denotes unchanged intron retention (Figure-3A). ECM genes with the most significant increase in intron expression were labeled and highlighted as red dots, and further exemplified by the track view of *Bmp7* and *Cthrc1* (Figure-3A, 3B). In these ECM genes, the increased intron retention correlated with the reduced exonic

expression, indicated by intronic expression and exonic expression respectively (Figure-3C). To validate the reproducibility of elevated intron retention in ECM transcripts, we expanded our analysis to incorporate four additional embryos from each group, isolated mature mRNA using poly(A)+ magnetic beads, and designed primers for RT-PCR analysis of intronic and exonic region of additional ECM genes. In line with our initial observations, intronic expression increased significantly, coupled with decreased exonic expression in *Dcn, Tnn, Lox, Loxl1, Matn2, Col14a1, Adamts12, Adamts2, Adam12,* and *Scara5* of the mutant CNCCs as compared to the control (Figure-3D). These findings demonstrate that ECM transcripts in *Prmt1* CKO mandibles exhibited elevated intron retention, which correlates with a reduction in mRNA expression levels. Next, we analyzed the intronic reads profile of GAG degradation genes, which is the top downregulated gene cluster in the *Prmt1* CKO group and clinically associated with craniofacial birth defects. Similar to ECM genes, the majority of GAG degradation genes showed elevation of intron retention indicated by scatter plot (Figure-3E). The increased intron retention of *St6galnac3* and *Galn11*, indicated by higher intronic expression, correlated with reduced mRNA expression, indicated by lower mRNA expression (Figure-3F, 3G). Upon further inspection of the retained introns in ECM and GAG degradation transcripts, retained introns in the significantly downregulated genes all contain PTCs within their open reading frames, suggesting PTC-triggered transcript degradation via NMD-mediated RNA surveillance system (Supplemental Table S2) (Wong et al. 2016). Altogether, these data indicate that *Prmt1* deletion elevates intron retention in matrix genes encoding ECM proteins and GAG degradation enzymes and downregulated their mRNA expression within CNCCs.

PRMT1 methylates SFPQ in CNCCs.

PRMT1 is a methyltransferase that catalyzes methylation of arginine (R) residues within RG/RGG/GAR repeats, which are motifs highly enriched in RNA-binding proteins, particularly splicing regulators (Thandapani et al. 2013). PRMT1 catalyzes asymmetric dimethylation of arginine (ADMA) residues, which regulates the turnover, subcellular localization, and splicing activity of splicing regulators (Smith et al. 2004). We therefore hypothesized that PRMT1 regulates intron retention via methylation of splicing regulators. Earlier studies by Graham et al. and our team have characterized the landscape of arginine

methylation governed by PRMT1 and identified a cohort of RNA processing proteins exhibiting PRMT1-dependent methylation (Hartel et al. 2019). Among these, six splicing regulators are highly expressed in embryonic CNCCs: SFPQ, SRSF1, EWSR1, TAF15, TRA2B, and G3BP1 (Figure-4A). To discern whether any of the splicing regulators are accountable for elevated IR in *Prmt1* deficient CNCCs, we conducted an in-silico analysis of the RNA-seq data of mandibular CNCCs from control and *Prmt1 CKO* embryos utilizing rMATS (multivariant analysis of transcript splicing, MATS), followed by rMAPS2 (RNA map analysis) (Hwang et al. 2020; Y. Wang et al. 2024). This integrated approach detects differential splicing events, and then generates visual maps illustrating the spatial distribution of RNA binding protein (RBP) motifs, and finally ranks RBPs based on motif density score and p-value (Park et al. 2016). Through this in-silico analysis, we identified a subset of RBPs with significantly differential binding in splicing events altered by PRMT1 ($p<0.00001$) (Supplemental Table S3). To pinpoint the most pertinent RBPs, we cross-referenced this significant RBP list with the previously identified six PRMT1-dependent and CNCC-enriched splicing regulators (Figure-4A). SFPQ emerged as the prime candidate, boasting the highest motif scores and the lowest P-values (Supplemental Figure-S1A-E, Supplemental Table S4).

Subsequently, we set out to determine whether SFPQ is methylated by PRMT1 in the craniofacial structures. Employing the proximity ligation assay (PLA), a technique designed for the in-situ detection of protein modifications, we probed the E13.5 embryo sections with anti-SFPQ and anti-methyl-arginine ADMA antibodies utilizing the PLA kit, which revealed the presence of methylated SFPQ as distinct green punctate staining. Within the craniofacial complex, a substantial presence of SFPQ methylation was observed in the mandibular processes at E13.5 (Figure-4B, 4Ca-d). In *Prmt1* CKO embryos, the signal of methyl-SFPQ showed marked diminishment within the CNCC-derived mesenchyme region (Figure-4Ce-h, 4D). The same analysis was conducted in the maxillary regions, which demonstrated robust SFPQ methylation in the control group and a dramatic reduction of methylation in *Prmt1* CKO embryos (Figure-4B, 4Ea-h, 4F). This CNCC-specific deletion of PRMT1 only reduced SFPQ methylation in the CNCCs, as non-CNCC lineages within the epithelium continued to display methyl-SFPQ signals (Figure-4Ga-f,

Supplemental Figure-S2). These findings establish SFPQ's methylation within the craniofacial complex as a PRMT1-dependent phenomenon.

PRMT1-induced methylation controls the activity of splicing regulators by modifying their stability, subcellular distribution, or RNA binding ability (Rho et al. 2007). We then proceeded to examine whether SFPQ protein expression was regulated by PRMT1 in the craniofacial complex. In control embryos, SFPQ exhibited prominent expression in CNCCs and mainly localized in the nucleus (Figure-4Ha-c, 4Ja-c). In *Prmt1*-deficient CNCCs, the protein's expression level witnessed a significant reduction within the mandibular and maxillary processes (Figure-4H-4K). A corroborating Western blot (WB) analysis using tissues from the craniofacial structures, including facial and anterior skull regions of embryonic heads, confirmed this decline in SFPQ expression (Figure-4L, 4M). Taken together, these data demonstrate that PRMT1 methylates SFPQ and promotes SFPQ protein expression in the craniofacial complex.

PRMT1-SFPQ pathway regulates the splicing of matrix genes in CNCCs.

To investigate the role of SFPQ in intron retention and gene expression within CNCCs, we isolated primary CNCCs from control (*Wnt1-Cre; Td*) embryos at E13.5 and enriched CNCCs by Td+ signal using cell sorting. Within a day, Td+ CNCCs were transfected with control or two independent SFPQ siRNAs to deplete SFPQ, and then poly(A)+ mRNA was purified followed by sequencing. To determine whether these primary isolated CNCCs retained CNCC characteristics after siRNA-mediated knockdown, the expression of CNCC markers *Twist1, Sox10, Msx1, Snai2,* and *Tfap2a* was examined (Achilleos and Trainor 2012; Ishii et al. 2012), which demonstrates that the CNCC marker expression in siRNA-transfected CNCCs after isolation from E13.5 embryos is comparable to that in fresh isolated CNCCs from E13.5 or E15.5 embryos (Supplemental Figure-S3), validating CNCC identity and supporting this method as a feasible approach for mechanistic study of CNCCs using developmental stage-specific embryos. Subsequently, using sequencing data from these siRNA-transfected CNCCs, we conducted bioinformatic analyses of differentially retained introns and differentially expressed genes between control and *Sfpq*-depleted CNCCs (Figure-5A, 5B). Genes with differential intron retention were assessed by IRI analysis, which revealed elevated intron retention in 397 (siSFPQ #1 vs. siCont) or 388 (siSFPQ #2 vs. siCont) genes and reduced intron retention

in 129 (siSFPQ #1 vs. siCont) or 151 (siSFPQ #2 vs. siCont) genes caused by *Sfpq* depletion (Figure-5A). Genes showing increased IR in *Sfpq* depleted CNCCs encompassed "proteoglycan metabolic process" and "proteoglycan biosynthesis process", which included GAG degradation genes, and "connective tissue development", which includes ECM genes (Figure-5Ca, 5Cb). *Sfpq* depletion also led to downregulation of 269 (siSFPQ #1 vs. siCont) or 232 (siSFPQ #2 vs. siCont) genes, among which 81 (siSFPQ #1 vs. siCont) and 60 (siSFPQ #2 vs. siCont) genes overlapped with elevated IR, respectively (Figure-5B). To determine intron retention co-regulated by SFPQ and PRMT1, we cross-analyzed genes with increased intronic expression in *Sfpq*-depleted CNCCs and *Prmt1*-deficient CNCCs and uncovered a partial overlap of a total of 64 genes (Figure-5D), represented by matrix genes *Col4a2, Adam12, Ntn1, App, St6galnac3, Galnt11 Galnt10,* and *Asph* (Figure-5E, 5F). The elevation of intronic retention in these genes also correlated with a reduction of gene expression (Figure-5G, 5H). GO analysis of differentially downregulated genes by both SFPQ-specific siRNAs (cluster 4 within the heatmap) further confirmed extracellular matrix, extracellular structure, and adhesion as the top categories (Figure-5I, 5J). These data indicate that depletion of SFPQ is sufficient to enhance IR and reduce gene expression for a cluster of matrix genes, suggesting loss of SFPQ as a regulatory mechanism for these genes in *Prmt1*-deficient CNCCs and SFPQ as a downstream effector for PRMT1 in CNCCs.

SFPQ depletion in CNCC enhanced intron retention in Wnt signaling and neuronal transcripts and reduced their mRNA expression.

More intriguingly, Wnt signaling emerged as the top pathway regulated by SFPQ in CNCCs (Figure-5C). *Sfpq* depletion increased the percentage of intron-retaining mRNAs of genes in this cluster and reduced total mRNA expression (Figure-6A). These Wnt signaling components-regulated by SFPQ encode both positive and negative regulators of Wnt signaling and encompass both canonical and non-canonical Wnt signal regulators. The SFPQ-regulated Wnt signaling components *Wwox* also exhibited increased IR and decreased expression in *Prmt1 CKO* CNCCs. Additionally, neuronal genes occupy a big fraction among SFPQ-regulated introns and genes in CNCCs, as "axon guidance", "axonogenesis", "neuron projection development", and "leaning or memory" emerged as the top biological processes among differentially

upregulated introns (Figure-5Ca, 5Cb), and "regulation of membrane potential", "learning or memory" "cognition" and "locomotory behavior" denotes differentially downregulated genes in *Sfpq*-depleted CNCCs (Figure-5I, 5J). In CNCCs, *Sfpq* depletion increased the percentage of intron-retaining mRNAs of these neuronal genes and reduced total mRNA expression (Figure-6B). Upon further inspection, "regulation of membrane potential" was also among identified among genes with decreased expression (38 out of 303 genes) in *Prmt1 CKO* CNCCs. Altogether, these findings demonstrate that, besides matrix genes, SFPQ regulates Wnt signaling components and neuronal gene expression via intron retention.

Disturbance of SFPQ activity impaired splicing of long genes and long introns.

