

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25

Dysregulation of co-transcriptional alternative splicing underlies CHARGE syndrome

Running Title: **Defective alternative splicing in CHARGE syndrome**

Catherine Bélanger^{1,7}, Félix-Antoine Bérubé-Simard^{1,7}, Elizabeth Leduc¹, Guillaume Bernas¹, Philippe M. Campeau², Seema R. Lalani³, Donna M. Martin⁴, Stephanie Bielas⁵, Amanda Moccia⁵, Anshika Srivastava⁵, David W. Silversides⁶ and Nicolas Pilon^{1*}

¹ Molecular Genetics of Development Laboratory, Department of Biological Sciences and BioMed Research Center, University of Quebec at Montreal (UQAM), Canada.

² Department of Pediatrics and CHU Sainte-Justine Research Centre, University of Montreal, Canada.

³ Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, USA.

⁴ Departments of Pediatrics and Human Genetics, University of Michigan Medical School, Ann Arbor, USA.

⁵ Departments of Human Genetics and Neuroscience, University of Michigan Medical School, Ann Arbor, USA.

⁶ Department of Veterinary Biomedicine, Faculty of Veterinary Medicine, University of Montreal, Canada.

⁷ These authors contributed equally

*Corresponding Author: pilon.nicolas@uqam.ca (NP)

KEY WORDS: Alternative splicing, Ago2, CHARGE syndrome, Chd7, Fam172a, Gene regulation, Mouse model, Neural crest cells, Rapamycin, Sex reversal.

26 **SUMMARY**

27 CHARGE syndrome is a severe developmental disorder with wide phenotypic variability, mainly caused
28 by mutations in *CHD7* (*Chromodomain Helicase DNA-binding protein 7*), known to encode a chromatin
29 remodeler. The genetic lesions responsible for *CHD7* mutation-negative cases are unknown, at least in part
30 because the pathogenic mechanisms underlying CHARGE syndrome remain poorly defined. Here, we
31 report the characterization of a mouse model for *CHD7* mutation-negative cases of CHARGE syndrome
32 generated by insertional mutagenesis of *Fam172a* (*Family with sequence similarity 172, member A*). We
33 show that *Fam172a* plays a key role in the regulation of co-transcriptional alternative splicing, notably by
34 interacting with Ago2 (Argonaute-2) and *Chd7*. Validation studies in a human cohort allow us to propose
35 that dysregulation of co-transcriptional alternative splicing is a unifying pathogenic mechanism for both
36 *CHD7* mutation-positive and -negative cases. We also present evidence that such splicing defects can be
37 corrected *in vitro* by acute rapamycin treatment.

38
39
40
41
42

43 **SIGNIFICANCE STATEMENT**

44 A timely diagnosis is key for both survival and quality of life of CHARGE syndrome patients.
45 Unfortunately, a diagnosis is often difficult to establish, in part because many patients test negative for
46 mutation of *CHD7*, the only gene associated with this condition to date. Identifying additional CHARGE-
47 associated genes would not only help resolve diagnosis issues but could also help in identifying common
48 pathogenic mechanisms, which in turn could lead to highly desirable curative interventions for all patients.
49 Here, *FAM172A* is reported as a new candidate gene for CHARGE syndrome. This discovery has allowed
50 us to reveal a molecular process that is dysregulated in both *CHD7* mutation-positive and -negative cases,
51 such defect being correctable *in vitro* with the well-known drug rapamycin.

52
53
54 \body

55
56
57
58

59 INTRODUCTION

60 CHARGE syndrome affects ~1/10, 000 newborns world-wide and has a very complex clinical presentation
61 (1). This phenotypic complexity is notably highlighted by the acronym CHARGE, which stands for
62 Coloboma of the eye, Hear defects, Atresia of choanae, Retardation of growth/development, Genital
63 abnormalities, and Ear anomalies. However, diagnosis of CHARGE syndrome does not depend on the
64 concomitant presence of all these characteristics, as each one varies from severe to absent in affected
65 children. It is thought that CHARGE syndrome is an underdiagnosed condition with the mildest forms
66 presenting with hypogonadotropic hypogonadism and additional features such as cleft palate, characteristic
67 craniofacial dysmorphisms, inner ear dysplasia and intellectual disability (2, 3). Familial cases have also
68 been reported and are characterized by extensive clinical variability with the transmitting parent often
69 being very mildly affected or even asymptomatic (4-6). Such a wide range of phenotypic presentations has
70 resulted in multiple revisions of the diagnostic criteria over the past decade (2, 3). These criteria have been
71 subdivided into “major” and “minor” features based on their predictive value, with the most recent
72 inclusion rule being two major and an unlimited number of minor features (2). It is noteworthy that the
73 multiple anomalies in CHARGE syndrome can be life-threatening and, consequently, about 30% of
74 affected children die before their 5th birthday (7). Survival and quality of life of these children are tightly
75 linked to age of diagnosis, which is hard to establish not only because of the variable clinical presentation
76 but also because an important subset of cases remains genetically unexplained (1, 2).

77 Heterozygous mutation of *CHD7* (*Chromodomain Helicase DNA-binding protein 7*) is currently the only
78 known genetic cause of CHARGE syndrome (8). Yet, depending on the diagnostic criteria used, up to
79 ~30% of patients do not test positive for *CHD7* mutations (2). Based on high levels of *CHD7* gene
80 expression in neural crest derivatives and proposed roles for *CHD7* in neural crest development, we
81 hypothesized that genetically undefined CHARGE patients may harbor pathogenic variants that affect the
82 integrity of the neural crest cell (NCC) transcriptome. Indeed, previous studies with cellular and mouse
83 models revealed that *CHD7* is a chromatin remodeler that interacts with the SWI/SNF complex for fine-
84 tuning the expression levels of multiple genes at the heart of the NCC gene regulatory network such as
85 *SOX9*, *TWIST1* and *SNAI1* (9-12). It is also important to note that activation of p53 appears as another
86 relevant event in the pathogenic cascade initiated by *CHD7* deficiency (13).

87 Intriguingly, *CHD7* has recently been proposed to be one of several chromatin factors that might influence
88 alternative splicing in mammalian cell lines of non-NCC origin (14). Many studies in such cell lines have
89 shown that chromatin remodelling and histone modifications can not only regulate transcription but also
90 impact alternative splicing by modulating the elongation rate of RNA polymerase II and/or by participating
91 to the recruitment of splicing factors (15-17). In this regard, it is noteworthy that one of the preferred

92 binding partners of CHD7 appears to be PARP1 [poly(ADP-ribose) polymerase 1] (9, 18), which has been
93 recently proposed to directly influence alternative splicing by interacting with chromatin-associated
94 proteins, pre-mRNAs and splicing factors (19). There is also compelling evidence that the interplay
95 between chromatin structure and alternative splicing in mammalian cells involves the RNA interference
96 machinery and most especially the Argonaute members AGO1 and AGO2 (20-22). Of particular interest
97 for CHARGE syndrome, human AGO2 (but not AGO1) and CHD7 have both been reported to interact
98 with the core proteins of the SWI/SNF chromatin-remodelling complexes BRG1 and BAF155 (9, 23).
99 Whether any of these observations is relevant for CHARGE syndrome is, however, unknown.

100 Via a forward genetic screen in mice, we report here the generation and detailed characterization of a new
101 mouse model for *CHD7* mutation-negative CHARGE syndrome. *Fam172a*, the gene disrupted in this
102 mouse model, codes for a nuclear specific Ago2-binding protein that appears to couple transcription with
103 alternative splicing. Analysis of *Chd7* mutant mice and cells from human patients further allow us to
104 suggest that problems with co-transcriptional alternative splicing are likely common to all cases of
105 CHARGE syndrome, and that these splicing problems can be corrected *in vitro* by short exposure to
106 rapamycin.

107

108

109

110 RESULTS

111 **The *Toupee* line is a model for CHARGE syndrome.** *Toupee* is the fourth mouse line issued from a
112 forward genetic screen aimed at identifying genes with key roles in NCCs (24). As with the other lines
113 issued from this screen (25-27), *Toupee* was generated via random insertion of a *tyrosinase* (*Tyr*) minigene
114 into the FVB/N genetic background (28) and was identified using incomplete rescue of pigmentation as an
115 indicator of NCC defects. The name *Toupee* was chosen in recognition of the white spot present on the
116 head of heterozygous animals, which otherwise do not show any overt phenotype (Fig.1A).

117 Intercrosses of *Toupee* heterozygotes produce almost fully depigmented homozygous animals (Fig.1A).
118 Born at approximate Mendelian ratios, about 20% of these homozygotes (*Toupee*^{Tg/Tg}) die before postnatal
119 day (P)25 (SI Appendix, Fig.S1A-B). As summarized in SI Appendix Table S1, *Toupee*^{Tg/Tg} animals also
120 display a complex phenotype mimicking both the “major” and “minor” features of CHARGE syndrome (2,
121 3). As seen in CHARGE patients, *Toupee*^{Tg/Tg} mice have different combinations of such features, which
122 also vary in their severity. Among the major features, *Toupee*^{Tg/Tg} animals display retinal coloboma
123 (Fig.1B), cleft palate (Fig.1C) and hypoplastic semi-circular canals (Fig.1D). As observed in other mouse
124 models of CHARGE syndrome (29, 30), such malformations of semi-circular canals are believed to be the
125 cause of hyperactive circling behavior (Movie S1). The most frequently observed minor features (i.e. in
126 more than 50% of animals) are retarded growth (Fig.1A and SI Appendix, Fig.S1C), genital anomalies
127 (Fig.1E-G) as well as malformation of the heart (Fig.1H) and cranial nerves (Fig.1I). Other notable but less
128 frequently observed minor features include hypoplasia of the thymus (SI Appendix, Fig.S2A), hypoplasia
129 of the olfactory bulbs associated with a decreased sense of smell (SI Appendix, Fig.S2B-C) as well as
130 diverse craniofacial malformations such as asymmetry of facial bones (SI Appendix, Fig.S2D), partial
131 atresia of the oropharynx (SI Appendix, Fig.S2E) and delayed closure of the fontanelles (SI Appendix,
132 Fig.S2F). On very rare occasions, we also found evidence of hypoplastic kidneys (SI Appendix, Fig.S2G),
133 malformed outer ears (SI Appendix, Fig.S2H) and exencephaly/hydrocephaly (SI Appendix, Fig.S2I) in
134 *Toupee*^{Tg/Tg} animals. Finally, we noted that the gastrointestinal tract of mutant animals found dead before
135 P25 was often filled with air bubbles (SI Appendix, Fig.S2J). Combined with the poor postnatal growth
136 and anomalies of the oropharynx and cranial nerves described above, this strongly suggests that a subset of
137 *Toupee*^{Tg/Tg} animals have feeding difficulties – another hallmark of CHARGE syndrome (31).

138 CHARGE syndrome-related genital anomalies include delayed puberty in both sexes, small uterus in
139 females, and cryptorchidism and small penis in males. *Toupee*^{Tg/Tg} mice differ from previously described
140 mouse models with *Chd7* mutations (32, 33) in that both females and males appear to be affected (Fig.1E-
141 F). Interestingly, we found that both male and female *Toupee*^{Tg/Tg} animals are subfertile (SI Appendix,
142 Table S2). We further noticed that the sex ratio is distorted towards females in the *Toupee*^{Tg/Tg} population.

143 This quite unexpected finding prompted us to verify the concordance between phenotypic and
144 chromosomal sex, which revealed that 25% of *Toupee*^{Tg/Tg} genetic males are phenotypic females (Fig.1G).
145 Although male-to-female sex reversal has not been reported in CHARGE syndrome, this phenotype
146 appears not to be specific to the *Toupee* line as it was also detected in ~12% of XY *Chd7*^{Gt/+} animals (SI
147 Appendix, Table S3) – a well-recognized mouse model of CHARGE syndrome that contains a gene
148 trapped allele of *Chd7* (30).

149 To further test whether *Toupee* is a model for CHARGE syndrome, we asked whether the *Toupee* allele
150 could genetically interact with the gene trapped allele of *Chd7* (SI Appendix, Table S3 and Fig.S3). In
151 comparison to corresponding single heterozygotes, *Toupee*^{Tg/+};*Chd7*^{Gt/+} double heterozygotes were found
152 to be markedly smaller at weaning age (SI Appendix, Fig.S3A) and to exhibit a higher frequency of
153 premature postnatal death, circling behavior and male-to-female sex reversal (SI Appendix, Table S3).
154 Coloboma as determined by incomplete closure of the choroidal fissure in e12.5 embryos was also found
155 to be much more severe in double heterozygous mutants (SI Appendix, Fig.S3B). Moreover, the lower
156 than expected number of such mutants at birth suggests that an important subset of them dies *in utero*, an
157 outcome known to occur in *Chd7* homozygous mutants (30). All these observations strongly suggest that
158 *Toupee* is a valid model for CHARGE syndrome.

159 **NCC development is globally affected in *Toupee*^{Tg/Tg} embryos.** Based on previous studies showing that
160 NCCs are a major cell population impaired in CHARGE syndrome (9, 11, 12, 34), we undertook a detailed
161 analysis of this cell lineage in *Toupee*^{Tg/Tg} embryos. The aim of this analysis was to determine which of the
162 main basic cellular processes (i.e. proliferation, survival, migration and/or differentiation) might be
163 affected. We chose to focus on the e10.5 developmental stage since it allows analysis of virtually all key
164 steps of NCC development in different anteroposterior regions of a single embryo, with the cranial region
165 having the “oldest” and the elongating posterior region having the “youngest” NCCs.

166 We first monitored the number of NCCs (Sox10⁺) undergoing apoptosis (_{act}Caspase3⁺) as well as those
167 actively proliferating (Ki67⁺) via immunofluorescence analyses of trunk cross-sections. We detected
168 significant variations in both processes, with proliferation being decreased (Fig.2A) and apoptosis being
169 increased (Fig.2B) in *Toupee*^{Tg/Tg} embryos. Of note, these defects were found not to be exclusive to NCCs
170 (SI Appendix, Fig.S4A-B). Interestingly, we further noticed that mutant NCCs were closer to the dorsal
171 neural tube than normal, suggesting that trunk NCC migration was also affected in *Toupee*^{Tg/Tg} embryos.
172 To directly verify this possibility, we transferred the *Toupee* allele onto a Gata4p[5kb]-RFP (G4-RFP)
173 transgenic background and followed movements of recently induced trunk NCCs via time-lapse imaging.
174 In accordance with the previously described expression pattern of the G4-RFP transgene (35), an anterior
175 to posterior wave of transgene activation was noted in control NCCs delaminating from the dorsal neural

176 tube and ventrally migrating through the somites ([Movie S2](#)). In contrast, *Toupee*^{Tg/Tg} NCCs were found to
177 accumulate in the vicinity of the dorsal neural tube, to migrate more slowly, and to oscillate rather than
178 persist in their ventrally-oriented migration ([Movie S3](#)). Quantification of NCC speed and directionality at
179 the leading edge of migration streams revealed that both parameters are significantly impaired in
180 *Toupee*^{Tg/Tg};G4-RFP explants ([Fig.2C-D](#)). Yet, these defects appeared stronger than expected from the
181 relatively mild phenotypic presentation of postnatal *Toupee*^{Tg/Tg} animals, suggesting that mutant NCCs
182 might still be able to reach some of their final destinations. Accordingly, in similar fashion to what was
183 previously reported for *Tcof1*^{+/-} embryos (36), we found that hindgut colonization by RFP-labeled
184 *Toupee*^{Tg/Tg} enteric NCCs of vagal origin is delayed at e13.5 but not at e15.5 [i.e. a day before and a day
185 after the end of the normal period of colonization; (37)] ([Fig.2E](#) and SI Appendix, [Fig.S4C-D](#)). These
186 observations indicate that some of the NCC migration defects detected in *Toupee*^{Tg/Tg} embryos at early
187 stages can be compensated for at later stages.

188 To complement our cellular analyses, we next evaluated the impact of the *Toupee* mutation on NCC
189 transcriptional profile using RNA sequencing (RNAseq). To this end, we again took advantage of the G4-
190 RFP transgene to specifically recover NCCs from whole e10.5 embryos (control G4-RFP vs
191 *Toupee*^{Tg/Tg};G4-RFP) by FACS prior to deep sequencing of total RNA depleted of ribosomal RNA.
192 Differential analysis of gene expression levels identified several thousand genes that are dysregulated at
193 least 1.5-fold in *Toupee*^{Tg/Tg} NCCs (3488 genes at a DESeq *P*-value cut-off of 0.01), with a bias towards
194 upregulated genes (61.7% upregulated vs 38.3% downregulated) ([Fig.2F](#) and [Dataset S1](#)). Of note, this bias
195 appears much stronger when considering only those genes dysregulated at least 4-fold (622 upregulated
196 genes vs 19 downregulated genes). Gene Ontology (GO) analysis of the 3488 genes list identified 132
197 enriched terms (with ontology level ≥ 5 and $P < 0.05$) that can be classified into 7 main categories (from
198 most to less significant): Metabolic processes; Cell differentiation and morphogenesis; Cell signaling; Cell
199 motility and transport; Control of cell number; Gene expression; and Nervous system development (SI
200 Appendix, [Fig.S5](#)). Interestingly, restricting our analysis to genes that form the constantly expanding NCC
201 gene regulatory network (38-41) revealed that 90.6% (96 out of 106 identified) of those found to be
202 affected are downregulated in *Toupee*^{Tg/Tg} embryos (SI Appendix, [Table S4](#)). Essentially every aspect of
203 NCC development is represented in this list of downregulated genes, from induction/specification to
204 anteroposterior patterning, delamination/migration and formation of specific cell types/structures
205 (peripheral neurons and glia, melanocytes, craniofacial skeleton and enteric nervous system).

206 ***Toupee* is a hypomorphic allele of *Fam172a*.** Using whole-genome sequencing, we localized the
207 transgene insertion site of the *Toupee* line in the last intron of *Fam172a* (*Family with sequence similarity*
208 *172, member A*) – a poorly characterized but highly conserved gene (93% identity with its human

209 orthologue on Chr.5q15) (Fig.3A). Expression analyses using RT-qPCR and immunofluorescence at
210 multiple embryonic stages revealed that *Fam172a* is normally widely expressed during development –
211 including prominent expression in neural tissues – and robustly downregulated (down to ~15% on average)
212 in *Toupee*^{Tg/Tg} embryos (Fig.3B-C and SI Appendix, Figs.S6-S7). This effect appears highly specific as no
213 significant change in gene expression was detected for the other genes flanking the insertion (*Pou5f2*,
214 *Nr2f1* and *A830082K12Rik*) regardless of the stage (e10.5, e13.5 and e15.5), region (cranial and trunk) or
215 cell population (NCCs and non-NCCs) analyzed (Fig.3B-C and SI Appendix, Fig.S6). Importantly, *ex vivo*
216 transfection of a Myc-tagged *Fam172a* expression vector in primary cultures of dissociated e10.5
217 *Toupee*^{Tg/Tg} embryos fully rescued the global proliferation defect previously identified (SI Appendix,
218 Fig.S4A), thereby confirming causality of the *Fam172a* mutation (Fig.3D).

219 Bioinformatics-based analysis of *Fam172a* protein sequences (using the meta sites MyHits, MOTIF and
220 MetaDBsite) notably predicted an esterase-like serine hydrolase motif (G-X-S-X-G) and a bipartite
221 Lys/Arg-rich nuclear localization signal both overlapping with a large domain originally described in *S.*
222 *pombe* Arb2 [Argonaute binding protein-2 (42)] (Fig.3E). *In vitro* labelling of WT and mutant MBP-
223 tagged *Fam172a* with a serine hydrolase-specific fluorescent probe (TAMRA-Fluorophosphonate)
224 confirmed serine hydrolase activity and identified the serine 294 as the nucleophilic residue (Fig.3F), while
225 double-immunofluorescence and co-immunoprecipitation (co-IP) in multiple cell types/tissues showed that
226 *Fam172a* is found in the vicinity of and physically interacts with the nuclear fraction of the Argonaute
227 member Ago2 (Fig.3G-H and SI Appendix, Fig.S8A-B). Of note, *Fam172a* is apparently not a general
228 Argonaute binding protein as no interaction was detected with Ago1 (SI Appendix, Fig.S8D-F). Moreover,
229 in accordance with the nuclear specificity of the *Fam172a*-Ago2 interaction, transfection of the psi-
230 CHECK2-let-7x8 luciferase reporter in primary cultures of dissociated e10.5 embryos revealed that the
231 *Toupee* mutation has no impact on cytoplasmic post-transcriptional gene silencing (SI Appendix, Fig.S8C).
232 Yet, *Fam172a* is present in the cytoplasm where – in accordance with the prediction of a relevant C-
233 terminal retention signal (HEEL) – high amounts are frequently detected in the endoplasmic reticulum (SI
234 Appendix, Fig.S7C).

235 **Characterization of *Fam172a* function suggests a role for alternative splicing in the pathogenesis of**
236 **CHARGE syndrome.** To follow up on our discovery of a nuclear-specific *Fam172a*-Ago2 interaction, we
237 first evaluated the ability of *Fam172a* to interact with chromatin and/or RNA via untargeted chromatin
238 (ChIP) and RNA (RIP) immunoprecipitation assays. Using Neuro2a cells to model the NCC lineage, we
239 found that *Fam172a*, like Ago2, can bind both chromatin and RNA (Fig.4A). Interestingly, co-IP assays in
240 presence of DNase and/or RNase further revealed that efficient formation of the *Fam172a*-Ago2 complex
241 requires at least one type of nucleic acid (Fig.4B and SI Appendix, Fig.S8G).

242 To gain more insight into Fam172a function, we next sought to identify its interactors in an unbiased
243 manner via pull-down assays coupled to mass spectrometry. Using stringent binding conditions and three
244 different fractions of Neuro2a cells (chromatin, nucleoplasm and cytoplasm), this analysis notably revealed
245 a marked enrichment for chromatin proteins and RNA splicing factors among the interacting partners of
246 MBP-tagged Fam172a (SI Appendix, [Tables S5-S7](#)). Taken together with our ChIP, RIP and co-IP data
247 described above, this finding strongly suggested that Fam172a might bridge the chromatin with the
248 alternative splicing machinery as previously suggested for both Chd7 (14) and Ago2 (14, 21) – not
249 identified in our proteomic screen most likely because the Fam172a-Ago2 interaction is impaired in the
250 high-salt binding conditions used (SI Appendix, [Fig.S8H](#)). This intriguing possibility prompted us to re-
251 analyze our transcriptome data for the presence of aberrant splicing events in *Toupee*^{Tg/Tg} NCCs. Using the
252 rMATS computational tool (43) to compare WT and mutant RNAseq data, we discovered that 1166
253 transcripts are aberrantly spliced in *Toupee*^{Tg/Tg} NCCs (using variation in inclusion level > 0.1 and $P < 0.01$
254 as cut-off values) ([Fig.4C](#) and [Dataset S2](#)). Among the different rMATS categories, the vast majority of
255 these affected splicing events were found to fall into either the Skipped exon (52.4%) or the Retained
256 intron (31.5%) categories. GO analysis of the 1166 abnormally spliced transcripts yielded 75 enriched
257 terms (with ontology level ≥ 5 and $P < 0.05$) that can be classified into 6 main categories fully overlapping
258 with those of the differential expression level analysis (from most to less significant): Metabolic processes;
259 Cell motility and transport; Cell signaling; Gene expression; Control of cell number; and Cell
260 differentiation and morphogenesis (SI Appendix, [Fig.S9](#)). Cross-comparison of both RNAseq analyses
261 revealed that 30% of all aberrantly spliced transcripts (350 transcript isoforms out of 1166 in total)
262 correspond to genes also affected at the transcriptional level, this latter group representing 7.4% of all
263 differentially expressed genes (259 single genes out of 3488 in total).

264 To determine whether dysregulation of alternative splicing might represent a common signature for
265 CHARGE syndrome, we then analyzed the splicing pattern of several previously reported Ago2 target
266 genes (21) in both the *Toupee*^{Tg/Tg} and the *Chd7*^{Gt/+} mouse models in comparison to WT. Because the
267 selected genes (*Cd44*, *Col5a3*, *Mical2* and *Ift74*) are all expressed in the developing brain ([www.brain-](http://www.brain-map.org)
268 [map.org](http://www.brain-map.org)), this analysis was performed using whole heads from e12.5 embryos. Remarkably, our RT-qPCR
269 data revealed that variable exons of *Cd44*, *Mical2* and *Ift74* are specifically underrepresented in both
270 *Toupee*^{Tg/Tg} and *Chd7*^{Gt/+} mutants ([Fig.4D](#) and SI Appendix, [Fig.S10A](#)). Among the four tested genes, only
271 *Col5a3* did not appear to be significantly affected at the alternative splicing level. However, in contrast to
272 *Cd44*, *Mical2*, and *Ift74*, global transcription of *Col5a3* as determined by the level of constant exons was
273 found to be robustly downregulated in both mutant lines (SI Appendix, [Fig.S10A](#)). As suggested by
274 previous work (14), we further found that Chd7 can regulate alternative splicing of a large number of
275 genes in different spatiotemporal contexts. Indeed, rMATS-based re-analysis of recently published

276 RNAseq data from P7 granule neuron progenitors (18) revealed dysregulated alternative splicing of 227
277 and 252 transcripts upon heterozygous or homozygous loss of *Chd7*, respectively (using variation in
278 inclusion level > 0.1 and $P < 0.05$ as cut-off values) (Datasets S3-S4 and SI Appendix, Fig.S10B). These
279 aberrant splicing events were found to affect 208 and 226 genes, respectively, 60 of them being detected in
280 both datasets. Given the limited impact on gene expression levels previously reported in the original
281 analysis (only 151 dysregulated genes; most likely because of low sequencing depth) (18), the presence of
282 over 200 splicing defects appears highly significant in this case. It is also interesting to note that these
283 splicing anomalies are distributed into the different rMATS categories in a way very similar to what we
284 observed with *Toupee*^{Tg/Tg} NCCs (Fig.4C), with the Skipped exon and Retained intron categories again
285 being overrepresented (SI Appendix, Fig.S10B).

