

Supplemental information

**Haploinsufficiency of the Sin3/HDAC corepressor
complex member *SIN3B* causes a syndromic
intellectual disability/autism spectrum disorder**

Xenia Latypova, Marie Vincent, Alice Mollé, Oluwadamilare A. Adebambo, Cynthia Fourgeux, Tahir N. Khan, Alfonso Caro, Monica Rosello, Carmen Orellana, Dmitriy Niyazov, Damien Lederer, Marie Deprez, Yline Capri, Peter Kannu, Anne Claude Tabet, Jonathan Levy, Emmelien Aten, Nicolette den Hollander, Miranda Splitt, Jagdeep Walia, Ladonna L. Immken, Pawel Stankiewicz, Kirsty McWalter, Sharon Suchy, Raymond J. Louie, Shannon Bell, Roger E. Stevenson, Justine Rousseau, Catherine Willem, Christelle Retiere, Xiang-Jiao Yang, Philippe M. Campeau, Francisco Martinez, Jill A. Rosenfeld, Cédric Le Caignec, Sébastien Küry, Sandra Mercier, Kamran Moradkhani, Solène Conrad, Thomas Besnard, Benjamin Cogné, Nicholas Katsanis, Stéphane Bézieau, Jeremie Poschmann, Erica E. Davis, and Bertrand Isidor

SUPPLEMENTARY DATA

Supplemental Case Reports:

Individual 1 was 20 years old at the last clinical assessment, and he presented with autism spectrum disorder (ASD). He was born to non-consanguineous parents at 39 weeks gestational age with normal birth parameters. His development was borderline abnormal; he walked at 18 months but had no speech delay. The diagnosis of ASD with attention deficit hyperactivity disorder (ADHD) was made at the age of 18 years by a pediatric psychiatrist. Intelligence quotient (IQ) testing using the Wechsler Adult Intelligence Scale (WAIS) showed heterogeneous capacities with a global IQ of 96. He had a difficult school career but attended normal school. He has received methylphenidate treatment since the age of 9 years. He tires easily, has a general slowness and has frequent bouts of hypersomnia. His social relationships are very poor. A psychomotor evaluation at 17 years concluded global motor difficulties, dyspraxia, attention deficiency, temporospatial disorientation and relational difficulties. He also has migraines, myopia, and growth retardation (weight: 39 kg [-4SD]; height: 163 cm [-2SD]; occipital frontal circumference [OFC] 57 cm [0SD]). Dysmorphic features were also noted, including synophrys, micrognathia, small and low-set ears, and brachydactyly. Otherwise he had no epilepsy, and cerebral magnetic resonance imaging (MRI) was normal. Chromosomal microarray identified a *de novo* microdeletion of 1.02 Mb in 19p13.11 (GRCh37/hg19; chr19:16848440-17871985), encompassing 29 protein coding genes, including *SIN3B*.

Individual 2 was last assessed clinically at 11 years of age. She was born at term after an uneventful pregnancy. Her birth weight was 2140 g (<3rd percentile), and birth length was 46 cm (<3rd percentile). She had a mild ventricular septal defect (VSD), but did not require surgery. She had recurrent bronchitis until 3 years of age. She sat alone at 1 year old, walked at 32 months old and spoke her first words at 3 years old. She had temper tantrums as a toddler, but her behavior normalized once she started speaking. Nevertheless, she was treated with low doses of methylphenidate for hyperactive behavior. She never experienced epilepsy. She had strabismus (esotropia) and amblyopia with a mild hypermetropia. Fundus examination was normal. Auditory evoked potentials, electroencephalograms (EEG) and brain MRI are normal. Her familial history was also normal. Clinical examination showed a low frontal hairline, bilateral epicanthus, thick eyebrows, synophrys, downturned corners of the mouth, bifid uvula, fifth finger clinodactyly, brachydactyly and genu recurvatum. Growth measurements were normal (weight: +1SD, height: -0.5SD, OFC: -0.5SD). At 49 months of age, her development was calculated to be equivalent to a healthy child at 24 months (Brunet

Lezine scale). Her IQ, as assessed by the Wechsler Intelligence Scale for Children (WISC) is 41. By chromosomal microarray, we detected a ~1.52 Mb *de novo* CNV deletion encompassing *SIN3B* (GRCh37/hg19; chr19:15978604-17500427).

Individual 3 is a male last assessed at 2 years and 6 months of age. He is the third and last child of non-consanguineous healthy parents originating from Algeria. Left cleft lip and palate were discovered on a fetal ultrasound during the second trimester of pregnancy, and fetal brain MRI was unable to visualize olfactory bulbs. The family declined invasive clinical investigations. He was born at 37 weeks, with normal growth parameters (birth weight: 2850 g [22nd percentile]; birth length: 48 cm [28th percentile]; birth OFC: 32.5 cm [10th centile]). Left cleft lip/palate and olfactory bulb agenesis were confirmed at birth. Postnatal examination revealed micropenis and right iris coloboma with anterior segment dysgenesis. Echocardiography showed membranous VSD, and cerebral MRI a showed a hypoplastic corpus callosum in addition to olfactory bulb agenesis. Endocrine investigations were normal for thyroid, corticotropic and gonadotropic function. He had global developmental delay, sat at 17 months, walked at 28 months, and spoke his first words at 24 months. Postnatal microcephaly was observed with OFC at -3SD at 29 months old; weight and height were within normal average range for age. Targeted panel sequencing of genes involved in hypogonadotropic hypogonadism was negative, including analysis of the *KALI* and *CHD7* genes. Microarray revealed two *de novo* microdeletions: the first located in 19p13.11 (GRCh37/hg19; chr19:16599950-17469382) including *SIN3B* (~869 kb), and the other located in 18p11.31 (GRCh37/hg19; chr18:3192682-4854252) with a size of ~1.6 Mb.

Individual 4 was 8 years old at her last clinical assessment. The pregnancy was uneventful and resulted in a standard delivery without neonatal complications. She was followed in genetics clinic for a spectrum of phenotypes including neurodevelopmental disorder, bifid uvula, hypotonia, strabismus, and umbilical hernia; she underwent surgery for the latter two phenotypes. Early development of motor milestones was within normal limits. She was walking by 13 months and speaking by 2 years. Learning problems were noticed in kindergarten. Neurobehavioral examination performed at the age of 8 years and 10 months described hyperactivity and impulsivity, low frustration tolerance, anxiety and poor coordination. She also displayed hypersalivation as well as difficulties with articulation and pronunciation. Particular problems were noticed for tasks that demand visual-spatial processing. These difficulties require a significant one-on-one support during training. Her brain MRI showed mild tonsillar ectopia. Microarray

testing revealed a *de novo* ~877 kb deletion at the 19p13.11 locus encompassing *SIN3B* (GRC37/hg19; chr19:16456955-17333482).

Individual 5 was 3 years 10 months old at his last clinical examination. He is twin 2 of a dichorionic diamniotic pregnancy born after induction at 37 weeks with breech presentation. He weighed 2190 g (2-9th centile). He did not require resuscitation and was discharged on the day of delivery. He was treated for bronchiolitis at 4 weeks of age and was noted to have a heart murmur. He was diagnosed with tetralogy of Fallot, which was repaired electively at 10 months of age. Chromosomal microarray showed a *de novo* ~427 kb deletion at 19p13.11 (GRC37/hg19; chr19:16652215-17079033). His development was mildly delayed, crawling at 15 months and walking independently at 24 months. At 4 years 10 months there are no developmental concerns and he is making good progress in mainstream school. He has been discharged from Speech and Language Therapy. He has persistent drooling which is controlled by Sialanar. His twin brother had normal developmental milestones.

