

Supplemental Data

Dual Molecular Effects of Dominant *RORA* Mutations

Cause Two Variants of Syndromic Intellectual

Disability with Either Autism or Cerebellar Ataxia

Claire Guissart, Xenia Latypova, Paul Rollier, Tahir N. Khan, Hannah Stamberger, Kirsty McWalter, Megan T. Cho, Susanne Kjaergaard, Sarah Weckhuysen, Gaetan Lesca, Thomas Besnard, Katrin Öunap, Lynn Schema, Andreas G. Chiocchetti, Marie McDonald, Julitta de Bellescize, Marie Vincent, Hilde Van Esch, Shannon Sattler, Irman Forghani, Isabelle Thiffault, Christine M. Freitag, Deborah Sara Barbouth, Maxime Cadieux-Dion, Rebecca Willaert, Maria J. Guillen Sacoto, Nicole P. Safina, Christèle Dubourg, Lauren Grote, Wilfrid Carré, Carol Saunders, Sander Pajusalu, Emily Farrow, Anne Boland, Danielle Hays Karlowicz, Jean-François Deleuze, Monica H. Wojcik, Rena Pressman, Bertrand Isidor, Annick Vogels, Wim Van Paesschen, Lihadh Al-Gazali, Aisha Mohamed Al Shamsi, Mireille Claustres, Aurora Pujol, Stephan J. Sanders, François Rivier, Nicolas Leboucq, Benjamin Cogné, Souphatta Sasorith, Damien Sanlaville, Kyle Retterer, Sylvie Odent, Nicholas Katsanis, Stéphane Bézieau, Michel Koenig, Erica E. Davis, Laurent Pasquier, and Sébastien Küry

Chr15q21.3q22.2

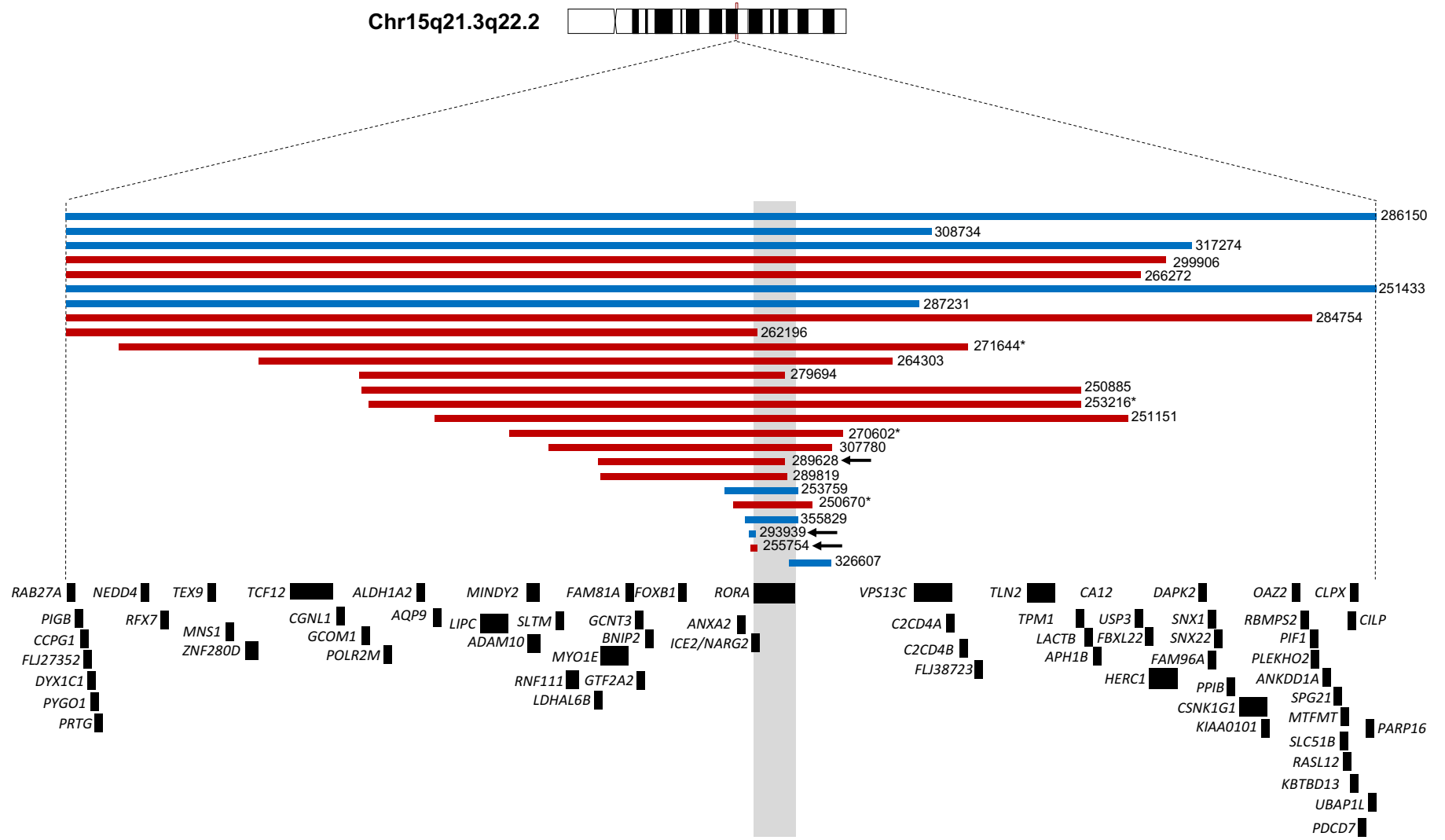


Figure S1

Figure S1. Schematic of copy number variants in the DECIPHER database encompassing the *RORA* locus on 15q22. Copy number loss (deletion) is shown in red; copy number gain (duplication) is shown in blue. DECIPHER individual case numbers are provided. Individuals included in our study are marked by arrows, and individuals reported previously (Yamamoto et al., 2014) are indicated by asterisks. Vertical gray shaded bar indicates the *RORA* locus. Database query was conducted in December 2017.