286 Focussing on the PMA-inducible *Cd44* gene model, ChIP assays in Neuro2a cells then showed that
287 Fam172a and Chd7 are – like Ago2 (21) – both normally present on transcribed chromatin regions
288 containing alternative splice sites (Fig.4E). Importantly, using RNA immunoprecipitation of chromatin
289 (RNA-ChIP), a similar pattern of interaction was also observed on corresponding regions of the chromatin-
290 associated pre-mRNA (Fig.4F). Moreover, co-IP data showed that Chd7 can physically interact with both
291 Fam172a and Ago2 in low-stringency conditions (Fig.4G and SI Appendix, Fig.S8I). Interestingly, this
292 analysis further showed that the presence of exogenous _{Myc}Fam172a seems to promote the Chd7-Ago2
293 interaction (Fig.4G). Together, these results strongly suggest that CHARGE syndrome-related
294 malformations are caused not only by defective transcription (as generally thought) but also by
295 dysregulated alternative splicing of genes that are commonly targeted by Fam172a, Ago2 and Chd7.
296 Although the exact relationship between these three proteins is currently unknown, we propose a model in
297 which Fam172a appears to be required for stabilizing protein-protein interactions at the chromatin-
298 spliceosome interface (Fig.4H).

299 **Alternative splicing defects are common in CHARGE syndrome patients, and these defects can be**
300 **corrected with acute rapamycin treatment.** To test our findings in humans, we generated
301 lymphoblastoid cell lines (LCLs) from genetically and phenotypically heterogeneous CHARGE syndrome
302 patients (SI Appendix, Table S8) and their unaffected parents. All patients were enrolled based on *CHD7*
303 mutation-negative status according to clinical genetic testing and all individuals were characterized by
304 exome sequencing. Careful analysis of *CHD7* sequences revealed a likely deleterious variant
305 (c.5050+1G>T) in one patient while the other six patients were confirmed as *CHD7* mutation-negative,
306 including a mother-child pair. As previously reported for other familial cases with or without *CHD7*
307 mutation (4-6), the mother (ear malformations and choanal atresia) was found to be less affected than her
308 child (ear malformations, choanal atresia, retarded growth, genital hypoplasia and a small kidney). Very

309 interestingly, exome sequencing data revealed the presence of two rare *FAM172A* variants in this mother-
310 child pair. The variant c.682G>C (p.Glu228Gln; frequency of 1.7e-05 in ExAC browser) was detected in
311 both the mother and child as well as in the unaffected maternal grand-mother, while the second variant,
312 c.916C>T (p.Arg306*; not found in ExAC browser), was only detected in the child. No paternal sample
313 was available for this analysis. Of note, the co-occurrence of two rare and likely deleterious variants in
314 single patients has also been reported for *CHD7* (44, 45), including at least one confirmed case of
315 compound heterozygosity (46). To experimentally evaluate the pathogenicity of the identified *FAM172A*
316 variants, we introduced the same mutations in the corresponding conserved residues of murine Fam172a.
317 In accordance with their location in the Arb2 domain, both variants were found to prevent the formation of
318 a Fam172a-Ago2 complex in co-IP assays, and the nonsense mutation was further found to affect Fam172a
319 protein levels (SI Appendix, Fig.S11A). Both variants also appeared to impact the subcellular distribution
320 of Fam172a (SI Appendix, Fig.S11B). Taken together with the finding of the c.682G>C variant in another
321 mother-child pair of CHARGE patients from a different cohort (SI Appendix, Table S9), these results
322 support the hypothesis that *FAM172A* deficiency may play a role in the pathogenesis of a subset of *CHD7*
323 mutation-negative cases of CHARGE syndrome. Intriguingly, obvious signs of coloboma were not
324 observed in any of these four *FAM172A*-affected individuals, suggesting they might even form a distinct
325 group of people with CHARGE (or CHARGE-like) syndrome. No convincing candidate gene could be
326 identified in the other four *CHD7* mutation-negative patients.

327 Although our cohort was relatively small, it allowed us to directly test the hypothesis that dysregulation of
328 alternative splicing is common to all cases of CHARGE syndrome irrespective of their *CHD7* or
329 *FAM172A* mutation status. To this end, we used the same approach as described above for mice and
330 evaluated alternative splicing of *CD44*, *COL5A3*, *MICAL2* and *IFT74* in our collection of LCLs via RT-
331 qPCR. Strikingly, the splicing patterns of all four genes were found to be severely dysregulated in all 7
332 patients when compared to their unaffected parents (Fig.5A-C and SI Appendix Fig.S12). For most of the
333 tested splicing events, the same trend could be noted regarding the inclusion/exclusion of a given exon.
334 However, no global rule could be observed. Indeed, tested variable exons appeared preferentially excluded
335 for *CD44* and preferentially included for *COL5A3*, while either inclusion or exclusion was promoted in the
336 case of *MICAL2* and *IFT74*. All these results thus support the idea that impairment of co-transcriptional
337 alternative splicing is a unifying pathogenic mechanism for all cases of CHARGE syndrome. Interestingly,
338 our RT-qPCR analyses also revealed that *FAM172A* expression is decreased in a majority of patients
339 within this cohort, thereby further supporting its role in the pathogenesis of CHARGE syndrome (SI
340 Appendix, Fig.S13A).

341 We then reasoned that our set of LCLs could serve in addition as a tool to explore the possibility of
342 developing a small molecule-based therapy for CHARGE syndrome. As a proof of concept, we evaluated
343 rapamycin, a well-known TOR inhibitor for which analogs (i.e. rapalogs) have been approved by the U.S.
344 Food and Drug Administration for multiple clinical conditions (47). One of the best known roles of the
345 nutrient-sensing TOR pathway is to stimulate ribosome biogenesis by activating the transcription of
346 ribosomal protein genes (48). Accordingly, one of the main effects of acute rapamycin treatments is the
347 downregulation of ribosomal protein gene expression, which in turn promotes splicing of other pre-
348 mRNAs due to relief of competition for a likely limiting pool of splicing factors (49). Since rapamycin-
349 mediated repression of ribosomal protein gene expression was reported to occur within a few minutes in
350 yeast (49), we treated LCLs with a moderate dose of rapamycin (10 μ M) during 30 minutes and then
351 analyzed alternative splicing of *CD44*, *COL5A3*, *MICAL2* and *IFT74* as described above (Fig.5D). Of note,
352 *CD44* is a known mTOR target gene (50), meaning that absolute levels of *CD44* constant exons also
353 allowed us to confirm efficacy of the rapamycin treatment (SI Appendix, Fig.S13B). Remarkably, such
354 acute rapamycin treatment was sufficient for positively impacting the splicing defects detected in all
355 patients of our cohort (Fig.5D). Both previously downregulated and upregulated splicing events were
356 found to be corrected, as indicated by the splicing fold change levels that became closer to the levels
357 observed in their respective unaffected parents (i.e. splicing fold change of 1). Analysis of *FAM172A*
358 expression in this context revealed a very modest increase, if any, upon rapamycin treatment (SI Appendix,
359 Fig.S13B).

360 These most encouraging results finally prompted us to directly evaluate the *in vivo* potential of rapamycin
361 as therapeutic agent for CHARGE syndrome. To this end, we administered a moderate dose of rapamycin
362 (1mg/kg) via IP injection to pregnant females from *Toupee*^{Tg/Tg} intercrosses twice a day between e9.5-
363 e11.5 (i.e. during the peak of NCC migration), and then analyzed e12.5 embryos for the presence of
364 coloboma – the phenotype with highest penetrance in *Toupee*^{Tg/Tg} animals (SI Appendix, Table S1). Very
365 interestingly, this approach allowed us to decrease the incidence of coloboma by 50% (Fig.5E). It should
366 however be noted that, although the rapamycin dose that we used was lower than the previously described
367 teratogenic dose (3mg/kg), most treated embryos did not look healthy, showing signs of growth retardation
368 and/or resorption (SI Appendix, Fig.S13D). We thus conclude that rapamycin is a promising drug for the
369 treatment of CHARGE syndrome but that more work will be required in order to balance therapeutic
370 benefits with adverse effects.

371

372

373

374 DISCUSSION

375 The *Toupee* insertional mutant line is the first viable mouse model for *CHD7* mutation-negative cases of
376 CHARGE syndrome. Extensive characterization of this mouse line notably allowed us to unveil a
377 previously overlooked partially penetrant male-to-female sex reversal phenotype, which we found to be
378 present in a subset of *Chd7* mutants as well. We further discovered that *Fam172a*, the *Toupee* mutated
379 gene, codes for a nuclear-specific binding partner of Ago2 that appears to couple chromatin structure with
380 alternative splicing. Finally, by analyzing a genetically heterogeneous cohort of CHARGE patients, we
381 found that dysregulation of co-transcriptional alternative splicing could well be the pathogenic mechanism
382 underlying all cases of CHARGE syndrome. This finding is expected to help guiding gene discovery for
383 *CHD7* mutation-negative cases and, most importantly, opens up the possibility of eventually developing
384 small molecule-based therapeutic interventions for this devastating pediatric disease.

385 **Fam172a, a highly conserved protein with enigmatic serine hydrolase activity.** Using different
386 experimental approaches, we confirmed the presence of four functional domains/motifs in murine
387 *Fam172a*: an Arb2-like domain, a bi-partite NLS, an esterase-like serine hydrolase motif and an ER
388 retention signal. With the exception of the ER retention signal, all these domains/motifs are also present in
389 *Fam172a* orthologs from evolutionary distant species such as human (417 aa; 93% overall identity with
390 mouse ortholog), zebrafish (415 aa; 71% overall identity with mouse ortholog), lamprey (311 aa; 47%
391 overall identity with mouse ortholog) and *C. elegans* (313 aa; 33% overall identity with mouse ortholog).
392 The C-terminal ER retention signal (HEEL) appears as a “recent” evolutionary gain, being present in
393 mouse and human orthologs but not in those from zebrafish, lamprey and *C. elegans*. From an evolutionary
394 point of view, this suggests that *Fam172a* is thus especially important in the nucleus, where our data
395 collectively reveal that it can connect the alternative splicing machinery to specific transcribed chromatin
396 regions (Fig.4H).

397 Since *Fam172a* is devoid of any predicted DNA or RNA binding domain, its ability to target specific
398 transcribed regions is most likely mediated by some of the chromatin- and spliceosome-associated proteins
399 that compose most of the *Fam172a* interactome (SI Appendix, Fig.S5 and Table S5). In fact, all of our
400 observations point to the chromatin-spliceosome interface as the preferred site of *Fam172a*'s action in the
401 nucleus, a notion further supported by our findings that the *Fam172a*-Ago2 interaction is destabilized by
402 DNase/RNase treatments (Fig.4B and SI Appendix, Fig.S8G). Of note, given the particular enrichment in
403 relevant gene expression-associated GO terms in our NCC RNAseq data (SI Appendix, Figs.S5 and S9),
404 *Fam172a* also appears to regulate transcription and/or splicing of several genes through indirect means (i.e.
405 through regulation of genes encoding other regulatory proteins). For example, the expression levels or
406 splicing of many genes encoding splicing accessory factors of the hnRNP family (*Hnrnpa2b1*, *Hnrnp2*,

407 *Hnrnp3*, *Hnrnpk* and *Hnrnpull1*) are dysregulated in *Toupee*^{Tg/Tg} NCCs (Datasets S1 and S2). Although we
408 cannot currently exclude the possibility that Fam172a might also regulate transcription and splicing
409 independently of each other, the existence of indirect splicing targets represents a very plausible reason as
410 to why gene expression levels and splicing patterns only partially overlap in our NCC RNAseq data.

411 More work will definitely be required to determine the nature of the interactions within the Fam172a
412 interactome (i.e. direct protein-protein or indirect as part of larger complexes) and the role played by the
413 Arb2-like domain in this regard. Determining which part of Fam172a function relies on its intriguing
414 esterase-like serine hydrolase activity as well as identifying its physiological substrate(s) will represent
415 other exciting – and most likely challenging – lines of future research. Like other “metabolic” serine
416 hydrolases, esterases form a structurally heterogeneous group of enzymes for which very few endogenous
417 substrates have been reported (51). To the best of our knowledge, CIB [CCG1/TAF_{II}250-interacting factor
418 B] (52), also named ABHD14B [α/β -hydrolase domain protein 14B] (53), is the only other known
419 example of a nuclear protein with esterase-like serine hydrolase activity (54). Although both CIB (54) and
420 Fam172a (this work) are similarly involved in transcription-related processes, their respective
421 physiological substrates are expected to be different. Indeed, the sequences flanking the nucleophilic serine
422 in CIB (S-X-S-X-S) differ substantially from the consensus observed in Fam172a (G-X-S-X-G). Moreover,
423 we failed to detect *in vitro* cleavage of *p*-nitrophenyl butyrate by purified MBP-tagged Fam172a (not
424 shown) whereas this general hydrolase substrate was reported cleavable by CIB (54).

425 **Is CHARGE syndrome a spliceosomopathy?** While the term neurocristopathy is appropriate to highlight
426 NCCs as the main cell population affected in CHARGE syndrome, the term spliceosomopathy should also
427 be considered to highlight this potential underlying pathogenic mechanism. Defined by the presence of
428 germline mutation of spliceosome-associated proteins, spliceosomopathies recently emerged as a subgroup
429 of rare diseases that notably include neurodegenerative conditions such as *retinitis pigmentosa*, spinal
430 muscle atrophy and amyotrophic lateral sclerosis as well as multiple craniofacial disorders such as Nager
431 syndrome, cerebrocostomandibular syndrome, Richieri-Costa-Pereira syndrome, Burn-McKeown
432 syndrome and the Guion-Almeida type of mandibulofacial dysostosis (55, 56). As observed in CHARGE
433 syndrome, most spliceosomopathies are believed to be caused by *de novo* dominant mutations although
434 familial cases as well as examples with autosomal recessive inheritance have also been reported (55).
435 Moreover, in accordance with all being considered neurocristopathies as well (57, 58), the phenotypic
436 presentation of the craniofacial disorders mentioned above can be very similar to CHARGE syndrome.
437 One especially striking example is the Guion-Almeida type of mandibulofacial dysostosis, which is caused
438 by mutation of the spliceosomal GTPase-encoding gene *EFTUD2* (59). Like CHARGE syndrome, clinical
439 diagnosis of the Guion-Almeida type of mandibulofacial dysostosis may be difficult to establish due to

440 variable phenotypic presentation characterized by different combinations of both major (external and inner
441 ear anomalies, choanal atresia/cleft palate) and minor (facial asymmetry, micrognathia, intellectual
442 disability, growth retardation, heart defects, oesophageal atresia, genito-urinary defects) features of
443 CHARGE syndrome (55). In line with such an extensive overlap between CHARGE syndrome and the
444 Guion-Almeida type of mandibulofacial dysostosis, *EFTUD2* mutations have been reported in 7 patients
445 initially referred for possible CHARGE syndrome (60). Of note, the wide range of possible differential
446 diagnosis for craniofacial spliceosomopathies is also highlighted by the fact that *EFTUD2* mutations have
447 been reported in patients initially diagnosed for Nager (61) and Feingold (60) syndromes.

448 Dysregulation of alternative splicing as a common pathogenic mechanism for CHARGE syndrome is in
449 agreement with previous findings regarding the contribution of activated p53. Indeed, while p53
450 expression and activity have been shown to be increased upon *CHD7* deficiency in both mouse and human
451 cells, *p53* heterozygosity only partially rescues the malformations found in *Chd7*-null mouse embryos (13).
452 Moreover, *p53* knockdown completely failed to rescue *chd7* loss in zebrafish (62) while p53 protein levels,
453 localization and activity – as deduced from the unaffected expression levels of its CHARGE-associated
454 downstream targets *Noxa*, *Perp* and *Dr5* (13) – do not appear to be affected in *Toupee*^{Tg/Tg} embryos (SI
455 Appendix, Fig.S14 and Dataset S1). Considering that p53 may be activated upon splicing impairment (63),
456 these observations thus suggest that p53 activation might contribute partially to the pathogenic cascade
457 downstream of dysregulated splicing and only in a context-dependent manner (i.e. in some *CHD7*
458 mutation-positive cases). Importantly, our findings are also in agreement with previous studies showing
459 that knockdown of *kdm2b* can partially rescue *chd7* loss in zebrafish (62). Indeed, given the prominent role
460 for H3K36 methylation in promoting alternative splicing (15, 64, 65), it is reasonable to think that
461 deficiency of a H3K36 demethylase such as *Kdm2b* could help compensate for reduced splicing efficiency.
462 Along these lines, it is further noteworthy that the *Drosophila* ortholog of *Chd7* (*Kismet*) has been reported
463 to promote H3K36 methylation by *Ash1* (66) – the *Drosophila* ortholog of *Ash1l* which we found to be
464 part of the *Fam172a* interactome (SI Appendix, Table S5).

465 Like other genes coding for spliceosome-associated proteins, *Fam172a* appears almost ubiquitously
466 expressed in the developing embryo (Figs.3C and SI Appendix, Fig.S7A-B). Alternative splicing is a very
467 pervasive process (67), raising the question as to why mutation of spliceosome genes results in tissue-
468 specific malformations. One possibility would be that some cell types are more vulnerable to splicing
469 defects than others, as suggested for retinal neurons in the context of *retinitis pigmentosa* (68). An
470 explanation for such differential vulnerability could be that, although spliceosome-associated proteins are
471 ubiquitous, their relative proportion may fluctuate as a function of tissue-specific variations in expression
472 levels. These differences can even be functionally amplified through competition between pre-mRNAs that

473 have different affinities for limiting splicing factors (49). In this context, the phenotypic outcome would
474 also depend on compensatory and/or antagonistic activity of available splicing regulatory proteins on their
475 cognate *cis*-acting RNA elements (69). Another important aspect to consider is that many genes are
476 believed to exert tissue-specific functions through tissue-specific transcript isoforms, which may also
477 depend on the tissue-specific activity of alternative enhancers, silencers and promoters (67). Considering
478 all of the above, it is not surprising that deficiency of a given splicing regulatory protein can lead to tissue-
479 specific perturbations of a large repertoire of transcripts, as reported for *SMN* mutation-dependent cases of
480 spinal muscular atrophy (70). In conclusion, regardless of the exact mechanism involved, NCCs appear as
481 vulnerable as retinal and motor neurons to splicing defects.

482 **Towards the development of small molecule-based treatment strategies for CHARGE syndrome.** In
483 the current study, we demonstrated that an acute rapamycin treatment is sufficient to correct alternative
484 splicing defects in LCLs from CHARGE patients and partially rescue coloboma in *Toupee*^{Tg/Tg} embryos,
485 thereby providing the first hope that small molecule-based strategies can be envisaged for CHARGE
486 syndrome and other spliceosomopathies. While we recognize that the specific therapeutic value of
487 rapamycin may depend on the relative weight of transcription *versus* splicing defects in the pathogenesis of
488 CHARGE syndrome (which is currently impossible to determine), it is important to bear in mind that
489 splicing defects are sufficient by themselves to cause neurocristopathy-related malformations (55).
490 Therefore, the correction of splicing defects in the context of CHARGE syndrome can only be
491 advantageous. As highlighted by our first *in utero* attempt (Fig.5E and SI Appendix, Fig.S13D), additional
492 work will clearly be required to determine to what extent rapamycin treatments can be effective and safe at
493 the organismal level. In theory, both *in utero* and early postnatal treatments are possible (71, 72) if care is
494 taken not to administer teratogenic doses (73). Feasibility of *in utero* treatments is notably supported by *ex*
495 *vivo* studies with chick embryos, which revealed that a 48h exposure to a low dose of rapamycin (200nM)
496 is well-tolerated by wild-type NCCs (72). Since pre-natal diagnosis of CHARGE syndrome is particularly
497 difficult to establish (74), it is also very interesting to note that an early post-natal treatment with
498 rapamycin can correct NCC-related craniofacial bone defects in mice (71). Apart from issues related to
499 dosage, frequency and duration of rapamycin treatments, other outstanding questions include: What is the
500 global impact of rapamycin treatment on the NCC transcriptome? Do acute treatments have long-lasting
501 effects? Can all or only a subset of CHARGE syndrome-related malformations be prevented/corrected? Is
502 it possible to correct some these malformations/dysfunctions postnatally? Both *Toupee* and *Chd7* mutant
503 mouse lines will be especially useful to answer these very important questions.

504

505 MATERIALS AND METHODS

506 **Animals.** Experiments involving mice were performed following the biomedical research guidelines of the
507 Canadian Council of Animal Care (CCAC). Further details about animal ethics approval could be found in
508 SI Methods. The *Toupee* transgenic mouse line was generated by standard pronuclear injection of a
509 previously described *tyrosinase* minigene in FVB/N zygotes (26), while the *Chd7^{Gt/+}* transgenic mouse line
510 has been described previously (30). Details about epistasis studies can be found in SI Methods. For some
511 studies in embryos, the *Toupee* allele was backcrossed on the *Gata4p[5kb]-RFP* transgenic background
512 (FVB/N) in which the *DsRed2* fluorescent marker is expressed by NCCs (35). Embryos were all generated
513 by natural mating and staged by considering noon of the day of vaginal plug detection as embryonic day
514 (e) 0.5. Except for studies involving time-lapse imaging or FACS, control and mutant embryos to be
515 compared were obtained from the same litter and processed in parallel. Olfaction tests were performed in
516 accordance with a previously described protocol (32) on a total of 20 adult mice (10 males and 10 females;
517 aged between 8 to 12 weeks) per genotype (control FVB/N vs *Toupee^{Tg/Tg}*). Further details about olfaction
518 tests can be found in SI Methods.

519 **Analysis of prenatal and postnatal tissues.** Haematoxylin-eosin staining of 10µm paraffin-embedded
520 tissue sections, immunofluorescence staining of 30µm cryosections or dissociated cells, neurofilament
521 immunohistochemistry of whole embryos, alizarin red-alcian blue staining of skeletons, and
522 microcomputed tomography analysis of cranial bones were performed as previously described (25, 26, 75,
523 76). Details about all antibodies used in this study, as well as image acquisition, can be found in SI
524 Appendix [Table S10](#) and SI Methods, respectively.

525 **Ex vivo time-lapse imaging of NCCs.** Live imaging of RFP-labelled NCCs was adapted from a
526 previously described suspension culture technique (27). Migration speed and persistence were calculated
527 for 10 individual cells from at least 4 embryos per genotype. See SI Methods for further details.

528 **FACS and RNA extraction.** FACS-mediated recovery of RFP-labelled NCCs was adapted from
529 previously described protocols (27) as described in SI Methods. RNA extraction from FACS-recovered
530 cells or whole embryonic tissues was performed using the RNeasy Plus Purification Kit (Qiagen) in
531 accordance with the manufacturer's protocol.

532 **Genotyping PCR and RT-qPCR.** The *Toupee* allele of *Fam172a* and the gene-trapped allele of *Chd7*
533 were genotyped by PCR using standard Taq DNA polymerase (Feldan) and primers flanking the respective
534 insertion/deletion. Chromosomal sexing was performed by PCR amplification of the male-specific *Zfy*
535 (423 bp) and the *Smcx/Smcy* paralogous gene pair, the later generating a single 300 bp amplicon in XX
536 genomes that is combined with a second 280 bp amplicon in XY genomes due to different intronic lengths.

537 RT-qPCR analyses were performed on 50 ng of total RNA using the Ssofast EvaGreen Supermix and
538 C1000 Touch thermal cycler (BioRad) in accordance with the manufacturer's protocol. The *Gapdh* gene
539 was used for normalization of absolute expression levels while constant exons of selected genes were used
540 for normalization of splicing events. Details about the primers used for RT-qPCR can be found in SI
541 Appendix [Table S11](#) while all other primers are listed in SI Appendix [Table S12](#).

542 **High throughput genome and transcriptome sequencing.** Whole genome and transcriptome sequencing
543 was adapted from previously described protocols (27) as described in SI Methods. Differential analysis of
544 transcript levels was performed using the DESeq and edgeR packages whereas differential analysis of
545 alternative splicing was performed using rMATS. Previously published RNAseq data of P7 granule neuron
546 progenitors [WT, *Chd7*-het and *Chd7*-null; single-end 50-bp reads, about 20 million reads per sample (18)]
547 were retrieved from the GEO database (<https://www.ncbi.nlm.nih.gov/geo/>) with accession numbers
548 GSM1857543 to GSM1857549. *P*-values were corrected via the Benjamini-Hochberg method. The GO
549 analyses were performed using GOToolBox (<http://genome.crg.es/GOToolBox/>) and REVIGO
550 (<http://revigo.irb.hr/>).

551 **Plasmid constructs and mutagenesis.** The complete 1251 bp *Fam172a* ORF (Ensembl transcript ID:
552 ENSMUST00000163257.7) was amplified by RT-PCR from an e12.5 FVB/N embryo head using the
553 Superscript II Reverse Transcriptase (ThermoFisher Scientific) and the Platinum Taq DNA polymerase
554 (ThermoFisher Scientific) in accordance with the manufacturer's instructions. Following cloning in the
555 pGEM-T vector (Promega) and validation by Sanger sequencing, the *Fam172a* ORF was sub-cloned into
556 the pIRES2-EGFP mammalian expression vector (Clontech; native or modified in house to include an N-
557 term Myc tag) as well as into the pMAL-c5X vector (New England Biolabs). Site-directed mutagenesis of
558 the Arb2 domain (p.Glu229Gln and p.Arg307*) and the serine hydrolase motif (p.Ser294Ala) was
559 performed in relevant *Fam172a* expression vectors using a previously described PCR-based approach (77).
560 Details about the primers used for cloning and mutagenesis of *Fam172a* can be found in SI Appendix
561 [Table S12](#). The psi-CHECK2-let-7x8 vector was as previously described (78).

562 **Cell culture and transfection.** Propagation of Neuro2a (N2a) and COS7 cell lines as well as cell
563 transfection using Genejuice reagents (Novagen) were performed as previously described (79). For
564 primary culture of e10.5 embryonic cells, details can be found in SI Methods. For co-localization studies,
565 dissociated cells from FVB/N embryos were analyzed by immunofluorescence after 16h of culture. For
566 rescue of the *Toupee* proliferation defect, dissociated cells from *Toupee*^{Tg/Tg} embryos were transfected with
567 1µg of _{Myc}Fam172a-IRES-EGFP or empty IRES-EGFP expression vector 2h after plating and analyzed via
568 immunofluorescence after another 48h of culture.

