

1 **Rare *CRHR2* and *GRM8* variants identified as candidate factors associated with eating**
2 **disorders in Japanese patients.**

3

4 **Akira Oka^{1,2*}, Shinji Hadano^{1,2,4}, Mahoko Takahashi Ueda³, So Nakagawa^{1,2,4}, Gen Komaki⁵,**
5 **Tetsuya Ando^{6,7}**

6

7 ¹Department of Molecular Life Sciences, Division of Basic Medical Science and Molecular
8 Medicine, Tokai University School of Medicine, 143 Shimokasuya, Isehara, Kanagawa 259-1193,
9 Japan

10 ²The Institute of Medical Sciences, Tokai University, 143 Shimokasuya, Isehara, Kanagawa,
11 259-1193, Japan

12 ³Department of Genomic Function and Diversity, Medical Research Institute, Tokyo Medical and
13 Dental University, Bunkyo, Tokyo, 113-8510 Japan

14 ⁴Micro/Nano Technology Center, Tokai University, Hiratsuka, Kanagawa, Japan

15 ⁵ Faculty of Medical Science, Fukuoka International University of Health and Welfare,
16 Momochihama, Sawara-ku, Fukuoka 814-0001, Japan

17 ⁶Department of Psychosomatic Medicine, Faculty of Medicine, School of Medicine, International
18 University of Health and Welfare, 4-3 Kozunomori, Narita, Chiba 286-8686, Japan

19 ⁷Department of Behavioral Medicine, National Institute of Mental Health, National Center of
20 Neurology and Psychiatry, 4-1-1 Ogawa-higashi, Kodaira, Tokyo 187-8553, Japan

21 * Corresponding author

22 Email: oka246@is.icc.u-tokai.ac.jp

23

24 **Keywords.** Eating disorder, *CRHR2*, *GRM8*, whole exome sequencing, affected sib-pair

25

1 **Abstract**

2 Eating disorders (EDs) are a type of psychiatric disorder characterized by pathological eating and
3 related behavior and considered to be highly heritable. The purpose of this study was to explore rare
4 variants expected to display biological functions associated with the etiology of EDs. We performed
5 whole exome sequencing (WES) of affected sib-pairs corresponding to disease subtype through their
6 lifetime and their parents. From those results, rare single nucleotide variants (SNVs) concordant with
7 sib-pairs were extracted and estimated to be most deleterious in the examined families. Two
8 non-synonymous SNVs located on corticotropin releasing hormone receptor 2 (*CRHR2*) and
9 glutamate metabotropic receptor 8 (*GRM8*) were identified as candidate disease susceptibility factors.
10 The SNV of *CRHR2* was included within the cholesterol binding motif of the transmembrane helices
11 region, while the SNV of *GRM8* was found to contribute to hydrogen bonds for an α -helix structure.
12 *CRHR2* plays important roles in the serotonergic system of dorsal raphe nuclei, which is involved
13 with feeding and stress-coping behavior. Moreover, *GRM8* modulates glutamatergic
14 neurotransmission, and is also considered to have effects on dopaminergic and adrenergic
15 neurotransmission. Further investigation regarding the biological function of these variants may
16 provide an opportunity for elucidate the pathogenesis of EDs.

17

18

1 **1. Introduction**

2 Eating behavior can be severely affected by eating disorders (EDs), including anorexia nervosa (AN)
3 and bulimia nervosa (BN). The main manifestations of AN are dietary restriction, fasting, excessive
4 exercise, or other weight control or loss behaviors as resistance to maintenance of normal body
5 weight. On the other hand, BN is characterized by repeated binge eating and inadequate
6 compensatory behaviors to avoid weight gain. Manifestations shared by these disorders include
7 aspiration for weight loss, fear of weight gain, and self-evaluation unduly influenced by body shape
8 and weight. Moreover, twin and familial studies demonstrated that genetic predisposition to AN and
9 BN is shared, at least in part [1-3]. The Diagnostic and Statistical Manual of Mental Disorders
10 (DSM-5) provides diagnostic criteria for a subgroup of AN defined as restricting type (AN-R),
11 which include severe dietary restriction without regular binge eating and/or purging behaviors, as
12 well as binge-eating/purging type (AN-BP) with regular binge eating and/or purging behaviors.

13 A recent GWAS for AN in European identified eight significant common variants [4].
14 However, that study captured only 1.7% of the phenotypic variance, thus additional subjects would
15 be required for future replication studies [4]. On the other hand, a sequencing study with linkage
16 analysis was performed using two large families with AN and BN, and those findings suggested that
17 each rare missense variant in estrogen-related receptor α (*ESRRA*) and histone deacetylase 4
18 (*HDAC4*) is associated with EDs [3]. *ESRRA* null mice also display behavioral deficits relevant to
19 EDs such as AN [5]. Moreover, *HDAC4*^{A778T} mice carrying the missense variant demonstrate several
20 ED-related feeding and behavioral deficits in a gender-dependent manner [6, 7].

21 Swedish twin study for EDs demonstrated that concordance rates in monozygotic twins
22 were higher than dizygotic twins within the same disease subtype, on the other hand, the rates were
23 similar in the different subtype [1]. Therefore, in the current study, to use subjects expected to be
24 higher heritability we selected seven affected sib-pairs corresponding to disease subtype through
25 their lifetime for whole exome sequencing (WES) for exploring rare variants functionally related to
26 the pathogenesis of EDs. However, the statistical power was quite insufficient to perform genetic
27 association and linkage analysis. Therefore, we explored the most deleterious coding variants shared

1 with sib-pairs in each family.

2

3 **2. Materials and methods**

4 *2.1 Patients and families*

5 Upon approval of the experimental procedures from the relevant ethical committees at the National
6 Institute of Mental Health, National Center of Neurology and Psychiatry (number: A2013-054) and
7 Tokai University (number: 14I-60), we obtained written informed consent in accordance with the
8 Declaration of Helsinki from 30 sib pairs and their family members prior to collection of DNA
9 samples. Among all subjects we selected 7 sib pairs corresponding to disease subtype through their
10 lifetime for reduction of sample heterogeneity and elevation of likelihood of identification of
11 susceptibility variants [8], though family P8 included a brother who displayed a different subtype
12 (**Fig. 1 and Supplementary Table 1**). The IDs of the samples in this paper were reassigned for the
13 experiment, therefore these IDs could not identify individuals. They are 15 individuals affected with
14 EDs and 11 unaffected individuals, all of Japanese origin. They are also three ANR and three BN
15 families, and one family displayed a diagnostic crossover from ANR to BN. Of the cases, each was
16 diagnosed according to the DSM-IV criteria in Japan. DSM-IV eating disorder diagnosis was made
17 by expert of psychosomatic doctors or psychiatrists expertized on eating disorders. DNA was
18 extracted using a QIAamp DNA blood kit (QIAGEN, Hilden, Germany).

19

20 *2.2 Genomic library construction and sequencing*

21 For exon fragment capture and sequencing, an Agilent SureSelect Target Enrichment, v. 5 (50 Mbp),
22 was used. Sequence analysis was performed using an Illumina Genome Analyzer IIX, HiSeq2000, or
23 HiSeq2500 platform.

24

25 *2.3 Sequencing data analysis*

26 Reads that passed quality control were mapped to the reference genome (UCSC Genome Browser
27 assembly GRCh37/hg19, <http://genome.ucsc.edu/>) with a Burrows-Wheeler Aligner, v. 0.5.9.

1 Potential PCR duplicates were flagged with Picard MarkDuplicates, v. 1.88
2 (<http://picard.sourceforge.net/>). Genome Analysis Toolkit v. 2.2-8, was used to perform map quality
3 score recalibration and variant detection. Single nucleotide variants (SNVs) and indels were then
4 annotated for functional consequences at the gene using the ANNOVAR. Predictions of variants with
5 risk were also performed using ANNOVAR, with non-synonymous variants (LJB, v. 3.0) employed
6 for the effects on protein function (SIFT, PolyPhen, MutationTaster, Mutation Assesor, LRT,
7 FATHMM) and evolutionary conservation (GERP++, phyloP, Siphy). Criteria used for deleterious
8 variants in each prediction are shown following:
9 SIFT (sift): D: deleterious, PolyPhen 2 HDIV: D: Probably damaging, LRT: D: Deleterious,
10 MutationTaster: A (disease_causing_automatic) or D (disease_causing), MutationAssessor: H: high
11 FATHMM: D: deleterious, GERP++ >5, phyloP >2.0, SiPhy >15

12

13 *2.4 Confirmation of SNVs*

14 For PCR and sequencing (Greiner Bio-one) of the SNV of *CRHR2*, forward
15 (CCTGGCAGGGGAGAAGAGC) and reverse (CCCCAAGCTGCCTCCTGACA) primers, and
16 for the SNV of *GRM8*, forward (GCAACTCCAAGTCATCCATTTTCTTCA) and reverse
17 (TGGAGCGAATTGCTCGGGATT) primers, were used. Sequencing reactions were carried out
18 using a BigDye Terminator Cycle Sequencing kit, v. 3.1 (Thermo Fisher Scientific). Automated
19 electrophoresis was performed with an ABI PRISM 3730 Genetic Analyzer (Thermo Fisher
20 Scientific).

21

22 **3. Results**

23

24 *3.1 WES*

25 We obtained rare 4,032 SNVs, 113 splice site variants and 76 insertion/deletion variants (allele
26 frequency <0.001) from a total of 633,030 variants (**Fig. 2**). Initially, we attempted to identify
27 recessive and compound variants concordant with the affected sib-pairs, however none were found

1 among these rare variants. Next, heterozygote variants concordant with affected sib-pairs were
2 extracted as dominant candidates. Finally, we estimated the most deleterious SNV in each family by
3 use of six functional prediction (SIFT, PolyPhen, MutationTaster, Mutation Assessor, LRT,
4 FATHMM) and three conservation-based (GERP++, phyloP, Siphy) methods [9].

5 A previous large exome sequencing study indicated a strong inverse relationship between average
6 SNV age and number of methods needed to predict a variant as deleterious, thus SNVs predicted to
7 be deleterious by multiple methods probably undergo a more intense purifying selection and are very
8 rare in populations [10], suggesting them as candidates for disease pathogenesis. Moreover, most
9 very-rare variants arose within the past 5–10,000 years and are population-specific [11]. Therefore,
10 381 SNVs and 4 *de novo* SNVs were evaluated (**Fig. 2, Supplementary Table 2**) based on the
11 number of methods that predicted a variant to be deleterious and a feature of the gene. Among those,
12 only two were predicted to be deleterious by all nine methods, an SNV in corticotropin releasing
13 hormone receptor 2 (*CRHR2*) noted in family P15 and an SNV in glutamate metabotropic receptor 8
14 (*GRM8*) noted in family P4 (**Table 1 and Fig.1-2**). These were considered likely to be specific for a
15 Japanese population, though a single allele for *CRHR2* was previously observed in Europeans (**Table**
16 **1**). In addition, SNVs supported by eight of the methods were observed in the P7, P8, and P11
17 families, while those supported by seven of the methods were observed in the P3 and P5 families.
18 Among the 10 genes, *CRHR2*, *GRM8*, and glutamate ionotropic receptor
19 α -amino-3-hydroxy-5-methyl-4-isoxazole propionic acid (AMPA) type subunit 2 (*GRIA2*) were
20 indicated to have higher levels of expression in the brain and related tissues (**Table. 1**).