SFPQ predominantly binds to intronic regions and regulates splicing (Hosokawa et al. 2019). We therefore used published CLIP-seq dataset generated from embryonic mouse brain to determine whether these aberrantly retained introns are SFPQ targets (Hosokawa et al. 2019). The retained *Ptk7* intron 1 encompassed a much higher density of SFPQ binding peaks when compared to the spliced intron 6 within the same gene (Figure-7A). We further conducted analysis for all the introns at the genomic level comparing SFPQ binding peaks among introns with elevated retention to introns with reduced retention or no changes and demonstrated that introns with elevated retention in *Sfpq*-depleted CNCCs showed significantly higher enrichment of SFPQ binding peaks (Figure-7B). We further analyzed SFPQ CLIP-seq peaks at the gene level, comparing genes with elevated IR to genes with no IR changes. Our analysis showed that genes containing elevated IR in *Sfpq*-depleted CNCCs exhibited a trend towards higher presence of SFPQ binding peaks based on this embryonic brain dataset (Figure-7C). These data suggest that introns with aberrant IR elevation in *Sfpq*-depleted CNCCs are direct splicing targets bound by SFPQ.

To understand the molecular basis of how SFPQ targets these matrix, Wnt and neuronal genes and introns, we evaluated characteristics shared among SFPQ-regulated genes, and recognized that they are either long genes (>100kd in length) or retain long introns (>10kd in length). Of note, Wnt pathway gene *Wwox* and matrix genes *St6galnac3* are 913kd and 526kd in length, and their retained introns are extremely long, 639kd and 215kd, respectively, in contrast to a median intron size of 0.6 ~ 2.4kb in the mouse genome (Hong, Scofield, and Lynch 2006). In the human and mouse genome, long genes mostly occur because of

long introns (Breschi, Gingeras, and Guigó 2017; Lopes et al. 2021). We then plotted the length of SFPQ-regulated genes against genes across the genome and uncovered that the genes with increased IR in *Sfpq* depleted CNCCs have a median length of ~100kb, which is much longer than genes in the control groups, where the medial length is around 10kb (Figure-7D). These findings suggest that SFPQ-regulated splicing facilitates regulation of long gene expression during craniofacial development.

SFPQ depletion reduces matrix and Wnt signaling gene expression via IR-induced NMD.

In *Sfpq*-depleted CNCCs, retained introns contain PTCs within their open reading frames, suggesting PTC-triggered NMD as the mechanism of degradation (Supplemental Table S5). To test intron-regulated mRNA decay, we utilized the ST2 cell line, which is a mesenchymal cell line frequently used in the mechanistic study of ossification. SFPQ-regulated intron splicing and mRNA expression for *Col4a2, St6galnac3,* and *Ptk7* is conserved in ST2 cells where *Sfpq* depletion caused an increase in the percentage of intron-retaining mRNAs*,* and a concurrent decrease in their transcript levels (Figure-7E). To determine whether *Col4a2, St6galnac3,* and *Ptk7* mRNA were degraded through intron-induced NMD, we inhibited NMD with a chemical inhibitor, NMDI14, which targets key components of the NMD machinery, Upf1, and Smg5/7 (Martin et al. 2014). NMDI14 induced accumulation of intron 4 of *Col4a2,* intron 1 of *St6galnac3,* and intron 1 of *Ptk7* (Figure-7F), indicating that mRNA decay of intron-retaining transcripts occurs through the NMD machinery. To assess the impact of NMD on the total level of these mRNAs, we used PCR primers that detect the fully spliced transcripts and observed that inhibition of NMD restored and increased the levels of these mRNAs in *Sfpq*-depleted ST2 cells (Figure-7G). These data indicate that depletion of *Sfpq* causes aberrant intron retention that accelerates mRNA decay through NMD.

DISCUSSION

Here we revealed prevalent intron retention within CNCCs during embryonic development and propose intron-regulated expression of long genes as a causal mechanism for craniofacial deformity. This study demonstrated PRMT1-SFPQ pathway-regulated intron retention and subsequent control of matrix and Wnt pathway component gene expression during craniofacial development. Genetic deletion of *Prmt1* or depletion of *Sfpq* in CNCCs caused aberrant intron retention in genes encoding matrix proteins and Wnt

pathway components, which triggers NMD to degrade these transcripts. Matrix and Wnt pathway genes regulated by this mechanism share a common feature as long genes with a median length of 100kb, suggesting higher susceptibility of these genes to perturbation of the PRMT1-SFPQ pathway or splicing defects during craniofacial development. IR has been linked to diseases like cancer and autoimmune diseases (Dvinge and Bradley 2015). This study on PRMT1-regulated IR in controlling the expression of matrix genes in CNCCs, to our knowledge, is the first characterization of IR in craniofacial development.

Many of the PRMT1-SFPQ pathway regulated matrix and Wnt signaling genes are linked to congenital conditions. For example, in our study, the long gene *Wwox* exhibited increased retention of long introns (introns 3 and 4) and decreased expression upon *Prmt1/Sfpq* deficiency. In human patients, pathogenic variants of *WWOX* with large deletion within long introns have been associated with epileptic encephalopathy syndrome manifesting shared facial phenotype (Dvinge and Bradley 2015). This deletion within these long introns may affect SFPQ binding or splicing that impairs PRMT1-SFPQ pathway-regulated splicing of human *WWOX* genes. Another SFPQ target, *Ptk7* is associated with neural tube defects and scoliosis (Berger, Wodarz, and Borchers 2017). *SFPQ* itself is linked to neurodegenerative diseases including amyotrophic lateral sclerosis, frontotemporal dementia, and Alzheimer's disease, and was recently identified by a transcriptional study as a genetic regulator in CNCCs (Feng et al. 2021).

A main feature shared by SFPQ-regulated genes in CNCCs is gene length, with a median length of 100kb in contrast to a median length of ~10kb across the genome. In each SFPQ-regulate gene, introns exhibiting the highest retention also tend to be longer introns, for example, intron 4 of *Col4a2* which displayed the most dramatic increase in intronic reads within *Col4a2* boasts a length of 45kb. Increased IR in neuronal and immune cells has been proposed to facilitate a rapid response to external stimuli, allowing a time frame shorter than that required for de novo transcription (Ni et al. 2016; Mauger, Lemoine, and Scheiffele 2016). Thus, during craniofacial morphogenesis, IR represents a developmental process that enables rapid production of proteins when they are required. Long genes over 100kb in length require two to several hours for transcription, which presents a challenge for the fast tempo during embryogenesis. This intron-regulated mechanism allows cells to control the rate of production post-transcriptionally, and to achieve

rapid protein translation without waiting for the time period required for transcription. Long genes regulated by this mechanism are highly enriched in matrix pathways, and responsible for the production of building blocks for bone, cartilage, and connective tissues. Since spliceosomopathies preferentially affect the craniofacial skeleton and limb, intron retention-regulated matrix expression provides mechanistic insights for the higher susceptibility of these tissue types to spliceosome dysfunction.

PRMT1 is a critical regulator of RNA splicing. In mammalian cells, it methylates over forty splicing factors to modulate protein stability and subcellular localization, thereby influencing their accessibility to mRNA (Thandapani et al. 2013; Wu et al. 2022). PRMT1-catalyzed arginine methylation also affects their RNA binding affinity and specificity, altering their activity and the splicing product (Fedoriw et al. 2019; Fong et al. 2019). Therefore, it is not surprising that the disruption of this methyltransferase leads to splicing defects. However, the elevation of IR in this study presented a distinct role for PRMT1 in CNCCs. IR has been linked to diseases like cancer and autoimmune diseases (Dvinge and Bradley 2015). This study on PRMT1-regulated IR in controlling the expression of matrix genes in CNCCs, to our knowledge, is the first characterization of IR in craniofacial development, adding developmental insights to our knowledge of the roles of IR in pathophysiology. Manipulation of SFPQ expression (knockdown with 50% efficiency) demonstrated that SFPQ shares around 10% overlap with PRMT1 in IR-regulated neural crest gene expression. Given that PRMT1 methylates over 40 splicing factors in mammals and many of these splicing factors are abundantly expressed in embryonic CNCCs, we expect that PRMT1-governed splicing will be mediated by multiple splicing regulators besides SFPQ. The functional network of splicing regulation during craniofacial morphogenesis awaits further investigation.

We recognized SFPQ as a key PRMT1 substrate in CNCCs based on an unbiased in silico analysis and further validation in the mandibular and maxillary primordium. This aligns with previous work demonstrating that PRMT1 is responsible for the association of SFPQ with mRNA in mRNP complexes in mammalian cells (Snijders et al. 2015). Our findings on altered neuronal genes in *Sfpq*-depleted CNCCs also echoed previous reports in which SFPQ regulates many neuronal genes involved in axon extension, branching, viability, and synaptogenesis via direct binding to their intronic regions (Ruskin, Zamore, and

Green 1988; Cosker et al. 2016) and loss of SFPQ function compromises splicing patterns, especially the accurate splicing of long introns in embryonic brain and AML samples (Luisier et al. 2018; Taylor et al. 2022; Thomas-Jinu et al. 2017). Findings in this study further revealed previously unappreciated roles of SFPQ in the regulation of matrix genes through IR, which bear functional significance in CNCCs and craniofacial skeleton formation.

Altogether, these findings demonstrate that the PRMT1-SFPQ pathway regulates intron retention in CNCCs and suggests intron retention in long introns or long genes as a mechanism that controls matrix and Wnt signaling gene expression during craniofacial morphogenesis.

MATERIAL AND METHODS

Animals

Prmt1fl/fl mice were generously provided by Dr. Stephane Richard (McGill University). *Wnt1-Cre; Prmt1*$^{fl/fl}$*; R26R* ^{tdTomato} and *Wnt1-Cre; R26R*^{tdTomato} mice were obtained by mating *Wnt1-Cre, R26R*^{tdTomato} mice (Jax #009107 and #007914) with *Prmt1fl/fl* mice. Animals were genotyped using established protocols (Chai et al., 2000; Yu et al., 2009) and all care and experiments followed USC's IACUC protocols.

Proximity Ligation Assay (PLA)

PLA was performed using a Duolink In Situ kit (Millipore Sigma Cat# DUO92101-1KT), SFPQ antibody (Everest Biotech Cat#EB09523), and pan-methyl arginine antibody, Asymmetric Di-Methyl Arginine (ADMA) (Cell Signaling Technology Cat# 13522) to detect methyl-SFPQ. Tissue sections were blocked with the Duolink blocking solution in a humidity chamber for 1 hour at 37ºC before incubating with the primary antibodies (diluted 1:50 in Duolink PLA probe diluent) overnight at 4ºC. The next day, tissue sections were incubated with the Goat-PLUS and Rabbit-MINUS PLA probes for 1 hour at 37ºC, followed by ligase and amplification solution. After the final washes, tissue sections were mounted with Duolink In Situ mounting media with DAPI to counterstain nuclei. Images were visualized under the Confocal microscope. The number of PLA signals that appear as punctuate green signals in the nuclei of the cells and the quantification was performed with CellProfille software.