569 **Co-IP, RIP, ChIP and RNA-ChIP.** Each IP-based assay in Neuro2a or COS7 cells was performed using
570 a confluent 100-mm plate previously transfected with either Fam172a- or _{Myc}Fam172a-IRES-EGFP
571 expression vector. For co-IP assays in cell lines and mouse tissues, the preparation of cell extracts as well
572 as western blotting were performed using previously described protocols (80), as detailed in SI Methods.
573 The contribution of nucleic acids was verified by pre-treatment of protein extracts with 100 µg/ml DNase
574 I (Sigma) and/or 100 µg/ml RNase A (Sigma) during 2h at 4°C. Untargeted ChIP and RIP experiments
575 were both adapted from previously described protocols (81, 82), as described in SI Methods. Targeted
576 ChIP and RNA-ChIP experiments of *Cd44* variable regions were performed on 3x10⁶ Neuro2A cells using
577 a previously described protocol (21), as described in SI Methods.

578 **MBP fusion proteins and serine hydrolase assay.** Details about MBP, _{MBP}Fam172a and
579 _{MBP}Fam172a^{S294A} protein production, as well as protein purification can be found in SI Methods. Serine
580 hydrolase activity was evaluated using the ActivX TAMRA-FP serine hydrolase probe (ThermoFisher
581 Scientific) in accordance with the manufacturer's instructions (SI Methods). TAMRA fluorescence was
582 directly detected in the gel using the NightOWL LB983 imaging system (Berthold), while loaded proteins
583 were detected via standard silver nitrate staining.

584 **Affinity purification coupled to tandem mass spectrometry analysis.** Neuro2a cell extracts were
585 fractionated as previously described (83). Samples were air-dried and sent to the Proteomics Discovery
586 Platform of the *Institut de recherches cliniques de Montreal* (IRCM). Details about protein inclusion
587 criteria could be found in SI Methods. Peptides were identified using the Mascot 2.5.1 search engine
588 (Matrix Science) and the UniProt_Mus_Musculus_txid10090 database. Results were analyzed using the
589 Scaffold 4 software (Proteome Software Inc.).

590 **Human studies.** Families provided informed consent on studies approved by the respective institutional
591 review board of the Baylor College of Medicine (experimental cohort for this study) and the University of
592 Michigan Medical School (replication cohort for this study). Exome sequencing and analysis was done as
593 previously described (84). Chromosomal sex was determined from sequencing results according to the
594 presence of the *SRY* gene. Lymphoblastoid cell lines were established (by Epstein-Barr virus infection) and
595 maintained as described previously (85). For each RT-qPCR experiment, about 3x10⁵ cells were used per
596 cell line (details about primers used can be found in SI Appendix [Table S11](#)). For alternative splicing
597 rescue experiments, cells were treated with either 10 µM rapamycin (Sigma) or vehicle only (ethanol) for
598 30 minutes just before being processed for RT-qPCR.

599 **Statistics.** Where applicable, data are presented as the mean ± SEM with the number of independent
600 biological replicates (n) indicated in the figure and/or legend. GraphPad Prism software version 6.0 was

601 used to determine significance of differences via the 2-tailed Student's *t*-test. Differences were considered
602 statistically significant when *P*-values were less than 0.05.

603

604 **AUTHOR CONTRIBUTIONS**

605 Conceived and designed the experiments: CB FABS NP. Performed the experiments: CB FABS EL GB.
606 Analyzed the data: CB FABS NP. Contributed reagents/materials/analysis tools: PMC SRL DMM SB AM
607 AS DWS. Drafted the manuscript: CB FABS NP. Revised the manuscript: PMC SRL DMM DWS NP.

608

609 **ACKNOWLEDGEMENTS**

610 The authors wish to thank D. Raiwet (U. of Montreal) for the phenotypic screening of the *Toupee* line, O.
611 Souchkova (UQAM) for help with mouse husbandry and breeding, D. Flipo (UQAM) for assistance with
612 confocal imaging and FACS analyses, R. Moreau's laboratory (UQAM) for help with microcomputed
613 tomography, and Dr M. Simard for kindly providing the psi-CHECK2-let-7x8 vector. The authors also
614 wish to acknowledge Dr J.W. Belmont and P. Hernandez (Baylor College of Medicine) for generously
615 providing the lymphoblastoid cell lines. The Proteomics Discovery Platform of the *Institut de recherches*
616 *cliniques de Montréal* (IRCM) is thanked for the mass spectrometry analyses whereas the Massively
617 Parallel Sequencing Platform as well as the Bioinformatics Platform of the McGill University and *Génom*
618 *Québec* Innovation Center are thanked for the deep sequencing and the analyses of sequencing results,
619 respectively. This work was funded by grants from the CHARGE Syndrome Foundation (USA) and the
620 Canadian Institutes of Health Research [grant #376482] to NP. CB is also supported by a doctoral
621 scholarship from the *Fondation du Grand Défi Pierre Lavoie* (Canada). NP is a Senior research scholar of
622 the *Fonds de la recherche du Québec – Santé* as well as the recipient of the UQAM Research Chair on
623 Rare Genetic Diseases. No competing interests declared.

624

625

626

627 **REFERENCES**

628

- 629 1. Hsu P, *et al.* (2014) CHARGE syndrome: a review. *Journal of paediatrics and child health* 50(7):504-511.
- 630 2. Hale CL, Niederriter AN, Green GE, & Martin DM (2016) Atypical phenotypes associated with pathogenic
- 631 CHD7 variants and a proposal for broadening CHARGE syndrome clinical diagnostic criteria. *Am J Med*
- 632 *Genet A* 170(2):344-354.
- 633 3. van Ravenswaaij-Arts CM, Blake K, Hoefsloot L, & Verloes A (2015) Clinical utility gene card for: CHARGE
- 634 syndrome - update 2015. *Eur J Hum Genet* 23(11).
- 635 4. Delahaye A, *et al.* (2007) Familial CHARGE syndrome because of CHD7 mutation: clinical intra- and
- 636 interfamilial variability. *Clin Genet* 72(2):112-121.
- 637 5. Jongmans MC, *et al.* (2008) Familial CHARGE syndrome and the CHD7 gene: a recurrent missense mutation,
- 638 intrafamilial recurrence and variability. *Am J Med Genet A* 146A(1):43-50.
- 639 6. Lalani SR, *et al.* (2006) Spectrum of CHD7 mutations in 110 individuals with CHARGE syndrome and
- 640 genotype-phenotype correlation. *Am J Hum Genet* 78(2):303-314.
- 641 7. Zentner GE, Layman WS, Martin DM, & Scacheri PC (2010) Molecular and phenotypic aspects of CHD7
- 642 mutation in CHARGE syndrome. *Am J Med Genet A* 152A(3):674-686.
- 643 8. Vissers LE, *et al.* (2004) Mutations in a new member of the chromodomain gene family cause CHARGE
- 644 syndrome. *Nat Genet* 36(9):955-957.
- 645 9. Bajpai R, *et al.* (2010) CHD7 cooperates with PBAF to control multipotent neural crest formation. *Nature*
- 646 463(7283):958-962.
- 647 10. Bouazoune K & Kingston RE (2012) Chromatin remodeling by the CHD7 protein is impaired by mutations
- 648 that cause human developmental disorders. *Proc Natl Acad Sci U S A* 109(47):19238-19243.
- 649 11. Fujita K, Ogawa R, & Ito K (2016) CHD7, Oct3/4, Sox2, and Nanog control FoxD3 expression during mouse
- 650 neural crest-derived stem cell formation. *The FEBS journal* 283(20):3791-3806.
- 651 12. Schulz Y, *et al.* (2014) CHD7, the gene mutated in CHARGE syndrome, regulates genes involved in neural
- 652 crest cell guidance. *Hum Genet* 133(8):997-1009.
- 653 13. Van Nostrand JL, *et al.* (2014) Inappropriate p53 activation during development induces features of
- 654 CHARGE syndrome. *Nature* 514(7521):228-232.
- 655 14. Allemand E, *et al.* (2016) A Broad Set of Chromatin Factors Influences Splicing. *PLoS Genet* 12(9):e1006318.
- 656 15. Luco RF, Allo M, Schor IE, Kornblihtt AR, & Misteli T (2011) Epigenetics in alternative pre-mRNA splicing.
- 657 *Cell* 144(1):16-26.
- 658 16. Saint-Andre V, Batsche E, Rachez C, & Muchardt C (2011) Histone H3 lysine 9 trimethylation and
- 659 HP1gamma favor inclusion of alternative exons. *Nat Struct Mol Biol* 18(3):337-344.
- 660 17. Sims RJ, 3rd, *et al.* (2007) Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of
- 661 transcription postinitiation factors and pre-mRNA splicing. *Mol Cell* 28(4):665-676.
- 662 18. Feng W, *et al.* (2017) Chd7 is indispensable for mammalian brain development through activation of a
- 663 neuronal differentiation programme. *Nature communications* 8:14758.
- 664 19. Matveeva E, *et al.* (2016) Involvement of PARP1 in the regulation of alternative splicing. *Cell discovery*
- 665 2:15046.
- 666 20. Allo M, *et al.* (2009) Control of alternative splicing through siRNA-mediated transcriptional gene silencing.
- 667 *Nat Struct Mol Biol* 16(7):717-724.
- 668 21. Ameyar-Zazoua M, *et al.* (2012) Argonaute proteins couple chromatin silencing to alternative splicing. *Nat*
- 669 *Struct Mol Biol* 19(10):998-1004.
- 670 22. Kalantari R, Chiang CM, & Corey DR (2016) Regulation of mammalian transcription and splicing by Nuclear
- 671 RNAi. *Nucleic Acids Res* 44(2):524-537.
- 672 23. Carissimi C, *et al.* (2015) ARGONAUTE2 cooperates with SWI/SNF complex to determine nucleosome
- 673 occupancy at human Transcription Start Sites. *Nucleic Acids Res* 43(3):1498-1512.
- 674 24. Pilon N (2016) Pigmentation-based insertional mutagenesis is a simple and potent screening approach for
- 675 identifying neurocristopathy-associated genes in mice. *Rare Diseases* 4(1):e1156287.
- 676 25. Bergeron KF, *et al.* (2015) Male-Biased Aganglionic Megacolon in the TashT Mouse Line Due to
- 677 Perturbation of Silencer Elements in a Large Gene Desert of Chromosome 10. *PLoS Genet* 11(3):e1005093.

- 678 26. Bergeron KF, *et al.* (2016) Upregulation of the Nr2f1-A830082K12Rik gene pair in murine neural crest cells
679 results in a complex phenotype reminiscent of waardenburg syndrome type 4. *Disease models &*
680 *mechanisms* 9(11):1283-1293.
- 681 27. Soret R, *et al.* (2015) A collagen VI-dependent pathogenic mechanism for Hirschsprung's disease. *J Clin*
682 *Invest* 125(12):4483-4496.
- 683 28. Methot D, Reudelhuber TL, & Silversides DW (1995) Evaluation of tyrosinase minigene co-injection as a
684 marker for genetic manipulations in transgenic mice. *Nucleic Acids Res* 23(22):4551-4556.
- 685 29. Bosman EA, *et al.* (2005) Multiple mutations in mouse Chd7 provide models for CHARGE syndrome. *Hum*
686 *Mol Genet* 14(22):3463-3476.
- 687 30. Hurd EA, *et al.* (2007) Loss of Chd7 function in gene-trapped reporter mice is embryonic lethal and
688 associated with severe defects in multiple developing tissues. *Mamm Genome* 18(2):94-104.
- 689 31. Bergman JE, *et al.* (2010) Death in CHARGE syndrome after the neonatal period. *Clin Genet* 77(3):232-240.
- 690 32. Bergman JE, Bosman EA, van Ravenswaaij-Arts CM, & Steel KP (2010) Study of smell and reproductive
691 organs in a mouse model for CHARGE syndrome. *Eur J Hum Genet* 18(2):171-177.
- 692 33. Layman WS, Hurd EA, & Martin DM (2011) Reproductive dysfunction and decreased GnRH neurogenesis in
693 a mouse model of CHARGE syndrome. *Hum Mol Genet* 20(16):3138-3150.
- 694 34. Asad Z, *et al.* (2016) Rescue of neural crest-derived phenotypes in a zebrafish CHARGE model by Sox10
695 downregulation. *Hum Mol Genet* 25(16):3539-3554.
- 696 35. Pilon N, Raiwet D, Viger RS, & Silversides DW (2008) Novel pre- and post-gastrulation expression of Gata4
697 within cells of the inner cell mass and migratory neural crest cells. *Dev Dyn* 237(4):1133-1143.
- 698 36. Barlow AJ, Dixon J, Dixon MJ, & Trainor PA (2012) Balancing neural crest cell intrinsic processes with those
699 of the microenvironment in Tcof1 haploinsufficient mice enables complete enteric nervous system
700 formation. *Hum Mol Genet* 21(8):1782-1793.
- 701 37. Bergeron KF, Silversides DW, & Pilon N (2013) The developmental genetics of Hirschsprung's disease. *Clin*
702 *Genet* 83(1):15-22.
- 703 38. Lake JI & Heuckeroth RO (2013) Enteric nervous system development: migration, differentiation, and
704 disease. *American journal of physiology. Gastrointestinal and liver physiology* 305(1):G1-24.
- 705 39. Simoes-Costa M & Bronner ME (2015) Establishing neural crest identity: a gene regulatory recipe.
706 *Development* 142(2):242-257.
- 707 40. Nelms BL & Labosky PA (2010) *Transcriptional Control of Neural Crest Development*, Developmental
708 Biology, San Rafael (CA)).
- 709 41. Theveneau E & Mayor R (2012) Neural crest delamination and migration: from epithelium-to-mesenchyme
710 transition to collective cell migration. *Dev Biol* 366(1):34-54.
- 711 42. Buker SM, *et al.* (2007) Two different Argonaute complexes are required for siRNA generation and
712 heterochromatin assembly in fission yeast. *Nat Struct Mol Biol* 14(3):200-207.
- 713 43. Shen S, *et al.* (2014) rMATS: robust and flexible detection of differential alternative splicing from replicate
714 RNA-Seq data. *Proc Natl Acad Sci U S A* 111(51):E5593-5601.
- 715 44. Balasubramanian R, *et al.* (2014) Functionally compromised CHD7 alleles in patients with isolated GnRH
716 deficiency. *Proc Natl Acad Sci U S A* 111(50):17953-17958.
- 717 45. Bartels CF, Scacheri C, White L, Scacheri PC, & Bale S (2010) Mutations in the CHD7 gene: the experience of
718 a commercial laboratory. *Genetic testing and molecular biomarkers* 14(6):881-891.
- 719 46. Xu C, *et al.* (2016) Compound Heterozygote CHD7 Mutations in a Female with Kallmann Syndrome and
720 CHARGE Features. *Endocrine Society's 98th Annual Meeting and Expo:Abstract* OR21-25.
- 721 47. Li J, Kim SG, & Blenis J (2014) Rapamycin: one drug, many effects. *Cell metabolism* 19(3):373-379.
- 722 48. Martin DE, Soulard A, & Hall MN (2004) TOR regulates ribosomal protein gene expression via PKA and the
723 Forkhead transcription factor FHL1. *Cell* 119(7):969-979.
- 724 49. Munding EM, Shiue L, Katzman S, Donohue JP, & Ares M, Jr. (2013) Competition between pre-mRNAs for
725 the splicing machinery drives global regulation of splicing. *Mol Cell* 51(3):338-348.
- 726 50. Hsieh AC, *et al.* (2012) The translational landscape of mTOR signalling steers cancer initiation and
727 metastasis. *Nature* 485(7396):55-61.
- 728 51. Bachovchin DA & Cravatt BF (2012) The pharmacological landscape and therapeutic potential of serine
729 hydrolases. *Nature reviews. Drug discovery* 11(1):52-68.

- 730 52. Padmanabhan B, Kuzuhara T, Mizuno H, & Horikoshi M (2000) Purification, crystallization and preliminary
731 X-ray crystallographic analysis of human CCG1-interacting factor B. *Acta crystallographica. Section D,*
732 *Biological crystallography* 56(Pt 11):1479-1481.
- 733 53. Long JZ & Cravatt BF (2011) The metabolic serine hydrolases and their functions in mammalian physiology
734 and disease. *Chemical reviews* 111(10):6022-6063.
- 735 54. Padmanabhan B, Kuzuhara T, Adachi N, & Horikoshi M (2004) The crystal structure of CCG1/TAF(II)250-
736 interacting factor B (CIB). *J Biol Chem* 279(10):9615-9624.
- 737 55. Lehalle D, *et al.* (2015) A review of craniofacial disorders caused by spliceosomal defects. *Clin Genet*
738 88(5):405-415.
- 739 56. Scotti MM & Swanson MS (2016) RNA mis-splicing in disease. *Nat Rev Genet* 17(1):19-32.
- 740 57. Snider TN & Mishina Y (2014) Cranial neural crest cell contribution to craniofacial formation, pathology,
741 and future directions in tissue engineering. *Birth defects research. Part C, Embryo today : reviews*
742 102(3):324-332.
- 743 58. Trainor PA & Andrews BT (2013) Facial dysostoses: Etiology, pathogenesis and management. *American*
744 *journal of medical genetics. Part C, Seminars in medical genetics* 163C(4):283-294.
- 745 59. Lines MA, *et al.* (2012) Haploinsufficiency of a spliceosomal GTPase encoded by EFTUD2 causes
746 mandibulofacial dysostosis with microcephaly. *Am J Hum Genet* 90(2):369-377.
- 747 60. Gordon CT, *et al.* (2012) EFTUD2 haploinsufficiency leads to syndromic oesophageal atresia. *J Med Genet*
748 49(12):737-746.
- 749 61. Bernier FP, *et al.* (2012) Haploinsufficiency of SF3B4, a component of the pre-mRNA spliceosomal complex,
750 causes Nager syndrome. *Am J Hum Genet* 90(5):925-933.
- 751 62. Balow SA, *et al.* (2013) Knockdown of fbxl10/kdm2bb rescues chd7 morphant phenotype in a zebrafish
752 model of CHARGE syndrome. *Dev Biol* 382(1):57-69.
- 753 63. Allende-Vega N, *et al.* (2013) p53 is activated in response to disruption of the pre-mRNA splicing
754 machinery. *Oncogene* 32(1):1-14.
- 755 64. Sorenson MR, *et al.* (2016) Histone H3K36 methylation regulates pre-mRNA splicing in *Saccharomyces*
756 *cerevisiae*. *RNA biology* 13(4):412-426.
- 757 65. Zhu K, *et al.* (2017) SPOP-containing complex regulates SETD2 stability and H3K36me3-coupled alternative
758 splicing. *Nucleic Acids Res* 45(1):92-105.
- 759 66. Dorigi KM & Tamkun JW (2013) The trithorax group proteins Kismet and ASH1 promote H3K36
760 dimethylation to counteract Polycomb group repression in *Drosophila*. *Development* 140(20):4182-4192.
- 761 67. Wang ET, *et al.* (2008) Alternative isoform regulation in human tissue transcriptomes. *Nature*
762 456(7221):470-476.
- 763 68. Tanackovic G, *et al.* (2011) PRPF mutations are associated with generalized defects in spliceosome
764 formation and pre-mRNA splicing in patients with retinitis pigmentosa. *Hum Mol Genet* 20(11):2116-2130.
- 765 69. Matera AG & Wang Z (2014) A day in the life of the spliceosome. *Nature reviews. Molecular cell biology*
766 15(2):108-121.
- 767 70. Zhang Z, *et al.* (2008) SMN deficiency causes tissue-specific perturbations in the repertoire of snRNAs and
768 widespread defects in splicing. *Cell* 133(4):585-600.
- 769 71. Fang F, *et al.* (2015) Neural Crest-Specific TSC1 Deletion in Mice Leads to Sclerotic Craniofacial Bone Lesion.
770 *Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral*
771 *Research* 30(7):1195-1205.
- 772 72. Wang XY, *et al.* (2015) High glucose environment inhibits cranial neural crest survival by activating
773 excessive autophagy in the chick embryo. *Scientific reports* 5:18321.
- 774 73. Hentges KE, *et al.* (2001) FRAP/mTOR is required for proliferation and patterning during embryonic
775 development in the mouse. *Proc Natl Acad Sci U S A* 98(24):13796-13801.
- 776 74. Colin E, *et al.* (2012) Prenatal diagnosis of CHARGE syndrome by identification of a novel CHD7 mutation in
777 a previously unaffected family. *Prenat Diagn* 32(7):692-694.
- 778 75. Sanchez-Ferras O, *et al.* (2016) A direct role for murine Cdx proteins in the trunk neural crest gene
779 regulatory network. *Development* 143(8):1363-1374.

- 780 76. Sanchez-Ferras O, Bernas G, Laberge-Perrault E, & Pilon N (2014) Induction and dorsal restriction of Paired-
781 box 3 (Pax3) gene expression in the caudal neuroectoderm is mediated by integration of multiple pathways
782 on a short neural crest enhancer. *Biochim Biophys Acta* 1839(7):546-558.
- 783 77. Pilon N, *et al.* (2003) Porcine SRY promoter is a target for steroidogenic factor 1. *Biol Reprod* 68(4):1098-
784 1106.
- 785 78. Johnston M, Geoffroy MC, Sobala A, Hay R, & Hutvagner G (2010) HSP90 protein stabilizes unloaded
786 argonaute complexes and microscopic P-bodies in human cells. *Mol Biol Cell* 21(9):1462-1469.
- 787 79. Sanchez-Ferras O, *et al.* (2012) Caudal-related homeobox (Cdx) protein-dependent integration of canonical
788 Wnt signaling on paired-box 3 (Pax3) neural crest enhancer. *J Biol Chem* 287(20):16623-16635.
- 789 80. Beland M, *et al.* (2004) Cdx1 autoregulation is governed by a novel Cdx1-LEF1 transcription complex. *Mol*
790 *Cell Biol* 24(11):5028-5038.
- 791 81. Berube-Simard FA, Prudhomme C, & Jeannotte L (2014) YY1 acts as a transcriptional activator of Hoxa5
792 gene expression in mouse organogenesis. *PLoS One* 9(4):e93989.
- 793 82. Niranjana Kumari S, Lasda E, Brazas R, & Garcia-Blanco MA (2002) Reversible cross-linking combined with
794 immunoprecipitation to study RNA-protein interactions in vivo. *Methods* 26(2):182-190.
- 795 83. Lavalleye-Adam M, *et al.* (2013) Discovery of cell compartment specific protein-protein interactions using
796 affinity purification combined with tandem mass spectrometry. *Journal of proteome research* 12(1):272-
797 281.
- 798 84. Campeau PM, *et al.* (2014) The genetic basis of DOORS syndrome: an exome-sequencing study. *Lancet*
799 *Neurol* 13(1):44-58.
- 800 85. Ling PD & Huls HM (2005) Isolation and immortalization of lymphocytes. *Curr Protoc Mol Biol* Chapter
801 28:Unit 28 22.

802

803

804

805 **FIGURE LEGENDS**

806

807 **Figure 1: Major and minor features of CHARGE syndrome in *Toupee*^{Tg/Tg} mice.**

808 (A) Comparison between wild-type (WT), *Toupee*^{Tg/+} and *Toupee*^{Tg/Tg} animals at postnatal day (P) 25. (B)
809 Bright field images of e12.5 eyes showing incomplete closure of the choroidal fissure in *Toupee*^{Tg/Tg}
810 embryos (n=10 WT, n=14 *Toupee*^{Tg/Tg}). (C) H&E-stained sagittal sections of e18.5 heads (n=10 WT, n=14
811 *Toupee*^{Tg/Tg}), with asterisks indicating cleft palate. (D) Bright field images of Alizarin red- and Alcian
812 blue-stained inner ears from P25 mice (n=7 WT, n=11 *Toupee*^{Tg/Tg}). CC, common crus; AntSCC, anterior
813 semi-circular canal. (E) Overview of genital anomalies in P25 *Toupee*^{Tg/Tg} males (left panels,
814 cryptorchidism; middle panels, smaller androgen-sensitive seminal vesicles and penis; right panel,
815 normally sized testes). T, testes; B1, bladder (F) *Toupee*^{Tg/Tg} females present hypoplastic uterine horns at
816 P25 (left panel) as well as delayed opening of the vaginal cavity after P20 (right panel). (G) PCR-based
817 sexing of *Toupee*^{Tg/Tg} animals revealed male-to-female sex reversal for 25% of XY animals. (H) A subset
818 of *Toupee*^{Tg/Tg} mice display heart malformation (left panel, increased weight at P25; right panel,
819 hypertrophy of the left ventricle at e15.5). LV, left ventricle. (I) Whole-mount staining of cranial nerves
820 in e10.5 embryos using anti-neurofilament immunohistochemistry. *Toupee*^{Tg/Tg} embryos exhibit
821 supernumerary sprouting in the facial (VII) nerve (arrows) as well as extensive mingling between
822 glossopharyngeal (IX) and vagal (X) nerves (arrowheads). * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$ (Student's *t*-
823 test). Scale bar: 50 μm (B), 1 mm (C-D), 500 μm (H-I).

824

825

826 **Figure 2: Global impairment of NCC development in *Toupee*^{Tg/Tg} embryos.** (A-B) Quantification of
827 Ki67⁺ proliferating (A) and ^{act}Casp3⁺ apoptotic (B) NCCs (expressed in % of Sox10⁺ NCCs) in 30 μm
828 transverse sections of e10.5 embryos at hindlimb level (*Toupee*^{Tg/+} vs *Toupee*^{Tg/Tg}). (C-D) Quantification of
829 NCC migration speed (C) and movement persistence (D) in e10.5 embryos (WT;G4-RFP vs
830 *Toupee*^{Tg/Tg};G4-RFP). (E) Quantification of the extent of colon colonization by enteric NCCs (expressed in
831 % of colon length from cecum to anus) at e13.5 and e15.5 (*Toupee*^{Tg/+};G4-RFP vs *Toupee*^{Tg/Tg};G4-RFP).
832 (F) Volcano plot summarizing a RNAseq-based analysis of differential gene expression levels in e10.5
833 NCCs (WT;G4-RFP vs *Toupee*^{Tg/Tg};G4-RFP). Only genes modulated at least 1.5-fold with a *P*-value below
834 0.01 are displayed. * $P \leq 0.05$, *** $P \leq 0.001$ (Student's *t*-test).