Individual 6 was last assessed at 15 years of age. He was born at 35 weeks of gestation by vaginal delivery, from non-consanguineous healthy parents; he had a weight of 2400 g (50th percentile), length of 47 cm (50-75th percentile) and OFC of 33 cm (75th percentile). The parental ages at the time of his birth were 32 and 39 for the mother and father, respectively. On the paternal side of the family, there is a history of intellectual disability, autism and schizophrenia. His brother has learning disabilities and attention deficit disorder. At the time of the study, individual 6 was 10 years old with a weight of 43 kg (97th percentile), height of 145 cm (90-97th percentile) and OFC of 55.5 cm (90-97th percentile). He has mild dysmorphic facial features with a broad nasal root and a prominent forehead. The affected individual was hypotonic from the neonatal period and had motor delay, eventually achieving independent ambulation at 24 months. He has mild intellectual disability and Asperger syndrome with speech delay (speech beginning at 24 months), lack of gestures, stereotypies (flapping, swinging, hitting with objects, repeat words or sounds), gaze aversion, does not partake in group activities, lacks empathy and has restricted interests. Other neurological signs included conductive hearing loss and aggressive behavior. Brain MRI images showed defects in the corpus callosum and sub-ependymal nodular heterotopy. Targeted gene sequencing detected a *de novo* c.31delA, p.Ser11Alafs*11 *SIN3B* variant.

Individual 7 was adopted in infancy and no details of the pregnancy, birth, and early development are known. He was globally delayed in the preschool years and began having major motor seizures at age 3 years. He had severe speech impairment when he entered public school at age 6 years. After 4 years, he attended a multiple disability school where he remained until age 21 years. Evaluations there documented serious language impairment with echolalia and halting speech, although he could construct a full sentence. He had a shuffling gait. The IQ was 56. For the next 30 years, he lived in different community homes and at age 50 years he moved to a residential facility because of a general decline in health and the diagnosis of Parkinson disease. His seizures continued to require anticonvulsant medication. Examination at age 50 showed head circumference at the 56th centile, height below the 3rd centile and weight below the 3rd centile. He had frontal balding, large ears with attached earlobes, full and arched eyebrows and a hand tremor. Whole exome sequencing analysis identified a heterozygous single base pair deletion in *SIN3B* which was confirmed by Sanger sequencing (NM_015260.3:c.1579del (p.Arg527Glyfs*12)).

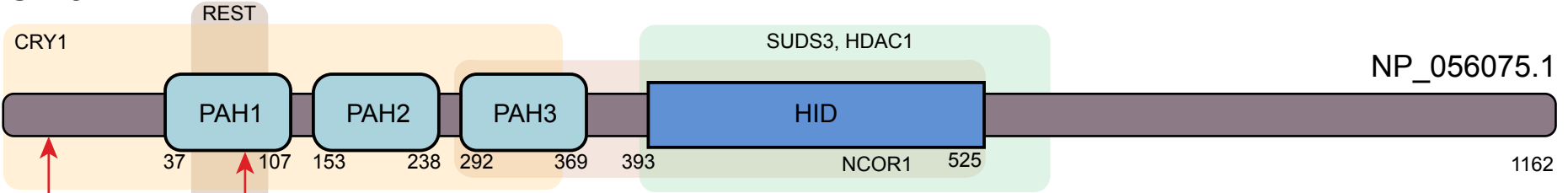
Individual 8 was 3 years old at the most recent exam. He is a male of northern European descent and was referred for genetics consultation for a suspected genetic etiology of his developmental delay, short stature and neurologic problems. He was born to a 34-year-old G2P2 mother and 36-year-old father, and the mother had gestational diabetes, which was well controlled. At 35 weeks, delivery was induced due to intrauterine growth restriction (IUGR) and oligohydramnios. His birth measurements were 1673 g (1st centile) and 41.3 cm (1st centile). He spent some time in the neonatal care unit when intestinal malrotation was discovered and was subsequently repaired surgically. His development has always been globally delayed. He walked at 22 months of age and says several words. At this age he was estimated to function at the level of an ~18-month-old according to physical therapists, occupational therapists, and speech therapists. He had hypotonia, difficulty chewing, sensory problems and oral aversion. He displayed no regression or waning energy. He had normal growth parameters: Weight: 15.7 kg (76th percentile), height: 94.5 cm (40th percentile), and OFC 49 cm (40th percentile). His body mass index (BMI) was consistent with the 90th percentile. He presents with plagiocephaly, prominent coronal suture, arched and full eyebrows, and small palpebral fissures. He had no signs of ASD and showed normal social and interactive abilities. Chromosomal microarray and baseline metabolic testing yielded normal results. No brain MRI was ordered.

His mtDNA testing showed a homoplasmic variant (m.3236A>G); this change was also homoplasmic in his mother who has fatigue, irritable bowel syndrome (IBS), gastroesophageal reflux disease (GERD) and migraines but no

history of delays; and in his brother who is 4 years old and has no developmental delay for his age. Consanguinity was denied. Exome sequencing identified a rare variant in *POLG* (MIM: 174763) and a rare variant in *TWNK/C10orf2* (MIM: 606075), however, these variants were considered unlikely causes of his neurodevelopmental phenotype because they were also present in his mother and his brother. He had a *de novo* missense variant in *SIN3B* (c.249C>G; p.Ile83Met).

Individual 9 was last examined at 5 years of age. He was born to a G3, P1 mother at 41 weeks gestation. His birth weight was 2920 g. His parents are non-consanguineous and of Jamaican background. Abnormalities were noted in his forearms since early in his life, and X-rays subsequently revealed bilateral proximal radio-ulnar synostosis. He started walking at eighteen months of age and first talked at 2 years of age. He has just completed kindergarten and appears to learn appropriately. However, he has poor eye contact and difficulties with self-regulation. He is noise sensitive and has frequent behavioral outbursts. He rarely initiates collaborative play and prefers to play by himself. He was diagnosed with ASD at severity level 1 and ADHD. He remains on no prescribed medications. On examination at age 5 years, his growth parameters are: weight: 26.9 kg, height: 120 cm and OFC: 55 cm (all >97th centile). Inspection of his facies reveals a broad forehead and high anterior hairline. He has an area of acanthosis nigricans over his back and one small hypopigmented mark. He has limited pronation and supination bilaterally. The following clinical investigations were reported as normal: SNP chromosomal array and fragile X testing, urine organic acids. Exome sequencing subsequently identified rare heterozygous variants in *CNOT1* (OMIM: 604917; NM_016284.5: c.4661A>G, p.Ile1621Val) and *SIN3B* (c.58G>A, p.Gly20Arg). Neither variant is present in the mother, but the father is unavailable for testing. The missense variant in *CNOT1* is rare (absent from gnomAD) and has been described previously in an individual affected with sporadic amyotrophic lateral sclerosis (ALS; MIM: 105400), who also had an alternate molecular basis for the ALS.