Tissue Processing and Immunofluorescence Staining

E13.5 control and *Wnt1-Cre; Prmt1*$^{fl/fl}$ mice embryos were fixed in 4% PFA overnight, dehydrated, embedded in O.C.T., sectioned at 8 µm thickness in sagittal orientation, and mounted on glass slides. The samples were submitted to antigen retrieval using Antigen Unmasking Solutions (Vector Laboratories Cat# H-3300-250) and washed and incubated in 0.5% Triton X-100 in PBS for 20 minutes for permeabilization. The samples were briefly washed with PBS, blocked with 10% goat serum for 1 hour, and then incubated with primary antibody SFPQ (Cell Signaling Technology Cat#23020, 1:100) at 4ºC overnight in a moisture chamber. After this, the samples were blocked with a secondary antibody (1:400, goat anti-rabbit IgG (H+L) Alexa Fluor 488 Cat# A-11008) for 1 hour followed by DAPI diluted 1:1000 in PBS for 10 minutes. Finally, the samples were mounted with mounting media (Electron Microscopy Science Cat# 1798510) and covered with cover glass for imaging using Keyence BZ-X800 and Leica DMI 3000B microscopes and quantification with CellProfille software.

Western Blotting

Heads of E13.5 control and *Wnt1-Cre; Prmt1*$^{fl/fl}$ mice were dissected, followed by removal of the brain. The tissue was homogenized with a pestle followed by lysis with RIPA Lysis and Extraction Buffer. The total protein concentration was determined by comparison with BSA standards. Twenty micrograms of total protein from each sample were loaded into each well of a 10% polyacrylamide gel. Western blot analysis was carried out as previously described (Iwata et al., 2012). Antibodies against PRMT1 (Cell Signaling, Cat# 2449; 1:1000), SFPQ (Cell Signaling Technology Cat#23020, 1:1000), and GAPDH (Cell Signaling, Cat# 97166; 1:5000) were used for Western blot. The samples were analyzed three times in independent blots and the bands were quantified by optical densitometry. The analysis was performed using ImageJ digital imaging processing software (ImageJ 1.48v, National Institutes of Health, Bethesda, MD, USA). The expression of each analyzed protein was normalized with GAPDH.

NCC collection and FACS analysis

The heads of E13.5 *Wnt1-Cre; R26R*^{tdTomato} and *Wnt1-Cre; Prmt1*$^{fl/fl}$*;R26R*^{tdTomato} mice were dissected and placed in a sterile tube containing Hanks' Balanced Salt Solution (Thermo Fisher Scientific Cat#

14175079). Following centrifugation, the supernatant was discarded, and the tissue was subjected to TrypLE (Thermo Fisher Scientific Cat# 12605010) incubation for 5-15 minutes at 37°C on a rotator. The resulting dissociated tissue was neutralized by the addition of fetal bovine serum (FBS) and then filtered through a 40µm strainer to obtain a single-cell suspension. After centrifugation, cells were resuspended in an appropriate volume of serum-free medium for FACS analysis. Sorting was performed based on tdTomato fluorescence (Excitation/emission: 554/582 nm), and td-positive cells were collected into separate tubes containing DMEM (Genesee Scientific Cat#25-501) supplemented with 20% FBS.

Primary CNCC culture and siRNA transfection

CNCCs were cultured in DMEM supplemented with 20% FBS in an incubator until they attach to the bottom. Then, reverse transfection with control siRNA (Qiagen Cat# 1027310), SFPQ siRNA #1 (Qiagen Cat#SI05783848), and SFPQ siRNA #2 (Qiagen Cat#SI05783876) at 40 nM using Lipofectamine RNAiMax transfection reagent (Invitrogen) for siRNA delivery were performed. 48h later, total RNA was isolated using TRIzol reagent (Invitrogen) following the manufacturer's protocols followed by mRNA isolation using NEBNext High-Input Poly(A) mRNA Magnetic Isolation Module (NEBNext® Cat#E3370S).

mRNA isolation and sequencing

Poly(A) mRNA isolation was extracted from total RNA by using the NEBNext High-Input Poly(A) mRNA Magnetic Isolation Module (NEBNext® Cat#E3370S). Five sets of primary isolated CNCCs from control (*Wnt1-Cre;ROSA*^{Td}) or *Prmt1 CKO* (*Wnt1-Cre; Prmt1*$^{fl/fl}$*; ROSA*^{Td}), and three sets of isolated CNCCs transfected with siControl or siSFPQs were sequenced at 40 million reads sequencing depth and 150 bp paired-end sequencing.

Bioinformatic Analysis of Differentially Regulated Genes (DEG) and differential Intron Retention (IR)

The sequencing quality of RNA-seq libraries was assessed by FastQC v0.11.8 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Low-quality bases and adapters were trimmed using Trim Galore (v0.6.7,https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). The reads were mapped to mouse genome mm10 using hisat2 (v2.1.0) (Kim et al. 2019). The mapped sam

files from hisat2 were converted to bam files which were then further turned to sorted bam files by samtools (v1.7) (Li et al. 2009). Mapped reads were then processed by htseq-count (v1.99.2) to calculate the read count of all genes (Anders, Pyl, and Huber 2015). The sorted bam files were then given to IRTools (https://github.com/WeiqunPengLab/IRTools/) to calculate the IR of each gene and the read count of genes' constitutive intronic regions (CIRs) and constitutive exonic regions (CERs). The expression level of a gene was expressed as TPM (Transcripts Per Kilobase Million) value. EdgeR was used to identify differentially expressed genes (DEGs) by requiring ≥ 1.5-fold expression changes and adjusted p-value < 0.05 (Robinson, McCarthy, and Smyth 2010). The intron retention index (IRI) of a gene is defined as the ratio of the overall read density of its CIRs to the overall read density of its CERs. Genes with 0< IRI <1 and CER expression greater than 1 were selected to identify differential IR genes. Student's t-test was used to identify differential IR genes by requiring ≥ 1.5-fold log2(IR) changes and p-value < 0.05. Gene Ontology analysis was done by DAVID (https://david.ncifcrf.gov/tools.jsp) and Metascape (https://metascape.org/gp/index.html#/main/step1) (Sherman et al. 2022; Zhou et al. 2019). The RNA-seq data are deposited to the Gene Expression Omnibus (GEO) database with accession number GSE266474.

rMATS and rMAPS Analysis

rMATS (https://github.com/Xinglab/rmats-turbo) and rMAPS2 (http://rmaps.cecsresearch.org/) were used to find potential RNA binding Proteins (RBPs) that contribute to differential splicing events (Hwang et al. 2020; Y. Wang et al. 2024). rMATS was first used to identify differential alternative splicing events between wild type and PRMT1 CKO samples by requiring FDR < 0.05. The output of rMATs were then fed to rMAPS2 to identify significant RBPs contributing to each type of Splicing events.

Quantitative PCR (qPCR)

For both experiments, Control, and Wnt1-Cre; Prmt1fl/fl mandibles, and siRNA-transfected CNCCs, mRNAs were quantified by real-time PCR with IQ Sybr Green Supermix (Bio-Rad) and normalized against GAPDH mRNA levels. Relative changes in expression were calculated using the ΔΔCt method. Primer sequences are listed in Supplemental Table S6.

Statistical analysis

Two-tailed Student's t-tests or Fisher's exact tests were applied for statistical analysis. For all graphs, error bars represent standard deviations. A p-value of <0.05 was considered statistically significant.

Author contributions: J.R and N.U. designed and performed experiments, analyzed data, drafted, and revised the manuscript; Q.C. and W.P. performed bioinformatic analysis of RNA-seq data; M.V., Y.C. & A. M. contributed to embryo collection and CNCC isolation; T.R. contributed to RT-PCR analysis, J.X., and W.P. concepted the study, supervised the performance of experiments and data analysis, and critically revised the manuscript.

Competing interests: The authors declare no competing interests.

Data Availability: RNA-seq data is deposited at the NCBI Gene Expression Omnibus (GSE 171630)

REFERENCES

Achilleos, A., and P. A. Trainor. 2012. "Neural crest stem cells: discovery, properties and potential for therapy." *Cell Res* 22 (2): 288-304. https://doi.org/10.1038/cr.2012.11.

Anders, S., P. T. Pyl, and W. Huber. 2015. "HTSeq--a Python framework to work with high-throughput sequencing data." *Bioinformatics* 31 (2): 166-9. https://doi.org/10.1093/bioinformatics/btu638.

Bélanger, C., F. A. Bérubé-Simard, E. Leduc, G. Bernas, P. M. Campeau, et al. 2018. "Dysregulation of cotranscriptional alternative splicing underlies CHARGE syndrome." *Proc Natl Acad Sci U S A* 115 (4): E620-e629. https://doi.org/10.1073/pnas.1715378115.

Berger, H., A. Wodarz, and A. Borchers. 2017. "PTK7 Faces the Wnt in Development and Disease." *Front Cell Dev Biol* 5: 31. https://doi.org/10.3389/fcell.2017.00031.

Blanc, R. S., and S. Richard. 2017. "Arginine Methylation: The Coming of Age." *Mol Cell* 65 (1): 8-24. https://doi.org/10.1016/j.molcel.2016.11.003.

Braunschweig, U., N. L. Barbosa-Morais, Q. Pan, E. N. Nachman, B. Alipanahi, et al. 2014. "Widespread intron retention in mammals functionally tunes transcriptomes." *Genome Res* 24 (11): 1774-86. https://doi.org/10.1101/gr.177790.114.

Breschi, A., T. R. Gingeras, and R. Guigó. 2017. "Comparative transcriptomics in human and mouse." *Nat Rev Genet* 18 (7): 425-440. https://doi.org/10.1038/nrg.2017.19.

Chai, Y., X. Jiang, Y. Ito, P. Bringas, Jr., J. Han, et al. 2000. "Fate of the mammalian cranial neural crest during tooth and mandibular morphogenesis." *Development* 127 (8): 1671-9. https://doi.org/10.1242/dev.127.8.1671. https://www.ncbi.nlm.nih.gov/pubmed/10725243.

Cosker, K. E., S. J. Fenstermacher, M. F. Pazyra-Murphy, H. L. Elliott, and R. A. Segal. 2016. "The RNA-binding protein SFPQ orchestrates an RNA regulon to promote axon viability." *Nat Neurosci* 19 (5): 690-696. https://doi.org/10.1038/nn.4280.

Dvinge, H., and R. K. Bradley. 2015. "Widespread intron retention diversifies most cancer transcriptomes." *Genome Med* 7 (1): 45. https://doi.org/10.1186/s13073-015-0168-9.

Fedoriw, A., S. R. Rajapurkar, S. O'Brien, S. V. Gerhart, L. H. Mitchell, et al. 2019. "Anti-tumor Activity of the Type I PRMT Inhibitor, GSK3368715, Synergizes with PRMT5 Inhibition through MTAP Loss." *Cancer Cell* 36 (1): 100-114.e25. https://doi.org/10.1016/j.ccell.2019.05.014.