21

Table 1. Most deleterious variants in each family estimated by functional prediction

Gene symbol	Family	No. of methods ^a	Chr	Position ^b	Substitution		SNV	Frequency		Gene description	Gene expression ^f
					Alt ^c	Amino acid		Japanese ^d	European ^e (non-Finnish)		
<i>CRHR2</i>	P15	9	7	30702316	T	p.G231R	rs569607645	0.000414	0.000009	Corticotropin-releasing hormone receptor 2	Pituitary, brain
<i>GRM8</i>	P4	9	7	126409942	C	p.Y445C	NA	0.000000	0.000000	Glutamate metabotropic receptor 8	Testis, brain
<i>LGR4</i>	P7	8	11	27390603	C	p.N556S	rs754383451	0.000000	0.000016	Leucine-rich repeat-containing G protein-coupled receptor 4	Ovary, pancreas
<i>MYH7B</i>		8	20	33583163	A	p.E951K	rs370896124	0.000535	0.000008	Myosin heavy chain 7B	Heart, skeletal muscle
<i>POC1B</i>		8	12	89866042	T	p.D155N	NA	0.000826	0.000000	POC1 centriolar protein homolog B	Heart, testis
<i>ACADS</i>	P8	8	12	121176662	T	p.R325W	rs121908006	0.000000	0.000026	Acyl-CoA dehydrogenase, C-2 to C-3 short chain	Liver, skeletal muscle
<i>ALAS1</i>	P11	8	3	52236598	T	p.P92L	rs370171958	0.000826	0.000015	Aminolevulinate, delta-, synthase 1	Adrenal gland, brain
<i>C3orf70</i>	P3	7	3	184800841	A	p.S236L	rs201783027	0.000413	0.000031	Chromosome 3 open reading frame 70	Colon, artery
<i>NR1P1</i>			21	16338385	T	p.R710H	rs139218138	0.000000	0.000008	Nuclear receptor interacting protein 1	Cervix, adipose
<i>GRIA2</i>	P5	7	4	158253975	T	p.T296I	rs200845438	0.000400	0.000015	Glutamate receptor, ionotropic, AMPA 2	Brain, pituitary

- 1 Abbreviations: NA, not applicable, ^aMaximum number of methods for predicting variant as deleterious in each family, ^bPhysical position in UCSC Genome Browser in Human Feb. 2009 Human GRCh37/hg19
- 2 Assembly, ^cAllele observed in this study, ^dAllele frequency in Japanese in Human Genetic Variation Database (<https://www.hgvd.genome.med.kyoto-u.ac.jp/>), ^eAllele frequency in European gnomAD (<https://gnomad.broadinstitute.org/>), ^fTop two tissues showing highest expression level in GTEx (release V8, 54 non-diseased tissue sites across nearly 1000 individuals).
- 3

1

2 *CRHR2* encodes a neuropeptide receptor for corticotropin-releasing factor (CRF), a critical
3 coordinator of the hypothalamic-pituitary-adrenal axis [12] and for three urocortins (UCN1–UCN3)
4 [13]. *CRHR2* is also a specific receptor for UCN2 and 3, which are associated with regulation of
5 stress, anxiety, and food intake [14–17]. *GRM8* encodes a G-protein coupled metabotropic glutamate
6 receptor that influences inhibition of the cyclic AMP cascade and regulates presynaptic glutamate
7 release [18], and also has been found to have genetic associations with psychiatric phenotypes,
8 including depression [19, 20] and eating behavior [18]. The two variants were also confirmed by
9 Sanger sequencing (**Supplementary Figure 1**). Thus, it is considered that *CRHR2* and *GRM8* are
10 potent candidate genes due to the pathogenesis of EDs in individuals included in this study.

11 *GRIA2* has a significant role in excitatory neurotransmission [21], and has been identified as a
12 causative gene of intellectual disability and developmental epileptic encephalopathy. Therefore, we
13 considered that *GRIA2* was a weaker candidate gene for EDs in this study, as the phenotypes
14 generated by mutations may not be concordant with EDs. Furthermore, among the 381 SNVs
15 considered to be candidates for dominant, all other SNVs were thought to be unlikely as causative
16 factors of EDs after considering the descriptions and locations of expression (**Table 1**).

17 Four *de novo*, five splice site SNVs, and five insertion and deletion variants concordant with
18 affected sib-pairs were identified as heterogeneous (**Supplementary Table 3 - 5**). However, in
19 consideration of their features, these fourteen genes are unlikely to be causative factors of EDs.

20

21 *3.2 Evaluation of impact of SNVs on CRHR2 and GRM8*

22 *CRHR2* is classified as a secretin subfamily among class B G protein-coupled receptors (GPCRs). A
23 GPCR features seven transmembrane helices (TM), three extracellular loops, and three intracellular
24 loops. The rare SNV in family P15 (rs569607645) was found located within the TM4 region. Gly
25 231, with a codon corresponding with the SNV conserved across species and among the class B
26 GPCRs (**Fig. 3**), and located within a highly conserved sequence motif, GWGxP, in class B GPCRs
27 [22]. Moreover, the substitution converts hydrophobic (glycine) into hydrophilic (arginine) amino

1 acid, implying that this SNV influences conformation and function of CRHR2.

2 Metabotropic glutamate (mGlu) receptors are members of the class C family of GPCRs that
3 modulate cellular responses to the excitatory neurotransmitter L-glutamic acid (L-Glu) in many
4 synapses in the CNS [23]. The amino terminal domain (ATD) of mGlu receptors contains an
5 orthosteric glutamate recognition site, which is highly conserved across the eight mGlu receptor
6 subtypes [24]. The rare SNV in family P4 was located within an α -helix structure of the ATD region.
7 Tyrosine 441, with a codon corresponding with the SNV, was found to be conserved across species
8 and hydrogen bonds to Asp 422 in an adjacent α -helix structure (**Supplementary Figure 2**) [23].
9 Therefore, the substitution of tryptophan to cysteine is likely to cancel the hydrogen bond. We also
10 evaluated the effect of the SNV on protein stability using X-ray crystal structure data (PDB ID
11 6BSZ) and ENCoM, which are used to predict the effect of mutations on thermostability and
12 dynamics, as well as to generate geometrically realistic conformational ensembles [25]. Those
13 results indicated that the conformation of GRM8 was expected to be destabilized (-0.822 kcal/mol).
14 Thus, the substitution may have an impact on the structure and functions of GRM8.

15

16 **4. Discussion**

17 *4.1 CRHR2*

18 Common SNVs in *CRHR2* were not detected in a large genome-wide association study for anorexia
19 nervosa and other psychiatric disorders [4]. On the other hand, WES with bipolar disorder families
20 identified a rare and functionally relevant nonsense variant within the intracellular C-terminal region
21 of *CRHR2* [26]. Thus, it is suggested that a rare SNV in *CRHR2* contributes to the pathogenesis of a
22 psychiatric disorder.

23 A study that used *Crhr2* knockout mice demonstrated that CRHR2 is essential for sustained
24 feeding suppression induced by UCNs, members of the CRF family of peptides [27]. In another
25 investigation, *Crhr2*-mutant mice showed decreased food intake following food deprivation [12].
26 Therefore, UCNs are potent for suppression of food intake by CRHR2-specific mediation [28].
27 Furthermore, *Crhr2*-mutant mice were also reported to have an anxiogenic phenotype and impaired

1 stress recovery [13]. These observations are related to serotonergic neurons. Serotonin (5-HT) is
2 produced by serotonergic neurons of the dorsal raphe nucleus (DRN) located within the brain stem
3 and regulates feeding behavior [29], while CRHR2 is specifically expressed in the DRN
4 serotonergic neuron of [30] and plays important roles in controlling serotonergic neuronal activity
5 [31]. Hammack et al. performed experiments within the rat DRN using UCN2, an agonist for the
6 highly selective CRHR2, as well as an antagonist for that and demonstrated that CRHR2 in the DRN
7 mediates behavioral consequences of uncontrollable stress [32]. Additionally, anxiety-related stimuli
8 by UCN2 were shown to lead to increased activation of serotonergic neurons within the DRN [33].
9 The DRN projects to the bed nucleus of the stria terminalis, which plays essential roles in threat
10 processing, and is responsible for such emotional states as fear and anxiety [34, 35]. The tone of the
11 DRN–5-HT system is regulated in a dynamic manner through CRHR2 activation, and either
12 decreased or increased depending on the level of endogenous or exogenous ligands [36]. Indeed, in
13 malnourished individuals suffering from AN the cerebrospinal fluid has reduced amounts of
14 5-hydroxyindoleacetic acid which is the major brain metabolite of 5-HT and is thought to reflect
15 extracellular 5-HT concentrations, indicating that abnormal activity of 5-HT system is related to EDs
16 pathogenesis. [37]. Thus, CRHR2 mediates not only food intake, but also passive coping behavior
17 and depression-like responses triggered by uncontrollable stress [13] via the serotonergic neuron in
18 the DRN.

19 Structural analysis demonstrated that the motif included with rs569607645 is a
20 cholesterol-binding site, and that artificial mutations in that site reduce the activation potency of
21 UCN1 and CRF to CRHR2, suggesting that the site plays a role in stabilizing peptide binding [38].
22 Moreover, bound cholesterol in the structure of CRHR2 contribute to GPCR signaling, indicating
23 that the GPCR signalosome carries out its function in the cholesterol-rich lipid raft [38]. Thus, the
24 substitution of G231R provided by the rare SNV rs569607645 impairs the functions of CRHR2,
25 which may decrease the coping mechanism in response to stress [13] and/or feed suppression [28] in
26 patients with EDs.

27 ED patients have higher rates of anxiety disorders as compared with normal individuals,

1 suggesting that both disorders might share common etiological factors, which can increase the
2 susceptibility of an individual to either [39]. Stress is prominently involved in the pathogenesis and
3 development of EDs and anxiety disorders. Moreover, coping mechanisms such as a positive attitude,
4 planning, and social support seem to be impaired in ED patients who often do not cope well with
5 emotional distress, implicating that such mechanisms may be disrupted [40]. Therefore, a decay of
6 CRHR2 functions may play a part in the pathogenesis of EDs.

7

8 4.2 *GRM8*

9 As mentioned above, associations between psychiatric disorders and *GRM8* has been indicated
10 with strong evidence in previous reports. On the other hand, a sequencing study for bipolar disorder
11 identified enrichment of damaging mutations in GPCRs including *CRHR2* and the *GRM* gene family,
12 with increased numbers of deleterious variants [26].

13 The mGluRs modulate glutamatergic neurotransmission, and are considered to also have effects
14 on dopaminergic and adrenergic neurotransmission, implicating their involvement in a number of
15 neurological disorders [41]. An expression study using animals indicated that *GRM8* may play a role
16 in feeding behavior and metabolism via the hypothalamic pathway [18, 42].