SIN3B



p.Gly20Arg
Individual 9

p.Ile83Met
Individual 8

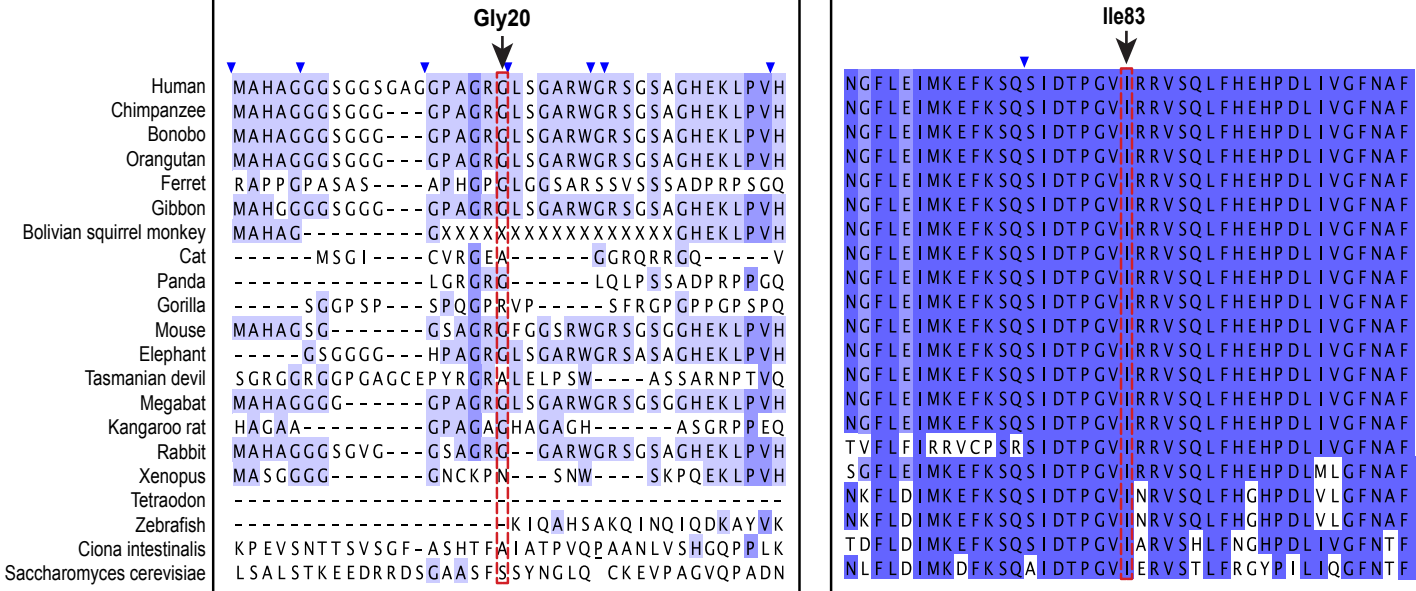


Figure S1

A

sin3b (*Danio rerio*)ENSDARG00000062472
Chromosome 8: 14,174,041-14,201,318
reverse strand

exon 13

gRNA

PAM

ACATGATCTTTAACTA**TGAGGACA**AAACAGATCCTGG**AGG**ACGCCGCTTCCCT
 TGTACTAGAAATTGAT**ACTCCTGTTTGTCTAGGACC**TCC**T**GCGGCGAAGGGA