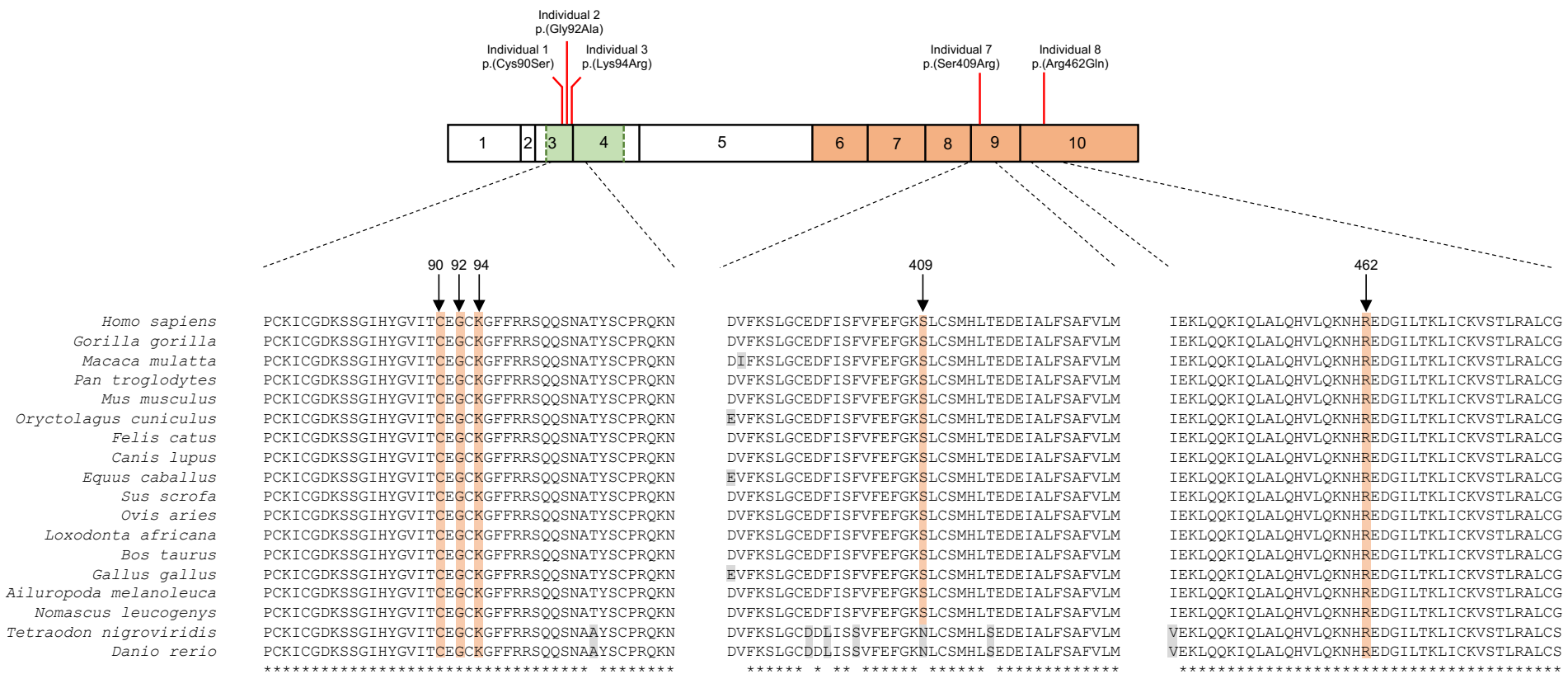


Figure S2

Figure S2. Multiple protein sequence alignment of RORA orthologs across vertebrate species show conservation of missense mutations detected in cases. Schematic of the RORa protein (NP_599023.1) and its domains (DNA binding domain, green; ligand binding domain, orange) and location of each of the missense variants reported in this study (top). Protein sequence alignments from 18 vertebrate species is shown (bottom). Orange shading in the alignment indicates RORA amino acids Cys90, Gly92, Lys94, Ser409, and Arg462. Grey shading indicates species-specific non-conserved amino acids. Asterisk indicates 100% conservation across mentioned species. Multiple sequence alignment was performed with ClustalW using the following UniProt identifiers: P35398, *Homo sapiens*; G3QE47, *Gorilla gorilla*; F7E9K9, *Macaca mulatta*; H2R8C9, *Pan troglodytes*; P51448, *Mus musculus*; G1T4B2, *Oryctolagus cuniculus*; M3WJ80, *Felis catus*; F1P607, *Canis lupus*; F6X794, *Equus caballus*; F1S075, *Sus scrofa*; W5QI24, *Ovis aries*; G3SLZ5, *Loxodonta africana*; F1N7R0, *Bos taurus*; F1NML9, *Gallus gallus*; G1LN19, *Ailuropoda melanoleuca*; G1RMH6, *Nomascus leucogenys*; H3CSJ1, *Tetraodon nigroviridis*; A7VL70, *Danio rerio*.

A

			90	92	94						
<i>Homo sapiens</i>	NR1F1/ROR-alpha	CKICGDKS	SGIHYGVITC	EGCKGFFRRS	QQSNA--TYS	CP	-RQKNCLIDRT	SRNRCQHCR	L	QK	
	NR1H3/LXR-alpha	CSVCGDKA	SGFHYNVLSC	EGCKGFFRRS	VIKGA--HYICH	-SGGHCPMDTY	MRRKCQECR	L	RK		
	NR1A1/TR-alpha	CVVCGDKA	TGYHYRCITC	EGCKGFFRRT	IQKNLHPTYS	CK	-YDSCVIDKI	TRNQCQLCR	F	KK	
	NR1I1/VDR	CGVCGDRA	TGFHFNAMTC	EGCKGFFRRS	MKRKA--LFT	CP	-FNGDCRITKD	NRRHQCACR	L	KR	
	NR1B1/RAR-alpha	CFVCQDKS	SGYHYGVSA	EGCKGFFRRS	IQKNM--VYT	CH	-RDKNCIINKV	TRNRCQYCR	L	QK	
	NR1C3/PPAR-gamma	CRVCGDKA	SGFHYGVHAC	EGCKGFFRRT	IRLKL--IYD	RC	--DLNCRIHKK	SRNKCQYCR	F	QK	
	NR1D1/Rev-erb-alpha	CKVCGDVA	SGFHYGVHAC	EGCKGFFRRS	IQQNI--QYK	RC	LKNENCIVRI	NRNRCQYCR	F	KK	
	NR4A2/NURR1	CAVCGDNA	ACQHYGVRTC	EGCKGFFKRT	VQKNA--KYV	CL	-ANKNCPVDR	RRNRCQYCR	F	QK	
	NR3A1/ER-alpha	CAVCNDYA	SGYHYGVWSC	EGCKAFFKRS	IQGHN--DYM	CP	-ATNQCTIDKN	RRKSCQACR	L	RK	
	NR3C3/PR	CLICGDEA	SGCHYGVLT	C	GSCKVFFKRA	MEGQH--NYL	CA	-GRNDIVDKI	RRKNCPCAR	L	RK
	NR3C2/MR	CLVCGDEA	SGCHYGVVT	C	GSCKVFFKRA	VEGQH--NYL	CA	-GRNDIIDKI	RRKNCPCAR	L	QK
	NR3C1/GR	CLVCSDEA	SGCHYGVLT	C	GSCKVFFKRA	VEGQH--NYL	CA	-GRNDIIDKI	RRKNCPCAR	L	RK
	NR3B1/ERR-alpha	CLVCGDVA	SGYHYGVAS	C	EACKAFFKRT	IQGSI--EYS	CP	-ASNECEITR	RRKACQYCR	F	TK
	NR5A1/SF-1	CPVCGDKV	SGYHYGLLTC	EGCKGFFKRT	VQNNK--HYT	CT	-ESQSKIDKT	QRKRCPCRF	C	QK	
	NR6A1/GCN1	CLICGDR	TGLHYGIISC	EGCKGFFKRS	ICNKR--VYR	CS	-RDKNCVMSR	QRNRCQYCR	L	LK	
	NR2C1/TR2	CVVCGDKA	SGRHYGAVTC	EGCKGFFKRS	IRKNL--VYS	CR	-GSKDCIINKH	HRNRCQYCR	L	QR	
	NR2B1/RXR-alpha	CAICGDRS	SGKHVGVYSC	EGCKGFFKRT	VRKDL--TYT	CR	-DNKCLIDKR	QRNRCQYCR	L	QK	
	NR2A1/HNF-4-alpha	CAICGDR	TGKHVGVYSC	DGCKGFFRRS	VRKNH--MYS	CR	-FSRQCVVDK	KRNQCRCR	L	KK	
	NR2F1/COUP1	CVVCGDKS	SGKHVGVYSC	EGCKSFFKRS	VRRNL--TYT	CR	-ANRNCPIDQH	HRNQCQYCR	L	KK	
	<i>Trichoplax adhaerens</i>	ERR-like XP_002117375.1	CLVCGDR	SGLHYGVLS	C	EGCKAFFKRS	IQSSV--AYT	CP	-SGSRCKVDK	RRKCCQACR	L
RXR-like XP_002109459.1		CSICGQRS	LRRHYGVYSC	EGCKGFFKRT	VRKNL--TYT	CR	-DNRNCDIDK	QRNRCQYCR	L	QK	
HNF4-like XP_002115810.1		CAICGDR	TGKHVGVYSC	DGCKGFFRRS	VRKNH--MYS	CR	-FSRQCVVDK	KRNQCRCR	L	KK	
COUP-like XP_002109806.1		CLICGDRS	NGRHYGVISC	EGCKGFFKRS	VRRNM--KYA	CT	CSANACKITKA	NRNQCQF	C	QK	