17

18 **5 Conclusions**

19 We found two rare SNVs as candidates for predisposition to EDs in affected sib-pair families. One
20 found located on *CRHR2* plays important roles in the serotonergic system in the DRN, and is
21 involved with feeding and stress-coping behavior, while the other SNV is located on *GRM8*, which
22 modulates cellular responses to the excitatory neurotransmitter in the CNS, and may play a role in
23 feeding behavior and metabolism via the hypothalamic pathway. Moreover, the present findings
24 indicated potential for both non-synonymous SNVs to have impact on conformation and functions of
25 molecules. However, this study only observed the deleterious SNVs in one family for each gene and
26 was not able to indicate that the SNVs explained how much of the genetic variance of EDs.
27 Furthermore, it is impossible to rule out all candidate variants. Therefore, these genes must be

1 confirmed by genetic analysis for additional families and sporadic cases and by functional analysis
2 in the future investigations.

3

4 **Data accessibility**

5 Data obtained in this study are available from the authors upon reasonable request.

6

7 **Author contributions**

8 AO conceived the studies, performed experiments, contributed to the data analysis, and drafted the
9 manuscript. SH, MU and SN were contributed to the data analysis. GK and TA were involved in
10 sample collection, editing the manuscript.

11

12 **Funding**

13 This work was supported by JSPS KAKENHI (16K09275 and 23390191).

14

15 **References**

16

17 [1] C.M. Bulik, L.M. Thornton, T.L. Root, et al., Understanding the relation between anorexia nervosa
18 and bulimia nervosa in a Swedish national twin sample, *Biol Psychiatry*, 67 (2010) 71-77.

19 [2] M. Strober, R. Freeman, C. Lampert, et al., Controlled family study of anorexia nervosa and bulimia
20 nervosa: evidence of shared liability and transmission of partial syndromes, *The American journal of*
21 *psychiatry*, 157 (2000) 393-401.

22 [3] H. Cui, J. Moore, S.S. Ashimi, et al., Eating disorder predisposition is associated with ESRRA and
23 HDAC4 mutations, *The Journal of clinical investigation*, 123 (2013) 4706-4713.

24 [4] H.J. Watson, Z. Yilmaz, L.M. Thornton, et al., Genome-wide association study identifies eight risk
25 loci and implicates metabo-psychiatric origins for anorexia nervosa, *Nature genetics*, 51 (2019)
26 1207-1214.

27 [5] H. Cui, Y. Lu, M.Z. Khan, et al., Behavioral disturbances in estrogen-related receptor alpha-null mice,
28 *Cell reports*, 11 (2015) 344-350.

29 [6] M. Lutter, M.Z. Khan, K. Satio, et al., The Eating-Disorder Associated HDAC4(A778T) Mutation
30 Alters Feeding Behaviors in Female Mice, *Biol Psychiatry*, 81 (2017) 770-777.

31 [7] K.C. Davis, K. Saito, S.R. Rodeghiero, et al., Behavioral Alterations in Mice Carrying Homozygous

- 1 HDAC4 (A778T) Missense Mutation Associated With Eating Disorder, *Frontiers in neuroscience*, 14
2 (2020) 139.
- 3 [8] D.E. Grice, K.A. Halmi, M.M. Fichter, et al., Evidence for a susceptibility gene for anorexia nervosa
4 on chromosome 1, *Am J Hum Genet*, 70 (2002) 787-792.
- 5 [9] X. Liu, X. Jian, E. Boerwinkle, dbNSFP v2.0: a database of human non-synonymous SNVs and their
6 functional predictions and annotations, *Human mutation*, 34 (2013) E2393-2402.
- 7 [10] W. Fu, T.D. O'Connor, G. Jun, et al., Analysis of 6,515 exomes reveals the recent origin of most
8 human protein-coding variants, *Nature*, 493 (2013) 216-220.
- 9 [11] M. Nagasaki, J. Yasuda, F. Katsuoka, et al., Rare variant discovery by deep whole-genome
10 sequencing of 1,070 Japanese individuals, *Nature communications*, 6 (2015) 8018.
- 11 [12] T.L. Bale, A. Contarino, G.W. Smith, et al., Mice deficient for corticotropin-releasing hormone
12 receptor-2 display anxiety-like behaviour and are hypersensitive to stress, *Nature genetics*, 24 (2000)
13 410-414.
- 14 [13] M.J. Henckens, J.M. Deussing, A. Chen, Region-specific roles of the corticotropin-releasing
15 factor-urocortin system in stress, *Nature reviews. Neuroscience*, 17 (2016) 636-651.
- 16 [14] C. Yeh, C.H. Ting, M.L. Doong, et al., Intracerebroventricular urocortin 3 counteracts central acyl
17 ghrelin-induced hyperphagic and gastroprokinetic effects via CRF receptor 2 in rats, *Drug design,
18 development and therapy*, 10 (2016) 3281-3290.
- 19 [15] P. Cottone, V. Sabino, T.R. Nagy, et al., Centrally administered urocortin 2 decreases gorging on
20 high-fat diet in both diet-induced obesity-prone and -resistant rats, *International journal of obesity* (2005),
21 37 (2013) 1515-1523.
- 22 [16] T.M. Reyes, K. Lewis, M.H. Perrin, et al., Urocortin II: a member of the corticotropin-releasing
23 factor (CRF) neuropeptide family that is selectively bound by type 2 CRF receptors, *Proceedings of the
24 National Academy of Sciences of the United States of America*, 98 (2001) 2843-2848.
- 25 [17] K. Lewis, C. Li, M.H. Perrin, et al., Identification of urocortin III, an additional member of the
26 corticotropin-releasing factor (CRF) family with high affinity for the CRF2 receptor, *Proceedings of the
27 National Academy of Sciences of the United States of America*, 98 (2001) 7570-7575.
- 28 [18] M.T. Gast, A. Tonjes, M. Keller, et al., The role of rs2237781 within GRM8 in eating behavior, *Brain
29 Behav*, 3 (2013) 495-502.
- 30 [19] A. Terracciano, T. Tanaka, A.R. Sutin, et al., Genome-wide association scan of trait depression, *Biol
31 Psychiatry*, 68 (2010) 811-817.
- 32 [20] W. Li, K. Ju, Z. Li, et al., Significant association of GRM7 and GRM8 genes with schizophrenia and
33 major depressive disorder in the Han Chinese population, *Eur Neuropsychopharmacol*, 26 (2016)
34 136-146.
- 35 [21] D. Capauto, A. Colantoni, L. Lu, et al., A Regulatory Circuitry Between Gria2, miR-409, and
36 miR-495 Is Affected by ALS FUS Mutation in ESC-Derived Motor Neurons, *Molecular neurobiology*, 55
37 (2018) 7635-7651.

- 1 [22] K. Hollenstein, J. Kean, A. Bortolato, et al., Structure of class B GPCR corticotropin-releasing factor
2 receptor 1, *Nature*, 499 (2013) 438-443.
- 3 [23] Q. Chen, J.D. Ho, S. Ashok, et al., Structural Basis for (S)-3,4-Dicarboxyphenylglycine (DCPG) As
4 a Potent and Subtype Selective Agonist of the mGlu(8) Receptor, *Journal of medicinal chemistry*, 61
5 (2018) 10040-10052.
- 6 [24] J. Hao, Q. Chen, Insights into the Structural Aspects of the mGlu Receptor Orthosteric Binding Site,
7 *Current topics in medicinal chemistry*, 19 (2019) 2421-2446.
- 8 [25] V. Frappier, M. Chartier, R.J. Najmanovich, ENCoM server: exploring protein conformational space
9 and the effect of mutations on protein function and stability, *Nucleic acids research*, 43 (2015) W395-400.
- 10 [26] C. Cruceanu, J.F. Schmouth, S.G. Torres-Platas, et al., Rare susceptibility variants for bipolar
11 disorder suggest a role for G protein-coupled receptors, *Mol Psychiatry*, 23 (2018) 2050-2056.
- 12 [27] S.C. Coste, R.A. Kesterson, K.A. Heldwein, et al., Abnormal adaptations to stress and impaired
13 cardiovascular function in mice lacking corticotropin-releasing hormone receptor-2, *Nature genetics*, 24
14 (2000) 403-409.
- 15 [28] W. Pan, A.J. Kastin, Urocortin and the brain, *Progress in neurobiology*, 84 (2008) 148-156.
- 16 [29] C.P. Magalhães, M.F. de Freitas, M.I. Nogueira, et al., Modulatory role of serotonin on feeding
17 behavior, *Nutritional neuroscience*, 13 (2010) 246-255.
- 18 [30] J. Ren, A. Isakova, D. Friedmann, et al., Single-cell transcriptomes and whole-brain projections of
19 serotonin neurons in the mouse dorsal and median raphe nuclei, *eLife*, 8 (2019).
- 20 [31] M.W. Lieb, M. Weidner, M.R. Arnold, et al., Effects of maternal separation on serotonergic systems
21 in the dorsal and median raphe nuclei of adult male Tph2-deficient mice, *Behavioural brain research*, 373
22 (2019) 112086.
- 23 [32] S.E. Hammack, M.J. Schmid, M.L. LoPresti, et al., Corticotropin releasing hormone type 2 receptors
24 in the dorsal raphe nucleus mediate the behavioral consequences of uncontrollable stress, *The Journal of
25 neuroscience : the official journal of the Society for Neuroscience*, 23 (2003) 1019-1025.
- 26 [33] J.H. Fox, C.A. Lowry, Corticotropin-releasing factor-related peptides, serotonergic systems, and
27 emotional behavior, *Frontiers in neuroscience*, 7 (2013) 169.
- 28 [34] S. Ueda, M. Hosokawa, K. Arikawa, et al., Distinctive Regulation of Emotional Behaviors and
29 Fear-Related Gene Expression Responses in Two Extended Amygdala Subnuclei With Similar Molecular
30 Profiles, *Frontiers in molecular neuroscience*, 14 (2021) 741895.
- 31 [35] C.A. Marcinkiewicz, G. Bierlein-De La Rosa, C.E. Dorrier, et al., Sex-Dependent Modulation of
32 Anxiety and Fear by 5-HT(1A) Receptors in the Bed Nucleus of the Stria Terminalis, *ACS chemical
33 neuroscience*, 10 (2019) 3154-3166.
- 34 [36] L. Pernar, A.L. Curtis, W.W. Vale, et al., Selective activation of corticotropin-releasing factor-2
35 receptors on neurochemically identified neurons in the rat dorsal raphe nucleus reveals dual actions, *The
36 Journal of neuroscience : the official journal of the Society for Neuroscience*, 24 (2004) 1305-1311.
- 37 [37] W.H. Kaye, J.L. Fudge, M. Paulus, New insights into symptoms and neurocircuit function of

- 1 anorexia nervosa, Nature reviews. Neuroscience, 10 (2009) 573-584.
- 2 [38] S. Ma, Q. Shen, L.H. Zhao, et al., Molecular Basis for Hormone Recognition and Activation of
3 Corticotropin-Releasing Factor Receptors, Molecular cell, 77 (2020) 669-680.e664.
- 4 [39] E. Pallister, G. Waller, Anxiety in the eating disorders: understanding the overlap, Clinical
5 psychology review, 28 (2008) 366-386.
- 6 [40] N.T. Burkert, K. Koschutnig, F. Ebner, et al., Structural hippocampal alterations, perceived stress,
7 and coping deficiencies in patients with anorexia nervosa, The International journal of eating disorders,
8 48 (2015) 670-676.
- 9 [41] J.-Q. Wang, A.-L. Brownell, Development of Metabotropic Glutamate Receptor Ligands for
10 Neuroimaging, Current Medical Imaging Reviews, 3 (2007) 186-205.
- 11 [42] S.E. Higgins, L.E. Ellestad, N. Trakooljul, et al., Transcriptional and pathway analysis in the
12 hypothalamus of newly hatched chicks during fasting and delayed feeding, BMC genomics, 11 (2010)
13 162.