835

836

837 **Figure 3: Functional characterization of *Fam172a*, the *Toupee* causative gene.** (A) Schematic
838 representation of the *Toupee* transgene insertion site in cytoband C1 of chromosome 13 (adapted from the
839 Ensembl website), where ~10 copies of a *tyrosinase* minigene are inserted in a 2327bp deletion in the last
840 intron of *Fam172a* (red box). (B) RT-qPCR analysis of gene expression in e10.5 embryos (WT;G4-RFP vs
841 *Toupee*^{Tg/Tg};G4-RFP). Transcript levels of genes around the transgene insertion site were monitored in
842 FACS-recovered NCCs (+) and non-NCCs (-) from the head and the trunk (n=3 per condition). (C)
843 Immunofluorescence labeling of the Fam172a protein (red) in sagittal sections of e10.5 mouse embryos
844 (*Toupee*^{Tg/+} vs *Toupee*^{Tg/Tg}, n=3 per genotype). DAPI (blue) was used to counterstain nuclei. (D)
845 Quantification of proliferation in cultures of dissociated e10.5 embryos (WT vs *Toupee*^{Tg/Tg}; n=3 per
846 condition) after transfection with a _{Myc}Fam172a-expressing vector. (E) Schematic representation of the
847 functional domains of mouse Fam172a protein compared to its human orthologue. The serine in position
848 215 (highlighted in red) corresponds to the supernumerary amino acid between mouse (417aa) and human
849 (416aa) sequences. NLS, nuclear localization signal; Ser hydrolase, esterase-like serine hydrolase motif;
850 Arb2-like, domain homologous to yeast Arb2 (Argonaute binding protein 2); ER, endoplasmic reticulum
851 retention signal. (F) Hydrolase activity is demonstrated by covalent binding of a TAMRA-labelled
852 fluorophosphonate probe on _{MBP}Fam172a protein (upper panel). No binding is detected on _{MBP}Fam172a
853 bearing a S294A mutation nor on the MBP tag alone. The lower panel shows that silver-stained MBP-
854 tagged Fam172a proteins (~90 kDa; slightly higher in presence of fluorophosphonate probe) and MBP
855 alone (42kDa) were all present at the expected size in the same gel. (G) Co-IP assays using cytoplasmic
856 (Gapdh⁺) and nuclear (H3⁺) fractions of Neuro2a cells transfected with a _{Myc}Fam172a-expressing vector
857 (n=3). Inputs correspond to 10% of protein extracts used for IP. (H) Double immunofluorescence labeling
858 of Fam172a and Ago2 in dissociated cells obtained from WT e10.5 embryos (n=7), and counterstained
859 with DAPI. The lower right panel shows the overlap of Fam172a and Ago2 signals (Pearson's correlation
860 coefficient of 0.82). * $P \leq 0.05$ (Student's *t*-test). Scale bar: 500 μ m (C), 25 μ m (H).

861

862

863 **Figure 4: Both *Fam172a* and *Chd7* play a role in alternative splicing.** (A) Untargeted ChIP and RIP
864 assays in Neuro2A cells transfected with a *Fam172a*-expressing plasmid (n=3). (B) Co-IP assays using
865 RNase- and/or DNase-treated whole cell extracts of Neuro2A cells transfected with a _{Myc}Fam172a
866 expression vector. Inputs correspond to 10% of protein extracts used for IP (n=3 per condition). Impact of
867 each treatment on the integrity of proteins, DNA and RNA is shown in SI Appendix Fig.S8G. (C) Donut
868 chart showing the distribution of the 1166 differentially modulated alternative splicing events ($P < 0.01$;
869 variation in inclusion level ≥ 0.1) in *Toupee*^{Tg/Tg} e10.5 NCCs. Upward- and downward-pointing arrows
870 indicate splicing events that are over- and under-represented in *Toupee*^{Tg/Tg} e10.5 NCCs, respectively. (D)

871 RT-qPCR analysis of splicing events for *Cd44*, *Col5a3*, *Mical2* and *Ift74* in *G4-RFP* (WT),
872 *Toupee^{Tg/Tg};G4-RFP* (*Toupee^{Tg/Tg}*) and *Chd7^{Gt/+}* heads of e12.5 embryos (n=5 per genotype). Expression
873 levels of variable regions are normalized with levels of corresponding constant regions (indicated between
874 parentheses). **(E-F)** ChIP (E) and RNA-ChIP (F) assays of the PMA-inducible *Cd44* gene in Neuro2a cells
875 **(G)** Co-IP assays in Neuro2A cells transfected with empty or *MycFam172a*-expressing vector. Inputs
876 correspond to 10% of protein extracts used for IP (n=3 per condition). **(H)** Potential mode of action of
877 *Fam172a* and *Chd7* in Ago2-mediated alternative splicing (adapted from (21)). * $P \leq 0.05$, ** $P \leq 0.01$, ***
878 $P \leq 0.001$ (Student's *t*-test).

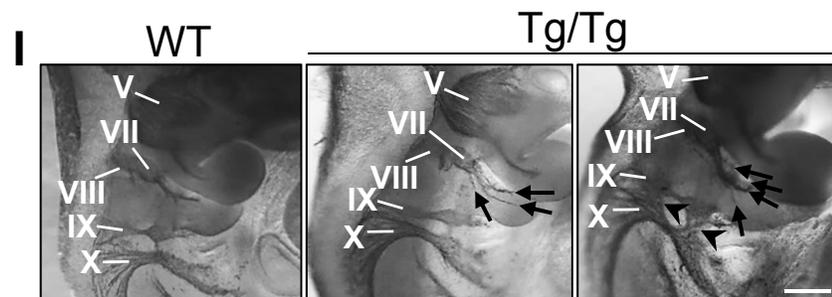
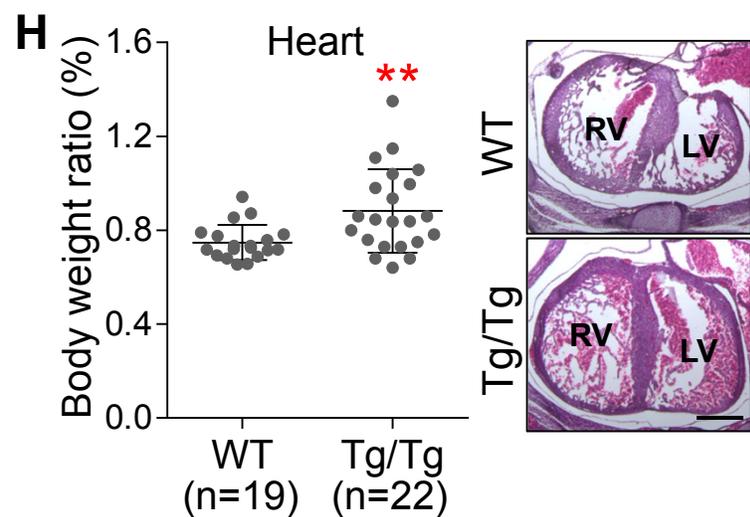
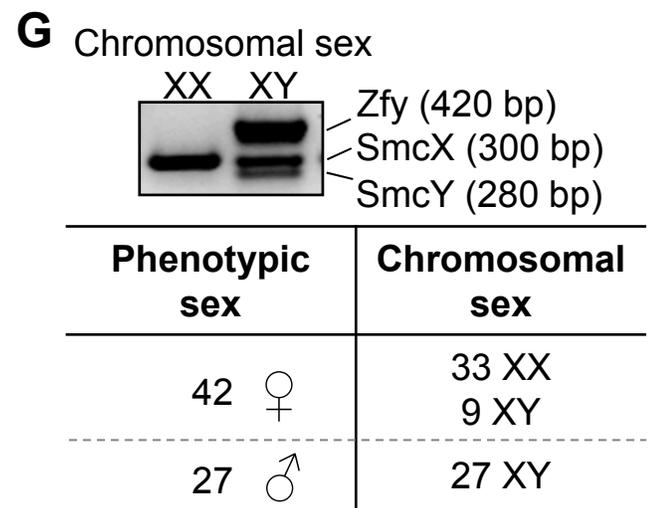
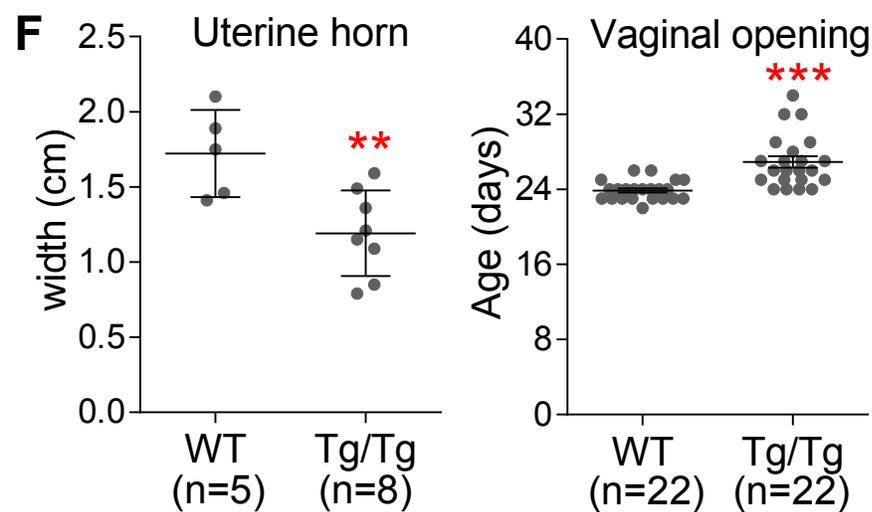
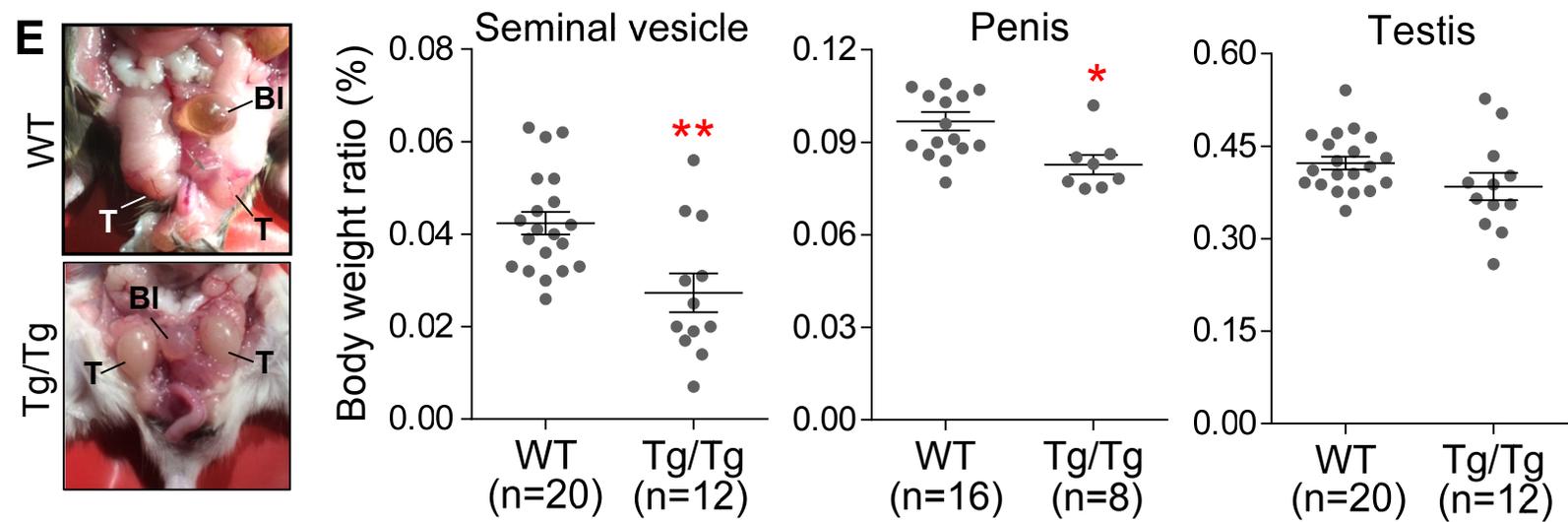
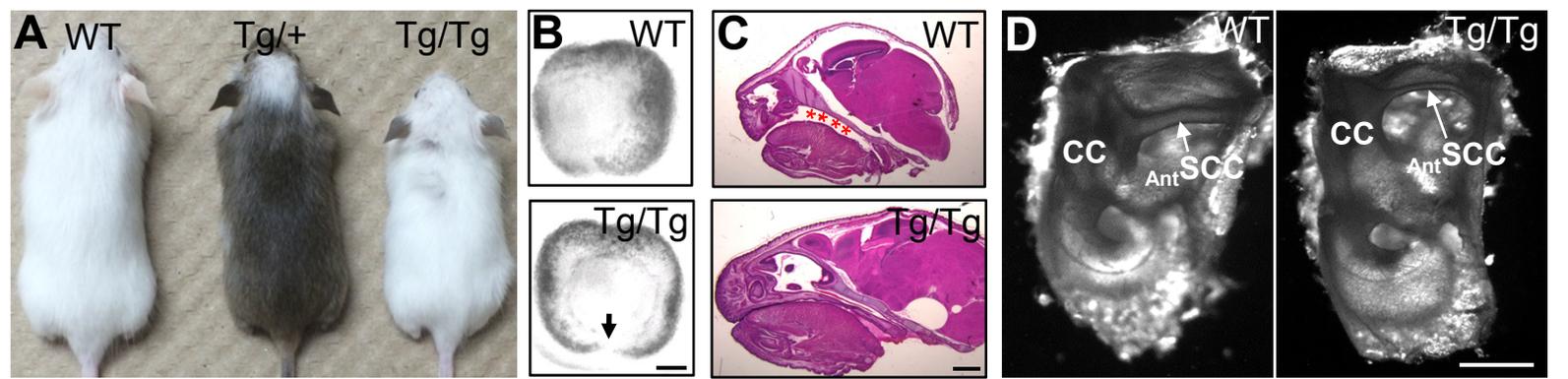
879

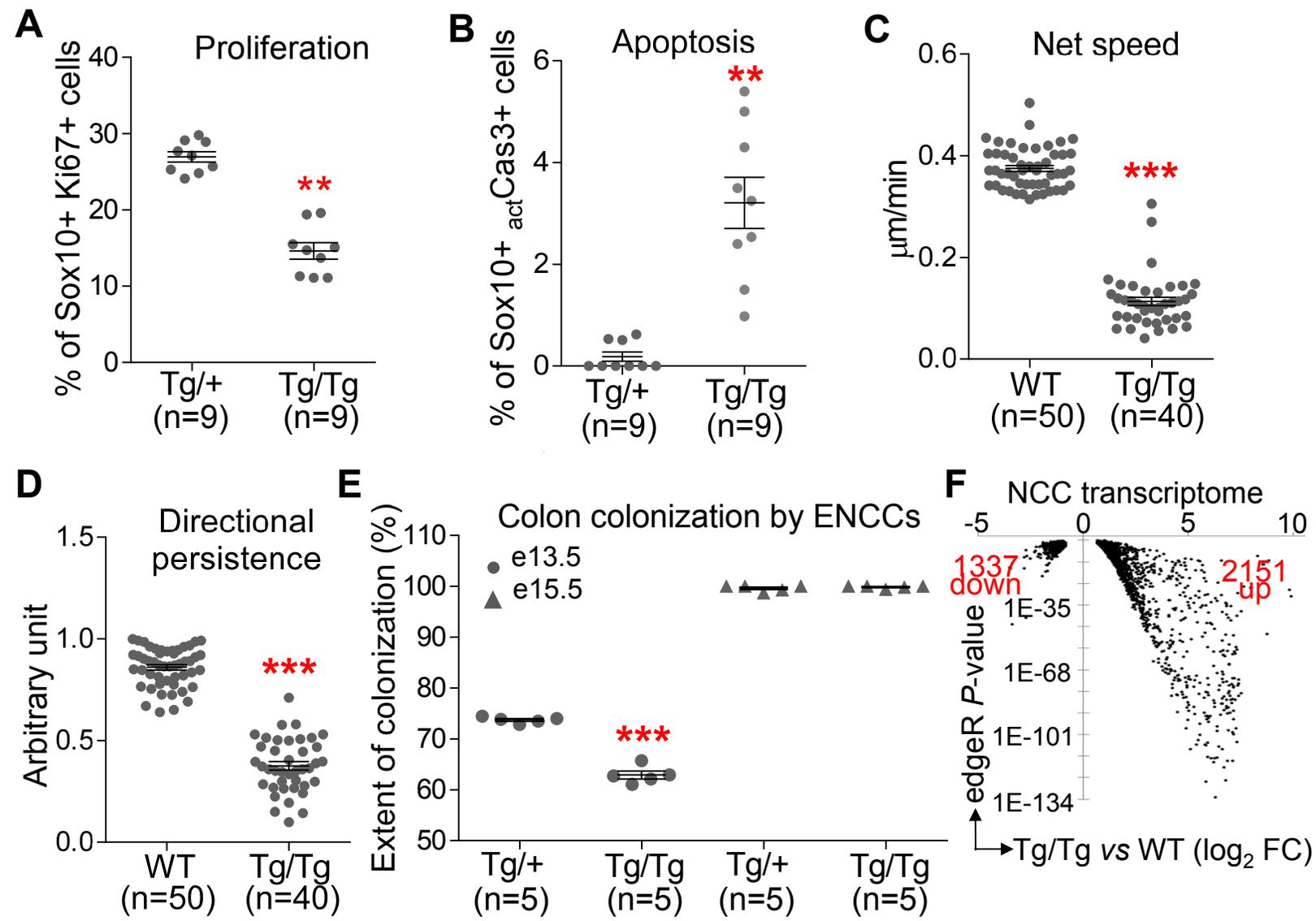
880

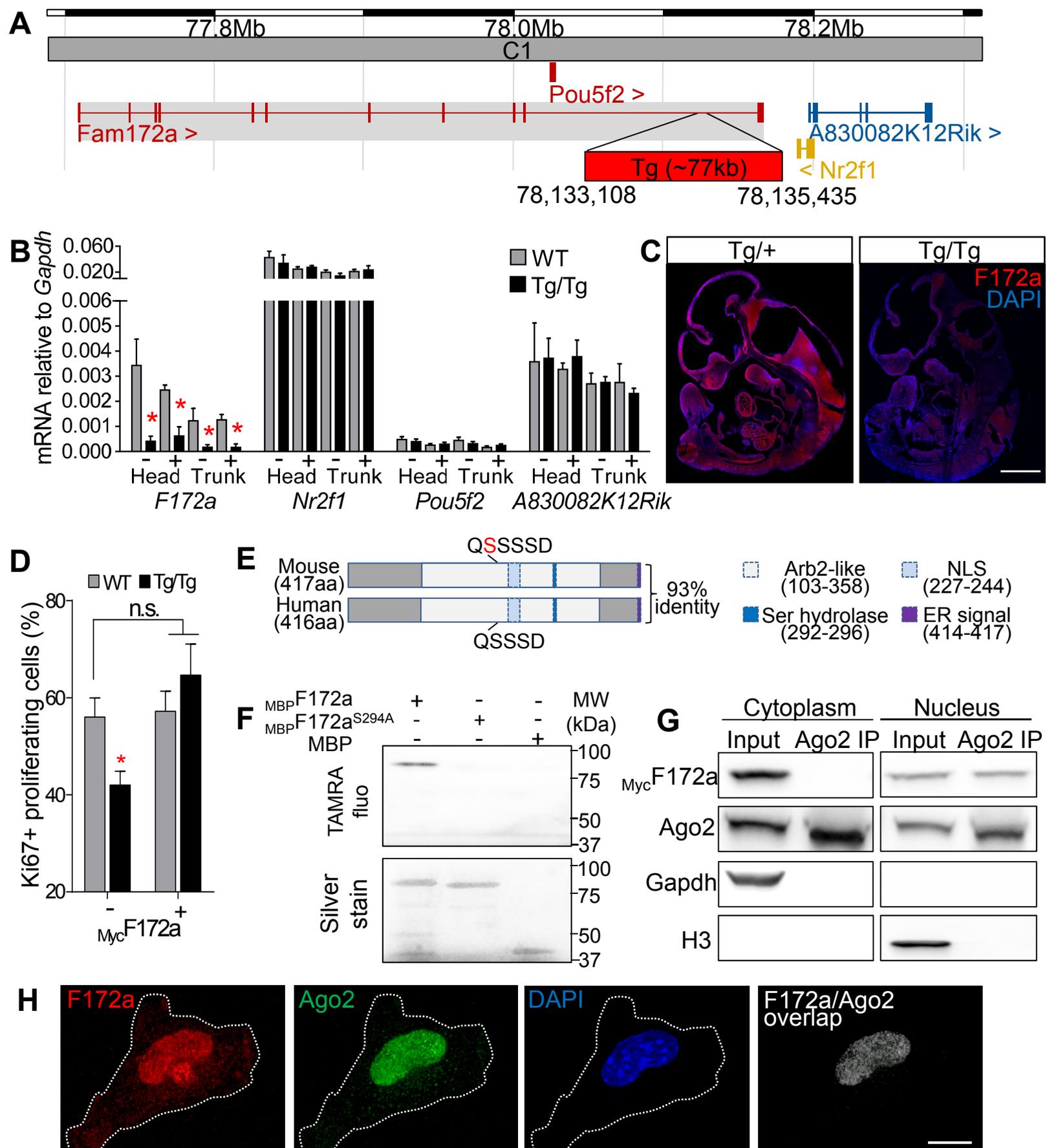
881 **Figure 5: Rapamycin-correctable dysregulation of alternative splicing in CHARGE syndrome**
882 **patients. (A-D)** RT-qPCR analysis of splicing events for *CD44*, *COL5A3*, *MICAL2* and *IFT74* in
883 lymphoblastoid cell lines. Expression levels of variable regions were normalized with levels of
884 corresponding constant regions (indicated between parentheses). Results for unaffected parents were
885 combined and used as reference value for calculation of splicing fold change (red dashed line). In panels A
886 to C, each graph depicts the results obtained for a given family (for each individual, n=9 from 3
887 independent experiments). Results for other families can be found in SI Appendix [Fig.S12](#), while detailed
888 information about each patient can be found in SI Appendix [Table S8](#). Panel D depicts the results obtained
889 after a 30-minute treatment with rapamycin (10 μ M) or vehicle only (ethanol), each vertically-aligned pair
890 of dots corresponding to a single CHARGE patient (for each individual, n=6 from 2 independent
891 experiments). In panel D, statistic tests refer to the difference between rapamycin and vehicle treatments. *
892 $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$, ns: not significant (Student's *t*-test). **(E)** Occurrence of coloboma in
893 e12.5 *Toupee^{Tg/Tg}* embryos (for each condition, n=8 embryos/16 eyes) following a 3-day *in utero*
894 exposition to rapamycin (1mg/kg) or vehicle (20% ethanol). For each phenotypic group (coloboma and
895 WT-like), the average width of choroidal fissure is indicated in their corresponding bar subdivision.