B

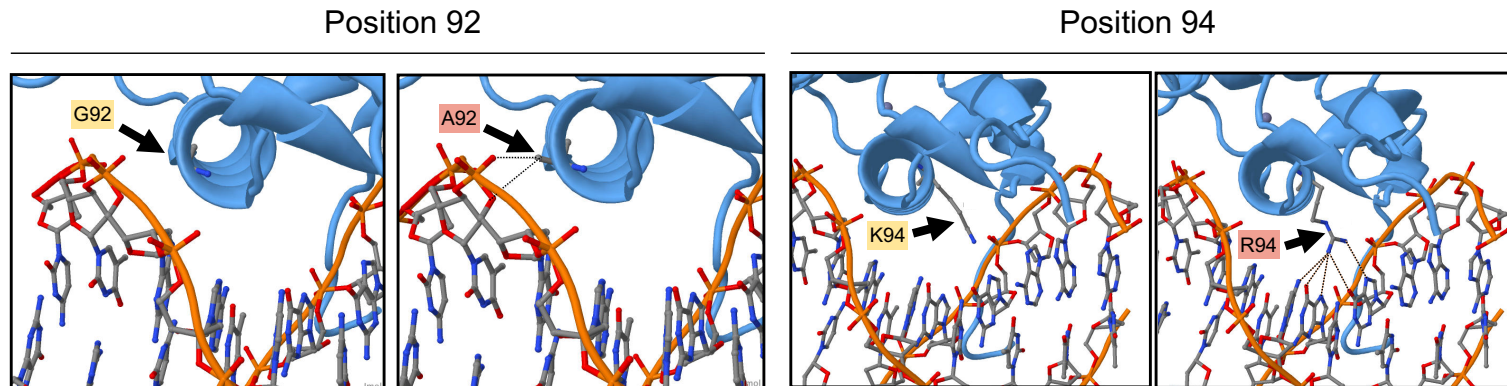


Figure S3

Figure S3. 3D modeling and sequence conservation of the first zinc finger motif of the RORa DNA-binding domain. (A) Sequence comparison of amino acids in the conserved DNA binding domain of human paralogous nuclear receptor proteins and of the four *Trichoplax adhaerens* (primitive metazoan) orthologs. The identifier following “NR” indicates the sub-family to which the nuclear receptors belong. Amino acids that match the consensus sequence are shown in bold and shaded gray. Residue numbering at the top of the alignment corresponds to RORa isoform a. The mutated amino acid cysteine (C) 90, glycine (G) 92, and lysine (K) 94 are indicated on top. Amino acids involved in the P-box are indicated by an asterisk (*). (B) Homology model of RORa DNA-binding domain in complex with RORE (ROR DNA response element consisting of the consensus core motif AGGTCA). RORa is depicted in blue as a backbone carbon CA trace and atoms are represented in a standard color scheme. Modeled interaction of RORa isoform a (NP_599023.1) wild type (WT) Gly92 and variant Ala92 (left) or WT Lys94 and variant Arg94 (right) in the recognition α -helix of the RORa DBD with the major groove of DNA. For either mutant, the putative clash between RORa DBD and the DNA are represented by black dotted lines. Clashes between the RORa DBD indicate either distances too short for van der Waals interaction (Ala92-phosphate backbone; left), or for hydrogen bonds (Lys94-guanine-guanine and complementary cytosine; right).