14

15 **FIGURE LEGENDS**

16

17 **FIGURE 1** Sib-pair families affected by EDs were examined using whole exome sequencing.
18 Individuals labeled with an ID number were sequenced. The disease phenotype of affected sib-pairs
19 in the P15 family was transferred from ANR to BN. The rare variants *GRM8* and *CRHR2* were
20 observed in the P4 and P15 families, respectively.

21 Abbreviations: ANR, anorexia nervosa restricting type BN, bulimia nervosa

22

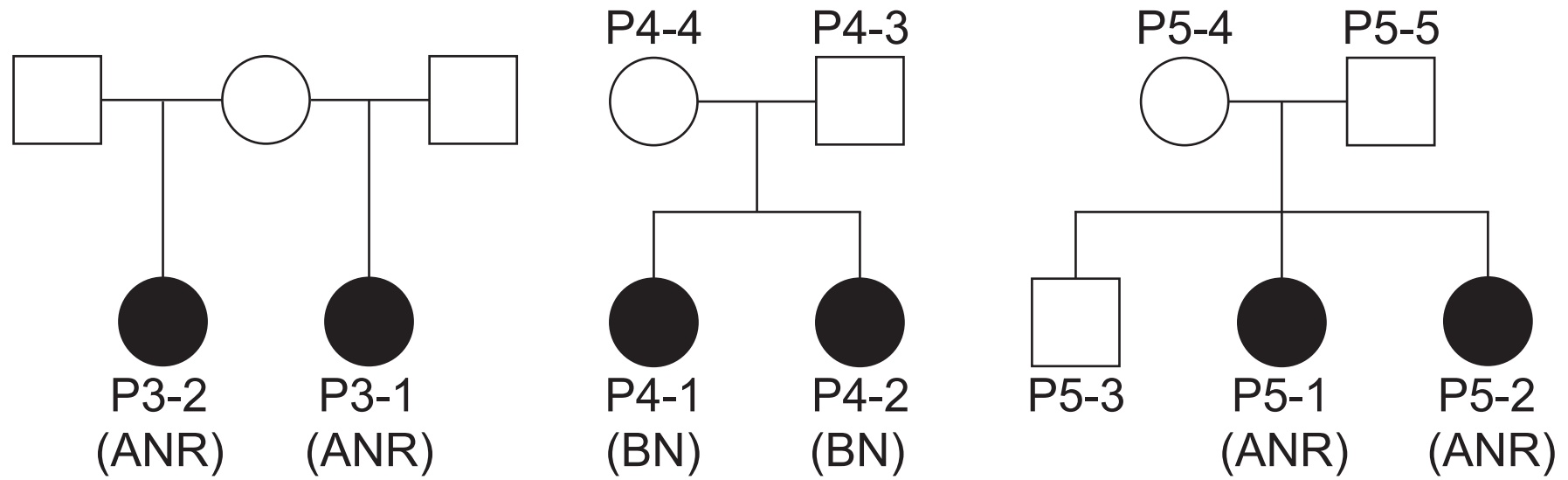
23 **FIGURE 2** Scheme for screening variants with whole exome sequencing.

24

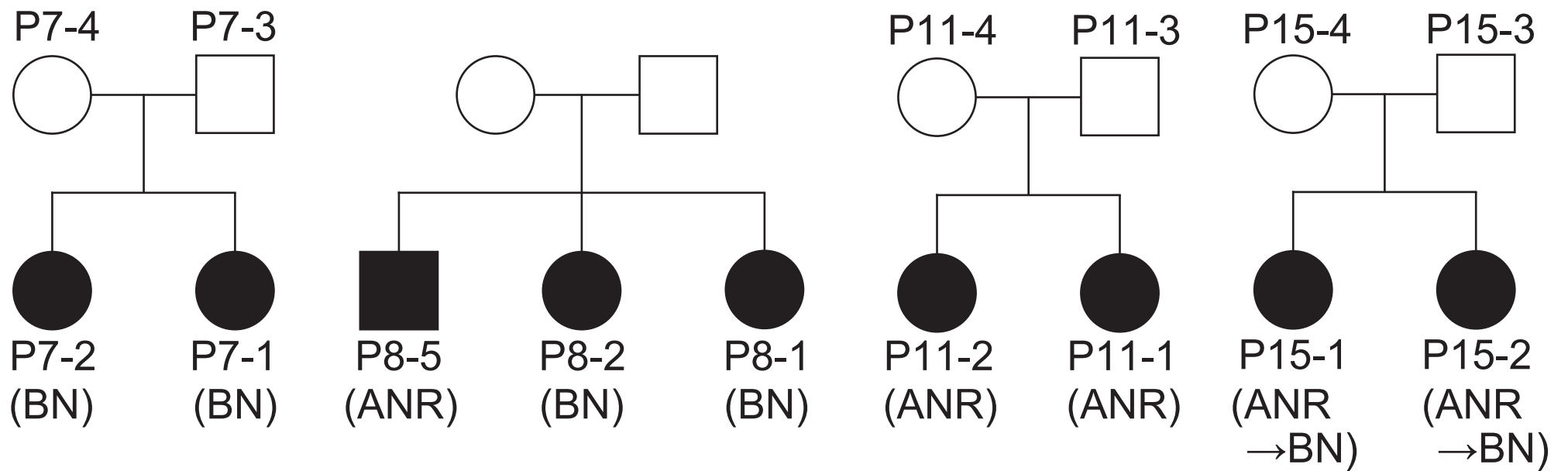
25

26 **FIGURE 3** Multiple amino acid sequence alignments of TM4 helix of human class B GPCRs and
27 *CRHR2* in various species. Upper alignment indicates TM4 of a selected set of human class B
28 GPCRs included with the GWG x P motif involved in cholesterol binding. Lower alignment
29 indicates evolutionarily conserved amino acids. The level of blue shading is shown according to
30 sequence conservation. Black square indicates amino acid position corresponding to human G231.