B

UI

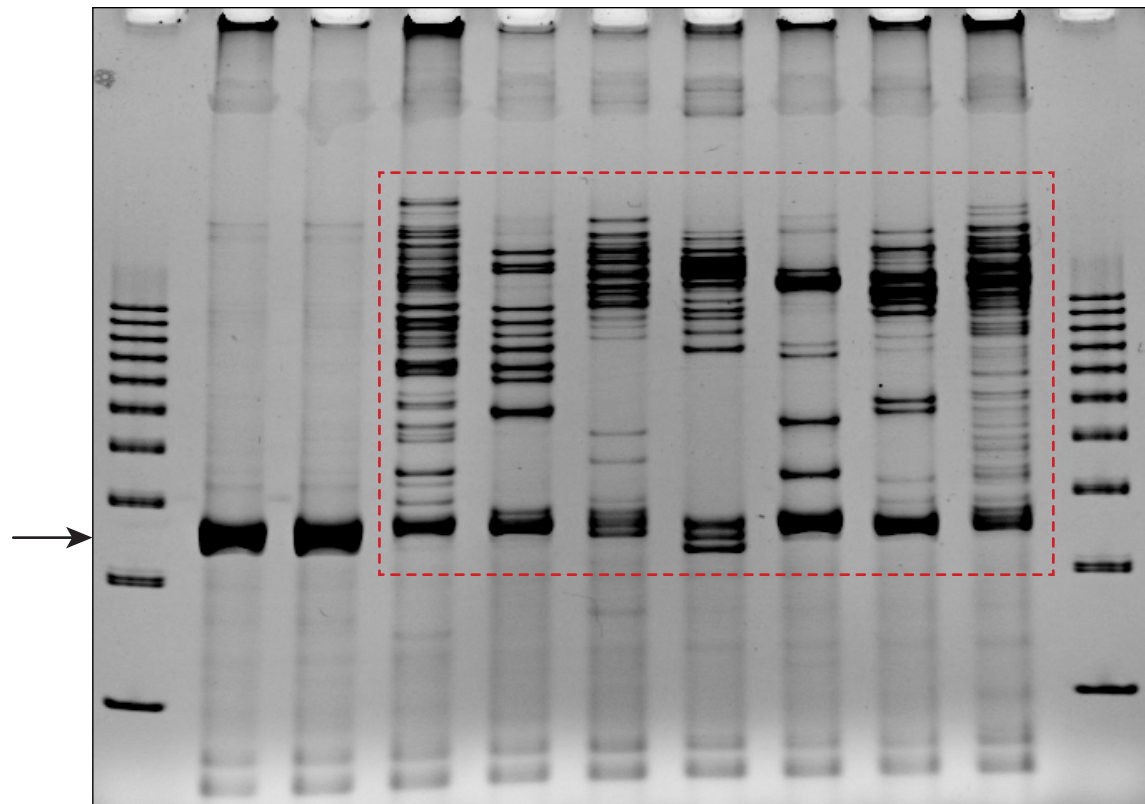
sin3b F0 mosaic CRISPR mutants

Figure S2

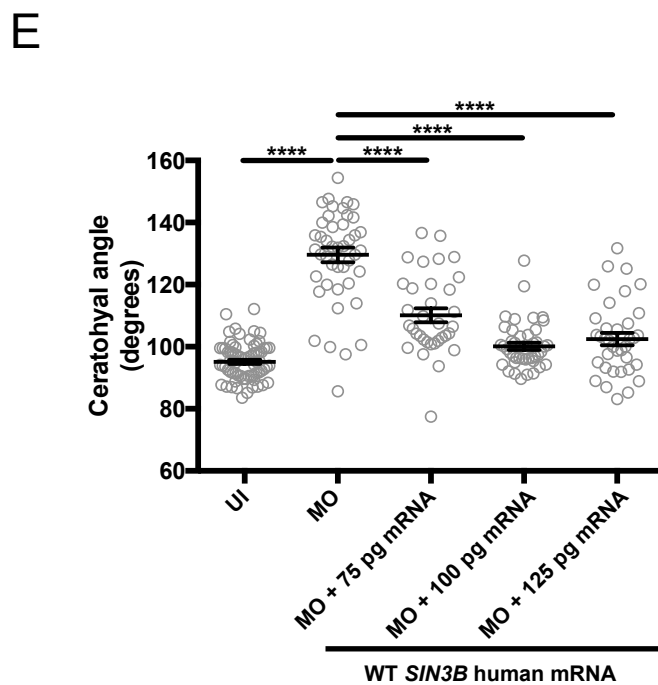
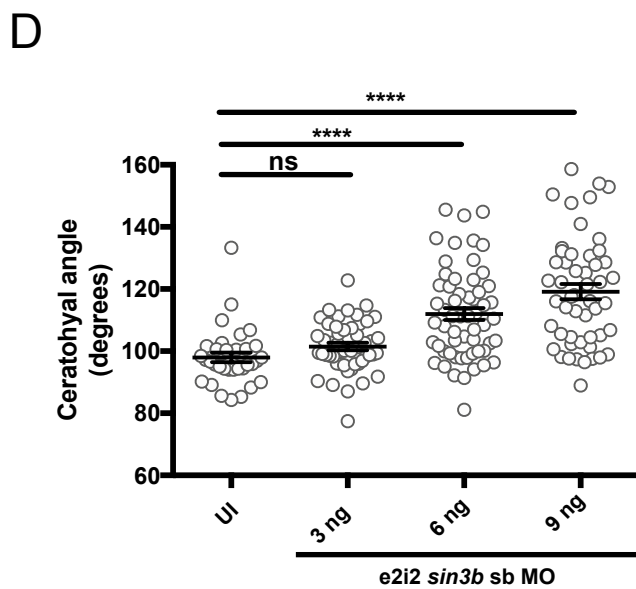
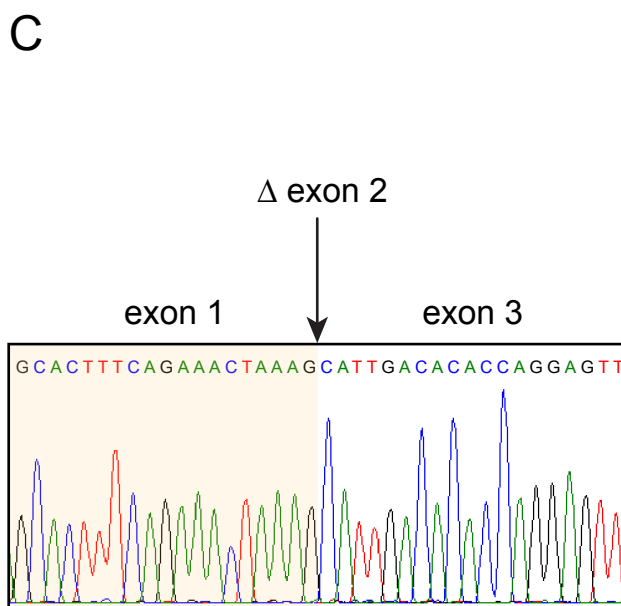
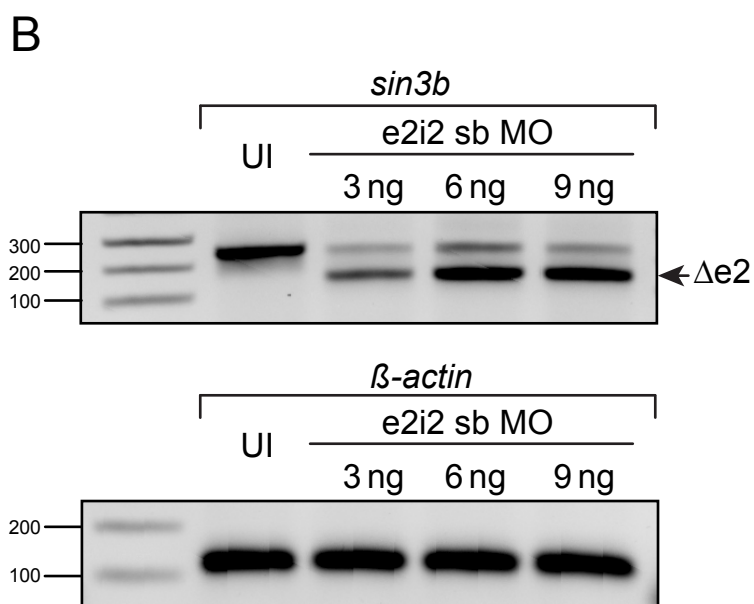
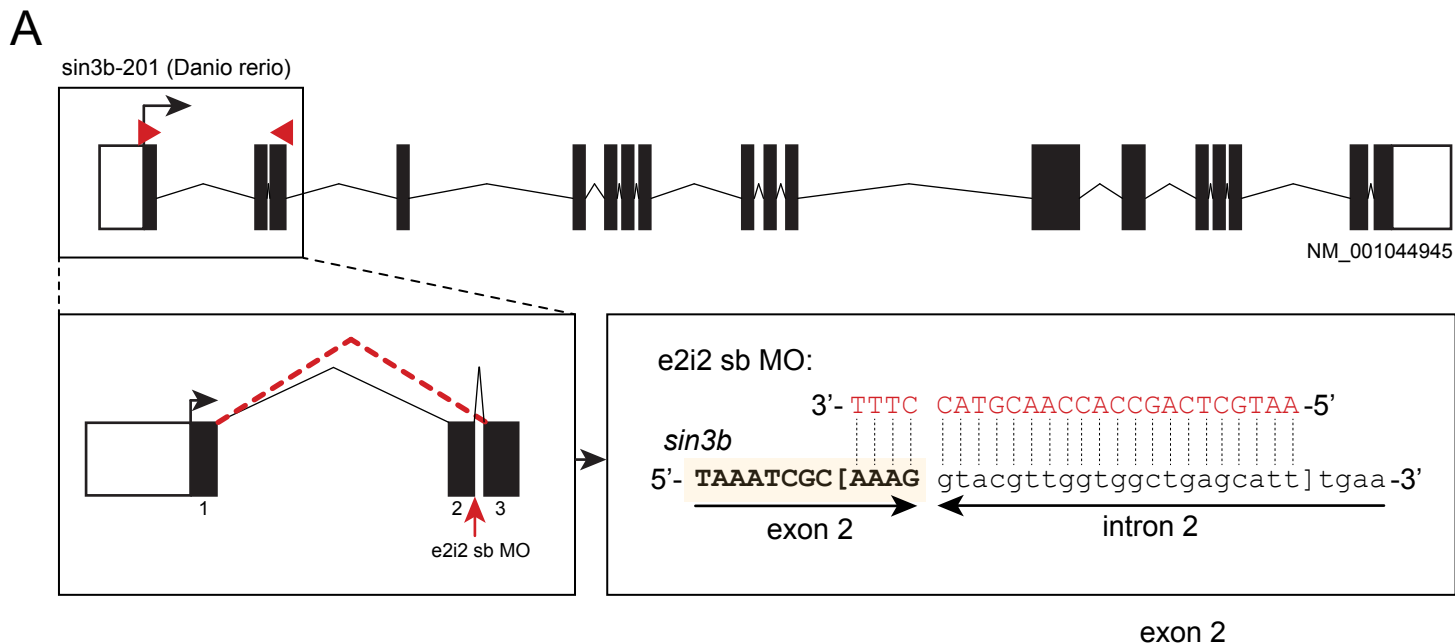
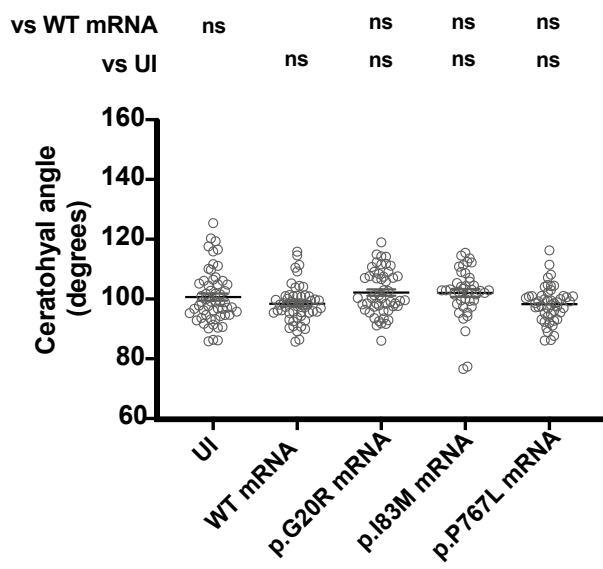
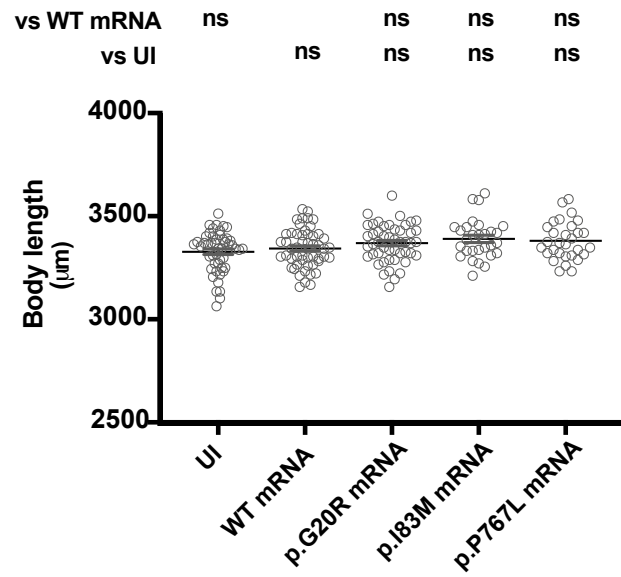