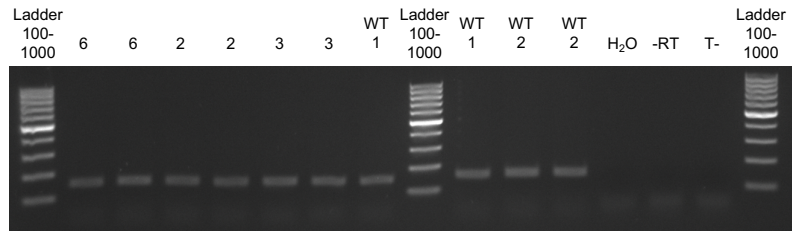
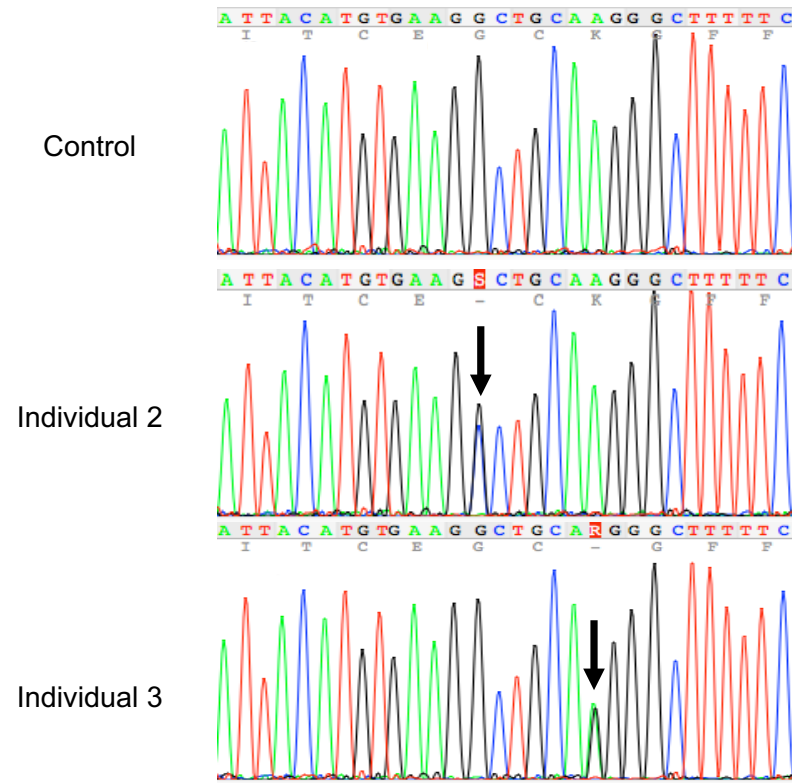
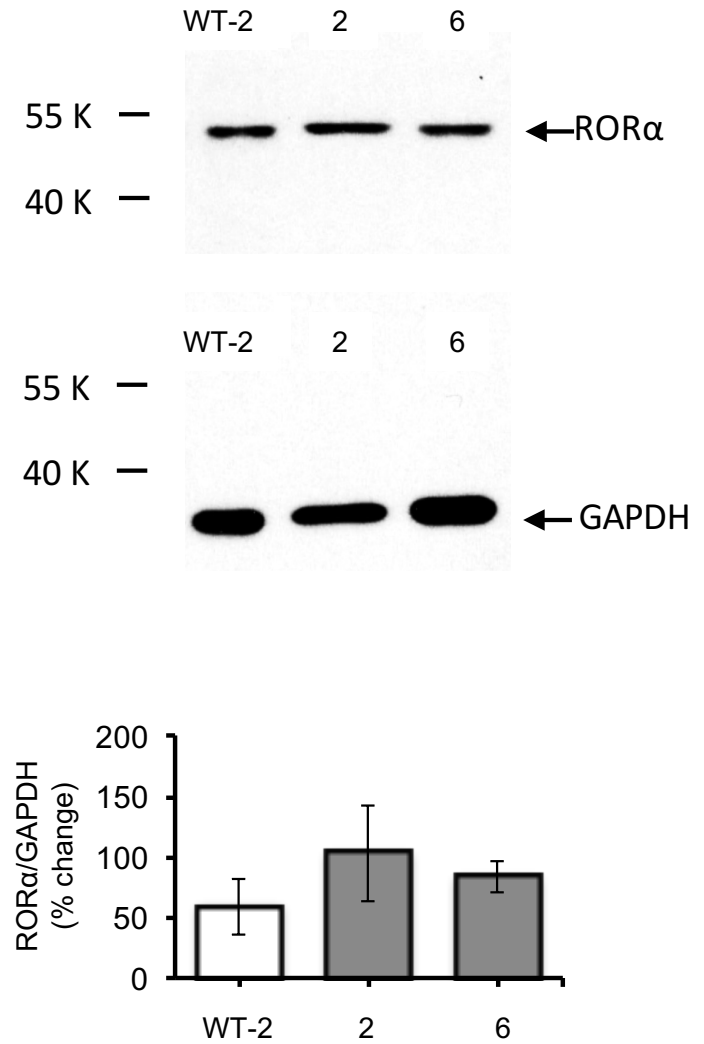
A**B****C****Figure S4**

Figure S4. Molecular and biochemical characterization of RORA variants. (A) RT-PCR analysis of a 152bp fragment of *RORA* cDNA encompassing the c.281A>G mutation in affected individuals 2, 3 and 6 and control individuals (WT1 and WT2). c.282G is the last nucleotide of exon 3 of *RORA* (NM_134261.2). Neither abnormal sized product nor semi-quantitative variation was observed, suggesting unimpaired splicing. (B) Sequencing of the *RORA* RT-PCR products, showing the presence of missense changes c.275G>C (p.Gly92Ala) and c.281A>G (p.Lys94Arg) in individuals 2 and 3, respectively. (C) Western blot analysis of protein lysates derived skin fibroblasts established from individual 2 and 6, and control individual WT2. The anti-ROR α antibodies were probed against fibroblast cell lysate and revealed a band of the expected size (55 kDa). No smaller product was detected, in particular no truncated fragment was detected for individual 6 bearing a heterozygous frame-shift mutation. The blot was reprobed with anti-GAPDH antibodies. The results show a moderate increase of ROR α levels in individual 2 and 6 when compared with wild-type levels as indicated by densitometry analysis. Quantification is based on 3 measurements (individuals 2 and 6) or 4 measurements (WT-2). A representative result of the Western blots is shown. Error bars represent standard error of the mean.