31



GRM8



CRHR2

Figure 1

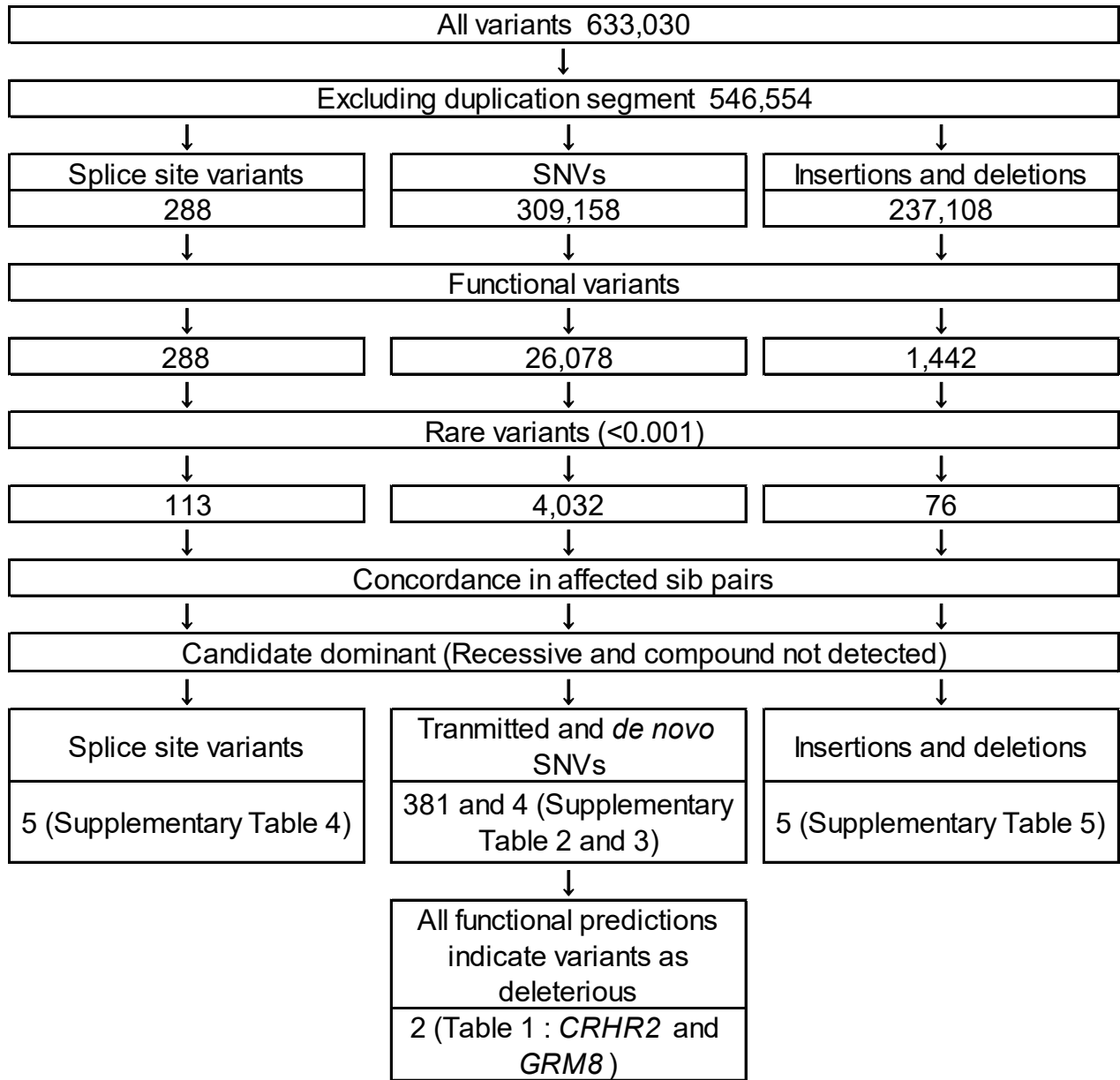


Figure 2

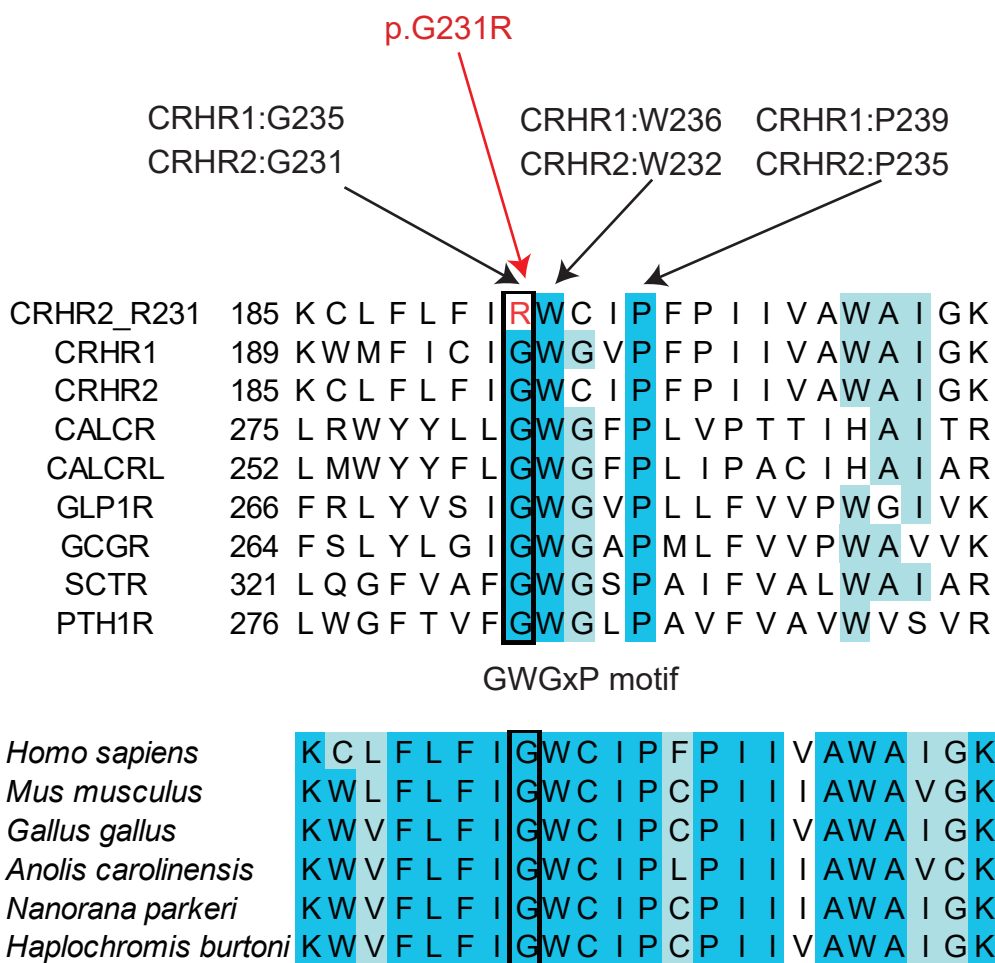
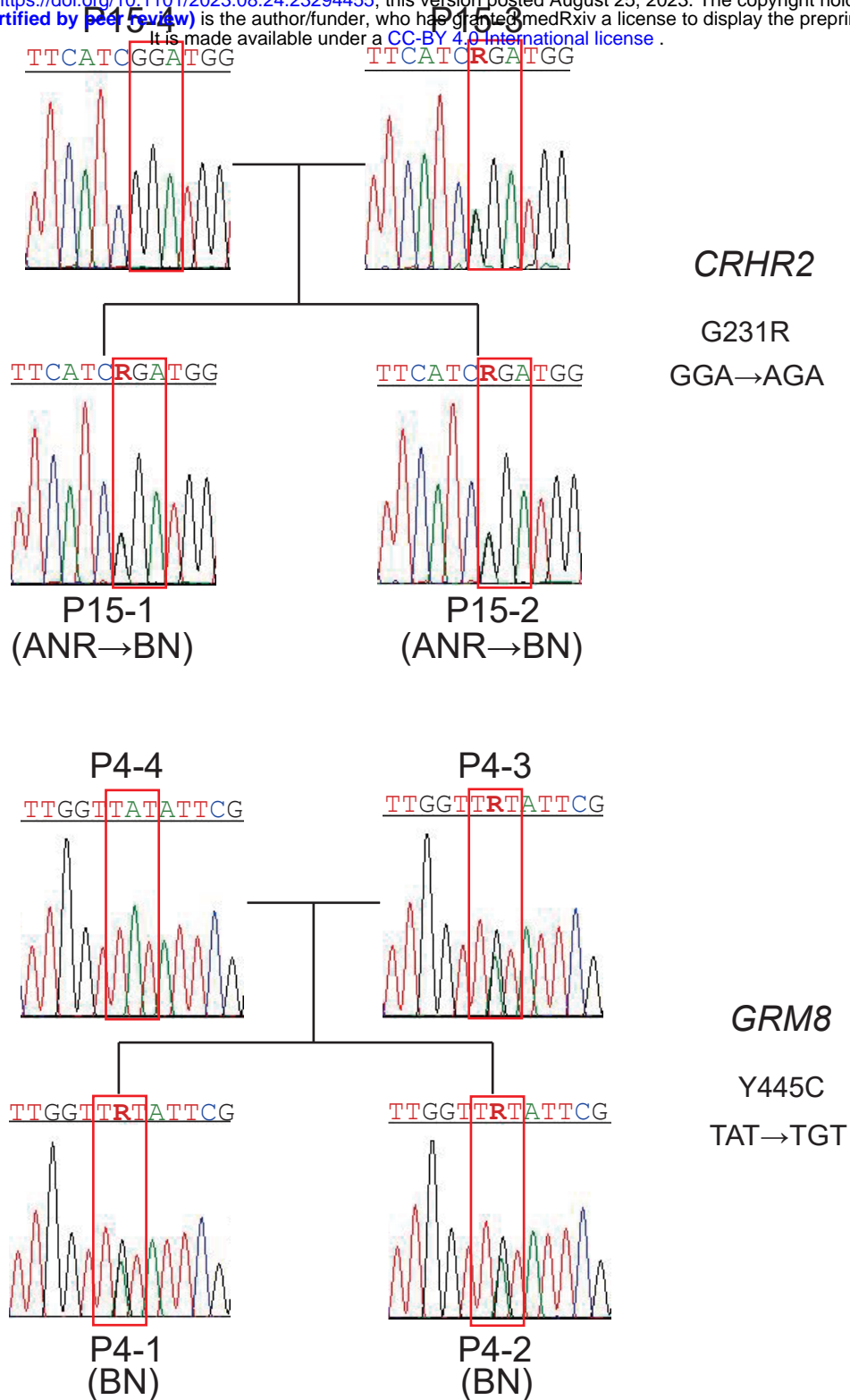


Figure 3

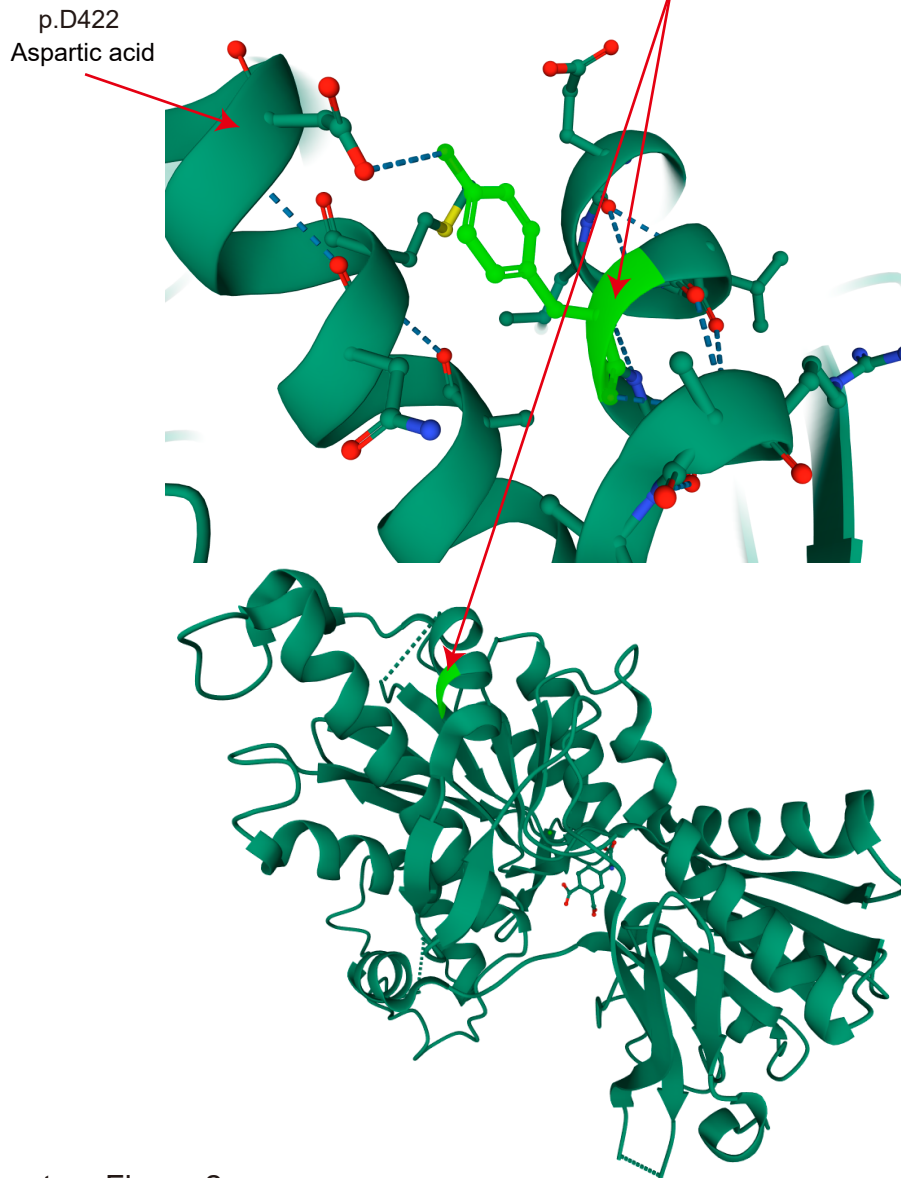


Supplementary Figure 1

Confirmation by Sanger sequencing for SNVs. Red squares indicate codons providing amino acid substitutions. Upper panel shows SNV of CRHR2 in P15 family. Lower panel shows SNV of GRM8 in P4 family.

medRxiv preprint doi: <https://doi.org/10.1101/2023.08.24.23294455>; this version posted August 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

<i>Homo sapiens</i>	S T I D G K E L L G Y I R A V N F N G S A G T
<i>Mus musculus</i>	V T I D G K E L L S Y I R A V N F N G S A G T
<i>Gallus gallus</i>	D G K E L L R A V N F N G S A G T
<i>Anolis carolinensis</i>	A T I D G K E L L S Y I R A V N F N G S A G T
<i>Nanorana parkeri</i>	K T I D G K E L L S Y I R A V N F N G S A G T
<i>Haplochromis burtoni</i>	A N I D G K E L L N H I R A V N F N G S A G T



Supplementary Figure 2

Multiple amino acid sequence alignment of GRM8 showing evolutionarily conserved amino acids and ATD structure of GRM8. Upper alignment indicates evolutionarily conserved amino acids. Black square indicates amino acid position corresponding to human Y445. Middle panel indicates ribbon diagram around Y445 (light green). Hydrogen bonds are depicted as blue dotted lines. Lower panel indicates ribbon representation for all ATDs including Y445 (light green). These ribbon diagram images were created using Mol* Viewer [1] with the PROTEIN DATA BANK (<https://www.rcsb.org/>) for PDB ID 6BSZ [2].