Figure S3

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

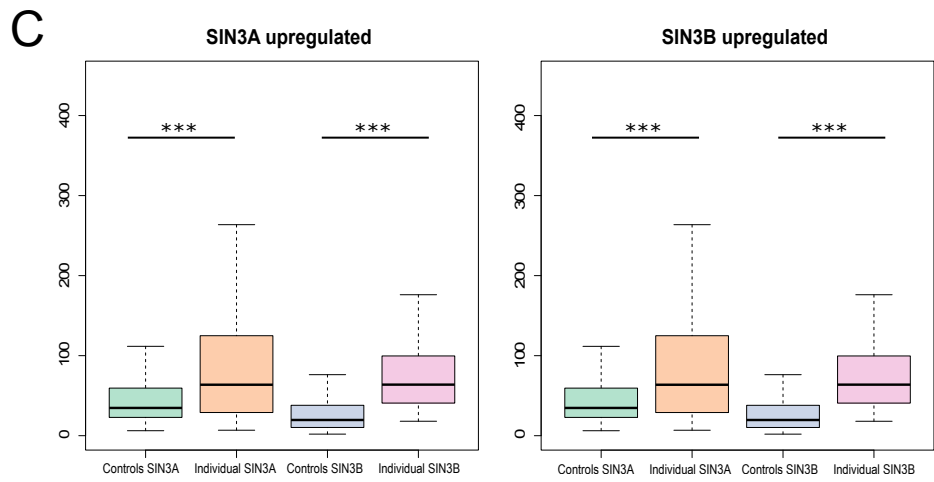
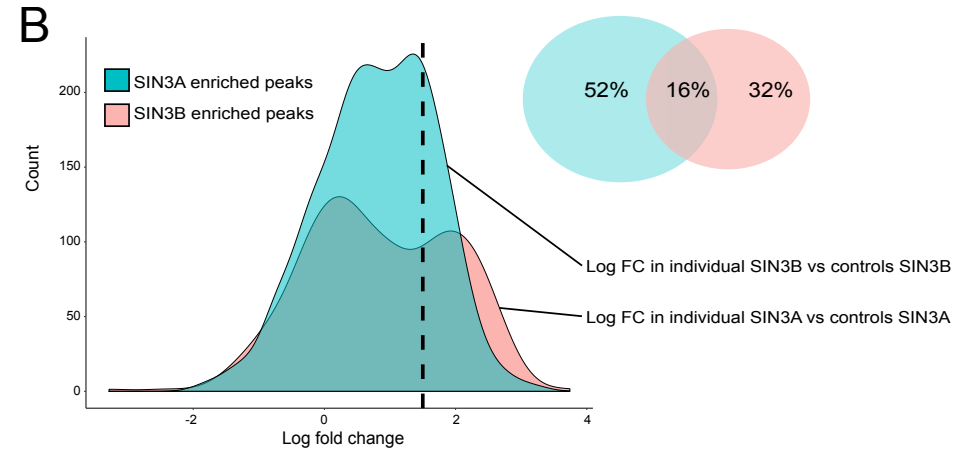
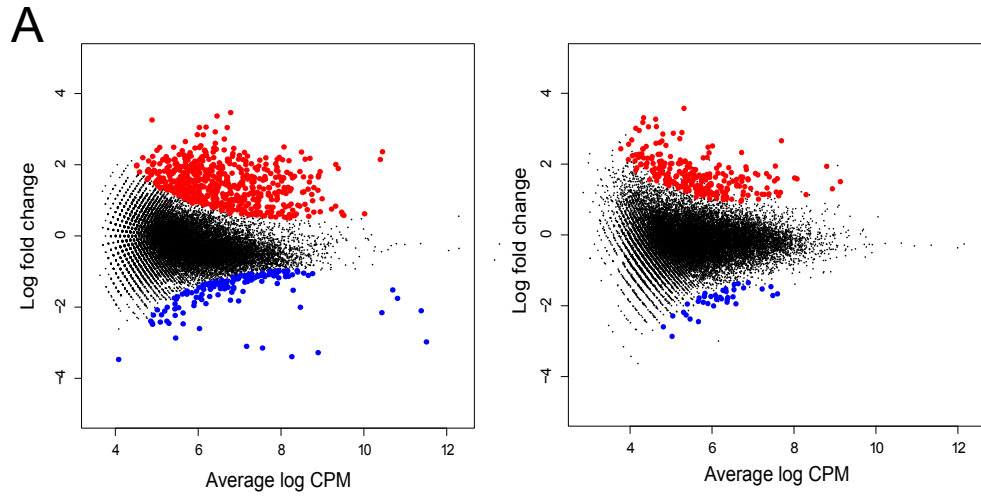


Figure S5

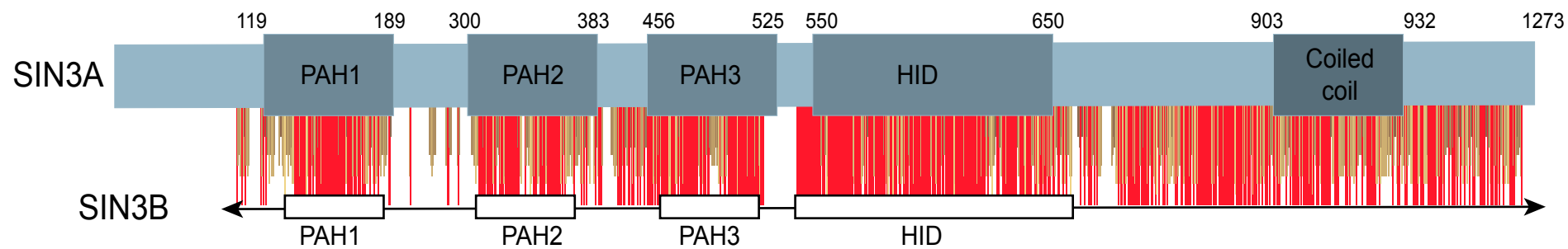


Figure S6

Supplementary Figure Legends:

Figure S1. Multiple sequence alignment of SIN3B orthologs from various eukaryotic species sorted by pairwise identity demonstrates conservation of residues affected by missense variants. Top, schematic of human SIN3B protein (NP_056075.1; see Figure 2A), with the location of SIN3B missense mutations indicated with vertical red arrows. Bottom, multiple sequence alignment (Clustal Omega algorithm) of 40-amino acid blocks encompassing residues Gly20, and Ile83; variants are indicated by black vertical arrows and dashed red boxes.

Figure S2. Efficient CRISPR/Cas9 targeting of the *sin3b* genomic locus in zebrafish F0 mosaic mutants. (A) Exon-intron structure of the zebrafish *sin3b* locus (genome assembly GRCz10; encodes GenBank ID: NP_001038410.1, 63% identical to human *SIN3B*). Coding regions, black boxes; untranslated regions, white boxes; lines, introns. Guide (g)RNA target sites are indicated by vertical arrow on exon 13. gRNA sequence and protospacer adjacent motifs (PAM) are shown. (B) Heteroduplex analysis indicates efficient targeting of the *sin3b* locus by the gRNA. gRNA/Cas9 cocktail was injected into the cell of one-cell-stage embryos, genomic DNA was extracted at 2 days post-fertilization, PCR flanking target sites was performed. Amplicons were denatured, reannealed slowly and migrated on a polyacrylamide gel (n=2 uninjected [UI] controls and 7 F0 mutants, respectively). Heteroduplexes are indicated by red box; black arrow indicates amplification fragment at expected size for wild-type PCR product. We sequenced 24 colonies per embryo from 3 F0 mutants to confirm >90% mosaicism.