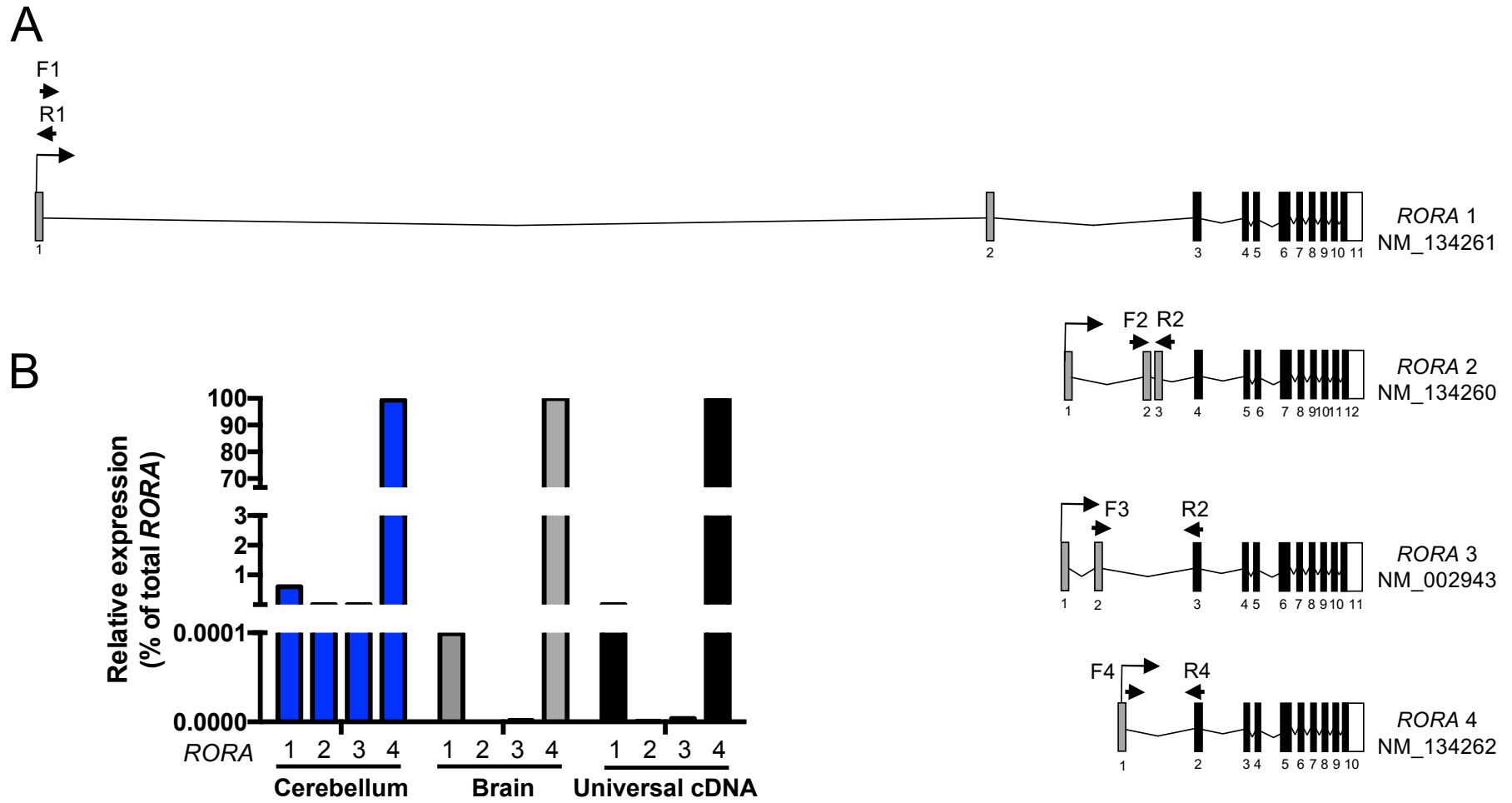
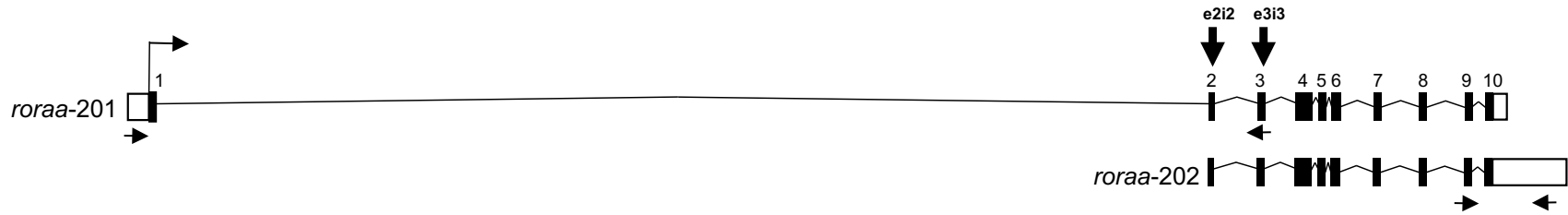


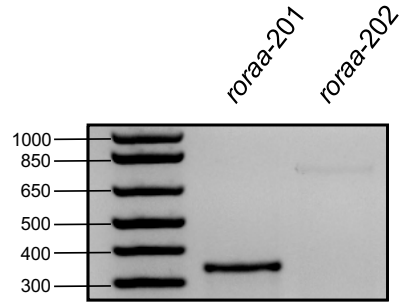
Figure S5

Figure S5. *RORA* encodes four protein coding splice variants. (A) Schematic of the four splice variants expressed from the human *RORA* locus with RefSeq annotations. Gray boxes, coding exons that are unique to each transcript; black boxes, coding exons that are shared among the four transcripts; white boxes, untranslated regions. Arrows indicate position of oligonucleotides used for qRT-PCR experiments. (B) *RORA* 4 is the major splice isoform expressed in the central nervous system. qRT-PCR was performed on control adult human cDNA and analyzed by the Δ Ct method with normalization to β -actin mRNA expression. Canonical isoform *RORA* 1 is weakly expressed, as well as *RORA* 2 and *RORA* 3.

A *rora* (*Danio rerio*)