- [1] D. Sehnal, S. Bittrich, M. Deshpande, et al., Mol* Viewer: modern web app for 3D visualization and analysis of large biomolecular structures, *Nucleic acids research*, 49 (2021) W431-w437.
- [2] Q. Chen, J.D. Ho, S. Ashok, et al., Structural Basis for (S)-3,4-Dicarboxyphenylglycine (DCPG) As a Potent and Subtype Selective Agonist of the mGlu(8) Receptor, *Journal of medicinal chemistry*, 61 (2018) 10040-10052.

Supplementary Table 1: Clinical features of affected sib pairs in this study.

Family ID	Individual ID	Gender	Subtype through lifetime	Height (cm)	Weight (kg)	
					Min.	Max.
P3	P3-1	F	ANR	158	NA	NA
	P3-2	F	ANR	151	NA	NA
P4	P4-1	F	BN	156	NA	NA
	P4-2	F	BN	161	NA	NA
P5	P5-1	F	ANR	156	28	NA
	P5-2	F	ANR	167	44	58
P7	P7-1	F	BN	162	NA	NA
	P7-2	F	BN	158	43	60
P8	P8-1	F	BN	152	NA	NA
	P8-2	F	BN	160	39	53
	P8-3	M	ANR	169	42	53
P11	P11-1	F	ANR	162	26	64
	P11-2	F	ANR	168	33	52
P15	P15-1	F	ANR→BN	162	33	57
	P15-2	F	ANR→BN	168	40	47

Gene symbol	Chr	Position ^a	Ref	Alt	cDNA	Amino acid	Exon	Transcript ID	SNV	Frequency ^d	No. of methods ^e
FAM213B	1	2518564	C	T	c.C188T	p.A63V	exon2	NM_001195736	NA	0.000000	2
PRDM16	1	3342747	C	A	c.C3242A	p.A1081D	exon14	NM_022114	NA	0.000000	5
MEGF6	1	3407118	G	A	c.C4600T	p.R1534W	exon37	NM_001409	NA	0.000000	2
DFFB	1	3784601	G	A	c.G566A	p.R189Q	exon4	NM_001282669	rs201973418	0.000415	7
C1orf167	1	11835556	C	T	c.C2053T	p.R685C	exon9	NM_001010881	rs41275456	0.000000	0
C1orf167	1	11838841	A	G	c.A2464G	p.S822G	exon11	NM_001010881	rs56001051	0.000000	0
C1orf167	1	11839034	G	A	c.G2548A	p.A850T	exon12	NM_001010881	rs55967531	0.000000	0
C1orf167	1	11845222	T	C	c.T3443C	p.V1148A	exon16	NM_001010881	rs55867221	0.000000	0
C1orf167	1	11847449	G	A	c.G3572A	p.R1191H	exon17	NM_001010881	rs3737967	0.000000	0
C1orf167	1	11848068	G	C	c.G3812C	p.C1271S	exon18	NM_001010881	rs1537514	0.000000	0
VPS13D	1	12309316	G	A	c.G484A	p.G162S	exon6	NM_015378	NA	0.000000	2
PADI1	1	17559348	C	T	c.C1196T	p.P399L	exon11	NM_013358	NA	0.000000	0
ALPL	1	21904130	G	A	c.G1333A	p.V445I	exon10	NM_001177520	NA	0.000000	1
HSPG2	1	22206699	G	T	c.C2247A	p.H749Q	exon17	NM_001291860	NA	0.000413	2
KIAA0754	1	39879446	C	T	c.C3509T	p.A1170V	exon1	NM_015038	NA	0.000828	0
CCDC30	1	43047046	C	T	c.C1081T	p.R361W	exon8	NM_001080850	NA	0.000000	4
HOOK1	1	60314066	C	T	c.C1009T	p.R337C	exon11	NM_015888	NA	0.000829	6
DOCK7	1	63050524	G	A	c.C1826T	p.P609L	exon16	NM_001272002	NA	0.000000	0
IL23R	1	67705932	G	C	c.G1116C	p.L372F	exon9	NM_144701	rs199542433	0.000414	6
IL23R	1	67705933	A	T	c.A1117T	p.I373F	exon9	NM_144701	rs201752419	0.000414	3
C1orf146	1	92709776	G	A	c.G163A	p.V55I	exon4	NM_001012425	NA	0.000000	2
CD101	1	117556339	A	G	c.A1153G	p.T385A	exon4	NM_001256106	NA	0.000428	0
HFE2	1	145416517	C	T	c.C184T	p.R62W	exon2	NM_213652	NA	0.000000	8
MRPL9	1	151735817	G	A	c.C139T	p.L47F	exon1	NM_001300733	NA	0.000414	4
NES	1	156639758	C	T	c.G4222A	p.E1408K	exon4	NM_006617	NA	0.000000	4
INSRR	1	156821447	G	A	c.C1075T	p.R359C	exon4	NM_014215	NA	0.000413	6
OR10R2	1	158449783	G	A	c.G116A	p.S39N	exon1	NM_001004472	rs139828408	0.000000	1
PYHIN1	1	158908991	C	G	c.C533G	p.P178R	exon4	NM_152501	rs79928064	0.000000	0
ARHGAP30	1	161018556	C	G	c.G1724C	p.R575T	exon8	NM_001287602	NA	0.000000	0
NOS1AP	1	162313766	G	A	c.G580A	p.G194S	exon6	NM_001164757	NA	0.000000	3
SELL	1	169677900	A	G	c.T169C	p.S57P	exon3	NM_000655	NA	0.000000	5
SUCO	1	172501689	G	C	c.G77C	p.G26A	exon1	NM_016227	NA	0.000000	1
TOR3A	1	179054915	G	A	c.G526A	p.V176M	exon3	NM_022371	NA	0.000000	3
SMG7	1	183515464	C	T	c.C2608T	p.P870S	exon16	NM_001174061	NA	0.000545	3
CFH	1	196709910	C	T	c.C2944T	p.P982S	exon18	NM_000186	rs149938052	0.000828	8
PIK3C2B	1	204408074	C	T	c.G3505A	p.E1169K	exon24	NM_002646	rs143689889	0.000000	5
FLVCR1	1	213031874	G	C	c.G80C	p.R27T	exon1	NM_014053	NA	0.000910	1
KCNK2	1	215408418	G	A	c.G1199A	p.S400N	exon7	NM_001017424	rs76873069	0.000000	6
EPRS	1	220206940	T	C	c.A292G	p.T98A	exon4	NM_004446	NA	0.000000	2
HLX	1	221053600	A	C	c.A401C	p.Q134P	exon1	NM_021958	rs201006102	0.000845	1
DNAH14	1	225211527	A	G	c.A1318G	p.M440V	exon11	NM_001145154	NA	0.000000	0
CDC42BPA	1	227216678	C	G	c.G3764C	p.R1255P	exon28	NM_014826	NA	0.000413	6
OBSCN	1	228402755	T	A	c.T1784A	p.F595Y	exon5	NM_001098623	NA	0.000000	2
LYST	1	235926144	C	G	c.G6129C	p.K2043N	exon22	NM_000081	NA	0.000000	2
ACTN2	1	236906343	G	A	c.G1255A	p.G419S	exon11	NM_001103	NA	0.000414	6
KIF26B	1	245850232	C	T	c.C3947T	p.S1316L	exon12	NM_018012	NA	0.000000	2
OR1C1	1	247921589	C	G	c.G120C	p.L40F	exon1	NM_012353	rs201462993	0.000000	0
WDR35	2	20138102	T	C	c.A1987G	p.S663G	exon18	NM_020779	NA	0.000826	5
APOB	2	21229457	G	T	c.C10283A	p.T3428N	exon26	NM_000384	NA	0.000413	0
ADCY3	2	25050816	C	T	c.G2387A	p.R796H	exon13	NM_004036	rs372560995	0.000413	2
HADHB	2	26502162	G	A	c.G745A	p.V249M	exon8	NM_001281512	NA	0.000413	1
SLC5A6	2	27430151	C	T	c.G368A	p.R123H	exon3	NM_021095	rs188933728	0.000000	5
BIRC6	2	32727916	G	C	c.G9262C	p.V3088L	exon48	NM_016252	NA	0.000000	7
BIRC6	2	32740551	G	T	c.G11063T	p.R3688M	exon55	NM_016252	NA	0.000000	7
DYNC2L1	2	44032327	C	T	c.C938T	p.A313V	exon12	NM_001193464	rs147707047	0.000413	5
FANCL	2	58449116	G	A	c.C335T	p.S112L	exon5	NM_001114636	NA	0.000413	4
BCL11A	2	60688494	C	T	c.G1553A	p.S518N	exon4	NM_018014	NA	0.000000	4
AFTPH	2	64779380	G	A	c.G772A	p.A258T	exon2	NM_001002243	NA	0.000826	0
EGR4	2	73519027	G	A	c.C1328T	p.A443V	exon2	NM_001965	NA	0.000000	0
ACTR1B	2	98274979	C	T	c.G568A	p.V190I	exon6	NM_005735	NA	0.000000	4
DARS	2	136700968	C	T	c.G103A	p.V35I	exon4	NM_001293312	rs140573664	0.000000	4
XIRP2	2	168103167	G	C	c.G4599C	p.L1533F	exon7	NM_001199144	NA	0.000000	4
PDE11A	2	178879115	G	A	c.C985T	p.R329X	exon2	NM_016953	rs188985665	0.000826	1
AOX1	2	201527649	C	T	c.C3500T	p.A1167V	exon31	NM_001159	rs35345784	0.000000	5
LANCL1	2	211319834	G	A	c.C404T	p.T135I	exon4	NM_001136574	NA	0.000000	0
CATIP	2	219232598	C	G	c.C1075G	p.P359A	exon10	NM_198559	NA	0.000414	0
MROH2A	2	234708903	A	T	c.A1511T	p.H504L	exon14	NM_001287395	rs140951259	0.000000	0
MROH2A	2	234740136	G	A	c.G4726A	p.A1576T	exon39	NM_001287395	rs879665	0.000000	0
COL6A3	2	238275503	A	G	c.T3506C	p.V1169A	exon8	NM_057166	NA	0.000000	8
AGXT	2	241808686	G	T	c.G265T	p.A89S	exon2	NM_000030	NA	0.000000	5
PASK	2	242066279	G	C	c.C2051G	p.P684R	exon10	NM_001252119	rs56372985	0.000827	1
PLCD1	3	38050766	A	G	c.T1666C	p.S556P	exon10	NM_001130964	NA	0.000000	3
SLC22A13	3	38316953	G	A	c.C899A	p.R300Q	exon5	NM_004256	rs145061420	0.000413	2
NME6	3	48338264	A	G	c.T111C	p.L4P	exon2	NM_001308434	NA	0.000000	6
ALAS1	3	52236598	C	T	c.C275T	p.P92L	exon3	NM_001304443	rs370171958	0.000826	8
CADPS	3	62467420	T	C	c.A3151G	p.M1051V	exon22	NM_003716	NA	0.000000	5
ATXN7	3	63981783	C	T	c.C1850T	p.T617M	exon8	NM_001128149	rs368650687	0.000413	1
SUCLG2	3	67578623	G	A	c.C350T	p.A117V	exon4	NM_001177599	NA	0.000000	5
OR5K3	3	98109652	T	A	c.T143A	p.L48X	exon1	NM_001005516	NA	0.000826	3