Figure S3. Knockdown efficiency of *sin3b* morpholino antisense oligonucleotide. (A) Exon-intron structure of the zebrafish *sin3b* locus (genome assembly GRCz10). Coding regions, black boxes; untranslated regions, white boxes; lines, introns; red arrows, primers used to assess exon 2 splicing in panel B. Below, left, a zoomed view of the targeted exon donor site shows normal and aberrant (red dashed line) splicing of exon 2. Right, base complementarity of e2i2 splice-blocking (sb) morpholino (MO) on the *sin3b* exon 2 donor splice site is indicated. Yellow box indicates exon 2. Sequence within brackets indicates targeted region. (B) e2i2 sb MO leads to exclusion of *sin3b* exon 2 (Δ e2). Agarose gel image shows RT-PCR results obtained by primers shown in panel A (red arrows). *b-actin* was amplified to control for mRNA integrity. UI, uninjected control. (C) Chromatograms of TOPO-cloned PCR product amplifying MO-targeted region. Aberrant splicing results in exclusion of exon 2 leading to a frameshifting event and premature stop codon (p.Pro103*). (D) *sin3b* morphants injected with increasing doses of e2i2 splice-blocking (sb) MO (from 3

to 9 ng) display a broadened CH angle at 3 days post-fertilization compared to controls. (E) Co-injection of *sin3b* e2i2 splice-blocking MO with three increasing doses of *SIN3B* WT human mRNA (from 75 to 125 pg) rescues this phenotype significantly when co-injected with 9 ng of e2i2 sb MO (n=34-72 larvae/batch, repeated). Statistical comparisons were performed with a non-parametric Kruskal-Wallis test. ****, $p < 0.0001$; ns, not significant. Error bars represent standard error of the mean.

Figure S4. Heterologous expression of *SIN3B* human mRNA (wild-type or variant) does not induce craniofacial defects or reduced body length in zebrafish larvae. We injected 100 pg *in vitro* transcribed wild-type (WT) or variant *SIN3B* mRNA (NM_015260.4) in *-1.4coll1a1:egfp* larvae and assessed ceratohyal angle or body length at 3 days post-fertilization. Variants identified in affected individuals encode p.Gly20Arg, p.Ile83Met; p.Pro767Leu is a negative control variant (rs117307745; 6 homozygotes in gnomAD) Transgenic signal was imaged live with the VAST BioImager™, and anatomical measurements were obtained using ImageJ, NIH (n=38-41 larvae/batch; See Figure 4). P-value is not significant (ns, non-parametric Kruskal-Wallis test) between mRNA-injected larvae and controls.

Figure S5. *SIN3A* and *SIN3B* share similar acetylation profiles. (A) MA plot showing differential acetylation peaks between the affected individual with the *de novo SIN3A* deletion and healthy members of his own family (left); and individual 1 with the *de novo SIN3B* deletion and healthy members of his own family (right). In PBMCs derived from both cases there is an overall hyperacetylation. (B) Distributions of log FC (Fold change) between affected individuals and healthy controls. For the blue curve, the peak set consists of *SIN3B* enriched peaks (log FC > 1.5) and the log FC of *SIN3A* vs his family is shown. For the pink curve, the logFC of the *SIN3B* is shown for *SIN3A* enriched peaks (log FC > 1.5). Venn diagram showing that 16% of peaks have a logFC > 1.5 in *SIN3A* individual vs his family and in *SIN3B* individual vs his family. (C) Box plots showing the peak heights in affected individuals and controls for either *SIN3A* up-regulated or *SIN3B* up-regulated peaks. y-axis, acetylation peak heights; ***, $p < 0.001$ (Wilcoxon rank-sum test).

Figure S6. *SIN3A* protein structure and amino acid conservation patterns with *SIN3B*. *SIN3A* protein (UniProtKB Q96ST3) contains three N-terminal PAH domains (brown), an HDAC1 interacting domain (HID) and a central coiled-coil domain. Vertical bars below the schematic indicate conservation with *SIN3B* (UniProtKB O75182,

Multiple-sequence alignment with Clustal Omega, conservation scores obtained with Jalview program). Red bars indicate identical amino acids between SIN3A and SIN3B.

Table S1. Exome sequencing or chromosomal microarray methodology for individuals 1-9

Individual ID	1	2	3	4	5	6	7	8	9
Contributing Research Center	Nantes France	Gosselies Belgium	Paris France	Austin/Houston USA	Newcastle UK	Valencia Spain	Greenwood USA	New Orleans USA	Toronto Canada
Variant	1.02 Mb (19:1684844-17871985)	1.5Mb (19:1597860-17500427)	869 kb (16599950-17469382)	877 kb (16456955-17333482)	427 kb (16652215-17079033)	c.31_31delA p.Ser11Alafs*11	c.1579delC p.Arg527Glyfs*12	c.249C>G, p.Ile83Met	c.58G>A p.Gly20Arg
WES approach	NA	NA	NA	NA	NA	Trios-Custom Panel (1256 genes)	Next Gen - Confirmed by capillary sequencing	Next Gen - Confirmed by capillary sequencing	Next Gen - Confirmed by capillary sequencing
Capture reagent	NA	NA	NA	NA	NA	Agilent SureSelect	Agilent clinical research exome V2	Agilent clinical research exome kit	IDT xGen Exome Research Panel v1.0
Sequencer	NA	NA	NA	NA	NA	Illumina HiSeq 2000	Illumina NovaSeq	Illumina HiSeq 2000	Illumina HiSeq
Microarray Platform	Agilent 60K	Agilent 60K	Agilent 60K	Agilent 180K (BCM V8.1)	Whole genome 60k oligo array (ISCA version2)	NA	NA	NA	NA

Abbreviations are as follows: ND, no data available; NA, not applicable

Table S3: *In silico* predictions of pathogenicity for *SIN3B* missense variants

Individual ID		Individual 8	Individual 9
<i>SIN3B</i> variant	Nucleotide	c.249C>G	c.58G>A
	Protein	p.Ile83Met	p.Gly20Arg
	Inheritance	<i>de novo</i>	ND
CADD PHRED v1.4	Score	24.4	23.1
SIFT	Prediction	Affect protein function	Tolerated
	Score	0.00	0.33
	Threshold	damaging <0.05	damaging <0.05
Polyphen-2 v2.2.2r398	Prediction	Probably damaging	Probably damaging
	Score	1.000	0.867
	Threshold	damaging >0.85	damaging >0.85
Mutation Taster	Prediction	Disease causing	Disease causing
	Score	10	125
	Threshold	0 to 215	0 to 215
Conservation	PhyloP	2.991 [flanking]	0.463 [flanking]
		1.963	2.411
		0.464 [flanking]	2.411 [flanking]
	phastCons	1 [flanking]	1 [flanking]
		1	1
		0.998 [flanking]	1 [flanking]

Abbreviations are as follows: CADD, combined annotation dependent depletion; ND, no data available; NA, not applicable; SIFT, sorting intolerant from tolerant.