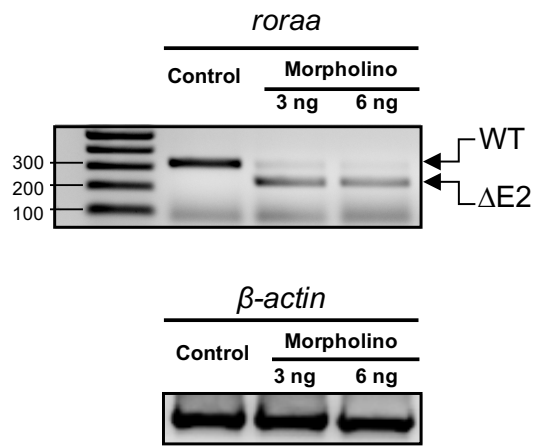


B



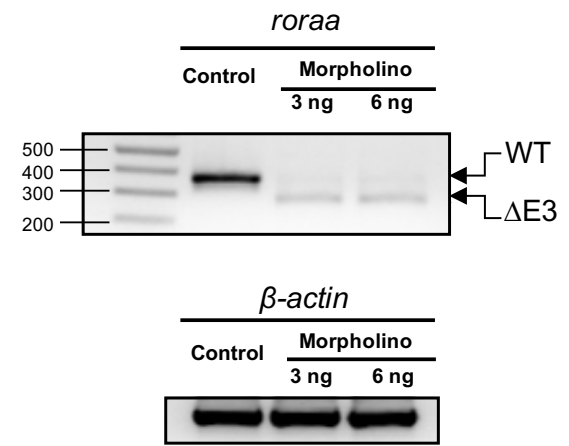
C

e2i2 splice-blocking morpholino



D

e3i3 splice-blocking morpholino



E

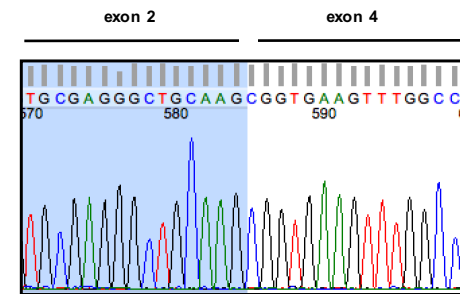
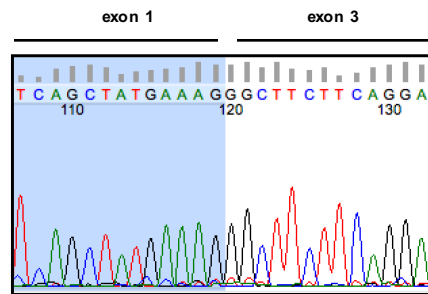
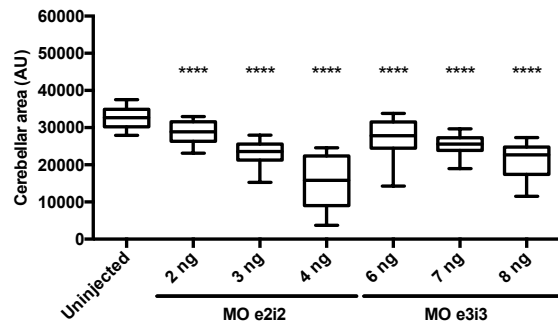
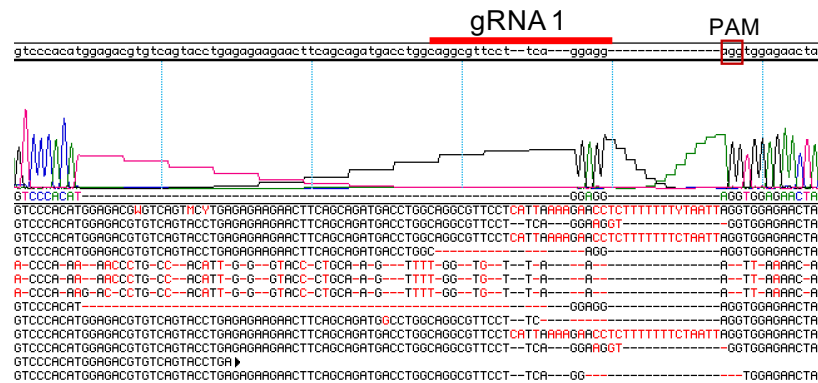
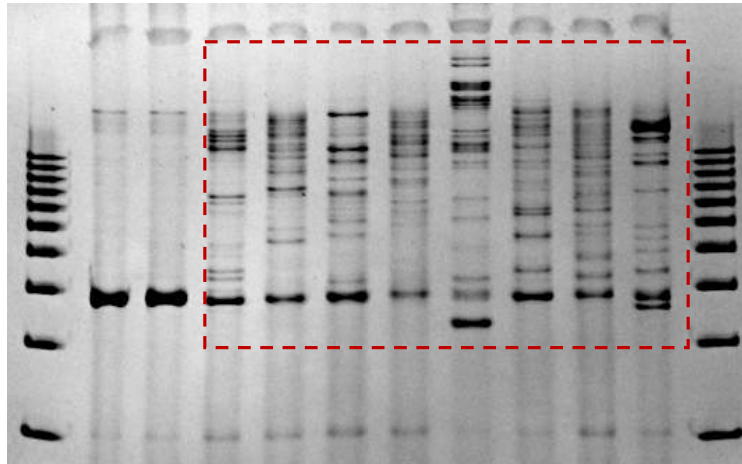


Figure S6

Figure S6. *rora* morpholino targeting efficiency. (A) Schematic of two zebrafish *rora* transcripts, *rora*-201 and *rora*-202 (genome assembly Zv10); coding regions, black boxes; untranslated regions, white boxes. Zebrafish *rora*-201 encodes GenBank ID: NP_001103637 (95% similar; 91% identical to human RORA 4). Vertical arrows indicate splice-blocking (sb) morpholino (MO) target sites on the donor sites of exons 2 and 3, respectively; horizontal arrows indicate RT-PCR primers used to generate amplicons in panel B. (B) *rora*-201 and *rora*-202 are detectable in zebrafish during larval development (3 days post-fertilization, dpf). Agarose gel image showing RT-PCR results obtained by using isoform-specific primers (shown in panel A). (C) Agarose gel images (top), and chromatograms of TOPO-cloned PCR product (bottom) demonstrate that the e2i2 sb MO induces exclusion of exon 2 (DE2) leading to a frameshift and introduction of a premature stop codon (p.Ile62*); WT, wild-type transcript. (D) Agarose gel images (top), and chromatogram of TOPO-cloned PCR product (bottom) shows that the e3i3 sb MO leads to skipping of exon 3 (DE3) leading to a frameshift and premature stop codon (p.Val88*). (E) *rora* e2i2 or e3i3 sb MOs lead to dose-dependent reduction of cerebellar size in 3 dpf larval batches immunostained with anti-acetylated tubulin antibody. Stars indicate *P* value compared to WT; ****, $p < 0.0001$. Error bars in (E) represent 5th and 95th Percentile.



B gRNA 1 UI *rora* F0 mutants



C gRNA 2 UI *rora* F0 mutants

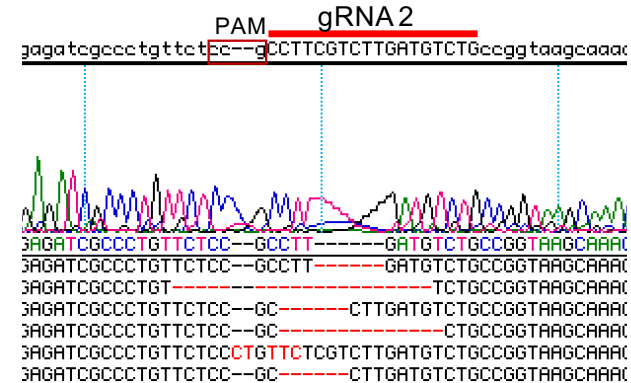
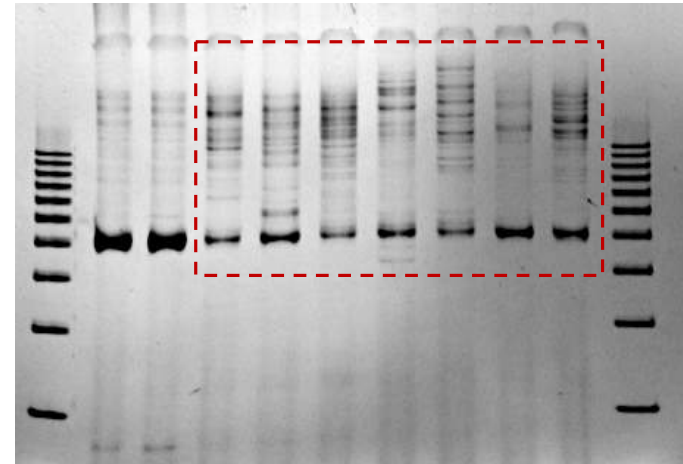


Figure S7

Figure S7. *rora* CRISPR/Cas9 genome editing efficiency. (A) Schematic of two zebrafish *rora* transcripts, *rora*-201 and *rora*-202 (genome assembly Zv10); coding regions, black boxes; untranslated regions, white boxes. guide (g)RNA target sites are by vertical arrows on exons 5 and 8. (B and C) Top, CRISPR/Cas9 targeting of *rora* F0 mutant embryos leads to small insertion and deletion events detectable by heteroduplex formation; PCR products flanking target sites were amplified, denatured, slowly reannealed and migrated on a polyacrylamide gel (n=2 uninjected [UI] controls and 8 or 7 F0 mutants for gRNA 1 and gRNA 2, respectively). Heteroduplexes are indicated by red boxes. Bottom, representative chromatograms indicating insertion and deletion events at the gRNA target. gRNA sequence and protospacer adjacent motifs (PAM) are shown. We sequenced 3 F0 mutants/gRNA and 24 colonies per embryo to confirm >90% mosaicism for both gRNA 1 and gRNA 2.