Feng, Z., Z. Duren, Z. Xiong, S. Wang, F. Liu, et al. 2021. "hReg-CNCC reconstructs a regulatory network in human cranial neural crest cells and annotates variants in a developmental context." *Commun Biol* 4 (1): 442. https://doi.org/10.1038/s42003-021-01970-0.

Fong, J. Y., L. Pignata, P. A. Goy, K. C. Kawabata, S. C. Lee, et al. 2019. "Therapeutic Targeting of RNA Splicing Catalysis through Inhibition of Protein Arginine Methylation." *Cancer Cell* 36 (2): 194-209.e9. https://doi.org/10.1016/j.ccell.2019.07.003.

Gou, Y., J. Li, O. Jackson-Weaver, J. Wu, T. Zhang, et al. 2018. "Protein Arginine Methyltransferase PRMT1 Is Essential for Palatogenesis." *J Dent Res* 97 (13): 1510-1518.

Gou, Y., J. Li, J. Wu, R. Gupta, I. Cho, et al. 2018. "Prmt1 regulates craniofacial bone formation upstream of Msx1." *Mech Dev* 152: 13-20. https://doi.org/10.1016/j.mod.2018.05.001.

Hartel, N. G., B. Chew, J. Qin, J. Xu, and N. A. Graham. 2019. "Deep Protein Methylation Profiling by Combined Chemical and Immunoaffinity Approaches Reveals Novel PRMT1 Targets." *Mol Cell Proteomics* 18 (11): 2149-2164. https://doi.org/10.1074/mcp.RA119.001625.

Hong, X., D. G. Scofield, and M. Lynch. 2006. "Intron size, abundance, and distribution within untranslated regions of genes." *Mol Biol Evol* 23 (12): 2392-404. https://doi.org/10.1093/molbev/msl111.

Hooper, J. E., K. L. Jones, F. J. Smith, T. Williams, and H. Li. 2020. "An Alternative Splicing Program for Mouse Craniofacial Development." *Front Physiol* 11: 1099. https://doi.org/10.3389/fphys.2020.01099. https://www.ncbi.nlm.nih.gov/pubmed/33013468.

Hosokawa, M., A. Takeuchi, J. Tanihata, K. Iida, S. Takeda, et al. 2019. "Loss of RNA-Binding Protein Sfpq Causes Long-Gene Transcriptopathy in Skeletal Muscle and Severe Muscle Mass Reduction with Metabolic Myopathy." *iScience* 13: 229-242.

Hwang, J. Y., S. Jung, T. L. Kook, E. C. Rouchka, J. Bok, et al. 2020. "rMAPS2: an update of the RNA map analysis and plotting server for alternative splicing regulation." *Nucleic Acids Res* 48 (W1): W300-w306. https://doi.org/10.1093/nar/gkaa237.

Ishii, M., A. C. Arias, L. Liu, Y. B. Chen, M. E. Bronner, et al. 2012. "A stable cranial neural crest cell line from mouse." *Stem Cells Dev* 21 (17): 3069-80. https://doi.org/10.1089/scd.2012.0155.

Jackson-Weaver, O., N. Ungvijanpunya, Y. Yuan, J. Qian, Y. Gou, et al. 2020. "PRMT1-p53 Pathway Controls Epicardial EMT and Invasion." *Cell Rep* 31 (10): 107739.

Jiang, X., S. Iseki, R. E. Maxson, H. M. Sucov, and G. M. Morriss-Kay. 2002. "Tissue origins and interactions in the mammalian skull vault." *Dev Biol* 241 (1): 106-16. https://doi.org/10.1006/dbio.2001.0487. https://www.ncbi.nlm.nih.gov/pubmed/11784098.

Kim, D., J. M. Paggi, C. Park, C. Bennett, and S. L. Salzberg. 2019. "Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype." *Nat Biotechnol* 37 (8): 907-915. https://doi.org/10.1038/s41587-019-0201-4.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, et al. 2009. "The Sequence Alignment/Map format and SAMtools." *Bioinformatics* 25 (16): 2078-9. https://doi.org/10.1093/bioinformatics/btp352.

Liu, Q., and G. Dreyfuss. 1995. "In vivo and in vitro arginine methylation of RNA-binding proteins." *Mol Cell Biol* 15 (5): 2800-8. https://doi.org/10.1128/MCB.15.5.2800. https://www.ncbi.nlm.nih.gov/pubmed/7739561.

Lopes, I., G. Altab, P. Raina, and J. P. de Magalhães. 2021. "Gene Size Matters: An Analysis of Gene Length in the Human Genome." *Front Genet* 12: 559998. https://doi.org/10.3389/fgene.2021.559998.

Luisier, R., G. E. Tyzack, C. E. Hall, J. S. Mitchell, H. Devine, et al. 2018. "Intron retention and nuclear loss of SFPQ are molecular hallmarks of ALS." *Nat Commun* 9 (1): 2010. https://doi.org/10.1038/s41467-018-04373-8.

Marquez, Y., J. W. Brown, C. Simpson, A. Barta, and M. Kalyna. 2012. "Transcriptome survey reveals increased complexity of the alternative splicing landscape in Arabidopsis." *Genome Res* 22 (6): 1184-95. https://doi.org/10.1101/gr.134106.111.

Martik, M. L., and M. E. Bronner. 2017. "Regulatory Logic Underlying Diversification of the Neural Crest." *Trends Genet* 33 (10): 715-727. https://doi.org/10.1016/j.tig.2017.07.015.

Martin, L., A. Grigoryan, D. Wang, J. Wang, L. Breda, et al. 2014. "Identification and characterization of small molecules that inhibit nonsense-mediated RNA decay and suppress nonsense p53 mutations." *Cancer Res* 74 (11): 3104-13.

Mauger, O., F. Lemoine, and P. Scheiffele. 2016. "Targeted Intron Retention and Excision for Rapid Gene Regulation in Response to Neuronal Activity." *Neuron* 92 (6): 1266-1278. https://doi.org/10.1016/j.neuron.2016.11.032.

Monteuuis, G., J. J. L. Wong, C. G. Bailey, U. Schmitz, and J. E. J. Rasko. 2019. "The changing paradigm of intron retention: regulation, ramifications and recipes." *Nucleic Acids Res* 47 (22): 11497-11513. https://doi.org/10.1093/nar/gkz1068.

Ni, T., W. Yang, M. Han, Y. Zhang, T. Shen, et al. 2016. "Global intron retention mediated gene regulation during CD4+ T cell activation." *Nucleic Acids Res* 44 (14): 6817-29. https://doi.org/10.1093/nar/gkw591.

Park, J. W., S. Jung, E. C. Rouchka, Y. T. Tseng, and Y. Xing. 2016. "rMAPS: RNA map analysis and plotting server for alternative exon regulation." *Nucleic Acids Res* 44 (W1): W333-8. https://doi.org/10.1093/nar/gkw410. https://www.ncbi.nlm.nih.gov/pubmed/27174931.

Plein, A., A. Fantin, and C. Ruhrberg. 2015. "Neural crest cells in cardiovascular development." *Curr Top Dev Biol* 111: 183-200. https://doi.org/10.1016/bs.ctdb.2014.11.006.

Rho, J., S. Choi, C. R. Jung, and D. S. Im. 2007. "Arginine methylation of Sam68 and SLM proteins negatively regulates their poly(U) RNA binding activity." *Arch Biochem Biophys* 466 (1): 49-57. https://doi.org/10.1016/j.abb.2007.07.017. https://www.ncbi.nlm.nih.gov/pubmed/17764653.

Robinson, M. D., D. J. McCarthy, and G. K. Smyth. 2010. "edgeR: a Bioconductor package for differential expression analysis of digital gene expression data." *Bioinformatics* 26 (1): 139-40. https://doi.org/10.1093/bioinformatics/btp616.

Ruskin, B., P. D. Zamore, and M. R. Green. 1988. "A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly." *Cell* 52 (2): 207-19. https://doi.org/10.1016/0092-8674(88)90509-0.

Ruta, V., V. Pagliarini, and C. Sette. 2021. "Coordination of RNA Processing Regulation by Signal Transduction Pathways." *Biomolecules* 11 (10). https://doi.org/10.3390/biom11101475.

Schmitz, U., N. Pinello, F. Jia, S. Alasmari, W. Ritchie, et al. 2017. "Intron retention enhances gene regulatory complexity in vertebrates." *Genome Biol* 18 (1): 216.

Sherman, B. T., M. Hao, J. Qiu, X. Jiao, M. W. Baseler, et al. 2022. "DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update)." *Nucleic Acids Res* 50 (W1): W216-w221. https://doi.org/10.1093/nar/gkac194.

Smith, W. A., B. T. Schurter, F. Wong-Staal, and M. David. 2004. "Arginine methylation of RNA helicase a determines its subcellular localization." *J Biol Chem* 279 (22): 22795-8. https://doi.org/10.1074/jbc.C300512200. https://www.ncbi.nlm.nih.gov/pubmed/15084609.

Snijders, A. P., G. M. Hautbergue, A. Bloom, J. C. Williamson, T. C. Minshull, et al. 2015. "Arginine methylation and citrullination of splicing factor proline- and glutamine-rich (SFPQ/PSF) regulates its association with mRNA." *Rna* 21 (3): 347-59. https://doi.org/10.1261/rna.045138.114.

Strobl-Mazzulla, P. H., M. Marini, and A. Buzzi. 2012. "Epigenetic landscape and miRNA involvement during neural crest development." *Dev Dyn* 241 (12): 1849-56. https://doi.org/10.1002/dvdy.23868.

Sun, X., Z. Liu, Z. Li, Z. Zeng, W. Peng, et al. 2023. "Abnormalities in intron retention characterize patients with systemic lupus erythematosus." *Sci Rep* 13 (1): 5141. https://doi.org/10.1038/s41598-023-31890-4.

Taylor, R., F. Hamid, T. Fielding, P. M. Gordon, M. Maloney, et al. 2022. "Prematurely terminated intron-retaining mRNAs invade axons in SFPQ null-driven neurodegeneration and are a hallmark of ALS." *Nat Commun* 13 (1): 6994. https://doi.org/10.1038/s41467-022-34331-4.

Thandapani, P., T. R. O'Connor, T. L. Bailey, and S. Richard. 2013. "Defining the RGG/RG motif." *Mol Cell* 50 (5): 613-23. https://doi.org/10.1016/j.molcel.2013.05.021.

Thomas-Jinu, S., P. M. Gordon, T. Fielding, R. Taylor, B. N. Smith, et al. 2017. "Non-nuclear Pool of Splicing Factor SFPQ Regulates Axonal Transcripts Required for Normal Motor Development." *Neuron* 94 (2): 322-336.e5. https://doi.org/10.1016/j.neuron.2017.03.026.

Tian, Y., Z. Zeng, X. Li, Y. Wang, R. Chen, et al. 2020. "Transcriptome-wide stability analysis uncovers LARP4-mediated NFκB1 mRNA stabilization during T cell activation." *Nucleic Acids Res* 48 (15): 8724-8739. https://doi.org/10.1093/nar/gkaa643.