Table S4. Primers used for *in vivo* modeling studies

aim	name	sequence
CRISPR_synthesis	sin3b_G_FOR	TAATACGACTCACTATAGTGAGGACAAACAGATCCTGG
CRISPR_synthesis	sin3b_G_REV	TTCTAGCTCTAAAAC CCAGGATCTGTTTGTCTCA
CRISPR_PCR1	sin3b_G_PCR1_FOR	CTAGCGCTTCTTTTTCCCTCT
CRISPR_PCR1	sin3b_G_PCR1_REV	TCAGTCTCGCTAAGCTCTCCAC
morpholino_e2i2_efficiency_PCR1	sin3b_e2i2_e3i3_PCR1-FOR	CTCACAGCAGCACTGCGAAGC
morpholino_e2i2_efficiency_PCR1	sin3b_e2i2_e3i3_PCR1-REV	GCCTGTACTCCGTCCTGCTCC

Table S5. Quality controls performed on Chip-Seq data. Samples with a fraction of reads in peaks (FRiP) <2% and/or non-redundancy fraction (NRF) <0.8 did not proceed for further analysis. One of the replicates from the father and one from the mother of the SIN3A family have been removed due to their poor quality (orange highlight).

Sample	All reads	Mapped reads	Unique reads	Peaks	FRIP	NRF
Sin3A_bro1	5947559	4836777	5438449	20558	4.5	0.9
Sin3A_bro2	5265336	3723635	4446974	18832	5.9	0.8
Sin3A_dad1	10655146	9294979	10204799	12009	0.8	1.0
Sin3A_dad2	5535869	3562419	4390943	50825	4.3	0.8
Sin3A_mom1	6304322	1302353	1474644	32073	12.0	0.2
Sin3A_mom2	7915320	5362653	6529163	17393	5.1	0.8
Sin3A_ind1	12921598	11571635	12529389	12604	2.7	1.0
Sin3A_ind2	10555157	9419854	10194104	14434	2.4	1.0
Sin3b_dad1	4653974	3478512	3773724	36416	4.6	0.8
Sin3b_dad2	4901383	4026913	4402745	16266	7.1	0.9
Sin3b_mom1	7308758	6235833	6855607	24961	14.1	0.9
Sin3b_mom2	5361496	4803034	5208508	30458	23.5	1.0
Sin3b_ind1	5502985	4592394	5005944	17765	8.0	0.9
Sin3b_ind2	5298111	4474531	4886599	18015	8.9	0.9

Supplementary methods:

Genetic analysis

To investigate potential tertiary structure alterations caused by nonsynonymous changes, we used template-based modeling by predicting the three-dimensional structure of SIN3B via the RaptorX online tool using SIN3B protein sequence (NP_056075.1) as input.¹ We performed annotations and measurements of distances between side chains using Pymol v2.0.7, with mutagenesis and measurements wizards.

Zebrafish embryo husbandry

All zebrafish work was performed in accordance with protocols approved by the Duke University Institutional Animal Care and Use Committee. Zebrafish adults (*Danio rerio*, wild-type strain ZDR, Aquatica BioTech) were maintained under continuous water flow and automatic control of a 14h/10h light/dark cycle, at 28°C. Zebrafish embryos were obtained by natural matings and reared at 28°C in a media composed of 0.3 g/L NaCl, 75 mg/L CaSO₄, 37.5 mg/L NaHCO₃, and 0.003% methylene blue. Phenotyping was performed at 3 days post fertilization (dpf) with the investigator blinded to experimental conditions.

***sin3b* CRISPR/Cas9 genome editing**

We designed a guide (g) RNA targeting *sin3b* (GRCz10: ENSDARG00000062472), using CHOPCHOP v2.^{2,3} gRNA was transcribed *in vitro* with the GeneArt precision gRNA synthesis kit (primer sequences available in Table S5; Thermo Fisher). We targeted the *sin3b* locus by microinjection into the cell of 1-cell stage embryos with 1 nl of cocktail containing 100 pg gRNA and 200 pg Cas9. Individual embryos (n=8) were collected at 1 dpf for DNA extraction to assess targeting efficiency. We PCR-amplified the gRNA target with primers located in flanking regions (Table S5), denatured the resulting product, and reannealed it slowly to form heteroduplexes (95°C for 5 min, ramped down to 85°C at -1°C/s and then to 25°C at -0.1°C/s). We performed polyacrylamide gel electrophoresis (PAGE) on a 20% precast 1 mm gel (Thermo Fisher) to visualize heteroduplexes as an indication of insertions and/or deletion events.⁴ To estimate mosaicism of F0 mutants, PCR products were cloned into a TOPO-TA vector (Thermo Fisher) and individual colonies (n=24) were Sanger sequenced (n=3 larvae/gRNA).

Transient *sin3b* suppression and heterologous expression experiments

With support from GeneTools, we designed a splice blocking (sb) morpholino (MO) targeting the splice donor site of exon 2 (e2i2; 5'- AATGCTCAGCCACCAACGTACCTTT-3') of *sin3b* (GeneTools, LLC; Figure S4), and injected 1 nl of MO into the yolk of zebrafish embryos at one-to-four cell stage. At one dpf, we harvested uninjected control and MO-injected larvae in Trizol (Thermo Fisher), extracted total RNA by isopropanol-mediated precipitation, and conducted first-strand cDNA synthesis with the QuantiTect Reverse Transcription kit (Qiagen). The targeted region of *sin3b* was PCR-amplified using primers complementary to flanking exons (Table S4), and amplicons were migrated by electrophoresis on a 1 % agarose gel; bands were extracted, gel purified using QIAquick gel extraction kit (Qiagen) and resulting clones were Sanger sequenced. The optimal MO dose for *in vivo* complementation experiments was determined by injection of three concentrations of MO (3, 6, 9 ng of e2i2). Gateway-compatible *SIN3B* (NM_015260.4) open reading frame (ORF) clone was provided by Genecopoeia and we transferred this ORF to a pCS2+ vector by LR clonase II-mediated recombination (Thermo Fisher) and validated sequences by Sanger sequencing. We introduced point mutations encoding variants identified in affected individuals (p.Gly20Arg, p.Ile83Met) or a negative control variant (p.Pro767Leu; rs117307745; 6 homozygotes in gnomAD, accessed June 2019) using site-directed mutagenesis as described.⁵ Linearized pCS2+ vectors containing WT or mutant ORFs were transcribed *in vitro* with the mMessage mMachine SP6 Transcription kit (Ambion).

Live imaging of zebrafish larvae

We performed a morphometric assessment of the zebrafish larval craniofacial structures and body length at 3 dpf using the Vertebrate Automated Screening Technology (VAST BioImager™, Union Biometrica). Live larvae were anesthetized with tricaine and loaded in a microcapillary for automated imaging using default parameters as described.⁶⁻⁸ Ceratohyal angle and body length were measured on ventral fluorescent and lateral bright field images, respectively, with ImageJ (NIH). Statistical analyses were performed using an unpaired bilateral Student's t-test (GraphPad Prism software).

Acetylated tubulin immunostaining

We performed standard whole-mount immunostaining on 3 dpf zebrafish larvae. Animals were anesthetized with tricaine, and fixed in Dent's solution. Primary detection was carried out with anti- α -acetylated tubulin antibody (T7451, mouse Sigma-Aldrich, 1:1000 dilution) and secondary detection was facilitated with Alexa Fluor 594 goat

anti-mouse IgG (ThermoFisher, 1:500 dilution). Dorsal images of fluorescent signal were acquired with an AZ100 microscope (Nikon) equipped with a Nikon digital sight black and white camera and NIS Elements software (Nikon). Commissural neurons were quantified by counting the number of axon tracts that cross the dorsal midline between the optic tecta as described.⁹

PBMC isolation and freezing

Peripheral blood mononuclear cells (PBMCs) from affected individuals and their family members were isolated by Ficoll gradient on EDTA-anticoagulated blood. Briefly, 5 mL of blood was diluted 1:2 in PBS. 5 mL of lymphocytes separation medium (Eurobio, catalog number CMSMSL0101) was added in another tube. The diluted blood was carefully added to the hand-leaned tube containing Ficoll to avoid mixing with Ficoll solution. Centrifugation was performed at 2000 rpm for 20 min at room temperature with medium acceleration and no brake. PBMCs were found at the white ring formed between the red cells at the bottom and the plasma at the top of the tube. PBMCs were collected and washed in 50 mL of PBS centrifuged at 1500 rpm for 10 min at 4°C. PBMCs were then resuspended in 5 mL PBS. After counting, PBMCs were distributed by 2M cells per 250 µL in a solution containing 10% dimethyl sulfoxide (DMSO) and 20% fetal bovine serum (FBS) in PBS. Samples were stored at -80°C until use. In this experiment, 2M cells in replicate samples per condition were used.