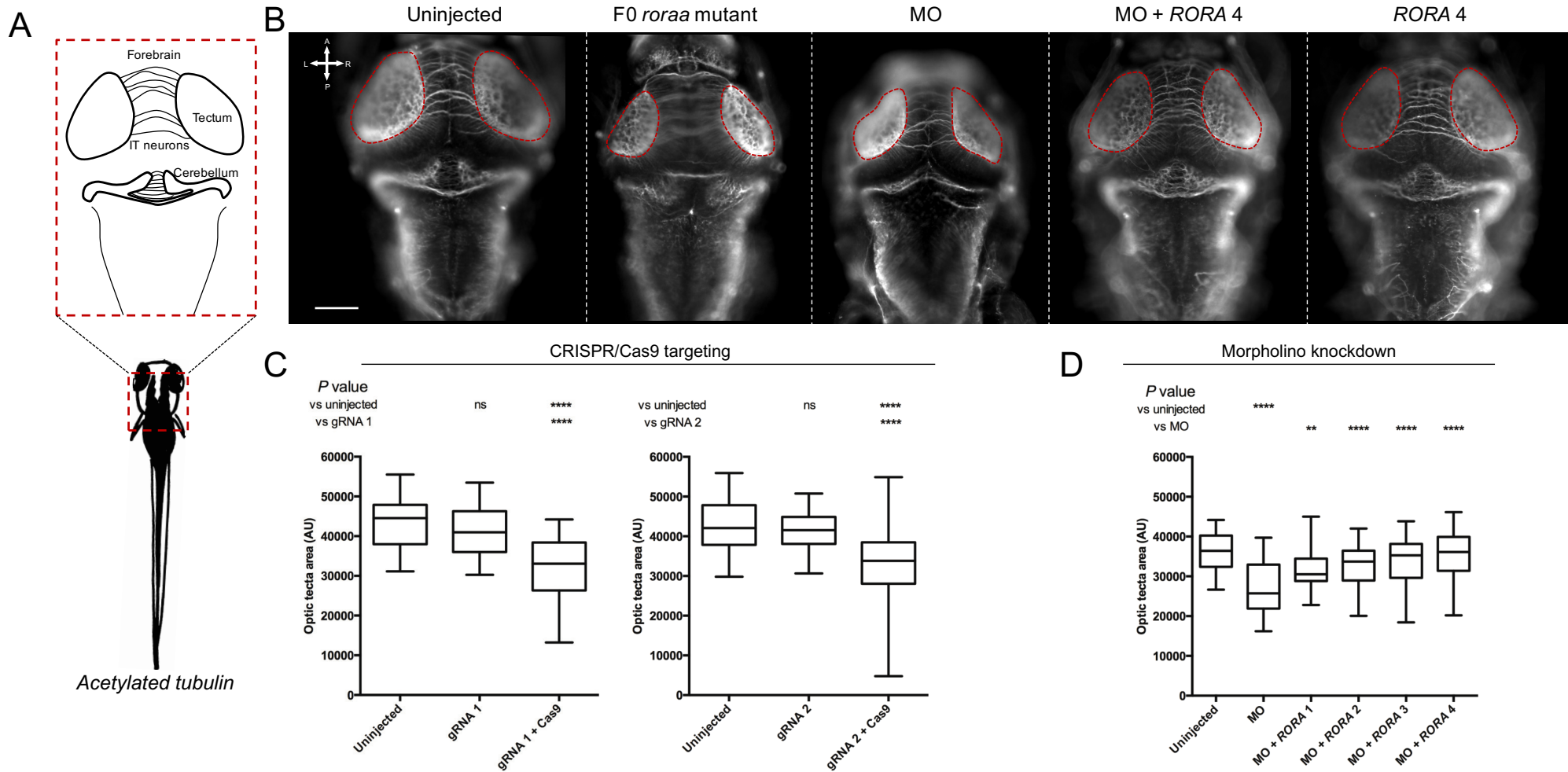


Figure S8

Figure S8. Disruption of *rora* in zebrafish larvae results in a reduction of optic tecta size. (A) Schematic of neuroanatomical structures painted with anti-acetylated tubulin antibody at 3 days post fertilization (dpf); IT, intertectal. (B) Representative dorsal images of acetylated tubulin immunostained larvae show that *rora* ablation causes neuroanatomical defects in CRISPR/Cas9 F0 mutants and morphants. Optic tecta size was measured as indicated by the dashed red outline on inset panels. Scale bar: 100 μ m. (C) Quantification of tecta area in larval batches is shown for two guide (g)RNAs targeting either *rora* exon 5 (gRNA 1) or exon 8 (gRNA 2) (D) *rora* morphants (injected with 3 ng morpholino; MO) display a neuroanatomical phenotype that can be rescued by four different wild type *RORA* mRNA transcripts: co-injection of *rora* e2i2 splice-blocking MO with *RORA* splice variants (*RORA* 1: NM_134261, *RORA* 2: NM_134260, *RORA* 3: NM_002943 and *RORA* 4: NM_134262). AU, arbitrary units. Stars indicate *P* value compared to uninjected controls (CRISPR/Cas9 and MO) or to morphants (MO + *RORA*). ****, $p < 0.0001$; **, $p < 0.01$; ns, not significant. Error bars in (C) and (D) represent 5th and 95th Percentile.