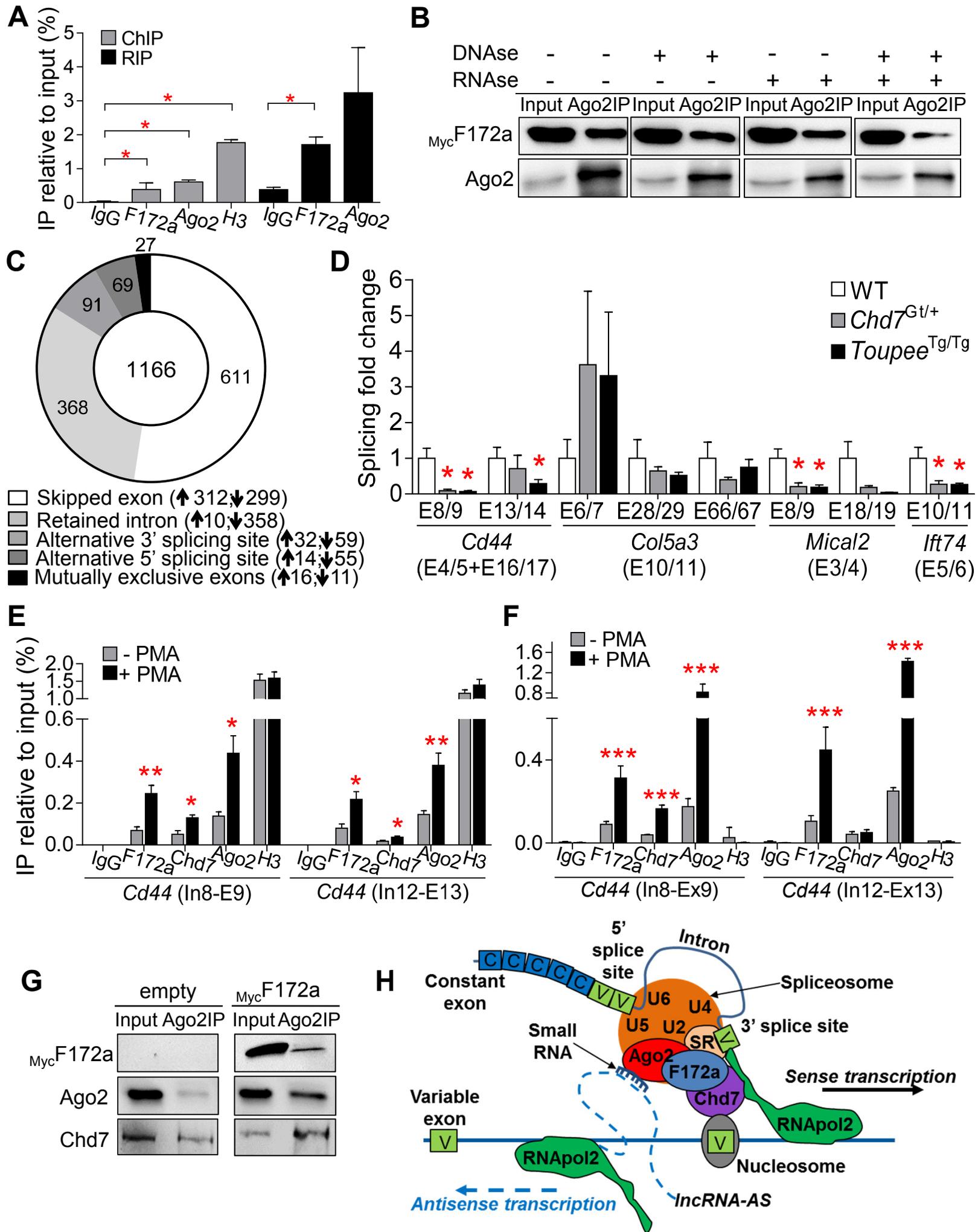
896

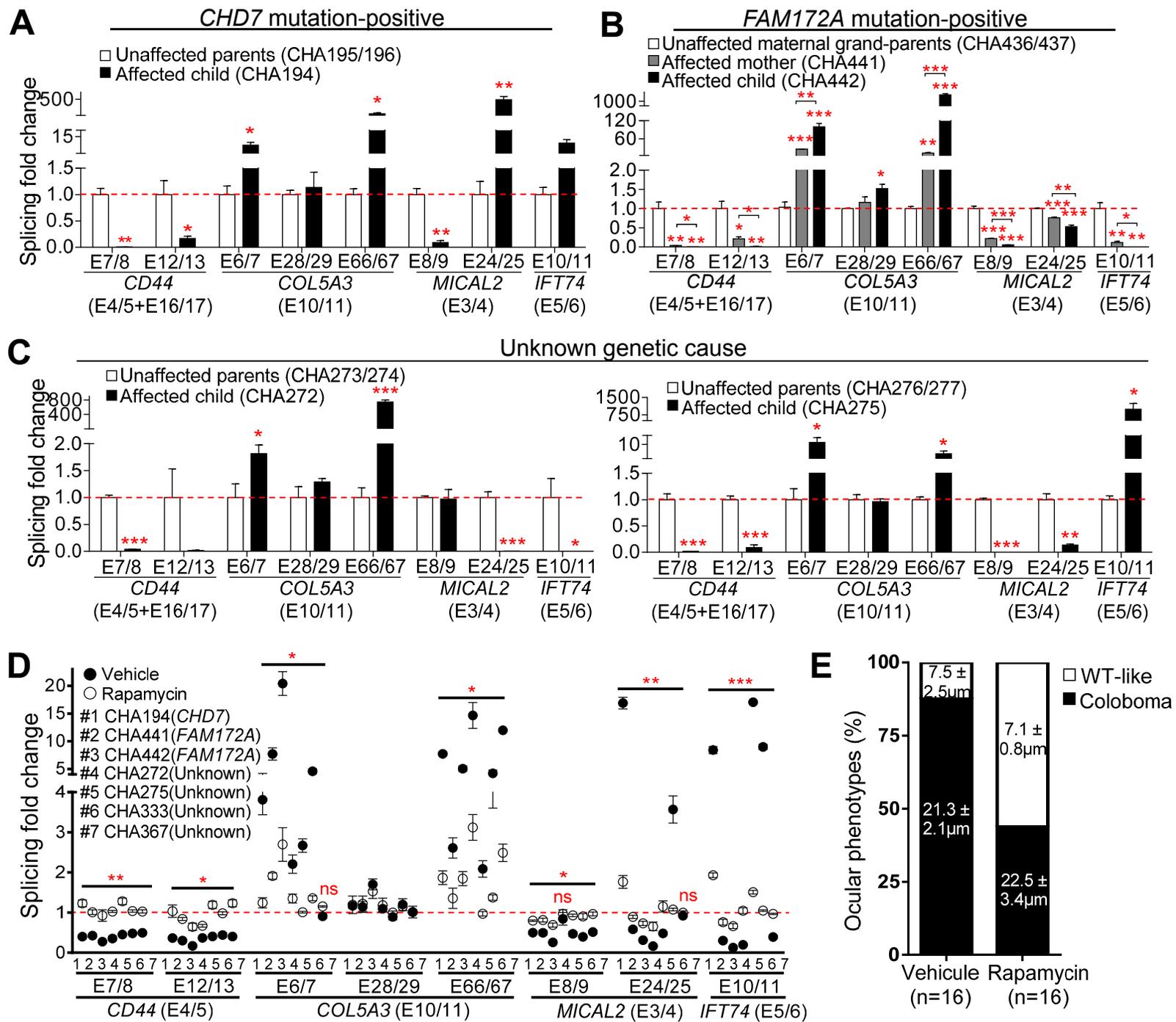
897











SI METHODS

Animals

Animal manipulation protocols were approved by the institutional ethics committee of the University of Quebec at Montreal (*Comité institutionnel de protection des animaux* (CIPA); Reference number: 650). For epistasis studies, *Toupee*^{Tg/Tg} females (FVB/N background) and *Chd7*^{Gt/+} males (C57BL/6-Sv129 mixed background) were intercrossed to generate *Toupee*^{Tg/+};*Chd7*^{Gt/+} double heterozygotes and corresponding *Toupee*^{Tg/+} single heterozygotes. *Chd7*^{Gt/+} single heterozygotes on the same hybrid background were generated by crossing the same *Chd7*^{Gt/+} males with wild-type FVB/N females.

For olfactory tests, after proper familiarization to the test conditions, mice were individually placed in a clean cage and exposed to a small piece of Whatman paper containing 100µl of water or 100µl of C57BL/6 urine (10⁻² dilution in water) during 2 min for each, with a 2 min interval between each exposition. The total number and cumulative duration of sniff bouts were recorded in order to express results as mean elapsed time per sniff bout.

Image acquisition

Image acquisition was performed with a Leica DFC 495 camera (for color images) or a Lumenera Infinity 2.0 camera (for black and white images) mounted on a Leica M205 FA stereomicroscope, with the exception of immunofluorescence analyses which were imaged with a Nikon A1 laser scanning confocal microscope. Image processing and analysis was all done with the ImageJ software, including for cell counting, measurement of choroidal fissure width as well as analysis of migration speed and persistence (see below). For Pearson correlation analysis, relevant images were deconvoluted and analyzed with the AutoQuant 3X program.

Ex vivo time-lapse imaging of NCCs

The posterior end of an e10.5 embryo (obtained from G4-RFP or *Toupee*^{Tg/Tg} ;G4-RFP intercrosses) containing pre-migratory NCCs was cut transversely behind the hindlimb bud and deposited on a small nitrocellulose filter (Millipore GSWP01300) so that a lateral side is attached to the filter and the other side is free to be imaged. The filter paper was then flipped (with tissue side down) onto paraffin rods streaked in parallel on a 60-mm cell culture dish (Corning). The samples were cultured in DMEM/F12 containing 10% FBS and penicillin/streptomycin, and incubated in a microscope incubation chamber (Okolab) under standard conditions (37°C, 5% CO₂) during 15h while 600 µm-thick stacks were acquired every 10 min (Nikon AIR confocal unit; 10X objective). Only NCCs migrating ventrally through the anterior half of somites were taken into account. Average speed (total tracked distance divided by total time) and persistence (straight line from start to finish divided by total tracked distance) were calculated from 40-min periods.

FACS

Embryonic tissues (whole e10.5 embryos for RNAseq; heads and trunks of e10.5, e13.5 and e16.5 embryos for RT-qPCR) obtained from G4-RFP or *Toupee*^{Tg/Tg} ;G4-RFP intercrosses were dissociated at 37°C with 1.3 mg/ml dispase II, 0.4 mg/ml collagenase and 0.1 mg/ml DNase I in EMEM medium. For each biological replicate, between 240,000-300,000 RFP+ and RFP- cells (from a total of 6-8 embryos at e10.5 or 1-2 embryos at e13.5 and e16.5) were collected using the BD FACSJazz cell sorter (BD Biosciences) and stored at -80°C until RNA extraction.

High throughput genome and transcriptome sequencing

Whole genome and transcriptome sequencing was performed at McGill University and Génome Québec Innovation Centre. This included library generation, sequencing using the HiSeq 2500 platform (Illumina) and bioinformatics analysis. For RNAseq, ribosomal RNA-depleted libraries (using 120 ng of total RNA as starting material) were generated and sequenced for three biological replicates per genotype (G4-RFP vs *Toupee*^{Tg/Tg}; G4-RFP) but only two per genotype met quality criteria for subsequent bioinformatics analysis. Paired-end sequences of 100-bp in length (215 million reads for the genomic DNA and between 51-84 million reads per transcriptome sample) were mapped onto the mm10 *Mus musculus* reference genome.

Cell culture and transfection

For primary culture of e10.5 embryonic cells, whole embryos were dissociated as described above for FACS and the resulting cell suspension (10^{-1} dilution in EMEM medium) was plated on gelatin-coated coverslips in 6-well plates. Dissociated cells were cultured in EMEM medium containing 10% FBS and penicillin/streptomycin under standard conditions (37°C and 5% CO₂).

Co-IP, RIP, ChIP and RNA-ChIP

Co-IPs were performed in defined buffer (20 mM Tris pH 8.0, 1 mM EGTA, 1.5 mM MgCl₂, 1 mM DTT, 1% Triton TX-100, 10% Glycerol, 1X Roche Complete protease inhibitors), containing either 25 or 200 mM NaCl. For untargeted ChIP and RIP experiments, cross-linked Neuro2a cells (using 1% PFA) were sonicated either in nuclear lysis buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 1% SDS, 1X Roche Complete protease inhibitors) for ChIP or in RIPA buffer (50 mM Tris-HCl pH 7.4, 2 mM EDTA, 150 mM NaCl, 1% NP-40, 0.5% Na-deoxycholate, 0.1% SDS, 1X Roche Complete protease inhibitors) for RIP. The nucleic acid content of purified ChIP and RIP samples was then assessed via Qubit fluorometric quantitation (ThermoFisher Scientific). For targeted ChIP and RNA-ChIP experiments, Neuro2a cells were treated with either vehicle (DMSO) or 25ng/ml Phorbol 12-myristate 13-acetate (PMA) for 4 hours. For RNA-ChIP, immunoprecipitated samples were also treated with 100 µg/ml DNase I during 2 hours at 4°C, followed by RNA extraction and reverse transcription.

MBP fusion proteins and serine hydrolase assay

MBP, MBP^{Fam172a} and MBP^{Fam172a^{S294A}} proteins were produced in BL21 bacteria via IPTG (0.3 mM) induction of relevant pMAL-c5X constructs. Bacterially-produced proteins were then purified via amylose affinity chromatography and eluted with column buffer (20 mM Tris-HCl pH7.4, 200 mM NaCl, 1 mM EDTA, 1X Roche Complete protease inhibitors) supplemented with 100 mM maltose. For serine hydrolase assay, each protein sample (7.5 µg) was mixed with the probe, incubated at room temperature in the dark for 30 minutes and migrated on a 10% SDS-polyacrylamide gel.

Affinity purification coupled to tandem mass spectrometry analysis

Neuro2a cell extracts were fractionated using buffer conditions specific to each fraction: cytoplasm (10 mM Tris-HCl pH 8.0, 0.34 M sucrose, 3 mM CaCl₂, 2 mM MgOAc, 0.1 mM EDTA, 0.5% NP-40, 1 mM DTT and 1X Roche Complete protease inhibitors), nucleoplasm (20 mM Hepes pH 7.9, 1.5 mM MgCl₂, 150 mM KOAc, 3 mM EDTA, 10% glycerol, 0.1% NP-40, 1 mM DTT and 1X Roche Complete protease inhibitors) and chromatin (150 mM Hepes pH 7.9, 1.5 mM MgCl₂, 150 mM KOAc, 10% glycerol, 15 U/ml Benzonase, 0.44 U/ml RNase A, 6.25 U/ml DNase I, 1X Roche Complete protease inhibitors). Each

Neuro2a cell fraction (1.25mg) was passed on an amylose resin column containing immobilized _{MBP}Fam172a or MBP alone. Following extensive washing, interacting proteins were co-eluted with _{MBP}Fam172a or MBP as described above for the purification of MBP fusion proteins. Eluates were then individually precipitated using 72% TCA, 0.075% Na-deoxycholate and 2.5X Tris-EDTA. Preparation of tryptic fragments, LC-MS/MS analysis on a LTQ Orbitrap Fusion Tribrid mass spectrometer (ThermoFisher Scientific) and peptide identification were performed at the Proteomics Discovery Platform of the *Institut de recherches cliniques de Montreal* (IRCM). Proteins identified via at least one peptide (with probability greater than 90%) were accepted as _{MBP}Fam172a interactors if they were enriched at least 1.5-fold in comparison to the MBP negative control and detected in at least two out of the three biological replicates per fraction.

Figure S1

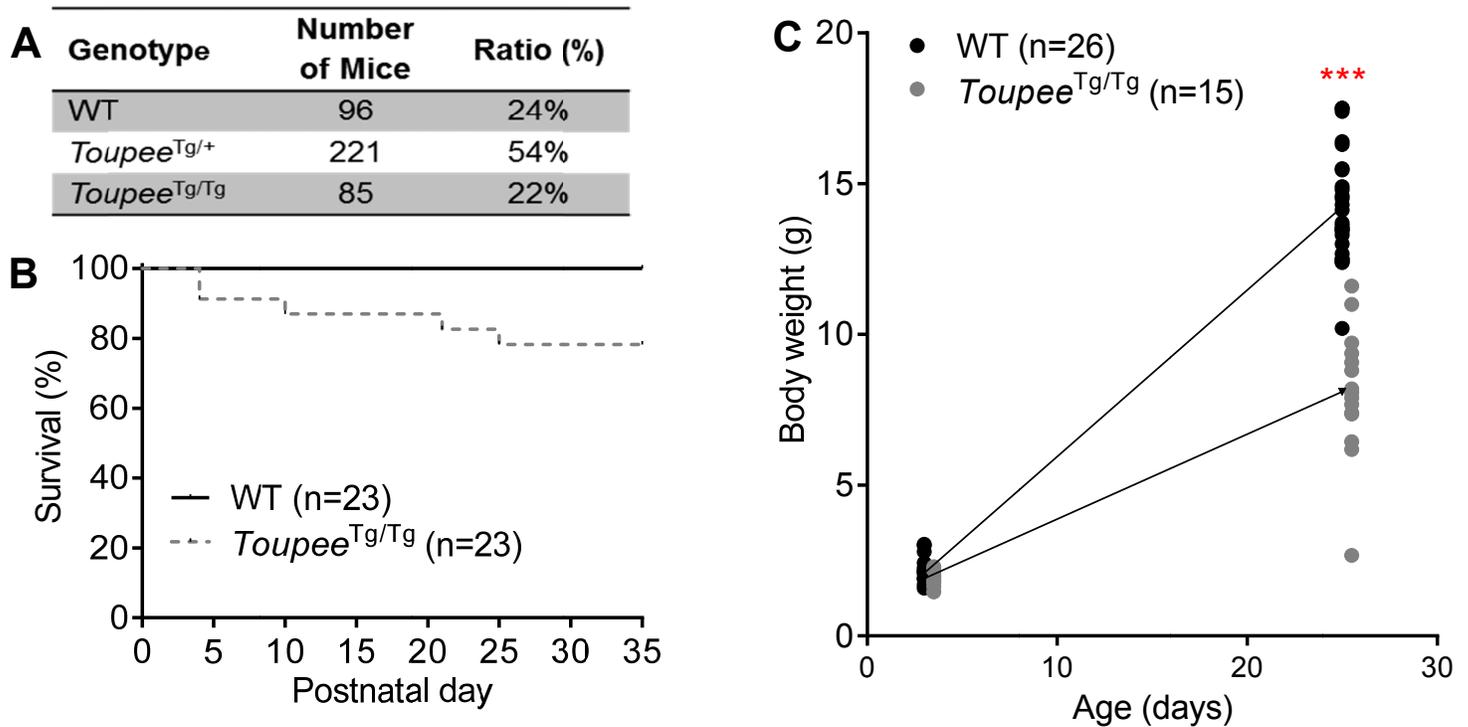


Figure S1. Early death and growth retardation in *Toupee*^{Tg/Tg} mice. (A) Genotype distribution of offspring from *Toupee*^{Tg/+} intercrosses (n=48 couples) at weaning. (B) Kaplan-Meier survival curve for WT and *Toupee*^{Tg/Tg} mice. (C) Body weight comparison of WT and *Toupee*^{Tg/Tg} mice at P3 and P25.

Figure S2

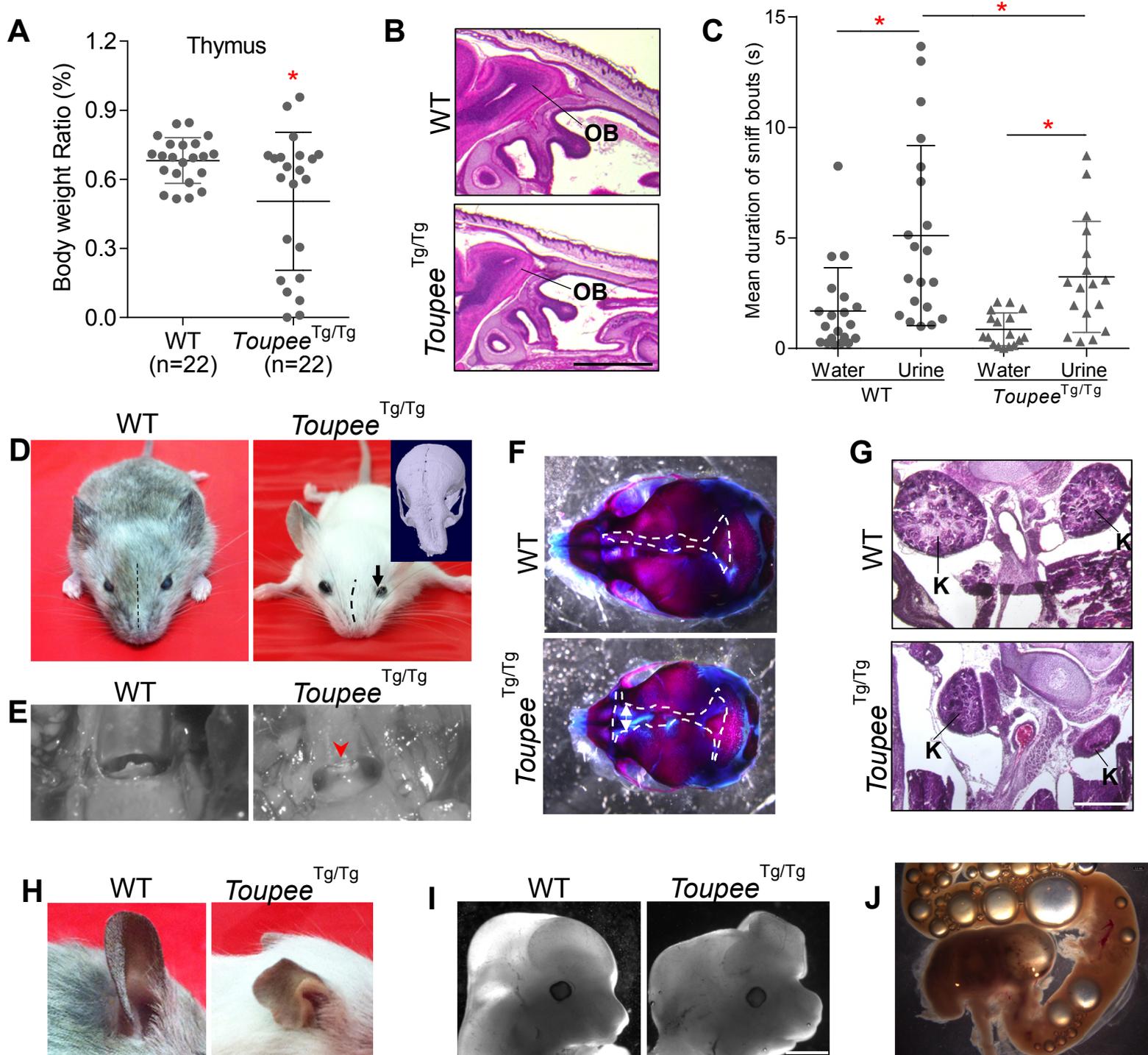


Figure S2. Less frequently observed minor features of CHARGE syndrome in *Toupee*^{Tg/Tg} mice. (A) A subset of *Toupee*^{Tg/Tg} mice exhibits a lower thymus/body weight ratio at P25. (B) H&E-stained sagittal sections of e18.5 heads (n=10 per genotype) revealing olfactory bulb hypoplasia in a subset of *Toupee*^{Tg/Tg} embryos. (C) Results of olfaction tests (n=20 mice per genotype) showing that adult *Toupee*^{Tg/Tg} mice are less efficient than WT mice at distinguishing urine from water. (D) A subset of *Toupee*^{Tg/Tg} mice displays facial asymmetry including unilateral eyelid ptosis (arrow) and twisted facial bones (see μ CT scan in the inset). (E) Frontal view of the oropharynx at P25 showing partial atresia (arrowhead) in a *Toupee*^{Tg/Tg} animal. (F) Alizarin red and alcian blue double-stained e18.5 skulls (n=10 per genotype) showing that closure of the fontanelles (delineated by dashed lines) is delayed in a subset of *Toupee*^{Tg/Tg} embryos (double-headed arrow). (G) H&E-stained cross-sections of e15.5 embryos showing kidney (K) hypoplasia in a *Toupee*^{Tg/Tg} animal. (H) The outer ear of *Toupee*^{Tg/Tg} animals is occasionally malformed. (I) Lateral view of e12.5 heads demonstrating exencephaly in a *Toupee*^{Tg/Tg} embryo. (J) Gastrointestinal track of a P24 *Toupee*^{Tg/Tg} animal filled with air bubbles. * $P < 0.05$ (Student's *t*-test). Scale bar: 1mm (B), 500 μ m (G).

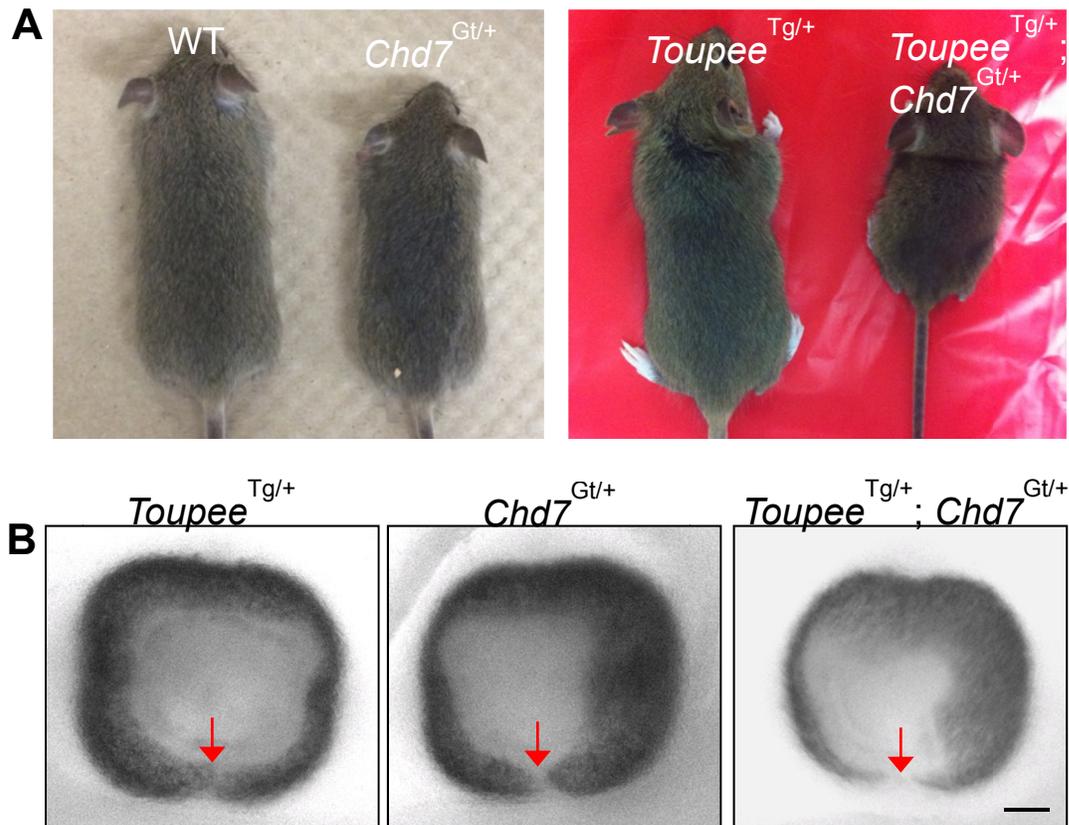


Figure S3. Genetic interaction between *Toupee*^{Tg} and *Chd7*^{Gt} alleles. (A) *Toupee*^{Tg/+}; *Chd7*^{Gt/+} double mutants are markedly smaller than corresponding single mutants at P25. (B) Representative bright field images of e12.5 eyes showing that the choroidal fissure (red arrows) is wider in double mutant embryos. Scale bar: 50 μ m.