Chromatin Immunoprecipitation (ChIP)- sequencing (seq) Experiments

We resuspended 2M PBMCs in 40 µL PBS. Cells were lysed for 20 min at 4°C in 40 µL 2X Lysis buffer (100 mM Tris-HCl pH 8, 300 mM NaCl, 2% Triton X-100, 0.2% sodium deoxycholate, 10mM CaCl₂) with protease inhibitor cocktail (catalog number P2714, Sigma Aldrich) and 5 mM sodium butyrate. Chromatin was fragmented with 300 units of Micrococcal nuclease (MNase; M0247S, New England Biolabs) per well for 10 minutes at 37°C. MNase digestion was stopped with 80 µL Lysis Dilution buffer (50 mM Tris-HCl pH 8, 150 mM NaCl, 1% Triton X-100, 50 mM EGTA, 1 mM EDTA, 0.1% sodium deoxycholate). After full speed centrifugation, supernatants were collected and filled up to 400 µL. 2.5% of each sample were pooled and kept as input. 2 µg of anti-H3K27ac (catalog number 39133, Active Motif) was added to each sample and rotated at 15 rpm at 4°C overnight. 25 µL G-protein dynabeads (Life Sciences) were added in each sample and submitted to rotation for 4 hours at 4°C. Beads were then washed twice with 200 µL of each of the following buffers (ice-cold): Wash buffer 1 (50 mM KOH HEPES pH 7.5, 150 mM NaCl,

2mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS), Wash buffer 2 (50 mM KOH HEPES pH 7.5, 300 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS), Wash buffer 3 (10 mM Tris-HCl pH 8, 250 mM LiCl, 1 mM EDTA, 0.5% NP40, 0.5% sodium deoxycholate). Final wash was performed in 200µL TE buffer (10 mM Tris-HCl pH 8, 1 mM EDTA). ChIP beads were eluted in 50 µL of ChIP elution buffer (50mM Tris-HCl pH7.5, 10mM EDTA, 1% SDS) with several cycles of 5 min incubation at 63°C / vortex for a total of 30 minutes prior to bead removal. ChIP and input samples were then digested with 250 µG/mL proteinase K (catalog number GEXPRK006R; Eurobio) in 50 µL TE buffer for 1 hour at 63°C. Samples were then filled up to 400 µL with TE buffer. DNA ChIP were then isolated by phenol chloroform isoamyllic acid (PCIAA) method and purified DNA was resuspended in 42 µL elution buffer (EB, 10mM Tris-HCl buffer pH8). Libraries were then prepared as described.^{10,11} Libraries were verified by Bioanalyzer (Caliper) for peaks sizes. Equimolar pools were performed prior QC and sequencing. NGS on Illumina high throughput NextSeq 500 (75 bp single-end) was performed.

ChIP-seq analysis

Single-end reads were mapped to the GRCh37 genome by the BWA algorithm¹² and duplicate reads (read-pairs mapping to the same genomic location) were collapsed. Reads mapping to non-canonical and mitochondrial chromosomes were also removed. For each sample, ChIP-seq peaks were detected using DFilter¹³ at a P-value threshold of 1×10^{-6} , and peaks were selected only if they are present in both duplicates. Poor quality samples (Fraction of reads in peaks [FRiP] < 2% and/or non-redundancy fraction [NRF] < 0.8) were discarded.¹³ A set of consensus peaks was then obtained by combining all the samples, thus obtaining a read count on these peaks using Bedtools.¹⁴ To perform Differential Peak Calling, differentially acetylated (DA) peaks were determined using edgeR^{15, 16} after a counts per million (cpm) normalization. For this analysis, only the peaks with at least 2 cpm in 8 of the 12 samples were retained and the DA peaks were defined with a Benjamini-Hochberg Q-value $\leq 5\%$. For data representation (PCA plot and heatmap) peaks were rlog transformed [Deseq2]. For boxplots, statistical analyses between affected individual and control acetylation rate either in the SIN3A up-regulated or SIN3B up-regulated set of peaks were performed on rlog normalized data with a Wilcoxon rank-sum test.

To determine gene ontology enrichment in up-regulated peaks, we used the GREAT tool by associating genomic regions with genes in a basal plus extension way according to GREAT parameters.¹⁷ Test regions, corresponding to the set of up-regulated peaks between the SIN3A individual + SIN3B individual vs controls, were compared to a

background set of peaks consisting of consensus peaks on which the differential analysis was performed. A minimum of 5 affected genes and a threshold of statistical significance of $FDR < 0.05$ were used, all non-redundant ontologies were selected.

Supplemental references

1. Källberg, M. *et al.* Template-based protein structure modeling using the RaptorX web server. *Nat. Protoc.* **7**, 1511–1522 (2012).
2. Montague, T. G., Cruz, J. M., Gagnon, J. A., Church, G. M. & Valen, E. CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic Acids Res.* **42**, W401–407 (2014).
3. Labun, K., Montague, T. G., Gagnon, J. A., Thyme, S. B. & Valen, E. CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. *Nucleic Acids Res.* **44**, W272–276 (2016).
4. Zhu, X. *et al.* An efficient genotyping method for genome-modified animals and human cells generated with CRISPR/Cas9 system. *Sci. Rep.* **4**, 6420 (2014).
5. Niederriter, A. R. *et al.* In vivo modeling of the morbid human genome using Danio rerio. *J. Vis. Exp. JoVE* e50338 (2013). doi:10.3791/50338
6. Frosk, P. *et al.* A truncating mutation in CEP55 is the likely cause of MARCH, a novel syndrome affecting neuronal mitosis. *J. Med. Genet.* **54**, 490–501 (2017).
7. Isrie, M. *et al.* Mutations in Either TUBB or MAPRE2 Cause Circumferential Skin Creases Kunze Type. *Am. J. Hum. Genet.* **97**, 790–800 (2015).
8. Shaw, N. D. *et al.* SMCHD1 mutations associated with a rare muscular dystrophy can also cause isolated arhinia and Bosma arhinia microphthalmia syndrome. *Nat. Genet.* **49**, 238–248 (2017).
9. Ansar, M. *et al.* Bi-allelic Variants in DYNC1I2 Cause Syndromic Microcephaly with Intellectual Disability, Cerebral Malformations, and Dysmorphic Facial Features. *Am. J. Hum. Genet.* **104**, 1073–1087 (2019).
10. Sun, W. *et al.* Histone Acetylome-wide Association Study of Autism Spectrum Disorder. *Cell* **167**, 1385–1397.e11 (2016).
11. Quail, M. A. *et al.* A large genome center’s improvements to the Illumina sequencing system. *Nat. Methods* **5**, 1005–1010 (2008).
12. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinforma. Oxf. Engl.* **26**, 589–595 (2010).
13. Kumar, V. *et al.* Uniform, optimal signal processing of mapped deep-sequencing data. *Nat. Biotechnol.* **31**, 615–622 (2013).
14. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinforma. Oxf. Engl.* **26**, 841–842 (2010).
15. McCarthy, D. J., Chen, Y. & Smyth, G. K. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* **40**, 4288–4297 (2012).
16. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinforma. Oxf. Engl.* **26**, 139–140 (2010).
17. McLean, C. Y. *et al.* GREAT improves functional interpretation of *cis*-regulatory regions. *Nat. Biotechnol.* **28**, 495–501 (2010).