ZNF80	3	13935594	G	A	c.A292D	p.C100E	exon11	NM_007136	NA	0.000826	4
ARHGAP31	3	119134525	A	A	c.A3749C	p.D1250A	exon12	NM_020754	NA	0.000000	2
GPR156	3	119885889	G	A	c.C2423T	p.T808M	exon9	NM_001168271	rs200926939	0.000413	5
SEMA5B	3	122634697	C	T	c.G1555A	p.G519R	exon12	NM_001256348	rs183425023	0.000000	0
UMPS	3	124456826	C	T	c.C722T	p.S241L	exon3	NM_000373	NA	0.000000	2
COL6A5	3	130190716	G	A	c.G7765A	p.A2589T	exon40	NM_001278298	rs73868680	0.000000	1
HLTF	3	148778571	A	T	c.T1235A	p.M412K	exon11	NM_003071	rs201653124	0.000000	1
ERICH6	3	150387158	G	A	c.C986T	p.P329L	exon12	NM_001308234	rs200581445	0.000413	1
GMPS	3	155633905	T	G	c.T1136G	p.L379R	exon9	NM_003875	NA	0.000000	5
SLC7A14	3	170198612	G	C	c.C1459G	p.P487A	exon7	NM_020949	NA	0.000000	7
USP13	3	179399676	G	C	c.G179C	p.G60A	exon2	NM_003940	NA	0.000000	5
C3orf70	3	184800841	G	A	c.C707T	p.S236L	exon2	NM_001025266	rs201783027	0.000413	7
SLC26A1	4	985463	T	C	c.A29G	p.Q10R	exon2	NM_022042	NA	0.000439	1
FAM53A	4	1646051	C	T	c.G901A	p.E301K	exon5	NM_001297435	NA	0.000000	0
HTT	4	3240332	A	G	c.A9050G	p.Y3017C	exon65	NM_002111	NA	0.000000	3
TMEM128	4	4249884	G	T	c.C46A	p.L16I	exon1	NM_001297551	rs6854167	0.000000	0
WFS1	4	6302786	G	T	c.G1264T	p.A422S	exon8	NM_001145853	NA	0.000413	4
LOC389199	4	7940903	A	G	c.A11G	p.Q4R	exon1	NM_203423	rs7695170	0.000000	1
LOC389199	4	7940915	A	T	c.C23T	p.P8L	exon1	NM_203423	rs7669645	0.000000	1
GPR78	4	8582821	A	C	c.A112C	p.T38P	exon1	NM_080819	NA	0.000000	1
BMP2K	4	79747277	C	A	c.C265A	p.L89I	exon2	NM_017593	NA	0.000414	6
HNRNPDL	4	83347711	C	G	c.G1097C	p.G366A	exon6	NM_031372	rs200123403	0.000826	6
SPARCL1	4	88403697	G	A	c.C1172T	p.T391M	exon6	NM_001291977	rs199501739	0.000000	4
ANK2	4	114288871	G	A	c.G4927A	p.V1643I	exon41	NM_020977	NA	0.000826	0
MAML3	4	140811296	G	A	c.C1294T	p.P432S	exon2	NM_018717	NA	0.000000	1
ELMOD2	4	141446643	C	G	c.C61G	p.L21V	exon2	NM_153702	rs201497167	0.000000	1
HHIP	4	145627795	A	G	c.A944G	p.H315R	exon5	NM_022475	NA	0.000000	6
PRMT9	4	148575205	G	C	c.C1504G	p.H502D	exon8	NM_001304458	NA	0.000000	5
PRMT9	4	148605066	C	G	c.G73C	p.V25L	exon1	NM_138364	NA	0.000000	4
LRBA	4	151788842	A	G	c.T2747C	p.L916S	exon22	NM_001199282	NA	0.000000	7
TMEM154	4	153601028	C	A	c.G58T	p.G20C	exon1	NM_152680	NA	0.000000	1
GRIA2	4	158253975	C	T	c.C887T	p.T296I	exon7	NM_000826	NA	0.000000	7
SH3RF1	4	170189972	C	G	c.G392C	p.R131T	exon2	NM_020870	rs200790356	0.000000	4
NEK1	4	170483332	C	A	c.G1036T	p.E346X	exon12	NM_001199400	NA	0.000000	1
SCRG1	4	174312559	G	C	c.C7G	p.L3V	exon2	NM_007281	NA	0.000829	0
CFAP97	4	186111851	C	T	c.G500A	p.C167Y	exon2	NM_001292033	rs182552516	0.000413	0
LRP2BP	4	186295608	C	T	c.G338A	p.G113E	exon4	NM_018409	NA	0.000413	7
AHR	5	434456	A	G	c.A1613G	p.K538R	exon11	NM_001242412	NA	0.000000	4
ICE1	5	5462961	A	T	c.A3514T	p.T1172S	exon13	NM_015325	NA	0.000000	0
TRIO	5	14143940	G	C	c.G106C	p.A36P	exon1	NM_007118	NA	0.000911	0
C5orf42	5	37167291	G	A	c.C7258T	p.L2420F	exon35	NM_023073	NA	0.000000	4
ZSWIM6	5	60840041	C	T	c.C3545T	p.A1182V	exon14	NM_020928	NA	0.000000	6
MAST4	5	65892790	G	A	c.G307A	p.A103T	exon1	NM_001164664	NA	0.000000	0
ANKRD31	5	74491679	G	C	c.C794G	p.P265R	exon7	NM_001164443	rs185973227	0.000000	0
POLK	5	74886182	A	G	c.A1273G	p.I425V	exon11	NM_016218	NA	0.000000	2
SLF1	5	94014564	A	G	c.A1879G	p.I627V	exon15	NM_032290	NA	0.000000	0
PPIC	5	122359668	T	G	c.A541C	p.T181P	exon5	NM_000943	rs149059275	0.000826	7
FBN2	5	127624229	C	T	c.G6658A	p.G2220S	exon53	NM_001999	NA	0.000000	2
FNIP1	5	131008060	C	G	c.G1993C	p.A665P	exon13	NM_001008738	NA	0.000000	1
P4HA2	5	131528750	G	A	c.C1555T	p.R519X	exon15	NM_001017973	rs200583507	0.000414	4
PCDHGB2	5	140740144	C	T	c.C442T	p.P148S	exon1	NM_018923	NA	0.000000	1
DPYSL3	5	146777371	C	T	c.G1661A	p.R554H	exon12	NM_001197294	rs188735247	0.000000	5
RANBP17	5	170380604	G	A	c.G1472A	p.R491H	exon13	NM_022897	rs139638431	0.000000	6
VAR2	6	30892159	C	T	c.C2075T	p.S692L	exon25	NM_001167733	rs138341222	0.000000	0
BTBD9	6	38256214	T	C	c.A1084G	p.I362V	exon7	NM_152733	NA	0.000000	3
DNAH8	6	38957803	A	G	c.A13069G	p.K4357E	exon88	NM_001206927	rs139250593	0.000828	2
CUL9	6	43167698	G	A	c.G3188A	p.R1063H	exon14	NM_015089	NA	0.000000	2
RRAGD	6	90121613	C	G	c.G100C	p.G34R	exon1	NM_021244	NA	0.000000	3
SLC22A16	6	110746218	C	T	c.G1592A	p.R531Q	exon8	NM_033125	NA	0.000000	2
RSPH4A	6	116953437	C	T	c.C1984T	p.P662S	exon6	NM_001010892	NA	0.000000	6
STXBP5	6	147599259	A	G	c.A757G	p.I253V	exon8	NM_001127715	NA	0.000826	2
SYNE1	6	152674804	C	T	c.G10957A	p.E3653K	exon68	NM_033071	NA	0.000000	6
TULP4	6	158850898	C	T	c.C512T	p.T171M	exon3	NM_001007466	NA	0.000419	4
C7orf50	7	1040177	T	C	c.A334G	p.K112E	exon4	NM_001134395	NA	0.000837	1
USP42	7	6194357	A	T	c.A3172T	p.R1058W	exon15	NM_032172	NA	0.000000	2
ABCB5	7	20739694	G	A	c.G938A	p.G313D	exon10	NM_178559	NA	0.000414	8
CRHR2	7	30702316	C	T	c.G688A	p.G230R	exon6	NM_001202482	rs569607645	0.000414	9
AMPH	7	38670924	C	A	c.G28T	p.A10S	exon1	NM_001635	NA	0.000000	2
PKD1L1	7	47886553	A	G	c.T5077C	p.F1693L	exon32	NM_138295	NA	0.000000	0
ZNF117	7	64439432	C	T	c.G517A	p.A173T	exon4	NM_015852	NA	0.000000	0
ZSCAN1	7	99655438	G	A	c.G517A	p.E173K	exon3	NM_145914	NA	0.000000	3
AP4M1	7	99703902	G	C	c.G1013C	p.R338P	exon13	NM_004722	NA	0.000000	2
Orai2	7	102087451	G	C	c.G486C	p.K162N	exon3	NM_001271819	NA	0.000000	2
GRM8	7	126409942	T	C	c.A1334G	p.Y445C	exon6	NM_000845	NA	0.000000	9
AOC1	7	150554973	A	G	c.A1415G	p.N472S	exon2	NM_001091	NA	0.000000	3
ABCB8	7	150737585	G	A	c.G988A	p.V330M	exon10	NM_001282293	NA	0.000000	4
CHPF2	7	150929872	C	T	c.C22T	p.P8S	exon1	NM_001284295	rs374377569	0.000000	1
ADAM28	8	24178758	G	A	c.G676A	p.E226K	exon8	NM_001304351	NA	0.000000	0
ADAM7	8	24349579	G	A	c.G1520A	p.R507H	exon14	NM_003817	NA	0.000000	4
PBK	8	27678103	G	T	c.C574A	p.P192T	exon6	NM_001278945	rs200662690	0.000826	1
ZFH4	8	77762550	C	T	c.C3916T	p.R1306W	exon9	NM_024721	NA	0.000414	1
ZFAND1	8	82626176	C	T	c.G457A	p.D153N	exon6	NM_001170796	NA	0.000000	0