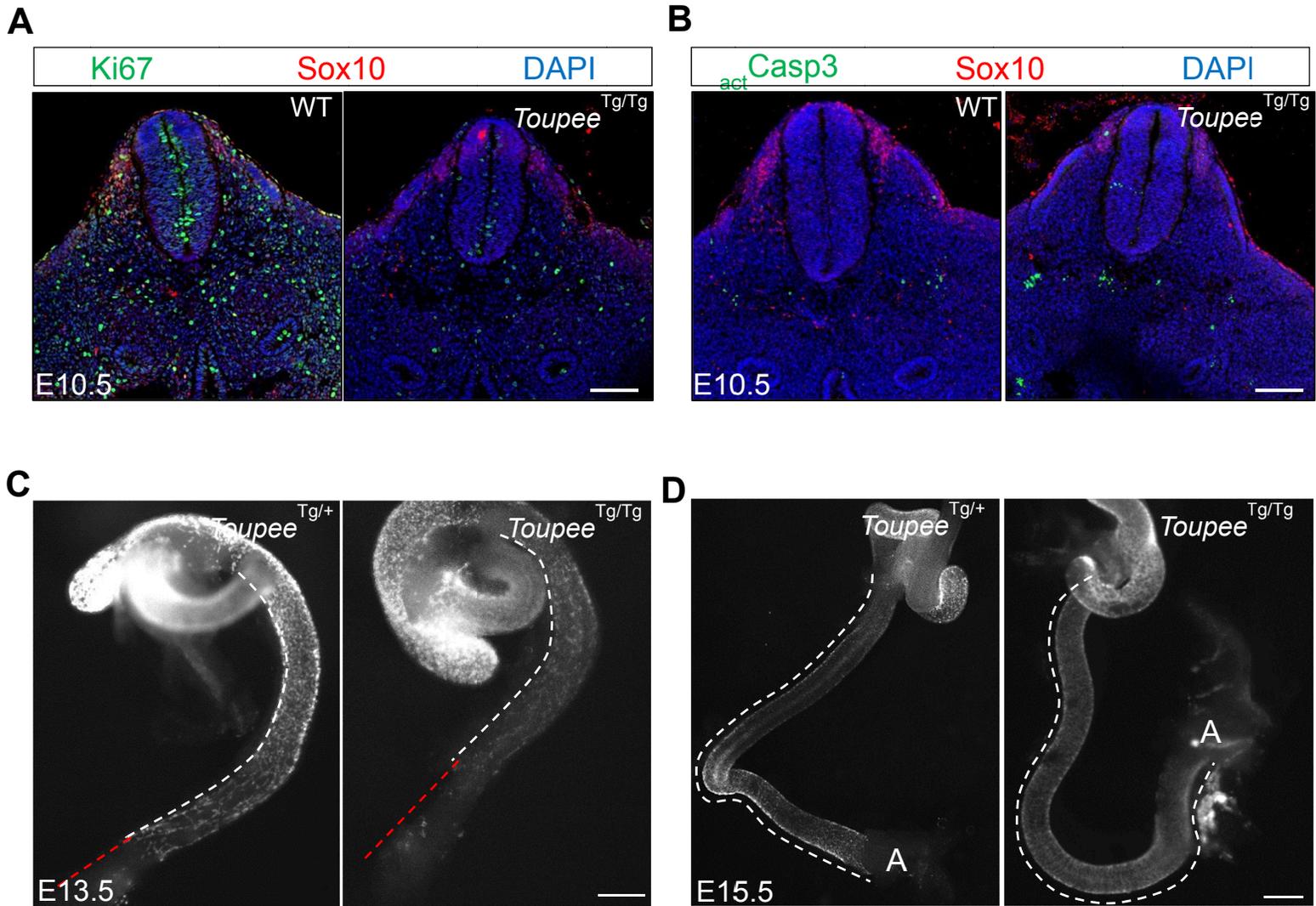


Figure S4. Defective NCC proliferation, survival and migration in *Toupee*^{Tg/Tg} embryos. (A-B) Representative immunofluorescence images used for quantification of NCC proliferation and survival in Fig.2A-B (n=9 per genotype). Cross-sections (30 μ m) of e10.5 embryos at the level of the hindlimb bud were double-labeled with anti-Sox10 (red) and anti-Ki67 or anti-activated Caspase3 (green), and counterstained with DAPI (blue). (C-D) Representative images used for quantification of hindgut colonization by fluorescently labeled (owing to G4-RFP transgene) enteric NCCs in Fig.2E. The white dashed lines delineate the colonized hindgut segment, whereas the red dotted lines delineate the uncolonized hindgut segment (for each developmental stage, n=5 per genotype). Scale bars: 100 μ m (A-B), 500 μ m (C) and 1mm (D).

Figure S5

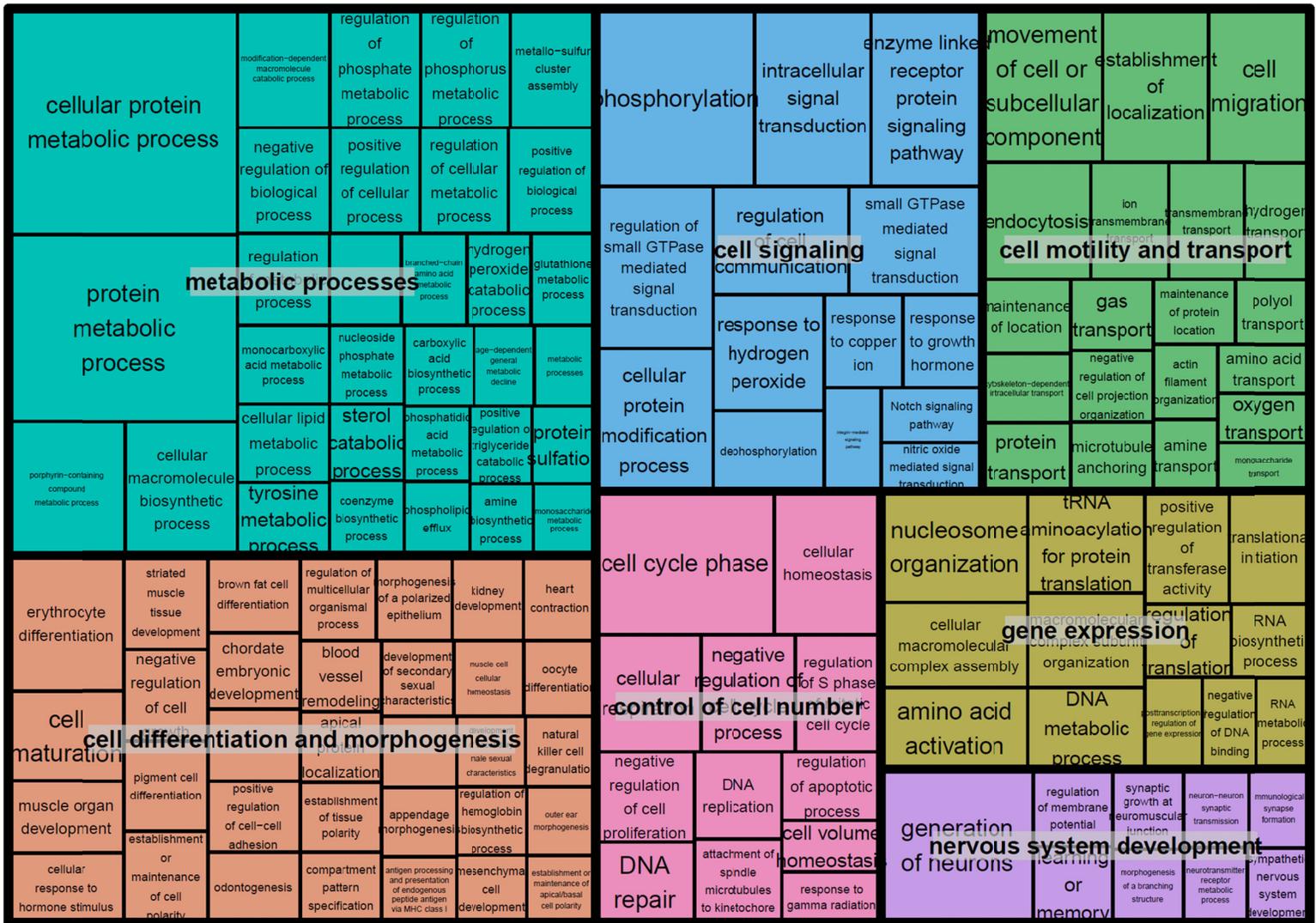


Figure S5. Enriched GO terms associated with the differentially expressed genes in *Toupee*^{Tg/Tg} NCCs. REVIGO TreeMap representation of the GO analysis of the 3488 modulated genes displayed in Fig.2F. The 132 enriched terms (ontology level ≥ 5 and $P < 0.05$) are distributed into 7 main categories: Metabolic processes; Cell differentiation and morphogenesis; Cell signaling; Cell motility and transport; Control of cell number; Gene expression; and Nervous system development. Size of boxes is proportional to statistical significance.

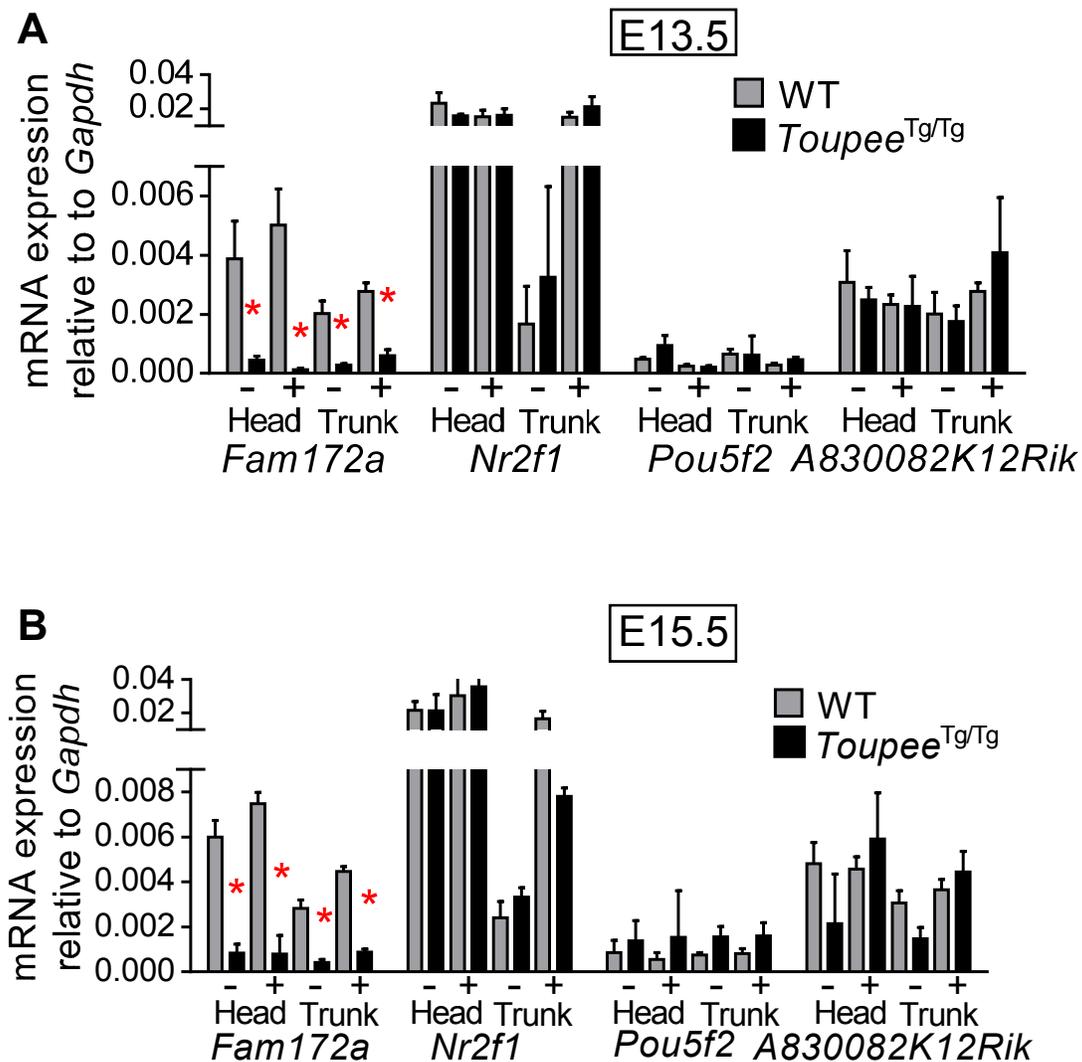


Figure S6. Specific downregulation of *Fam172a* expression in *Toupee*^{Tg/Tg} embryos. RT-qPCR analyses of gene expression in *G4-RFP* (WT) and *Toupee*^{Tg/Tg}*:G4-RFP* (*Toupee*^{Tg/Tg}) embryos at e13.5 (A) and e15.5 (B). Transcript levels of *Fam172a*, *Pou5f2*, *Nr2f1* and *A830082k12Rik* were monitored in FACS-recovered NCCs (+) and non-NCCs (-) from the head and the trunk (n=3 per condition). Expression levels were normalized relative to *Gapdh*. * $P \leq 0.05$.

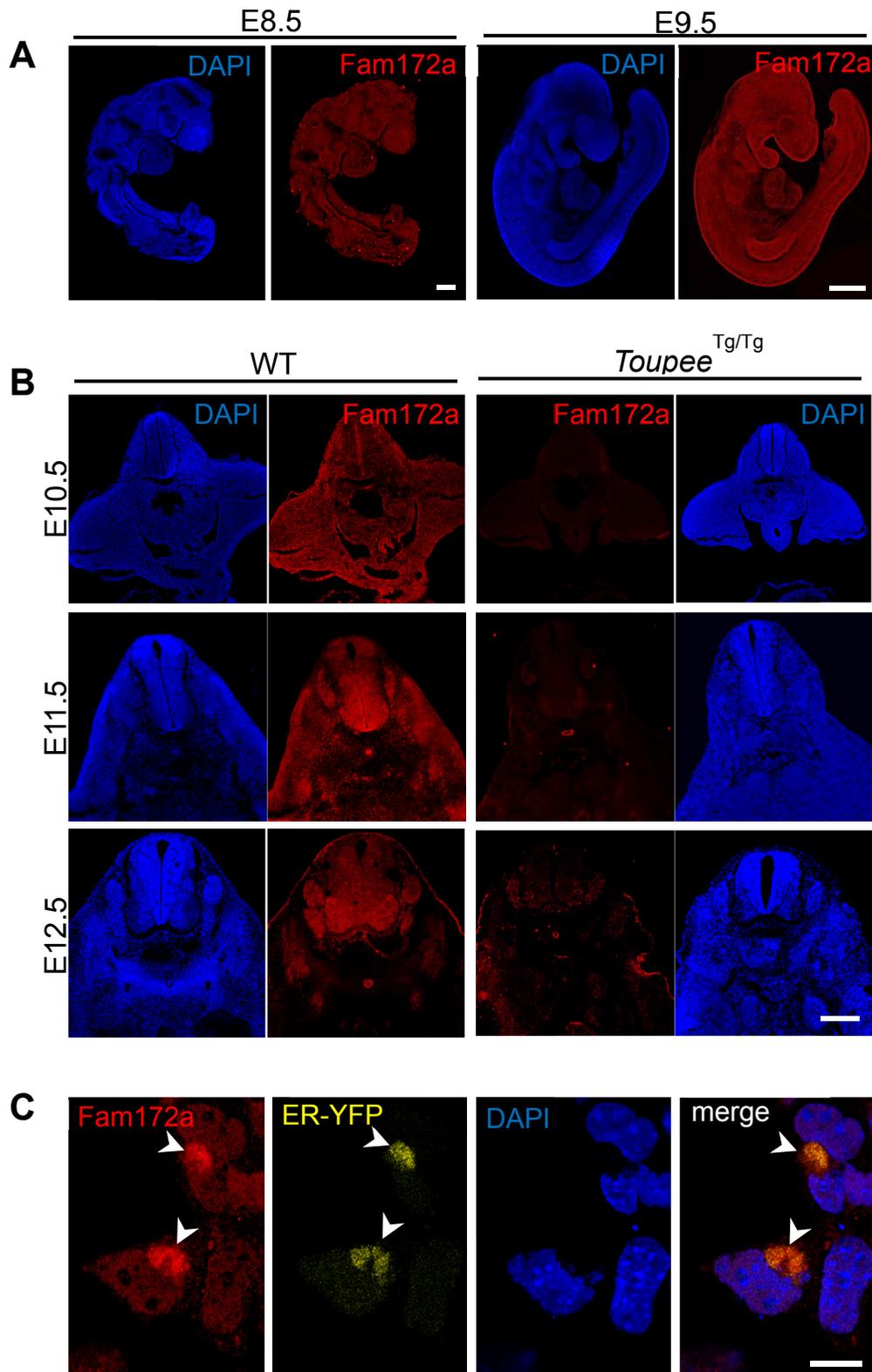


Figure S7. Analysis of Fam172a protein distribution in mouse embryos and COS7 cells. (A-B) Immunofluorescence analysis of the Fam172a protein in e8.5-9.5 WT embryos (A) and in cross-sections of e10.5-12.5 WT and *Toupee*^{Tg/Tg} embryos (B) at hindlimb level (n=6 per genotype). (C) Immunofluorescence labeling of endogenous Fam172a in COS7 cells transfected with the pEYFP-ER plasmid (Clontech) and counterstained with DAPI (n=6). White arrowheads indicate ER localization. Scale bar, 200 μ m (A-B), 25 μ m (C).

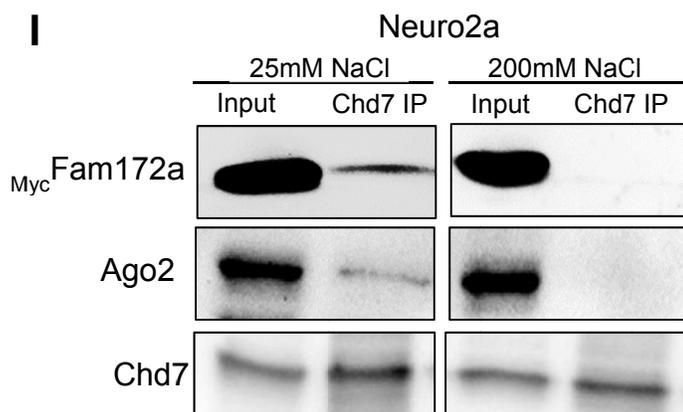
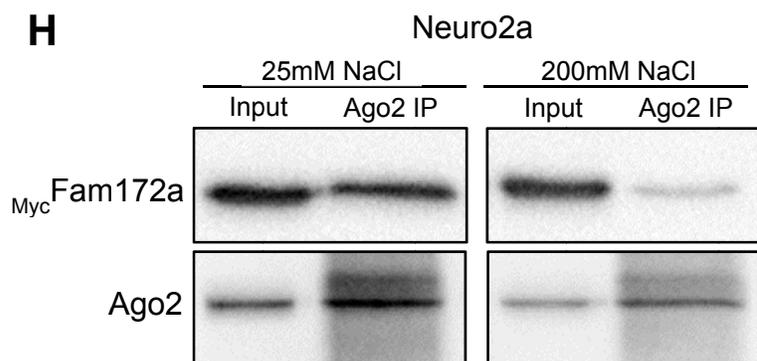
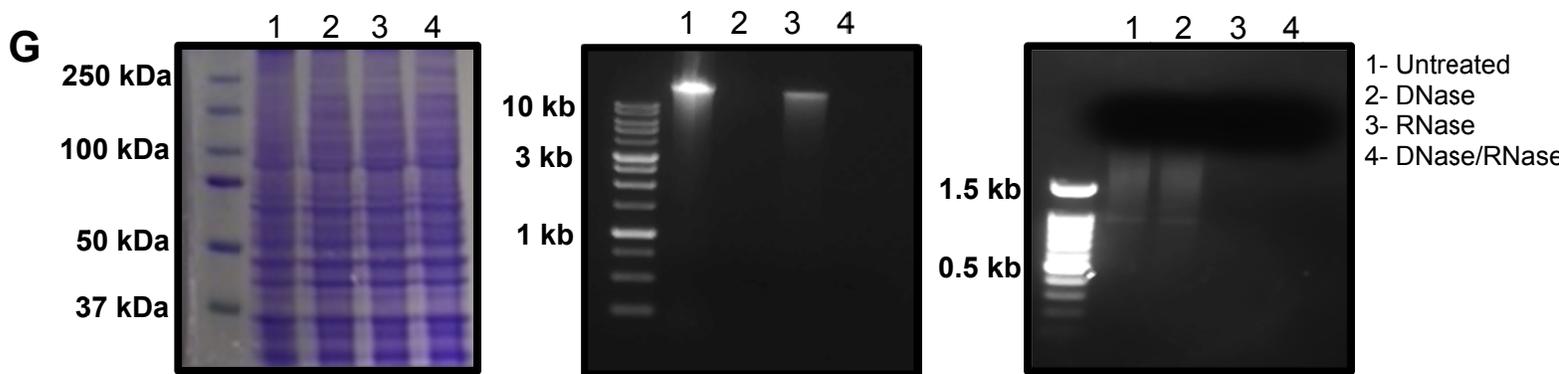
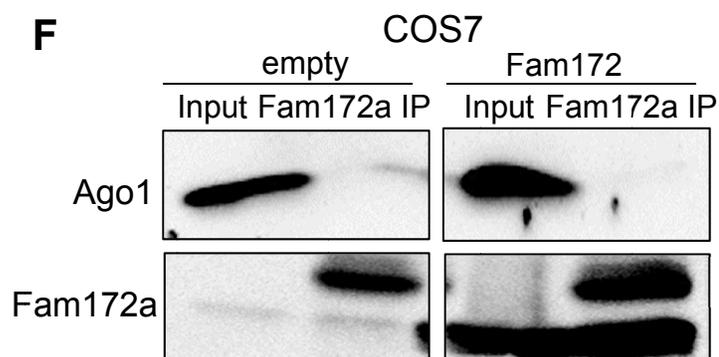
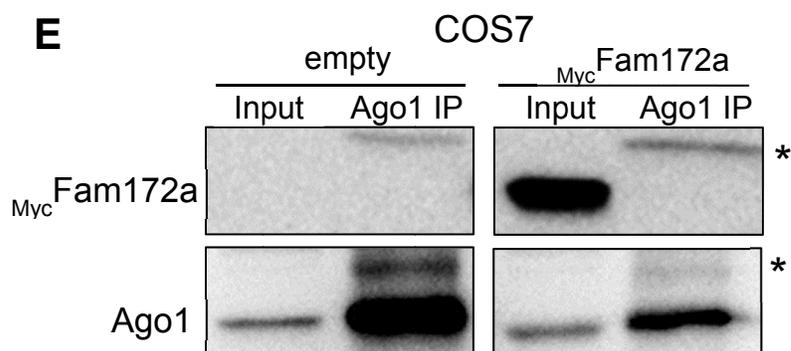
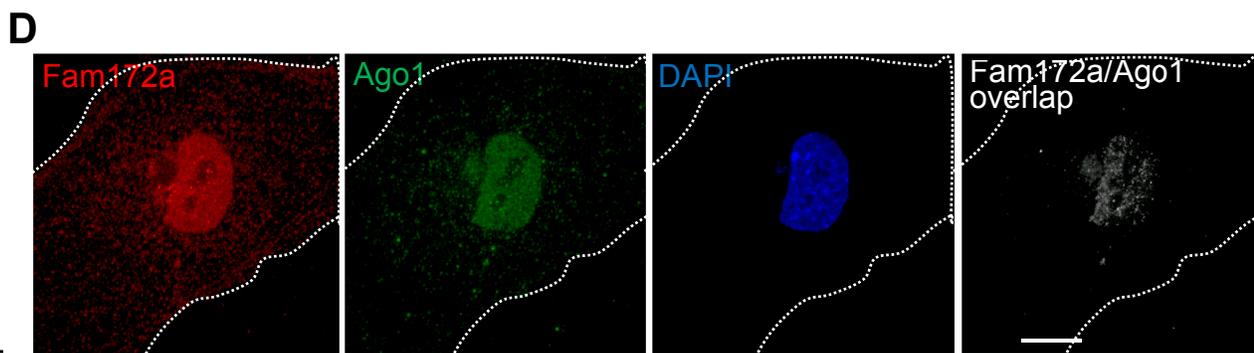
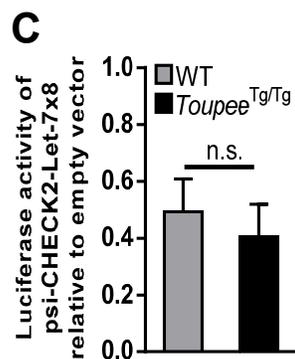
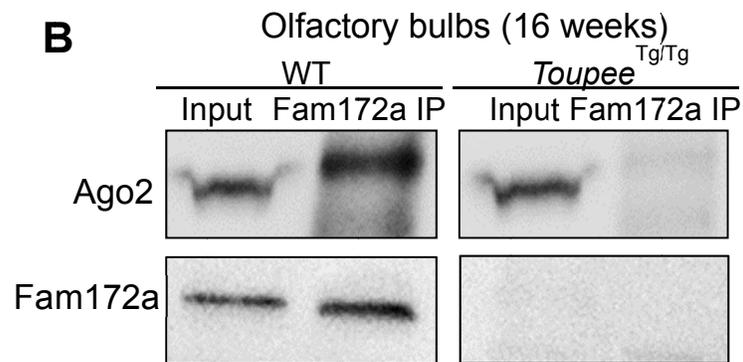
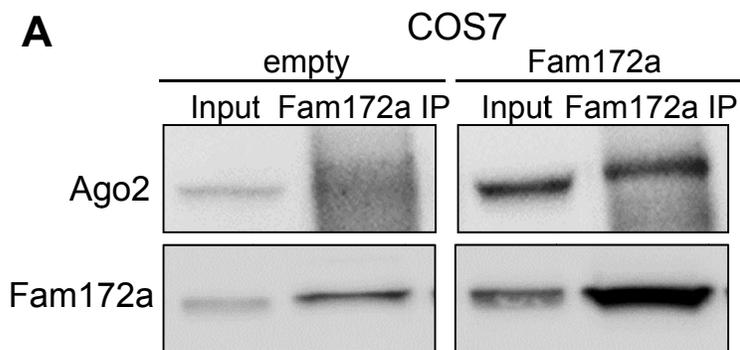


Figure S8. Fam172a physically interacts with Ago2 but not with Ago1. (A-B) Co-IP assays using whole cell extracts of COS7 cells (A; transfected with a Fam172a-expressing or empty vector) or of adult mouse olfactory bulbs (B; WT vs *Toupee*^{Tg/Tg}). IPs were performed with an anti-Fam172a antibody and analyzed by western blot using an anti-Ago2 antibody. Inputs correspond to 10% of protein extracts used for IP (n=3 per condition). **(C)** Luciferase activity from the psi-CHECK2-let-7x8 reporter relative to empty psi-CHECK2 vector in cultures of dissociated e10.5 embryos (WT vs *Toupee*^{Tg/Tg}). **(D)** Double immunofluorescence labeling of Fam172a (red) and Ago1 (green) in dissociated cells from WT e10.5 embryos (n=7). DAPI was used to counterstain nuclei (blue). The right panel shows the overlap of Fam172a and Ago1 signals (Pearson's correlation coefficient of 0.63) as determined with the AutoQuant 3X software. Scale bar: 25µm. **(E-F)** Co-IP assays using whole cell extracts of COS7 cells transfected with empty or _{Myc}Fam172a-expressing vector. IPs were performed with an anti-Ago1 (E) or anti-Fam172a (F) antibody and analyzed by western blot using anti-Ago1 and either anti-Myc (E) or anti-Fam172a (F) antibodies. Inputs correspond to 10% of protein extracts used for IP (n=3 per condition). The asterisk indicates the presence of mouse IgG heavy chain. **(G)** Impact of DNase and RNase treatments on the integrity of proteins (left panel), DNA (middle panel) and RNA (right panel) in whole-cell extracts, as described in Fig.4B. **(H-I)** Co-IP assays using whole cell extracts of Neuro2A cells (transfected with a _{Myc}Fam172a-expressing vector) in presence of different concentrations of NaCl. IPs were performed with an anti-Ago2 (H) or anti-Chd7 (I) antibody and analyzed by western blot using anti-Myc, anti-Ago2 and anti-Chd7 antibodies. Inputs correspond to 10% of protein extracts used for IP (n=3 per condition).