Wang, E. T., R. Sandberg, S. Luo, I. Khrebtukova, L. Zhang, et al. 2008. "Alternative isoform regulation in human tissue transcriptomes." *Nature* 456 (7221): 470-6. https://doi.org/10.1038/nature07509.

Wang, Y., Z. Xie, E. Kutschera, J. I. Adams, K. E. Kadash-Edmondson, et al. 2024. "rMATS-turbo: an efficient and flexible computational tool for alternative splicing analysis of large-scale RNA-seq data." *Nat Protoc* 19 (4): 1083-1104. https://doi.org/10.1038/s41596-023-00944-2.

Wong, J. J., A. Y. Au, W. Ritchie, and J. E. Rasko. 2016. "Intron retention in mRNA: No longer nonsense: Known and putative roles of intron retention in normal and disease biology." *Bioessays* 38 (1): 41-9. https://doi.org/10.1002/bies.201500117. https://www.ncbi.nlm.nih.gov/pubmed/26612485.

Wong, J. J., and U. Schmitz. 2022. "Intron retention: importance, challenges, and opportunities." *Trends Genet* 38 (8): 789-792. https://doi.org/10.1016/j.tig.2022.03.017.

Wu, H., Y. Zhang, S. Liu, D. Liu, A. Li, et al. 2022. "Protein Arginine Methyltransferase 1 and its Dynamic Regulation Associated with Cellular Processes and Diseases." *Protein Pept Lett* 29 (3): 218-230. https://doi.org/10.2174/0929866529666220124120208. https://www.ncbi.nlm.nih.gov/pubmed/35068392.

Zhou, Y., B. Zhou, L. Pache, M. Chang, A. H. Khodabakhshi, et al. 2019. "Metascape provides a biologist-oriented resource for the analysis of systems-level datasets." *Nat Commun* 10 (1): 1523. https://doi.org/10.1038/s41467-019-09234-6.

FIGURE AND LEGENDS

Figure 1. CNCC-specific deletion of *Prmt1* elevates intron retention in the embryonic mandibular process. (A) Expression levels of PRMT1-9 mRNAs in primary isolated cranial neural crest cells (CNCC) from *Wnt1-Cre; Rosa^{Td}* mouse embryonic heads at E13.5 and E15.5. TPM, transcript per million. (B) Diagram illustrating the isolation of CNCCs from embryonic mandibles, followed by poly(A)+ mRNA isolation and sequencing (n=4 in control and in *Prmt1 CKO*). (C) *Prmt1* deletion in CNCC caused changes in alternative splicing (AS). Changes in AS events were analyzed by rMATS using RNA sequencing data and significant changes in each type of AS were shown in a stacked bar chart. SE, skipped exon. IR, intron retention. MXE, mutually exclusive exons. A5SS, alternative 5' splice site. A3SS, alternative 3' splice site. a) The number of intron retention events that increased (Up, yellow color) or decreased (Down, brown color) in *Prmt1* deficient CNCCs. (D & E) Intron retention was prevalent in CNCCs and altered by *Prmt1* deletion. Track view of genes demonstrating intron retention and further elevated by *Prmt1* deletion, illustrated by *Pex12, Nkx3-2, Nkiras1, 1810026B05Rik* in *D* or reduced by *Prmt1* deletion, illustrated by *Tbx1* in *E*. (F) Quantification of intron expression in *Pex12, Nkx3-2, Nkiras1, 1810026B05Rik* and *Tbx1* by the percentage of intron-retaining mRNAs in each gene, calculated from IRI analysis based on RNA sequencing data. * $p<0.05$ *Prmt1* CKO vs Control. Control: *Wnt1-Cre; Rosa^{Td}*. *Prmt1 CKO: Wnt1-Cre; Prmt1$^{fl/fl}$; Rosa^{Td}*.

Figure 2. CNCC-specific *Prmt1* deletion reduced matrix gene expression in the developing mandibles. (A) Volcano plot illustrating upregulation of 160 and downregulation of 303 genes in the mandibular primordium of *Prmt1*-deficient embryos at E13.5, compared to control embryos. (B) Heatmap showing differential gene expression between control and *Prmt1*-deficient mandibles. The top downregulated matrix genes were labeled. (C) GO analysis of pathway enrichment in downregulated genes demonstrating glycosaminoglycan (GAG) degradation and extracellular matrix (ECM) organization as the top pathways. (D) Ingenuity Pathway Analysis suggesting connective tissue, bone, and cartilage development as affected

biological processes based on downregulated genes. Control: *Wnt1-Cre; Rosa*^{Td} (n=4). *Prmt1 CKO: Wnt1-Cre; Prmt1*$^{fl/fl}$*; Rosa*^{Td} (n=4).

Figure 3. PRMT1 regulates intron retention in ECM and GAG degradation genes. (A-C) Intron retention increased in the majority of ECM gene transcripts that were downregulated in *Prmt1* deficient embryos, as illustrated by a scatter plot based on intron retention index (IRI) analysis. The red line delineates where intron retention is unchanged. Genes with the top differential IR were represented by red dots and labeled. ECM gene transcripts *Bmp7* and *Cthrc1* demonstrating higher intron retention in *Prmt1*-deleted embryos were illustrated by track view in *B* and quantification of intronic (left) and exonic (right) expression in *C*. (D) Higher intron retention and lower mRNA level of ECM genes were validated in a panel by RT-PCR. Primers that span the intronic or intron-exon junction region were used to assess intronic expression. Primers that span the exonic region were used to examine exonic expression. (E-G) Intron retention increased in the majority of GAG degradation gene transcripts that were downregulated in *Prmt1* deficient embryos compared to control, as indicated by a scatter plot based on IRI analysis. The red line delineates where intron retention is unchanged. Genes were represented by black dots and labeled in red. GAG degradation genes *St6galnac3* and *Galnt11* demonstrating higher intron retention in *Prmt1*-deleted embryos were illustrated by (F) track view and (G) quantified for intronic (left) and exonic (right) expression. * $p<0.05$ *Prmt1* CKO vs Control. Control: *Wnt1-Cre; Rosa*^{Td}. *Prmt1 CKO: Wnt1-Cre; Prmt1*$^{fl/fl}$*; Rosa*^{Td}.

Figure 4. PRMT1 methylates SFPQ and regulates its expression. (A) Expression levels of splicing factors SFPQ, SRSF1, EWSR1, TAF15, TRA2B, HnRNPA1, WDR70, G3BP1 in primary isolated CNCCs from *Wnt1-Cre; Rosa*^{Td} mouse embryonic heads at E13.5 and E15.5. TPM, transcript per million. (B-F) SFPQ methylation diminished in the mandibular and maxillary processes of *Prmt1* deficient embryos at E13.5. B shows that sagittal sections of mouse tissues were used to analyze the mandibular (area outlined in red, bottom) and maxillary (area outlined in red, top) process. Methyl-SFPQ was detected by proximity ligation assay (PLA) in the embryonic mandible (C) and maxilla (E) of control and *Prmt1* deficient embryos. Green puncta indicated methyl-SFPQ, and nuclei were counterstained with DAPI (blue).

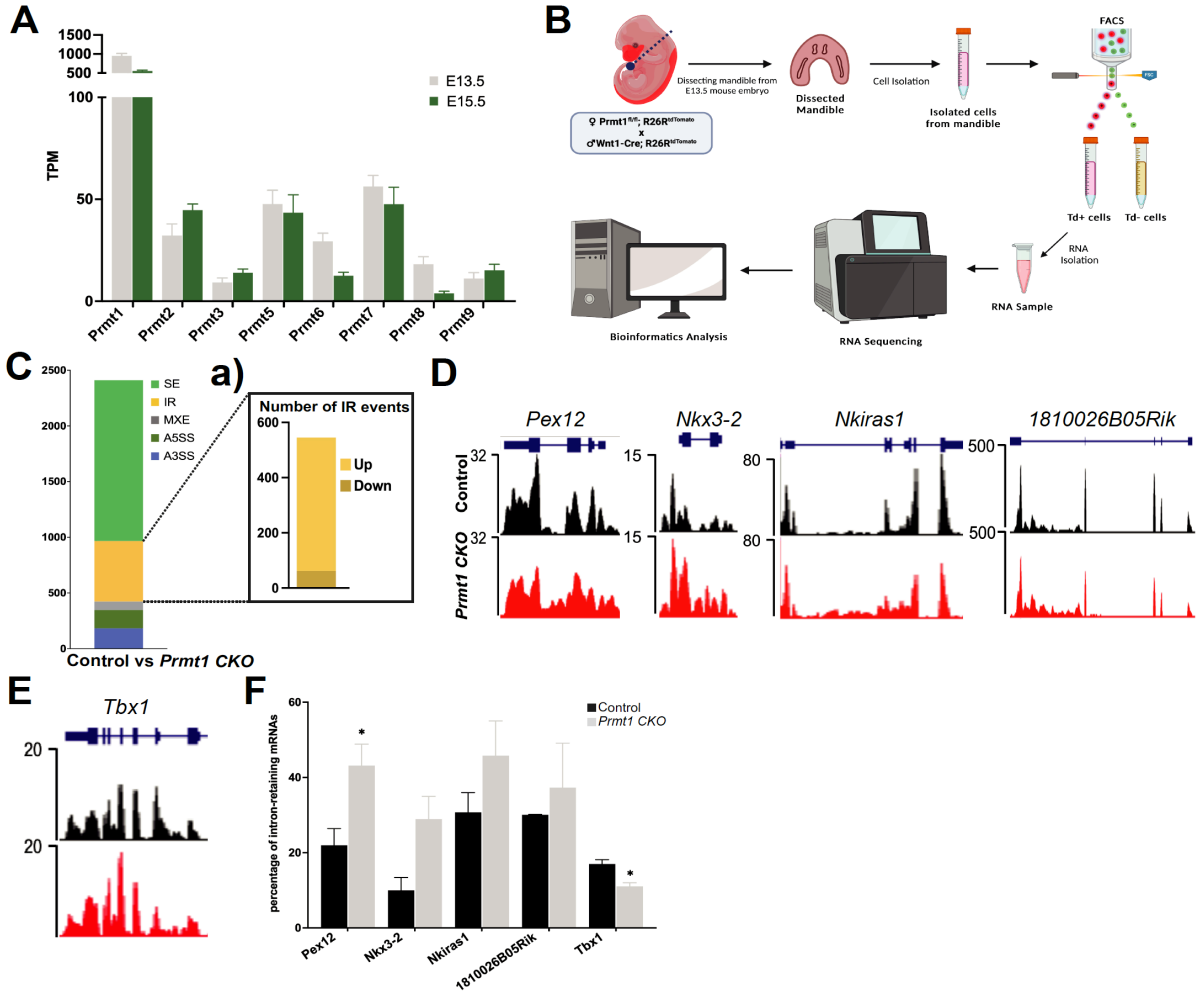
Representative images are shown for Control (Ca-Cd, Ea-Ed) and *Prmt1 CKO* (Ce-Ch, Ee-Eh). Higher magnification views in (Cd, Ed) and (Ch, Eh) illustrate methyl-SFPQ (green puncta) in CNCC nuclei, as indicated by white arrows. D&F show quantification of PLA puncta normalized to cell number in four biological replicates, presented as mean ± SEM. (G) SFPQ methylation signal remained robust in the epithelial region of craniofacial structures, while CNCC-specific deletion of *Prmt1* caused a drastic reduction in the craniofacial mesenchyme. Higher magnification of the mandibular tip showing mesenchyme and epithelium outlined with dotted lines in Control (Ga-Gc) and *Prmt1 CKO* (Gd-Gf). (H-K) SFPQ protein expression was significantly reduced in the mandibular and maxillary processes of *Prmt1* deficient embryos. The level of SFPQ protein was detected by immunostaining in the embryonic mandible (H) and maxilla (J) of control and *Prmt1* deficient embryos, and quantified in I and K. (L-M) SFPQ protein expression declined dramatically in the *Prmt1* deficient embryonic head, as detected by Western blotting (L). Quantification was conducted with ImageJ and shown in M. Mx, maxilla. Md, mandible. * $p<0.05$ *Prmt1* CKO vs Control. Control: *Wnt1-Cre; Rosa*^{Td}. *Prmt1 CKO: Wnt1-Cre; Prmt1*$^{fl/fl}$*; Rosa*^{Td}.