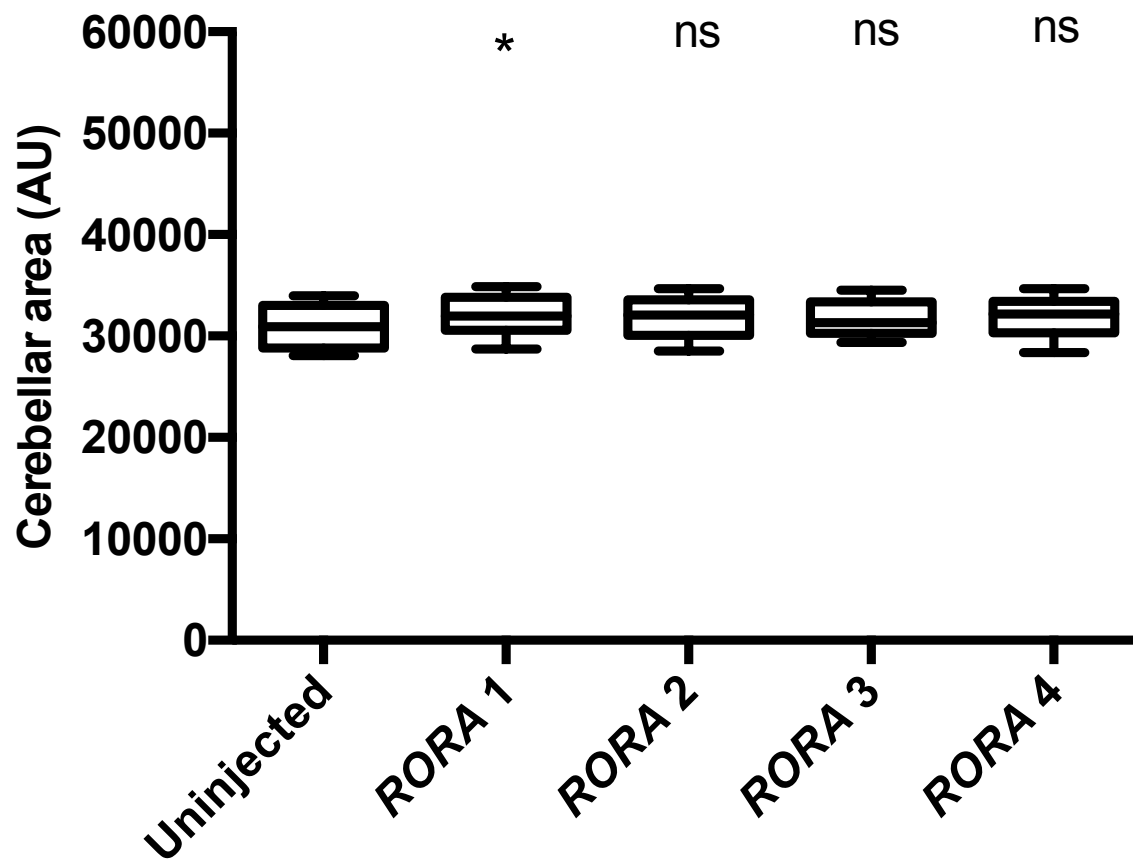


Figure S9

Figure S9. Ectopic expression of human WT *RORA* in zebrafish larvae. *In vivo* ectopic expression of *RORA* splice variants does not induce cerebellar abnormalities. Zebrafish embryos were injected with 200 pg of wild-type *RORA* mRNA corresponding to each of the four protein coding isoforms, stained with acetylated tubulin antibody, imaged and measured (see Figure 3). AU, arbitrary units. *P* values are not significant (ns) except for *RORA* 1: $P=0.03$ (bilateral t-test). Error bars represent 5th and 95th Percentile.

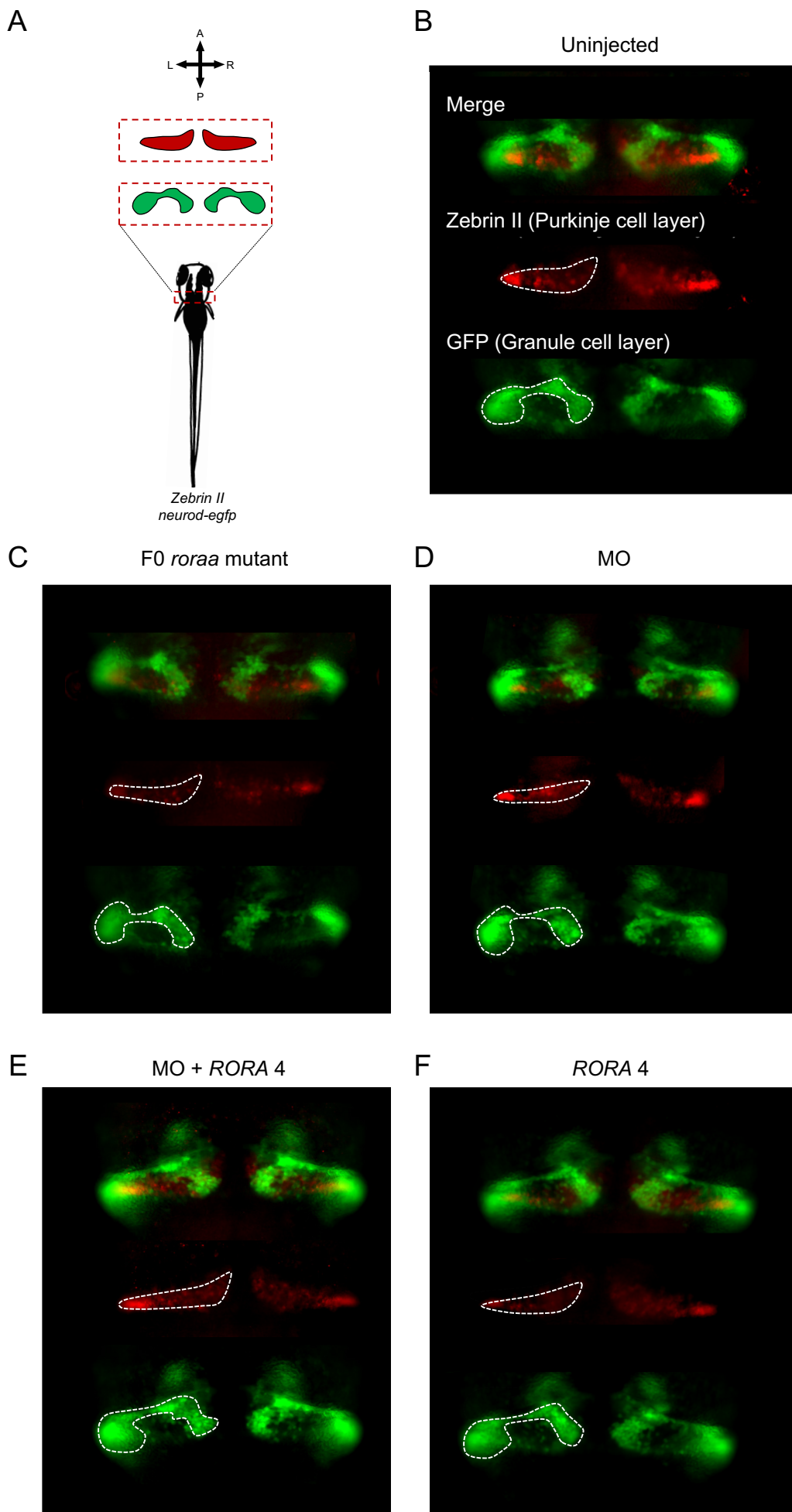


Figure S10

Figure S10. Disruption of *rora* in zebrafish larvae results in cerebellar hypoplasia driven by Purkinje and granule cell loss. (A) Schematic of cerebellar cell types assessed in 3 dpf larvae using either a *neurod:egfp* transgene (green) or anti-zebrin II immunostaining (red). Orientation is indicated with A, anterior; P, posterior; L, left; R, right. (B-F) Representative dorsal images show that reduction of Purkinje and granule cells contributes to cerebellar defects induced by *rora* targeting. Transgenic *neurod:egfp* larvae were fixed, immunostained with anti-zebrin II antibody (red), and the area comprised of each cell type was measured (as indicated in the schematic; see Figure 3E-H). Dashed white lines indicate measured area in (B) uninjected larvae; (C) larvae injected with CRISPR/Cas9 cocktail; (D) e2i2 splice-blocking morpholino (MO); (E) MO co-injected with wild-type *RORA* 4 mRNA; or (F) wild-type *RORA* mRNA alone.