Figure S10

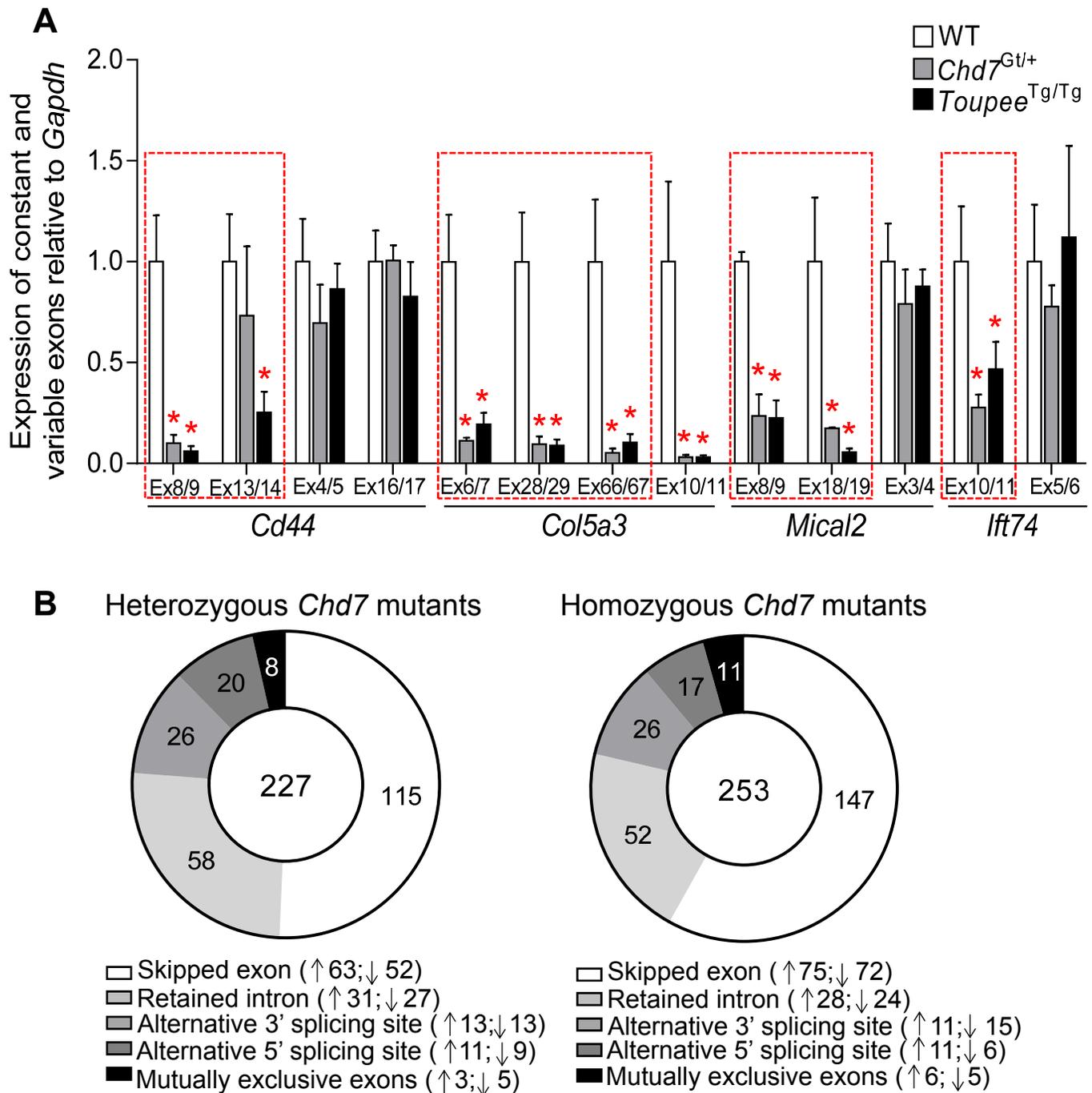


Figure S10. Detailed analysis of alternative splicing in *Toupee*^{Tg/Tg} and *Chd7*^{Gt/+} mutants. (A) RT-qPCR analysis of variably (delineated in red) and constantly expressed regions of 4 previously described Ago2-regulated genes (*Cd44*, *Col5a3*, *Mical2* and *Ifi74*) in the head of *G4-RFP* (WT), *Toupee*^{Tg/Tg}; *G4-RFP* (*Toupee*^{Tg/Tg}) and *Chd7*^{Gt/+} e12.5 embryos (n=5 per genotype). Transcript levels were normalized relative to *Gapdh*. * $P \leq 0.05$ (Student's *t*-test). **(B)** Donut chart showing the distribution of the differentially modulated alternative splicing events ($P < 0.05$; variation in inclusion level ≥ 0.1) in P7 granule neuron progenitors from heterozygous (227 events; Dataset S3) and homozygous (253 events; Dataset S4) *Chd7* mutants as determined by rMATS-based analysis of recently published RNAseq data (22). Upward- and downward-pointing arrows indicate splicing events that are over- and under-represented in *Chd7* mutants, respectively.

Figure S11

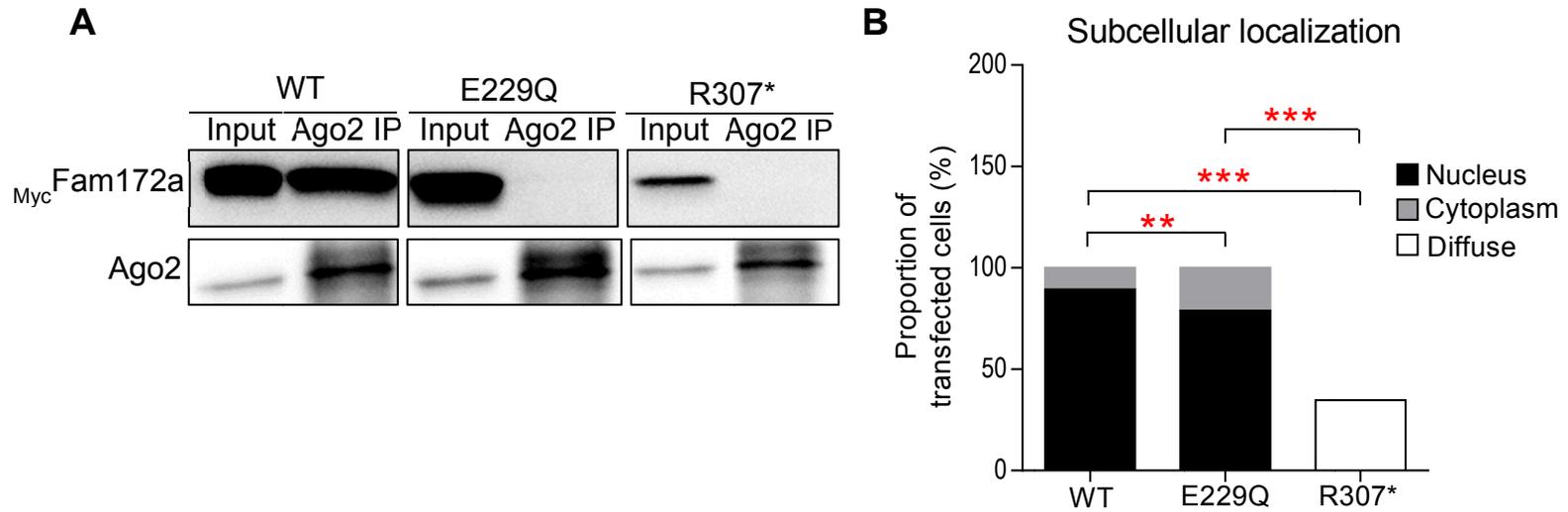


Figure S11: Variants of Fam172 identified in human CHARGE patients fail to interact with Ago2.

(A) Co-IP assays using whole cell extracts of Neuro2A cells transfected with MycFam172a -, $\text{MycFam172a}^{\text{E229Q}}$ - or $\text{MycFam172a}^{\text{R307*}}$ -expressing plasmids. Following Ago2 IP, the presence of Myc-tagged Fam172a and Ago2 proteins was revealed using anti-Myc and anti-Ago2 antibodies, respectively. Inputs correspond to 10% of protein extracts used for IP (n=3 per condition). **(B)** Quantification of Fam172a localization following anti-Myc immunofluorescence labeling of Neuro2A cells transfected with MycFam172a -, $\text{MycFam172a}^{\text{E229Q}}$ or $\text{MycFam172a}^{\text{R307*}}$ expression vectors. ** $P \leq 0.01$, *** $P \leq 0.001$ (Student's *t*-test)

Figure S12

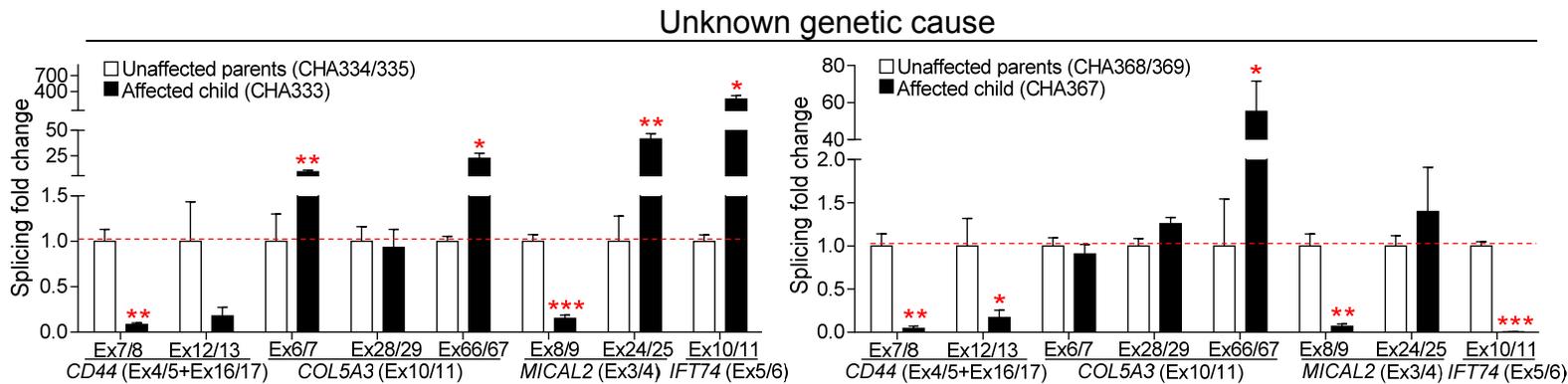


Figure S12: Dysregulation of alternative splicing in supplemental CHARGE syndrome patients. RT-qPCR analysis of splicing events for *CD44*, *COL5A3*, *MICAL2* and *IFT74* in lymphoblastoid cell lines. Expression levels of variable regions were normalized with levels of corresponding constant regions (indicated between parentheses). Results for unaffected parents were combined and used as reference value for calculation of splicing fold change (red dashed line). Each graph depicts the results obtained for a given family (for each individual, n=9 from 3 independent experiments). Results for other families can be found in [Fig.5A-C](#), while detailed information about each patient can be found in [Table S8](#). * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$ (Student's *t*-test).

Figure S13

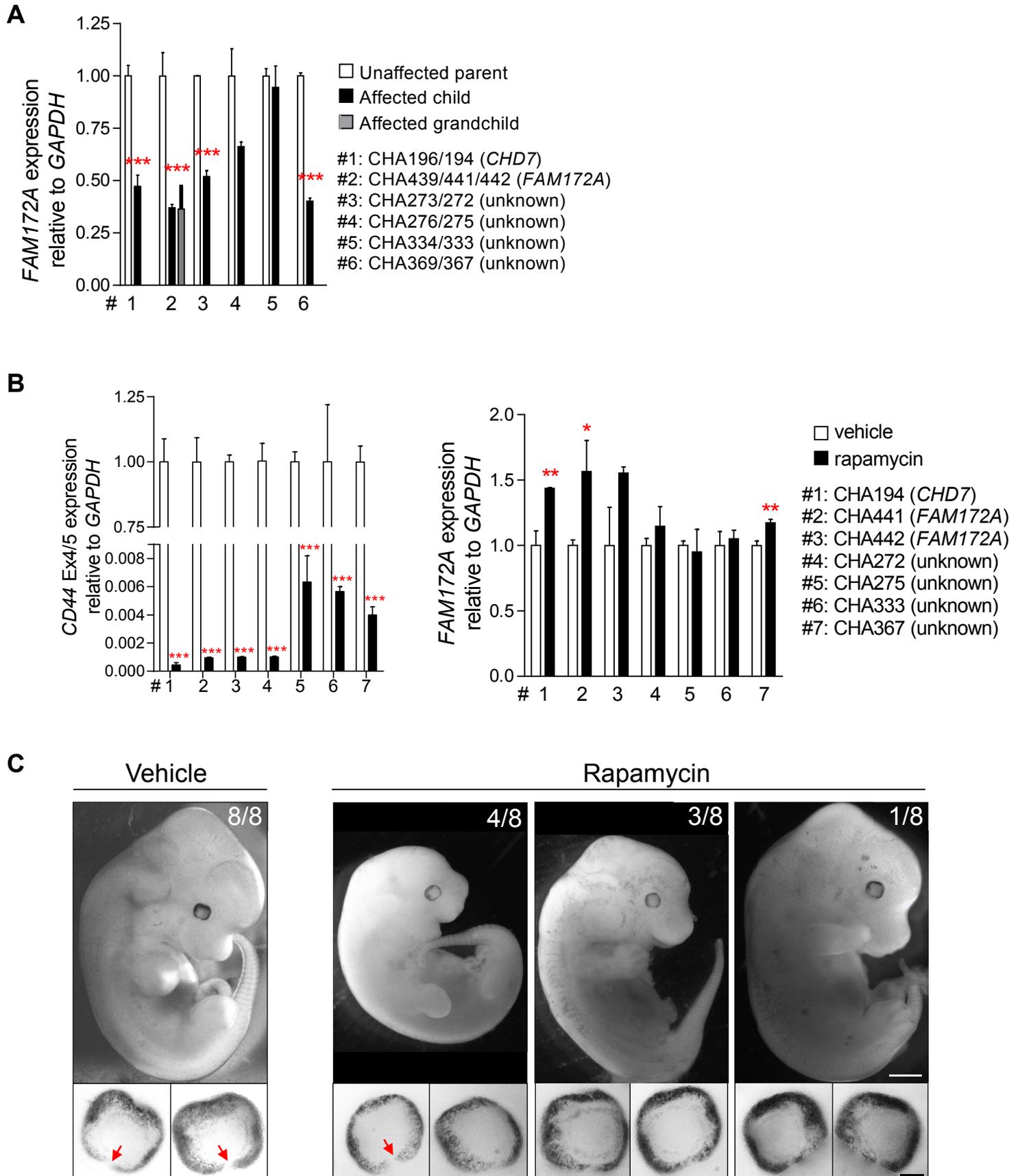


Figure S13: Analysis of rapamycin effects in human LCLs and *Toupee*^{Tg/Tg} mouse embryos. (A-B) *GAPDH*-normalized RT-qPCR analyses of *FAM172A* and *CD44* expression in human LCLs derived from CHARGE patients and/or one of their unaffected parents, in absence (A) or presence (B) of rapamycin. The decreased expression of *CD44* constant exons 4-5 demonstrates the efficacy of the rapamycin treatment (left panel in B). Note that basal expression of *FAM172A* is generally lower in CHARGE patients and that rapamycin has very little effect on it (right panel in B). **(C)** Bright field images of e12.5 *Toupee*^{Tg/Tg} embryos that were previously exposed to vehicle (20% ethanol) or rapamycin (1mg/kg) for 3 days *in utero* (n=8 embryos per condition). Exposition to rapamycin decreased the incidence of coloboma (red arrows in bottom panels), while it also increased the incidence of growth retardation and malformation/resorption. The ratios displayed in the top right corner of top panels refer to the number of embryos with similar morphology, which is not correlated with their ocular phenotype. Scale bar, 1mm (whole embryos) or 50µm (eyes).

Figure S14

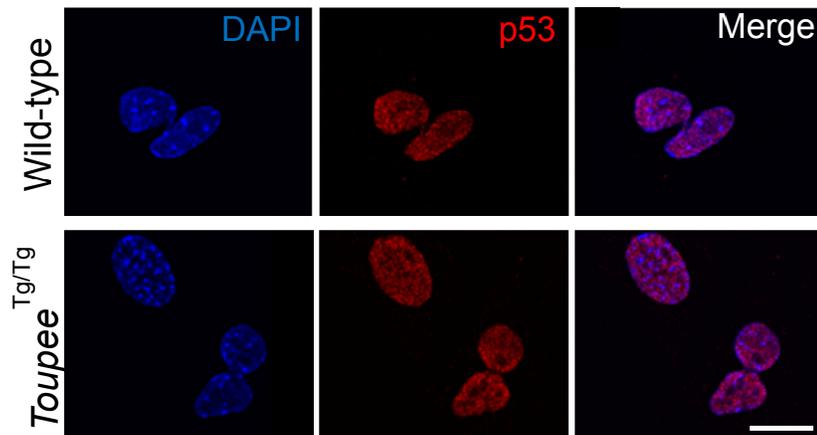


Figure S14: p53 expression appears unaffected in *Toupee*^{Tg/Tg} embryos. Immunofluorescence analysis of p53 protein distribution in dissociated cells from WT and *Toupee*^{Tg/Tg} e10.5 embryos. Neither the expression levels of p53 nor its localization appears affected in *Toupee*^{Tg/Tg} cells. Scale bar, 10 μ m

Table S1: Overview of major and minor features of CHARGE syndrome in *Toupee*^{Tg/Tg} mice

	Abnormality	Frequency	
Major	Coloboma of the eye	86% (n=14) ^a	
	Inner ear defects (circling behavior)	67% (n=61)	
	Cleft palate	21% (n=14)	
Minor	Heart malformation (left ventricular hypertrophy)	80% (n=10)	
	Cranial nerve anomalies (VII, IX and/or X)	60% (n=10)	
	Retardation of growth	61% (n=61)	
	Genital abnormalities	Males (cryptorchidism)	52% (n=27)
		Females (hypoplastic uterine horns)	58% (n=12)
	Thymus hypoplasia	36% (n=22)	
	Olfactory bulb hypoplasia	30% (n=10)	
	Craniofacial malformations	Delayed closure of fontanelles	30% (n=10)
		Twisted skull	18% (n=61)
		Oropharyngeal atresia	16% (n=61)
Kidney hypoplasia	8% (n=12)		

Note: ^a10/14 unilateral, 2/14 bilateral.

Table S2: *Toupee*^{Tg/Tg} mice are subfertile

Reproduction parameter	WT males WT females (n = 9 couples)	<i>Toupee</i>^{Tg/Tg} males WT females (n = 10 couples)	WT males <i>Toupee</i>^{Tg/Tg} females (n = 9 couples)
Average number of days before vaginal plug	1.33 ± 1.00	1.92 ± 0.79 ^{n.s.}	7.73 ± 6.60*
Average number of pups per litter	11.11 ± 1.76	7.44 ± 2.92**	7.00 ± 0.66*
Number of unproductive matings	0	1	5

Note: n.s. not significant; * $P < 0.05$; ** $P < 0.01$.

Table S3: Genetic interaction between *Toupee*^{Tg} and *Chd7*^{Gt} alleles

Genotype	<i>Chd7</i> ^{Gt/+} (n=29) ^a	<i>Toupee</i> ^{Tg/+} (n=32) ^b	<i>Toupee</i> ^{Tg/+} ; <i>Chd7</i> ^{Gt/+} (n=20) ^b
Phenotype			
Frequency at birth (expected)	49% (50%)	62% (50%)	38% (50%)
Weight at P25 (g)	11.40 ± 2.01	13.26 ± 2.14	6.69 ± 2.56
Premature postnatal death (≤P30)	3%	0%	20%
Circling behavior	27%	0%	45%
Male-to-female sex reversal (number of XY)	12% (16)	0% (15)	33% (12)
Width of choroidal fissure at e12.5 (µm) ^c	10.32 ± 8.16	4.23 ± 7.40	26.37 ± 27.31

Note: ^a From *Chd7*^{Gt/+} X WT FVB/N crosses

^b From *Chd7*^{Gt/+} X *Toupee*^{Tg/Tg} crosses

^c n=16 for *Chd7*^{Gt/+} and 34 for the others

Table S4: Dysregulation of the NCC gene regulatory network in *Toupee*^{Tg/Tg} NCCs

Category	Gene name
Effectors of neural crest induction signals	<i>Lef1, Lrp5, Fzd3, Fzd5*, Fzd7, Bmpr1a, Bmpr2, Smad1, Smad5, Notch1, Notch2, Notch3, Dll1, Hes5</i>
Neural crest specifiers	<i>Sox8, Sox10, Tfp2a, Tfp2b, Tfp2c, Pax3, Mycn, Meis3, Dnmt3a, Dnmt3b, Ets1, Cited4*, Nr2f2, Tead1, Tead2</i>
Neural crest AP patterning	<i>Ednra, Hoxb2, Hoxa3, Hoxb5</i>
Neural crest EMT and migration	<i>Lims1*, Cdh2, Cdh6, Mmp14, Adam10*, Rxrg, Gja1, Vangl1*, Vangl2, Prickle1, Prickle2, Daam2, Ptk7, Celsr1, Celsr2, Celsr3, Erbb2, Erbb3, Erbb4</i>
Neural crest pathfinding	<i>Ephb3, Epha7, Ephb2, Ephb4, Epha8*, Efna5, Nrp1, Nrp2, Plxna1, Plxna3, Plxna4, Plxnb1, Plxnc1, Sem3b, Sem3d, Sem3g, Sem4c, Sem4g, Sem6a, Robo1, Robo2, Slit2, Slit3</i>
Craniofacial differentiation	<i>Sox5, Alx3, Rara</i>
Melanocyte differentiation	<i>Mitf, Tyr, Dct, Mef2c, Kit, Nf1, Ctb2</i>
Neuronal differentiation	<i>Phox2b, Phox2a, Hand2, Ascl1, Th, Dbh, Runx1*</i>
Glial differentiation	<i>Plp1, Fabp7, Rela*</i>
Enteric nervous system	<i>Ednrb, Ret, Gfra1, Gfra2, Gfra3, Foxo3*, L1cam, Dcc, Kif26a, Lgi4, Prokr1*</i>

Note: Genes included were those whose expression was modulated at least 1.5-fold ($P \leq 0.01$). Asterisks indicate upregulated genes.