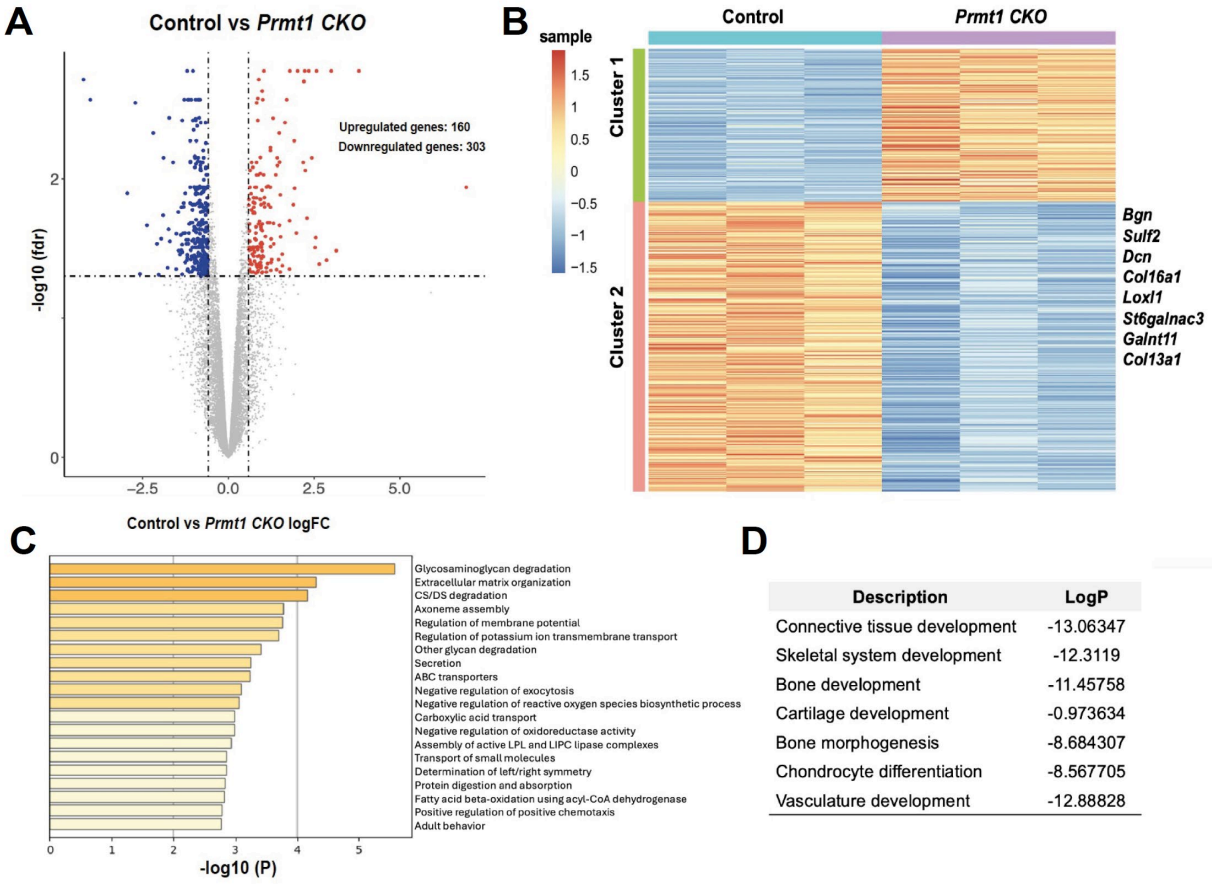
Figure 5. PRMT1-SFPQ pathway regulates matrix genes in CNCCs. (A) SFPQ depletion in CNCCs altered intron retention. IR events in genes showing increased (Up, yellow color) or decreased (Down, brown color) intron retention were illustrated by stacked bar graphs. CNCCs were transfected with two independent siRNAs targeting SFPQ, or control siRNA, followed by poly(A)+ mRNA extraction and RNA sequencing. (B) SFPQ depletion in CNCCs altered gene expression. Pie chart demonstrating differentially regulated genes in CNCCs with SFPQ depletion by two independent siRNAs. Shaded areas among the downregulated genes (orange) indicate overlap with IRI up. (C) GO analysis of genes with increased IR events in CNCCs with SFPQ depletion by siRNA#1 (a) or #2 (b). (D) Genes with increased IR events in SFPQ-depleted CNCCs demonstrated 8.28% (64 out of 773) overlap with *Prmt1* deficient CNCCs. (E-H) Matrix genes regulated by both *Prmt1* deletion and SFPQ depletion. Bar graph showing elevated IR (E, F) and reduced expression (G, H) in both SFPQ-depleted and *Prmt1* deletion CNCCs. TPM, transcripts per million. (I-J) Heatmap and GO analysis of SFPQ-regulated genes in CNCC. * $p<0.05$ siSFPQ vs siControl. siControl, CNCCs transfected with control siRNA. siSFPQ#1 and siSFPQ#2, CNCCs transfected with two

independent SFPQ siRNAs. * p<0.05 *Prmt1* CKO vs Control. Control, *Wnt1-Cre; Rosa*^{Td}. *Prmt1 CKO, Wnt1-Cre; Prmt1*$^{fl/fl}$*; Rosa*^{Td}.

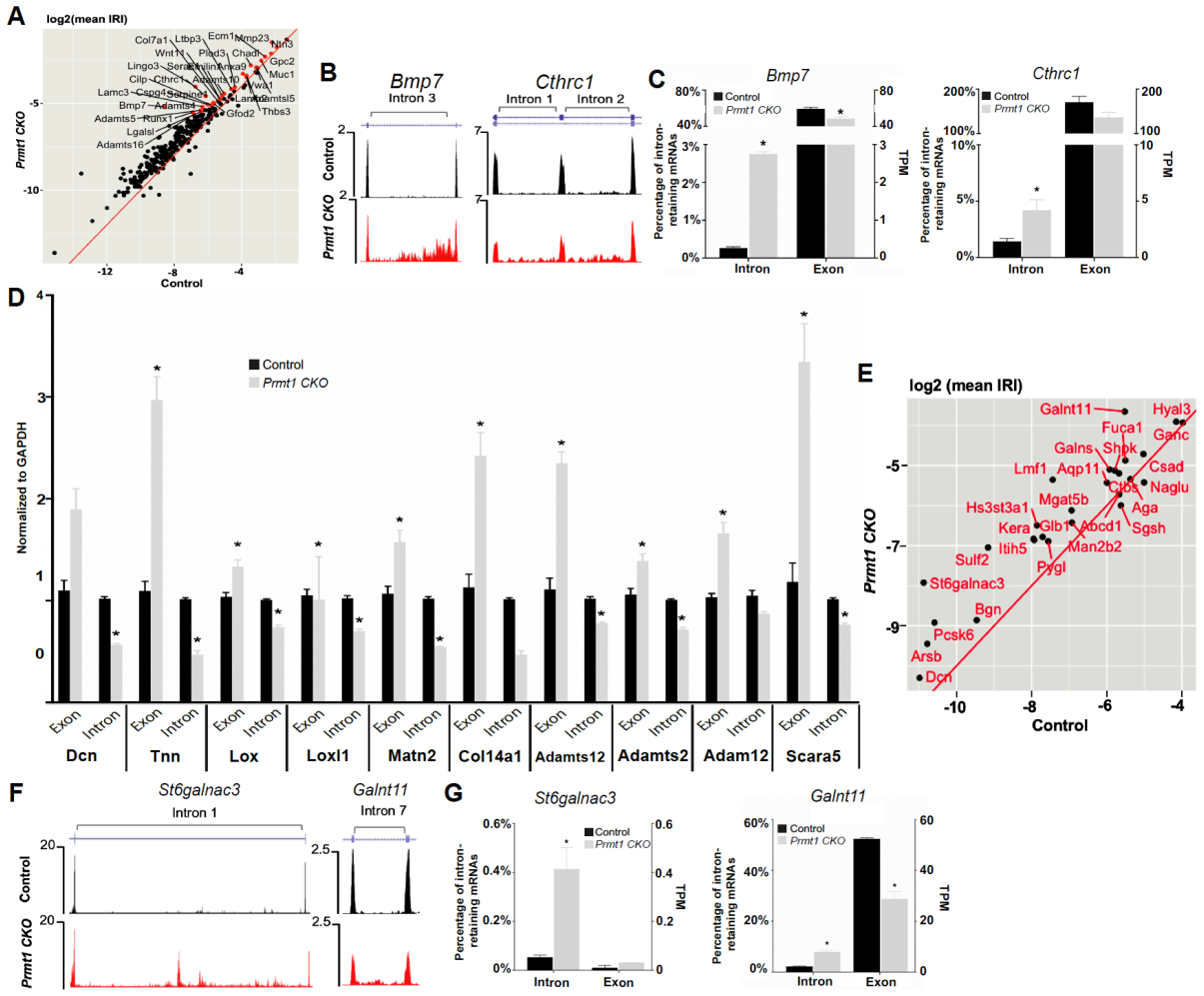
Figure 6. SFPQ regulates intron retention of Wnt signaling and neuronal genes in CNCCs. SFPQ depletion in CNCCs elevated intron retention and decreased mRNA expression of a group of Wnt signaling components (A) and neuronal genes (B). The levels of mRNA expression (left) and intron retention (right) were illustrated by a two-sided bar graph. TPM, transcripts per million. * p<0.01 siSFPQ vs siControl. siControl, CNCCs transfected with control siRNA. siSFPQ#1 and siSFPQ#2, CNCCs transfected with two independent SFPQ siRNAs.