Table S5: Fam172a interacting proteins in the chromatin fraction of Neuro2a cells

Category	Protein name	Gene name	Accession number	MW (kDa)	emPAI
Histones	Histone H4	Hist1h4a	P62806	11	10,000
	Histone H3.2	Hist1h3b	P84228	15	5,640
	Histone H2A type 2-A	Hist2h2aa1	Q6GSS7	14	5,143
	Core histone macro-H2A.1	H2afy	Q9QZQ8	39	2,480
	Histone H2AX	H2afx	P27661	15	1,577
	Histone H1.5	Hist1h1b	P43276	22	1,415
	Histone H1.1	Hist1h1a	P43275	21	1,167
	Histone H1.3	Hist1h1d	P43277	22	0,716
	Histone H1.2	Hist1h1c	P15864	21	0,673
	Histone H1.4	Hist1h1e	P43274	21	0,598
	H1 histone family, member X	H1fx	Q80ZM5	20	0,149
	Histone H1.0	H1f0	P10922	20	0,126
Histone regulators	Histone-binding protein RBBP7	Rbbp7	Q60973	47	0,138
	Histone-lysine N-methyltransferase ASH1L	Ash1l	Q99MY8	331	0,082
	Histone deacetylase 2	Hdac2	P70288	55	0,068
	Histone deacetylase 7	Hdac7	Q8C2B3	101	0,046
	Histone-lysine N-methyltransferase EHMT2	Ehmt2	Q9Z148	138	0,013
	JmjC domain-containing histone demethylation protein 2C	Jmjd1c	Q69ZK6	260	0,003
Other chromatin factors	Regulator of chromosome condensation	Rcc1	Q8VE37	44	0,374
	FACT complex subunit SSRP1	Ssrp1	Q08943	81	0,137
	Heterochromatin protein 1-binding protein 3	Hp1bp3	Q3TEA8	60	0,013
	DNA replication licensing factor MCM5	Mcm5	P49718	82	0,009
mRNA splicing	Serine/arginine-rich splicing factor 1	Srsf1	Q6PDM2	27	1,410
	Heterogeneous nuclear ribonucleoprotein A/B	Hnrnpab	Q99020	30	0,796
	Heterogeneous nuclear ribonucleoprotein K	Hnrnpk	P61979	50	0,698
	Heterogeneous nuclear ribonucleoprotein D0	Hnrnpd	Q60668	38	0,314
	Heterogeneous nuclear ribonucleoproteins C1/C2	Hnrnpc	Q9Z204	34	0,303
	Polypyrimidine tract-binding protein 1	Ptbp1	P17225	56	0,183
	Serine/arginine-rich splicing factor 7	Srsf7	Q8BL97	30	0,083
	ELAV-like protein 1	Elavl1	P70372	36	0,075
	NHP2-like protein 1	Nhp2l1	E9PZS4	5	0,055
	RNA-binding protein FUS	Fus	P56959	52	0,048
	Intron-binding protein aquarius	Aqr	Q8CFQ3	170	0,011
RNA processing	Ribosomal L1 domain-containing protein 1	Rsl1d1	Q8BVY0	50	0,242
	Pumilio domain-containing protein KIAA0020	Kiaa0020	Q8BKS9	72	0,051
	U3 small nucleolar RNA-associated protein 18 homolog	Utp18	Q5SSI6	61	0,012
Ribosome biogenesis	40S ribosomal protein S3	Rps3	P62908	26	2,542
	Nucleophosmin	Npm1	Q61937	32	2,540
	Nucleolin	Ncl	P09405	76	0,975
	60S ribosomal protein L18a	Rpl18a	P62717	20	0,700
	60S ribosomal protein L5	Rpl5	P47962	34	0,325
	60S ribosomal protein L6	Rpl6	P47911	33	0,311
	Nucleolar transcription factor 1	Ubtf	P25976	89	0,132
	40S ribosomal protein S25	Rps25	P62852	13	0,129
	60S ribosomal protein L22	Rpl22	P67984	14	0,053
	60S ribosomal protein L19	Rpl19	P84099	23	0,033
GTPase-related proteins	GTP-binding nuclear protein Ran	Ran	P62827	24	0,430
	GTP-binding protein Di-Ras2	Diras2	Q5PR73	22	0,076
	Arf-GAP with SH3 domain, ANK repeat and PH domain-containing protein 2	Asap2	Q7SIG6	106	0,016

Cytoskeleton-associated proteins	Actin, cytoplasmic 1	Actb	P60710	41	1,141
	Neuronal migration protein doublecortin	Dcx	O88809	40	0,256
	Junction plakoglobin	Jup	Q02257	81	0,155
	Echinoderm microtubule-associated protein-like 6	Eml6	Q5SQM0	217	0,008
	Desmoglein-1-alpha	Dsg1a	Q61495	114	0,007
Other	Casein kinase II subunit alpha	Csnk2a1	Q60737	45	0,247
	Elongation factor 1-alpha 1	Eef1a1	P10126	50	0,087
	Glyceraldehyde-3-phosphate dehydrogenase	Gapdh	P16858	35	0,095
	Chymotrypsinogen B	Ctrb1	Q9CR35	27	0,062
	Voltage-dependent anion-selective channel protein 3	Vdac3	Q60931	30	0,053
	Sorting nexin-29	Snx29	Q9D3S3	91	0,031
	Stress-70 protein, mitochondrial	Hspa9	P38647	73	0,021
	Fibrous sheath-interacting protein 2	Fsip2	A2ARZ3	784	0,019

Note: Proteins were included if enriched at least 1.5-fold in comparison to the MBP negative control and detected in at least two out of the three biological replicates of this analysis. The indicated emPAI (exponentially modified Protein Abundance Index) value corresponds to the average of the three biological replicates. Accession number is for the Uniprot database.

Table S6: Fam172a interacting proteins in the nucleoplasm fraction of Neuro2a cells

Category	Protein name	Gene name	Accession number	MW	emPAI
RNA binding /processing	Lupus La protein homolog	Ssb	P32067	47	0,217
	ATP-dependent RNA helicase A	Dhx9	O70133	149	0,090
	Nucleoprotein TPR	Tpr	F6ZDS4	273	0,053
	Exosome complex component MTR3	Exosc6	Q8BTW3	28	0,039
	tRNA (cytosine(34)-C(5))-methyltransferase	Nsun2	Q1HFZ0	85	0,030
	DAZ-associated protein 1	Dazap1	Q9JII5	43	0,025
	Heterogeneous nuclear ribonucleoprotein Q	Syncrip	Q7TMK9	69	0,016
	Bifunctional glutamate/proline--tRNA ligase	Eprs	Q8CGC7	170	0,008
mRNA splicing	Nuclease-sensitive element-binding protein 1	Ybx1	P62960	35	0,130
	Serine/arginine-rich splicing factor 3	Srsf3	P84104	19	0,114
	Pre-mRNA-splicing regulator WTAP	Wtap	Q9ER69	44	0,084
	Interleukin enhancer-binding factor 2	Ilf2	Q9CXY6	43	0,080
	TAR DNA-binding protein 43	Tardbp	Q921F2	44	0,039
	RNA-binding protein FUS	Fus	P56959	52	0,021
	Far upstream element-binding protein 2	Khsrp	Q3U0V1	76	0,018
DNA binding/repair	Histone-binding protein RBBP7	Rbbp7	Q60973	47	0,173
	Superoxide dismutase [Cu-Zn]	Sod1	P08228	15	0,140
	DNA damage-binding protein 1	Ddb1	Q3U1J4	126	0,029
	DNA dC->dU-editing enzyme APOBEC-3	Apobec3	Q99J72	50	0,023
	Tyrosine-protein kinase BAZ1B	Baz1b	Q9Z277	170	0,018
	Tumor suppressor p53-binding protein 1	Tp53bp1	P70399	211	0,002
Cytoskeleton-associated proteins	Tubulin beta-5 chain	Tubb5	P99024	49	0,546
	MARCKS-related protein	Marcksl1	P28667	20	0,088
	Peripherin	Prph	P15331	53	0,083
	WD repeat-containing protein 1	Wdr1	O88342	66	0,064
	Drebrin	Dbn1	Q9QXS6	77	0,021
	Echinoderm microtubule-associated protein-like 6	Eml6	Q5SQM0	217	0,005
	Spectrin beta chain, non-erythrocytic 1	Sptbn1	Q62261	274	0,004
	Microtubule-actin cross-linking factor 1	Macf1	Q9QXZ0	831	0,001
Other	Leukocyte elastase inhibitor B	Serp1nb1b	Q8VHP7	42	1,870
	Elongation factor 1-alpha 1	Eef1a1	P10126	50	0,360
	ATP synthase subunit alpha	Atp5a1	Q03265	59	0,299
	GTP-binding protein Di-Ras2	Diras2	Q5PR73	22	0,285
	40S ribosomal protein S3	Rps3	P62908	26	0,233
	Malate dehydrogenase, mitochondrial	Mdh2	P08249	35	0,165
	Fibrous sheath-interacting protein 2	Fsip2	A2ARZ3	784	0,110
	Myristoylated alanine-rich C-kinase substrate	Marcks	P26645	29	0,061
	Asparagine synthetase	Asns	Q61024	64	0,037
	E3 ubiquitin-protein ligase TRIM8	Trim8	Q99PJ2	61	0,019
	Teneurin-1	Tenm1	Q9WTS4	305	0,003

Note: Proteins were included if enriched at least 1.5-fold in comparison to the MBP negative control and detected in at least two out of the three biological replicates of this analysis. The indicated emPAI (exponentially modified Protein Abundance Index) value corresponds to the average of the three biological replicates. Accession number is for the Uniprot database.

Table S7: Fam172a interacting proteins in the cytoplasm fraction of Neuro2a cells

Category	Protein name	Gene name	Accession number	MW	emPAI	
RNA binding /processing	Bifunctional glutamate/proline--tRNA ligase	Eprs	Q8CGC7	170	0,259	
	Nuclease-sensitive element-binding protein 1	Ybx1	P62960	36	0,148	
	Eukaryotic translation initiation factor 2 subunit 1	Eif2s1	Q6ZWX6	36	0,039	
	RNA-binding protein FUS	Fus	P56959	52	0,015	
	Squamous cell carcinoma antigen recognized by T-cells 3	Sart3	Q9JLI8	109	0,007	
Nucleotide binding /processing	ATP synthase subunit beta, mitochondrial	Atp5b	P56480	56	1,565	
	ADP/ATP translocase 1	Slc25a4	P48962	32	1,384	
	Nucleoside diphosphate kinase A	Nme1	P15532	17	1,307	
	High mobility group protein B2	Hmgb2	P30681	24	0,523	
	ADP/ATP translocase 2	Slc25a5	P51881	286	0,518	
	Up-regulated during skeletal muscle growth protein 5	Usmg5	Q78IK2	5	0,410	
	ATP synthase subunit gamma, mitochondrial	Atp5c1	Q91VR2	32	0,126	
	GTP-binding protein Di-Ras2	Diras2	Q5PR73	22	0,074	
	5'-AMP-activated protein kinase catalytic subunit alpha-1	Prkaa1	Q5EG47	63	0,048	
	Arf-GAP with SH3 domain, ANK repeat and PH domain-containing protein 2	Asap2	Q7SIG6	106	0,015	
Protein processing /repair	Peptidyl-prolyl cis-trans isomerase A	Ppia	P17742	17	6,238	
	Prohibitin	Phb	P67778	29	2,208	
	Elongation factor 1-alpha 1	Eef1a1	P10126	50	1,083	
	Prohibitin-2	Phb2	O35129	33	0,825	
	Elongation factor 2	Eef2	P58252	95	0,624	
	Elongation factor 1-gamma	Eef1g	Q9D8N0	44	0,465	
	T-complex protein 1 subunit alpha	Tcp1	P11983	60	0,345	
	Aminoacyl tRNA synthase complex-interacting multifunctional protein 2	Aimp2	Q8R010	35	0,249	
	Thioredoxin	Txn	P10639	11	0,238	
	Prolyl endopeptidase	Prep	Q9QUR6	80	0,191	
	26S protease regulatory subunit 6B	Psmc4	P43686	47	0,058	
	Transcription factor BTF3	Btf3	Q64152	22	0,057	
	Alpha-2-HS-glycoprotein	Ahsg	P29699	37	0,028	
Metabolic processes	Glyceraldehyde-3-phosphate dehydrogenase	Gapdh	P16858	35	1,586	
	L-lactate dehydrogenase A chain	Ldha	P06151	36	1,005	
	Aspartate aminotransferase, mitochondrial	Got2	P05202	47	0,767	
	Glucose-6-phosphate isomerase	Gpi	P06745	62	0,412	
	Bifunctional purine biosynthesis protein PURH	Atic	Q9CWJ9	65	0,298	
	D-3-phosphoglycerate dehydrogenase	Phgdh	Q61753	56	0,290	
	Fatty acid-binding protein, epidermal	Fabp5	Q05816	15	0,255	
	Malate dehydrogenase, cytoplasmic	Mdh1	P14152	36	0,190	
	Inorganic pyrophosphatase	Ppa1	Q9D819	32	0,185	
	Cytosolic acyl coenzyme A thioester hydrolase	Acot7	Q91V12	42	0,164	
	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase	Ddost	O54734	49	0,115	
	Dihydrolipoyllysine-residue succinyltransferase	Dlst	Q9D2G2	48	0,086	
	Glycogen phosphorylase, brain form	Pygb	Q8CI94	96	0,027	
	Procollagen-lysine,2-oxoglutarate 5-dioxygenase 3	Plod3	Q9R0E1	84	0,025	
	Cytoskeleton-associated proteins	Clathrin heavy chain 1	Cltc	Q68FD5	191	0,220
		T-complex protein 1 subunit gamma	Cct3	P80318	60	0,197
Spectrin beta chain, non-erythrocytic 1		Sptbn1	Q62261	275	0,065	
Desmoplakin		Dsp	E9Q557	332	0,046	

Other	Leukocyte elastase inhibitor B	Serpinb1b	Q8VHP7	42	4,340
	14-3-3 protein zeta/delta	Ywhaz	P63101	27	1,015
	Voltage-dependent anion-selective channel protein 1	Vdac1	Q60932	32	0,519
	Macrophage migration inhibitory factor	Mif	P34884	15	0,480
	Voltage-dependent anion-selective channel protein 3	Vdac3	Q60931	30	0,306
	Voltage-dependent anion-selective channel protein 2	Vdac2	Q60930	31	0,113

Note: Proteins were included if enriched at least 1.5-fold in comparison to the MBP negative control and detected in at least two out of the three biological replicates of this analysis. The indicated emPAI (exponentially modified Protein Abundance Index) value corresponds to the average of the three biological replicates. Accession number is for the Uniprot database.

Table S8: Genetic and phenotypic information about the CHARGE syndrome patients included in this study

Patient ID (sex)	CHD7 variant	FAM172A variant	Possible transmission	Phenotypic presentation
CHA194 (XX female)	NM_017780.3: c.5050+1G>T	no	<i>De novo</i> dominant	Coloboma, choanal atresia, ear malformations, hearing loss, delayed puberty, short stature
CHA272 (XY male)	no	no	Recessive or <i>de novo</i> dominant	Coloboma, choanal atresia, ear malformation, hearing loss, heart malformation, swallowing problems, developmental delay
CHA275 (XX female)	no	no	Recessive or <i>de novo</i> dominant	Choanal atresia, Mondini malformation, hearing loss, growth retardation, swallowing problems
CHA333 (XX female)	no	no	Recessive or <i>de novo</i> dominant	Coloboma, choanal stenosis, hearing loss, ear malformation, heart malformation, short stature, swallowing problems
CHA367 (XX female)	no	no	Recessive or <i>de novo</i> dominant	Coloboma, hypoplastic optic nerves, unilateral choanal atresia, ear malformation, heart malformation, short stature, agenesis of corpus callosum, swallowing problems
CHA441^a (XX female)	no	NM_032042.5: c.682G>C	<i>De novo</i> dominant	Choanal atresia, small abnormal ears, moderate hearing loss, vestibular dysfunction
CHA442^a (XX female)	no	NM_032042.5: c.682G>C NM_032042.5: c.916C>T	Inherited from mother / <i>De novo</i> dominant	Choanal atresia, profound hearing loss, small left kidney, genital hypoplasia, delayed puberty, growth retardation, growth hormone deficiency, hypothyroidism

Note: ^a Mother (CHA441) and child (CHA442) from the same family.

Table S9: Genetic and phenotypic information about the FAM172A mutation-positive mother-child pair from the replication cohort

Family member	CHD7 variant	FAM172A variant	Possible transmission	CHARGE syndrome-related features
Mother	no	NM_032042.5: c.682G>C	<i>De novo</i> dominant	Choanal atresia, blocked tear ducts, severe hearing loss, growth hormone deficiency, hydronephrotic left kidney, hypothyroidism
Child	no	NM_032042.5: c.682G>C	Inherited from mother	Choanal atresia, blocked tear ducts, strabismus, abnormal ears, severe hearing loss, developmental delay, immune deficiency, hypothyroidism

Table S10: Antibodies used in this study

	Antibody Name	Dilution	Source	RRID
Primary Antibodies	Rabbit molyclonal Anti-Fam172a	WB (1:1000) IF (1:500) IP (2 µg/ml)	Abcam ab121364	AB_11127114
	Rat monoclonal Anti-AGO2	IF(1:1000)	Sigma SAB4200085	AB_10600719
	Rat monoclonal Anti-AGO1	WB (1:1000) IF (1:1000) IP (3 µg/ml)	Sigma SAB4200084	AB_10602786
	Rabbit monoclonal Anti-CHD7 (D3F5)	WB (1:1000) IP (3 µg/ml)	Cell Signaling #6505	AB_11220431
	Rabbit polyclonal Anti-AGO2	WB (1:1000) IF (1:500) IP (3 µg/ml)	Abcam ab32381	AB_867543
	Rabbit polyclonal Anti-Active Caspase 3	IF (1:500)	Abcam ab13847	AB_443014
	Goat polyclonal Anti-Sox10	IF (1:250)	Santa Cruz Sc-17342	AB_2195374
	Mouse monoclonal Anti-Neurofilament	IHC (1:500)	DSHB 2H3	AB_2618380
	Rabbit polyclonal Anti-H3	WB (1:2500) IP (1 µg/ml)	Abcam ab1791	AB_302613
	Rabbit polyclonal Anti-Ki67	IF (1:1000)	Abcam ab15580	AB_443209
	Mouse monoclonal Anti-p53	IF (1:250)	Cell Signaling #2524 (1C12)	AB_331743
	Mouse monoclonal Anti-Myc tag	WB (1:500) IF (1:100)	In house hybridoma (9E10)	-
Secondary Antibodies	Donkey Alexa Fluor 488 Anti-Rat IgG	IF (1:500)	Jackson ImmunoResearch 712-545-150	AB_2340683
	Donkey Alexa Fluor 594 Anti-Rabbit IgG	IF (1:500)	Jackson ImmunoResearch 711-585-152	AB_2340621
	Donkey Alexa Fluor 647 Anti-Mouse IgG	IF (1:500)	Jackson ImmunoResearch 715-605-150	AB_2340862
	Bovine Alexa Fluor 647 Anti-Goat IgG	IF (1:500)	Jackson ImmunoResearch 805-605-180	AB_2340885
	Donkey Anti-Rabbit IgG HRP	WB (1:10 000)	AbD Serotec Star124p	AB_615912
	Donkey Anti-Mouse IgG HRP	WB (1:10 000)	AbD Serotec Star117	AB_324488
	Goat Anti-Rat IgG HRP	WB (1:10 000)	Santa Cruz Sc-2032	AB_631755

Table S11: Oligonucleotide primers used for RT-qPCR

Oligo Name	Sequence (5'-3')
Fam172a Forward	AGG TGA CTG CTG TGG CAT TGA C
Fam172a Reverse	GGC TTC TGC GAG CTG CTC TT
FAM172A Forward	ACC GCC TCT TGA TTT TCC TGA
FAM172A Reverse	GCC TCG TAT CTT TTC TGG TTC C
Nr2f1 Forward	CTT TCA GGA ACA GGT GGA GAA GC
Nr2f1 Reverse	AGG AAC ACT GGA TGG ACA TGT AAG G
A830082K12Rik Forward	TTC AAT AGT TCT GCT GAA TGC TCC
A830082K12Rik Reverse	ATA AAG TAG GGA TCC GTT TCC AGT TGT GTC AAC
Pou5f2 Forward	AGT GTG GTT TTC TAA CCG GAG CCA
Pou5f2 Reverse	TCA GAG AAC TGA ACG CTA ACC CTG
Gapdh Forward	CTG TGG CGT GAT GGC CGT GG
Gapdh Reverse	CCT CAG TGT AGC CCA AGA TG
Cd44_ Ex8 (v4) Forward	TAC CCC AGT TTT TCT GGA TCA GG
Cd44_ Ex9 (v5) Reverse	GCC ATC CTG GTG GTT GTC TG
CD44_ Ex7 (v4) Forward	AGG CTG GGA GCC AAA TGA AG
CD44_ Ex8 (v5) Reverse	CCG GAT TTG AAT GGC TTG GG
Cd44_ Ex13 (v9) Forward	TGG AAG ACT TGA ACA GGA CAG G
Cd44_ Ex14 (v10) Reverse	GTT TTC GTC TTC TTC CGG CTC
CD44_ Ex12 (v9) Forward	GCT TCA GCC TAC TGC AAA TCC
CD44_ Ex13 (v10) Reverse	GCC TTC ATG TGA TGT AGA GAA GC
Cd44_ Ex4 Forward	ACA GAC CTA CCC AAT TCC TTC G
_ Ex5 Reverse	GGG TGC TCT TCT CGA TGG TG
CD44_ Ex4 Forward	TGC CCA ATG CCT TTG ATG GAC
CD44_ Ex5 Reverse	GAA GTG CTG CTC CTT TCA CTG
Cd44_ Ex16 Forward	GGA GTT CCC GCA CTG TGA C
Cd44_ Ex17 Reverse	GGT CTC CTC ATA GGA CCA GAA G
CD44_ Ex16 Forward	AGG TGG AGC AAA CAC AAC CTC
CD44_ Ex17 Reverse	GAA TCA AAG CCA AGG CCA AGA G
Col5a3_ EX6 Forward	CCC CAA AGA CGA TGA ACC AG
Col5a3_ EX7 Reverse	TCT CTG TCT CAG GGA TGT GGA
COL5A3_ EX6 Forward	TCA AGT CCA CCT CCT GAC TCC
COL5A3_ EX7 Reverse	TAG GAT CGT GGC ATT GAG TCC
Col5a3_ EX28 Forward	TAC CTC TGG TAA CCG GGG TCT C
Col5a3_ EX29 Reverse	CCT TTT GGT CCC TCA TCA CCC
COL5A3_ EX28 Forward	AGG GGG AGA AAG GCG AGA A
COL5A3_ EX29 Reverse	CCT TTT GGT CCC TCA TCA CCC
Col5a3_ EX66 Forward	GCA GCC CAT CAG AGG TTC AC
Col5a3_ EX67 Reverse	AAG AGG GTC TTC GCC TGT CC
COL5A3_ EX66 Forward	ACC TAC TCC TGC CAG AAT GC

COL5A3_EX67 Reverse	GGT CTT CGT CTG TCC TTT CCG
Col5a3_EX10 Forward	GGA CAG CAG TTT GAG GGG
Col5a3_EX11 Reverse	CAC GGT CCC CAG GGA AG
COL5A3_EX10 Forward	GGG CAG CAG TTT GAG GGA C
COL5A3_EX11 Reverse	GGG TCG CCA GGG AAT CCT
Ift74_EX10 Forward	ACA CTT CAG CAA CAG CTA GAT TC
Ift74_EX11 Reverse	GGC GAT CCC ATG CTT TTG TC
IFT74_EX10 Forward	TCA ACA ACA ATT GGA TTC ACA GAA C
IFT74_EX11 Reverse	TGA TCT CGA TGG GAC TCC AAC
Ift74_EX5 Forward	TGT ACA ACC AAG AAA ATT CAG TGT
Ift74_EX6 Reverse	TGT TGT AGT CTG CTA GTT GTC CT
IFT74_EX5 Forward	TGT CAT ATG AAA AGA GGG CTG AG
IFT74_EX6 Reverse	TGT TGT AGT CTG CTA GTT GTC C
Mical2_EX8 Forward	GGA CAG TAC CCA CTA CTT TGT C
Mical2_EX9 Reverse	CCG AAC ACA GCA GCA TCT CT
MICAL2_EX8 Forward	CTG CTC GAC AAA GGT GTC ATC
MICAL2_EX9 Reverse	CGG GCA TAG GAT AGC AGG TT
Mical2_EX18 Forward	CGG GTC TCA GGC ATA GGT AAG
Mical2_EX19 Reverse	AGC CGG TAC TTG GTT GAG TTC
MICAL2_EX24 Forward	CGT GTG TAC GTG ATG GAA CG
MICAL2_EX25 Reverse	GCT TCA ACT CTG CCC GTC TC
Mical2_EX3 Forward	GTG GCA CAA ACT GGA TAA GCG
Mical2_EX4 Reverse	GCA GGA CAT TGT TCC GGG AG
MICAL2_EX3 Forward	CCT GGA AAG CCA AAG CCC TG
MICAL2_EX4 Reverse	GTT GTT CCG GGA GAA GGA GTC

Table S12: Oligonucleotide primers used for mouse genotyping as well as for cloning and mutagenesis of *Fam172a*

	Oligo Name	Sequence (5'-3')
Genotyping	Fam172a_Mut_Forward	GGG AGT AAG TCC TAC CAA TGT TAA ATC
	Fam172a_Mut_Reverse	TGA CCT TCT AAA CAG TCC CAT ATC CCC
	Fam172a_WT_Forward	GAA GTG GGA ACA AAA CAC CCT TGG
	Fam172a_WT_Reverse	CTA ATG AGT TTG GGG AAA TTA TCA TAG
	Chd7_WT/Mut_Forward	GCT CGC TTT TCA ACC TCA TTA CG
	Chd7_WT_Reverse	ATC AAC AAT GCC AAC CCA GGC
	Chd7_Mut_Reverse	CCT CCG ATT GAC TGA GTC GC
	Smcxy Forward	TGA AGC TTT TGG CTT TGA G
	Smcxy Reverse	CCA CTG CCA AAT TCT TTG G
	Zfy Forward	GAC CAG ATT GTT GTG GAA GTA CAA G
	Zfy Reverse	CCA GTG TGT CTG AAG TGT CAG CTG
	Fam172a ORF and mutagenesis	<i>Fam172a</i> ORF (<i>Xho</i> I) Forward
<i>Fam172a</i> ORF (<i>Eco</i> RI) Reverse		GAA TTC TCA CAG CTC CTC GTG CTT GAT CC
p.Glu229Gln (mut1_Arb2)		GGG AAG CGG <u>C</u> AA AGG AGA GAT AAG
p.Arg307* (mut2_Arb2)		CTG ATG ATT CAA <u>I</u> GA GAA GCA GAT
p.Ser294Ala (mut_hydrolase)		GTC CTC CAT AG <u>G</u> <u>C</u> AT GAG CCA CAA A

OTHER SUPPORTING INFORMATION

Dataset S1: RNAseq data of *Toupee*^{Tg/Tg} NCCs analyzed with edgeR and DESeq (3488 genes; FC ≥ 1.5; DESeq $P \leq 0.01$).

Dataset S2: RNAseq data of *Toupee*^{Tg/Tg} NCCs analyzed with rMATS (1166 events; inclusion level ≥ 0.1; $P < 0.01$).

Sheet 1: Skipped exons (611 events)

Sheet 2: Retained introns (368 events)

Sheet 3: Alternative 3' splice site (91 events)

Sheet 4: Alternative 5' splice site (69 events)

Sheet 5: Mutually exclusive exons (27 events)

Dataset S3: RNAseq data of *Chd7*-het granule neuron progenitors analyzed with rMATS (227 events; inclusion level ≥ 0.1; $P < 0.05$).

Sheet 1: Skipped exons (115 events)

Sheet 2: Retained introns (58 events)

Sheet 3: Alternative 3' splice site (26 events)

Sheet 4: Alternative 5' splice site (20 events)

Sheet 5: Mutually exclusive exons (8 events)

Dataset S4: RNAseq data of *Chd7*-null granule neuron progenitors analyzed with rMATS (253 events; inclusion level ≥ 0.1; $P < 0.05$).

Sheet 1: Skipped exons (147 events)

Sheet 2: Retained introns (52 events)

Sheet 3: Alternative 3' splice site (26 events)

Sheet 4: Alternative 5' splice site (17 events)

Sheet 5: Mutually exclusive exons (11 events)