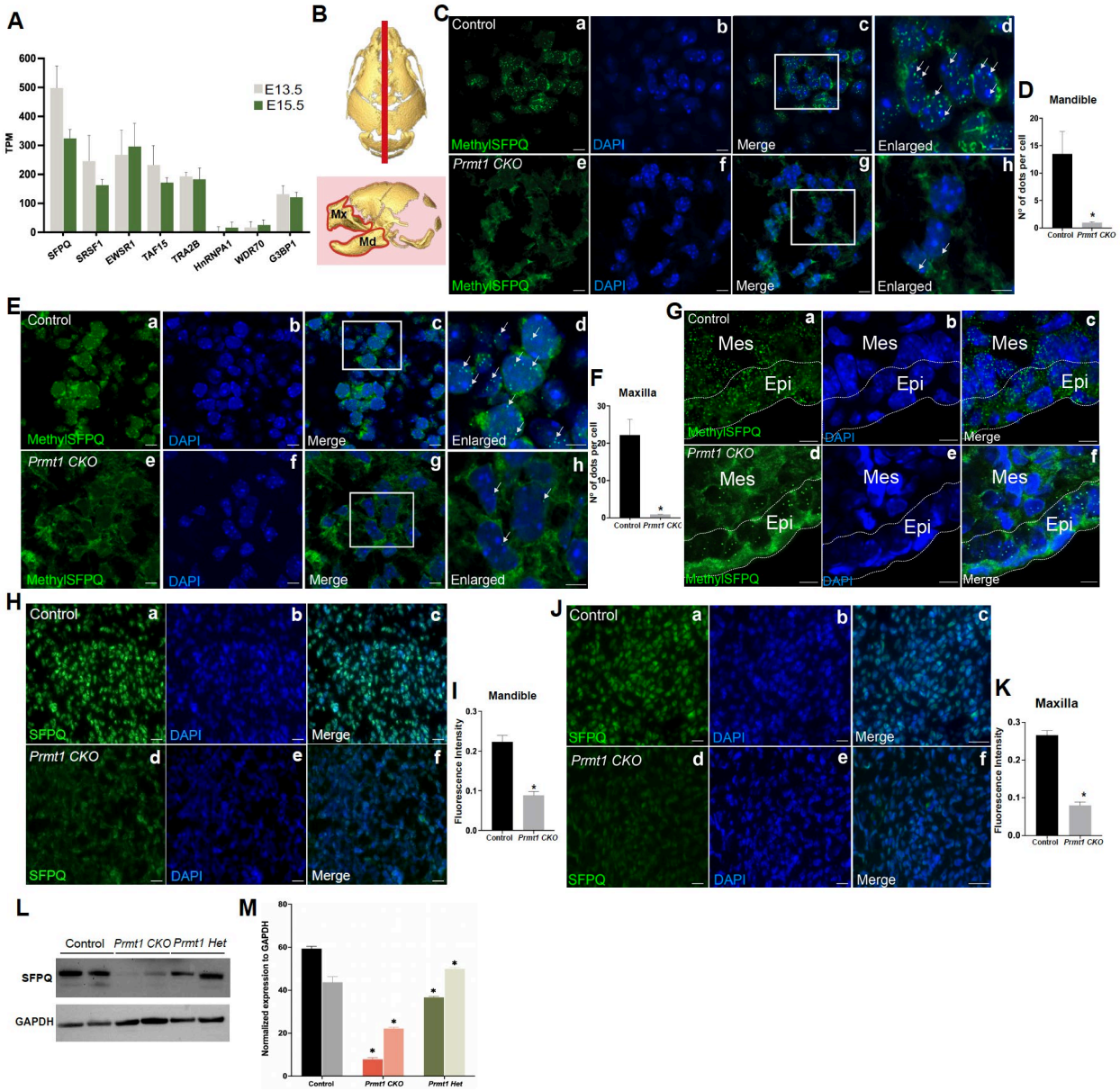
Figure 7: SFPQ depletion reduces long gene expression via intron retention triggered NMD. (A) SFPQ binding peaks were mapped to retained intron 1, but not spliced intron 6 of *Ptk7*. Peak distribution from published Sfpq CLIP-seq data using E13.5 brain (top) and track view of RNA-seq data using siCont or siSFPQ-transfected CNCCs (bottom) for *Ptk7,* with the retained *Ptk7* intron 1 and spliced intron 6 highlighted. (B) SFPQ binding peaks were significantly enriched in retained introns within CNCCs. Violin plot displaying the global distribution of enriched SFPQ binding peaks in introns with elevated retention compared to introns with reduced retention or no change. Peak number was normalized to intron length. P value was indicated at the top. (C) SFPQ binding peaks were preferentially enriched in genes with higher intron retention when compared to genes with no IR change. p = 0.07 using Fisher's exact test. (E) SFPQ-regulated genes were significantly longer. Violin plot displaying the distribution of length for genes showing increased intron retention (IR Up), decreased intron retention (IR Down), or unchanged intron expression (No Change). (F) SFPQ depletion promoted intron retention and reduced mRNA expression of *Col4a2, St6galnac3,* and *Ptk7* in ST2 cells. Bar chart showing RT-PCR analysis of intronic and exonic expression. * P<0.05 siSFPQ vs. siControl. (G-H) NMD inhibitor NMDI14 induced the accumulation of retained introns and mRNAs of *Col4a2, St6galnac3,* and *Ptk7* in ST2 cells. Bar chart showing RT-PCR analysis of intronic (G) and exonic expression (H) in DMSO or NMDI-treated cells. * P<0.05 NMDI vs. DMSO treatment of the same group.

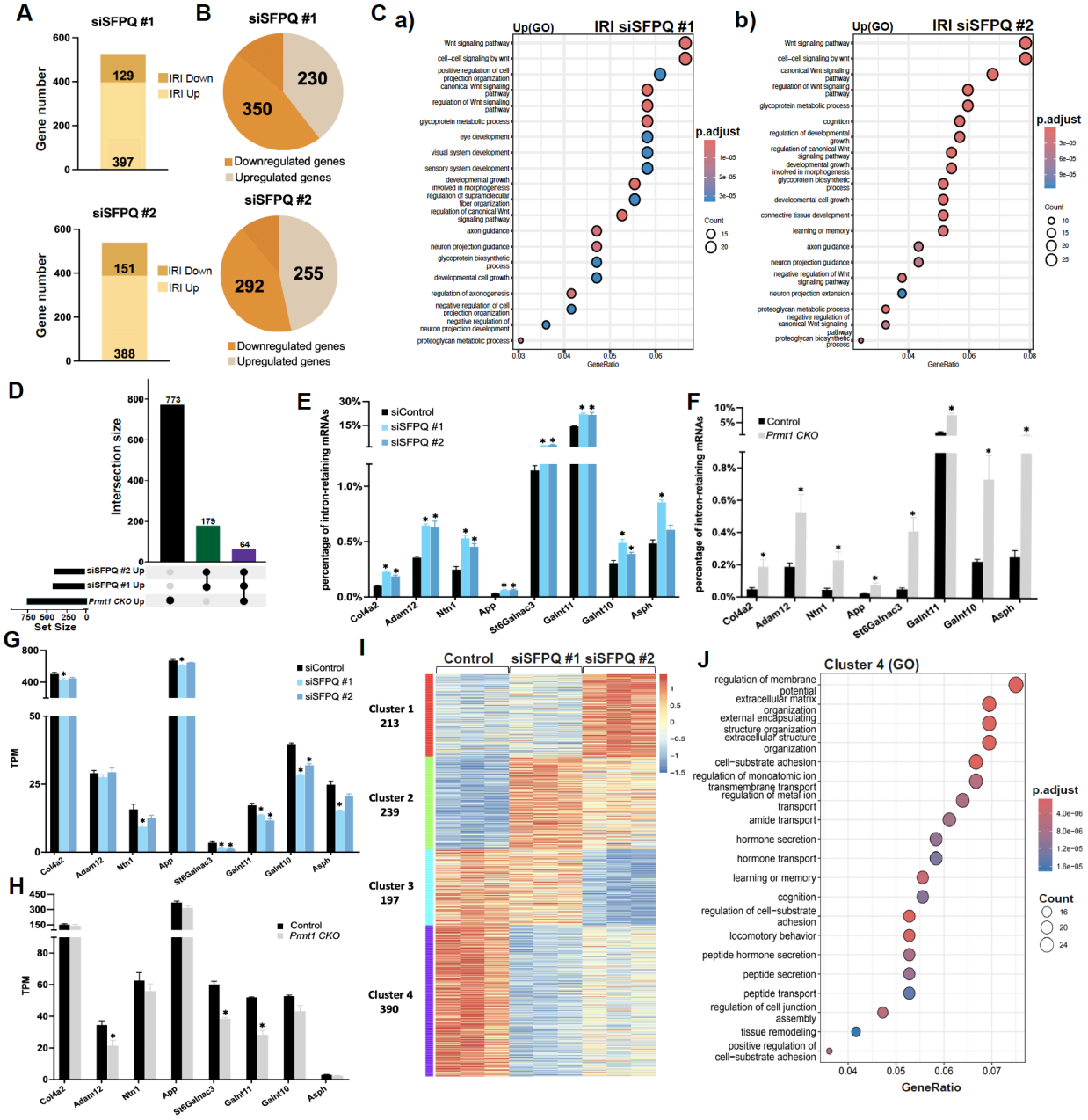

A
TPM
E13.5
E15.5
Prmt1
Prmt2
Prmt3
Prmt5
Prmt6
Prmt7
Prmt8
Prmt9
B
Dissecting mandible from E13.5 mouse embryo
Dissected Mandible
Cell Isolation
Isolated cells from mandible
FACS
Td+ cells
Td- cells
RNA Isolation
RNA Sample
RNA Sequencing
Bioinformatics Analysis
C
a)
SE
IR
MXE
A5SS
A3SS
Number of IR events
Up
Down
Control vs Prmt1 CKO
D
Pex12
Nkx3-2
Nkiras1
1810026B05Rik
Control
Prmt1 CKO
E
Tbx1
F
Control
Prmt1 CKO
percentage of intron-retaining mRNAs
Pex12
Nkx3-2
Nkiras1
1810026B05Rik
Tbx1

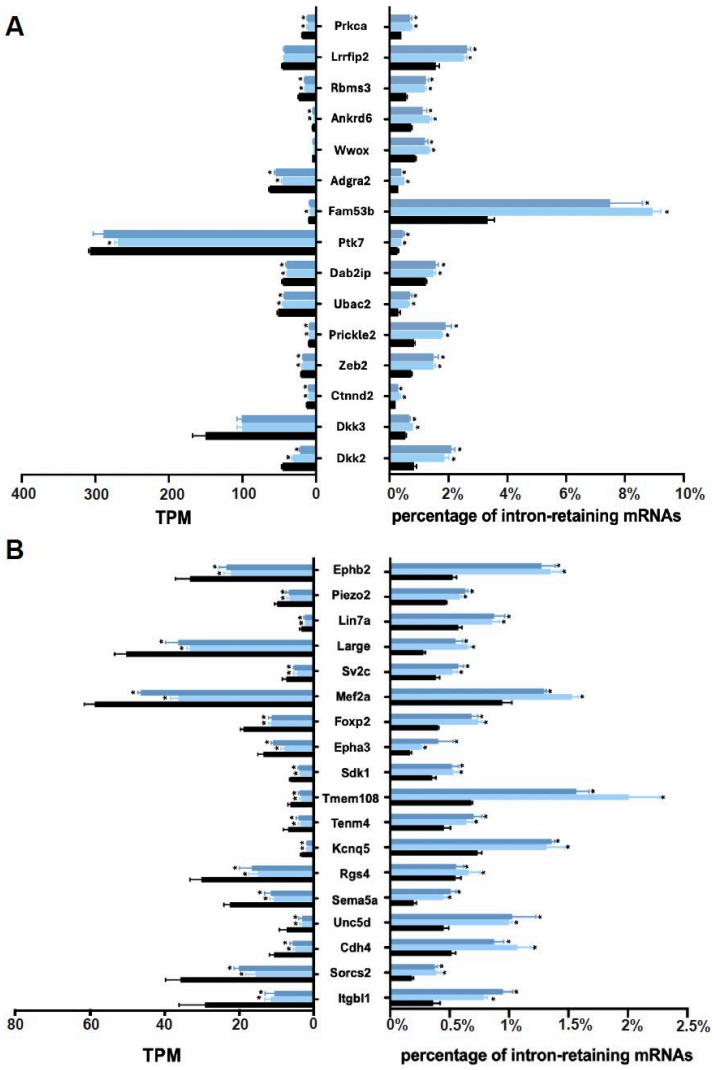


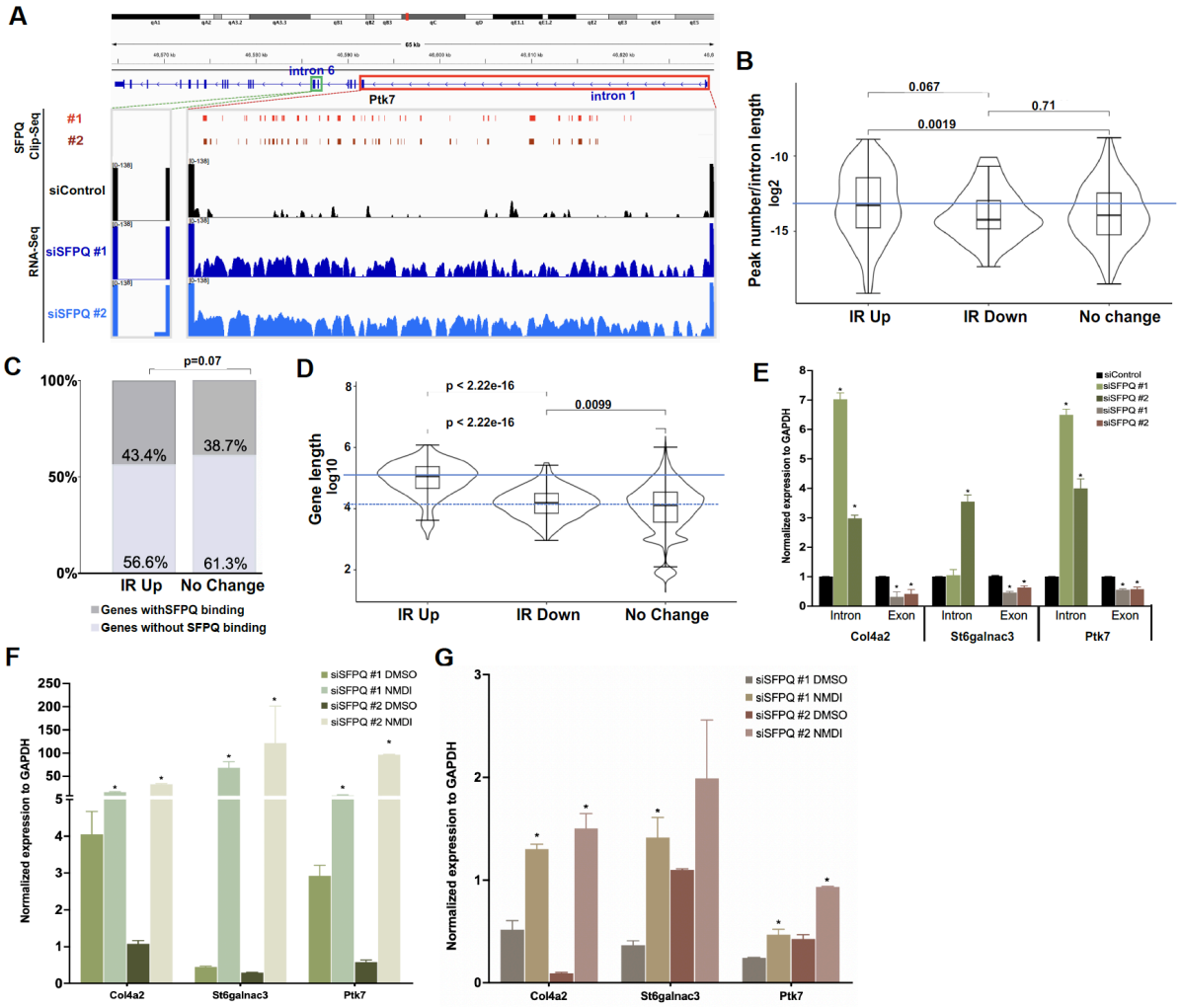
Description	LogP
Connective tissue development	-13.06347
Skeletal system development	-12.3119
Bone development	-11.45758
Cartilage development	-0.973634
Bone morphogenesis	-8.684307
Chondrocyte differentiation	-8.567705
Vasculature development	-12.88828


A
log2(mean IRI)
Prmt1 CKO
Control
B
Bmp7
Intron 3
Cthrc1
Intron 1
Intron 2
Control
Prmt1 CKO
C
Bmp7
Cthrc1
Control
Prmt1 CKO
Percentage of intron-retaining mRNAs
TPM
Intron
Exon
D
Normalized to GAPDH
Control
Prmt1 CKO
Exon
Intron
Dcn
Tnn
Lox
Loxl1
Matn2
Col14a1
Adamts12
Adamts2
Adam12
Scara5
E
log2 (mean IRI)
Prmt1 CKO
Control
Galnt11
Hyal3
Fuca1
Ganc
Galns
Shpk
Lmf1
Aqp11
Csad
Ctbs
Naglu
Hs3st3a1
Mgat5b
Aga
Kera
Glb1
Abcd1
Sgsh
Itih5
Man2b2
Sulf2
Pygl
St6galnac3
Bgn
Pcsk6
Arsb
Dcn
F
St6galnac3
Intron 1
Galnt11
Intron 7
Control
Prmt1 CKO
G
St6galnac3
Galnt11
Control
Prmt1 CKO
Percentage of intron-retaining mRNAs
TPM
Intron
Exon


A
TPM
E13.5
E15.5
SFPQ
SRSF1
EWSR1
TAF15
TRA2B
HnRNPA1
WDR70
G3BP1
B
Mx
Md
C
Control
Prmt1 CKO
MethylSFPQ
DAPI
Merge
Enlarged
D
Mandible
N° of dots per cell
E
F
Maxilla
G
Mes
Epi
H
SFPQ
I
Fluorescence Intensity
J
K
L
Control
Prmt1 CKO
Prmt1 Het
GAPDH
M
Normalized expression to GAPDH


A
siSFPQ #1
Gene number
129
397
IRI Down
IRI Up
siSFPQ #2
151
388
B
siSFPQ #1
230
350
Downregulated genes
Upregulated genes
siSFPQ #2
255
292
C a)
Up(GO)
IRI siSFPQ #1
p.adjust
Count
GeneRatio
b)
IRI siSFPQ #2
D
Intersection size
773
179
64
siSFPQ #2 Up
siSFPQ #1 Up
Prmt1 CKO Up
Set Size
E
siControl
siSFPQ #1
siSFPQ #2
percentage of intron-retaining mRNAs
Col4a2
Adam12
Ntn1
App
St6Galnac3
Galnt11
Galnt10
Asph
F
Control
Prmt1 CKO
percentage of intron-retaining mRNAs
G
TPM
H
I
Control
siSFPQ #1
siSFPQ #2
Cluster 1
213
Cluster 2
239
Cluster 3
197
Cluster 4
390
J
Cluster 4 (GO)
regulation of membrane potential
extracellular matrix organization
external encapsulating structure organization
extracellular structure organization
cell−substrate adhesion
regulation of monoatomic ion transmembrane transport
regulation of metal ion transport
amide transport
hormone secretion
hormone transport
learning or memory
cognition
regulation of cell−substrate adhesion
locomotory behavior
peptide hormone secretion
peptide secretion
peptide transport
regulation of cell junction assembly
tissue remodeling
positive regulation of cell−substrate adhesion
p.adjust
Count
GeneRatio


A
Prkca
Lrrfip2
Rbms3
Ankrd6
Wwox
Adgra2
Fam53b
Ptk7
Dab2ip
Ubac2
Prickle2
Zeb2
Ctnnd2
Dkk3
Dkk2
400 300 200 100 0
TPM
0% 2% 4% 6% 8% 10%
percentage of intron-retaining mRNAs
B
Ephb2
Piezo2
Lin7a
Large
Sv2c
Mef2a
Foxp2
Epha3
Sdk1
Tmem108
Tenm4
Kcnq5
Rgs4
Sema5a
Unc5d
Cdh4
Sorcs2
Itgbl1
80 60 40 20 0
TPM
0% 0.5% 1% 1.5% 2% 2.5%
percentage of intron-retaining mRNAs


A
intron 6
Ptk7
intron 1
SFPQ Clip-Seq
#1
#2
siControl
RNA-Seq
siSFPQ #1
siSFPQ #2
B
0.067
0.71
0.0019
Peak number/intron length log2
IR Up
IR Down
No change
C
p=0.07
100%
50%
0%
43.4%
38.7%
56.6%
61.3%
IR Up
No Change
Genes withSFPQ binding
Genes without SFPQ binding
D
p < 2.22e-16
p < 2.22e-16
0.0099
Gene length log10
IR Up
IR Down
No Change
E
siControl
siSFPQ #1
siSFPQ #2
siSFPQ #1
siSFPQ #2
Normalized expression to GAPDH
Intron
Exon
Col4a2
St6galnac3
Ptk7
F
siSFPQ #1 DMSO
siSFPQ #1 NMDI
siSFPQ #2 DMSO
siSFPQ #2 NMDI
Normalized expression to GAPDH
Col4a2
St6galnac3
Ptk7
G
siSFPQ #1 DMSO
siSFPQ #1 NMDI
siSFPQ #2 DMSO
siSFPQ #2 NMDI
Normalized expression to GAPDH
Col4a2
St6galnac3
Ptk7