medRxiv preprint doi: <https://doi.org/10.1101/2023.08.24.23294455>; this version posted August 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Gene	Chromosome	Position	Strand	RefSeq ID	Transcript ID	Exon	rsID	Distance	r2	
ZNF707	8	44776284	C	c.C958	p.C958	exon14	NM_001288808	NA	0.000000	
KANK1	9	731164	C	c.C2429	p.T810	exon4	NM_153186	rs114209322	0.000826	
RIC1	9	5756326	G	A	c.G1696A	p.A566T	exon15	NM_001206557	rs145216888	0.000827
UHRF2	9	6421108	C	A	c.C350A	p.A117D	exon2	NM_152896	NA	0.000000
MPDZ	9	13222290	C	T	c.G689A	p.S230N	exon6	NM_001261406	NA	0.000000
PSIP1	9	15486844	T	C	c.A374G	p.E125G	exon4	NM_021144	NA	0.000000
BNC2	9	16436081	T	C	c.A2111G	p.E704G	exon6	NM_017637	NA	0.000000
HAUS6	9	19063095	T	G	c.A1435C	p.K479Q	exon14	NM_001270890	rs147787866	0.000000
C9orf72	9	27562432	C	T	c.G547A	p.V183I	exon4	NM_001256054	NA	0.000422
TMEM8B	9	35842551	G	A	c.G116A	p.R39H	exon6	NM_001042590	NA	0.000000
SHB	9	38068006	T	C	c.A637G	p.T213A	exon1	NM_003028	NA	0.000880
PRUNE2	9	79323007	C	T	c.G4183A	p.E1395K	exon8	NM_001308047	NA	0.000000
NOL8	9	95085735	C	T	c.G190A	p.D64N	exon3	NM_017948	rs151033269	0.000827
FBP2	9	97355952	C	T	c.G57A	p.M19I	exon1	NM_003837	NA	0.000000
ERCC6L2	9	98684666	C	T	c.C1412T	p.A471V	exon8	NM_001010895	rs201546231	0.000829
TBC1D2	9	100962538	C	T	c.G1196A	p.S399N	exon6	NM_001267572	NA	0.000000
GALNT12	9	101570348	C	A	c.C368A	p.P123Q	exon1	NM_024642	NA	0.000000
ARPC5L	9	127639182	C	G	c.C425G	p.S142C	exon4	NM_030978	rs150162831	0.000828
FAM102A	9	130707095	C	T	c.G574A	p.V192M	exon6	NM_203305	NA	0.000000
SLC27A4	9	131117979	T	C	c.T1678C	p.C560R	exon12	NM_005094	NA	0.000827
PHYHD1	9	131702696	G	A	c.G443A	p.R148Q	exon7	NM_001100877	rs200845438	0.000413
TTF1	9	135277447	C	G	c.G762C	p.Q254H	exon2	NM_007344	NA	0.000000
DPH7	9	140449957	C	T	c.G1093A	p.A365T	exon9	NM_138778	NA	0.000000
SFMBT2	10	7423832	A	G	c.T29C	p.M10T	exon2	NM_001018039	rs141352828	0.000000
RBP3	10	48388554	G	A	c.C2324T	p.T775M	exon1	NM_002900	rs201724324	0.000415
PPA1	10	71977620	G	T	c.C245A	p.A82E	exon4	NM_021129	NA	0.000000
TBATA	10	72541674	G	A	c.C160T	p.R54W	exon4	NM_152710	NA	0.000414
CDH23	10	73538046	G	A	c.G5168A	p.R1723H	exon38	NM_022124	rs189361642	0.000000
FBXW4	10	103384550	G	A	c.C788T	p.T263M	exon6	NM_022039	NA	0.000000
KNDC1	10	135020716	G	A	c.G3655A	p.D1219N	exon20	NM_152643	NA	0.000000
SIGIRR	11	406932	G	C	c.C790G	p.Q264E	exon8	NM_001135053	NA	0.000434
RASSF7	11	562229	C	T	c.C275T	p.A92V	exon3	NM_001143993	NA	0.000000
TMEM80	11	704536	G	A	c.G742A	p.A248T	exon6	NM_001276274	rs140750324	0.000000
OTOG	11	17593709	T	C	c.T2074C	p.S692P	exon17	NM_001277269	rs7106548	0.000000
OTOG	11	17667439	G	C	c.G8726C	p.W2909S	exon55	NM_001277269	rs11024357	0.000000
LGR4	11	27390603	T	C	c.A1667G	p.N556S	exon18	NM_018490	NA	0.000000
DCDC1	11	31327261	C	T	c.G655A	p.D219N	exon6	NM_181807	NA	0.000000
ELP4	11	31703564	C	T	c.C1373T	p.T458I	exon10	NM_001288726	NA	0.000000
PAX6	11	31814879	G	A	c.C1139T	p.S380L	exon9	NM_001310159	rs3026384	0.000000
ACCS	11	44104752	A	G	c.A1145G	p.H382R	exon13	NM_001127219	rs201202479	0.000833
MYBPC3	11	47360898	C	T	c.G2125A	p.D709N	exon21	NM_000256	NA	0.000000
OR5D18	11	55587892	C	T	c.C787T	p.P263S	exon1	NM_001001952	NA	0.000000
OR5W2	11	55682040	A	G	c.T19C	p.S7P	exon1	NM_001001960	NA	0.000000
OR9Q2	11	57958338	G	A	c.G376A	p.V126M	exon1	NM_001005283	NA	0.000413
BEST1	11	61727452	C	A	c.C776A	p.P259H	exon7	NM_001300786	NA	0.000416
ATG2A	11	64669448	C	T	c.G4105A	p.G1369S	exon29	NM_015104	rs143404952	0.000000
PCNXL3	11	65403904	C	T	c.C5636T	p.P1879L	exon34	NM_032223	NA	0.000000
SART1	11	65744148	C	T	c.C1768T	p.R590C	exon14	NM_005146	NA	0.000538
LRP5	11	68207302	G	C	c.G2663C	p.G888A	exon21	NM_001291902	NA	0.000000
TPCN2	11	68848873	C	T	c.C1595T	p.P532L	exon18	NM_139075	NA	0.000000
PLEKHB1	11	73362899	G	A	c.G257A	p.R86H	exon3	NM_001130036	rs184116184	0.000426
FAT3	11	92086054	T	G	c.T776G	p.I259R	exon1	NM_001008781	NA	0.000000
KDM4E	11	94758987	A	G	c.A266G	p.K89R	exon1	NM_001161630	NA	0.000000
PGR	11	100933406	C	T	c.G1984A	p.V662I	exon4	NM_000926	rs150584881	0.000413
MMP13	11	102816478	T	G	c.A1212C	p.R404S	exon9	NM_002427	NA	0.000414
OR10S1	11	123847636	A	G	c.T763C	p.C255R	exon1	NM_001004474	NA	0.000000
OR10G7	11	123908916	C	A	c.G793T	p.D265Y	exon1	NM_001004463	NA	0.000000
KLRC3	12	10568304	A	G	c.T677C	p.I226T	exon6	NM_002261	NA	0.000000
TAS2R20	12	11149658	A	G	c.T817C	p.S273P	exon1	NM_176889	rs150448607	0.000413
PDE3A	12	20522343	G	C	c.G125C	p.G42A	exon1	NM_000921	NA	0.000424
HDAC7	12	48189043	C	G	c.G1097C	p.S366T	exon10	NM_001098416	NA	0.000000
ASIC1	12	50468007	G	A	c.G640A	p.G214S	exon1	NM_001256830	NA	0.000000
MARS	12	57894189	G	A	c.G1177A	p.A393T	exon10	NM_004990	rs141340466	0.000827
ARHGEF25	12	58010651	G	T	c.G1717T	p.A573S	exon15	NM_182947	NA	0.000000
SRGAP1	12	64437300	C	T	c.C746T	p.A249V	exon6	NM_020762	NA	0.000000
SRGAP1	12	64536271	G	A	c.G3077A	p.R1026H	exon22	NM_020762	NA	0.000000
GNS	12	65110525	A	G	c.T1655C	p.L552P	exon14	NM_002076	NA	0.000000
TBC1D30	12	65237176	G	C	c.C949C	p.E317Q	exon8	NM_015279	NA	0.000430
OSBPL8	12	76763472	A	C	c.T2059G	p.C687G	exon19	NM_001003712	rs199664136	0.000000
POC1B	12	89866042	C	T	c.G337A	p.D113N	exon4	NM_001199777	NA	0.000826
UHRF1BP1L	12	100452762	G	A	c.C2293T	p.R765W	exon14	NM_015054	rs369358032	0.000826
ACACB	12	109577466	G	A	c.G256A	p.G86S	exon1	NM_001093	rs371386144	0.000000
OAS1	12	113346391	C	G	c.C231G	p.D77E	exon2	NM_001032409	NA	0.000000
MED13L	12	116457747	G	A	c.C656T	p.T219M	exon6	NM_015335	rs191743300	0.000000
ACADS	12	121176662	C	T	c.C973T	p.R325W	exon8	NM_000017	rs121908006	0.000000
LRRC63	13	46840996	A	C	c.A1439C	p.Y480S	exon9	NM_001282460	rs41284167	0.000000
BIVM-ERCC5,ERCC5	13	103515408	G	T	c.G1909T	p.A637S	exon8	NM_000123	NA	0.000000
PARP2	14	20820490	A	G	c.A584G	p.Y195C	exon7	NM_001042618	rs369902077	0.000826
RNASE11	14	21052549	C	A	c.G85T	p.E29X	exon3	NM_145250	NA	0.000826
TTC6	14	38091460	G	C	c.G239C	p.R80P	exon3	NM_001310135	NA	0.000000
TTC6	14	38256794	T	C	c.T3436C	p.F1146L	exon17	NM_001310135	rs75245849	0.000000
KLHL28	14	45414389	C	T	c.G785A	p.R262H	exon2	NM_001308112	rs199672203	0.000826

SAVI	14	51711693	C	I	A	c.A1877	p.L951F	exon13	NM_021818	NA	0.000000	3
CHURC1.CHURC1-FNTB	14	65381181	C	A	C	c.G49	p.R17C	exon1	NM_0202559	NA	0.000414	2
PLEKHH1	14	68038492	G	A	C	c.G1458A	p.M486I	exon10	NM_020715	NA	0.000833	1
BDKRB1	14	96730381	T	A	C	c.T362A	p.L121X	exon3	NM_000710	NA	0.000000	3
AMN	14	103394845	G	A	C	c.G290A	p.G97D	exon4	NM_030943	NA	0.000906	1
PACS2	14	105849743	C	T	C	c.C1673T	p.A558V	exon16	NM_001100913	NA	0.000000	3
FMN1	15	33446344	T	C	C	c.A772G	p.K258E	exon4	NM_001277313	rs61744870	0.000000	0
EMC4	15	34520042	C	G	C	c.C350G	p.S117X	exon3	NM_001286420	NA	0.000413	5
CASC5	15	40937249	C	A	C	c.C5853A	p.N1951K	exon15	NM_144508	NA	0.000000	0
NUSAP1	15	41657631	C	T	C	c.C620T	p.T207I	exon6	NM_001243144	NA	0.000000	1
VPS39	15	42454330	T	C	C	c.A2360G	p.N787S	exon23	NM_015289	NA	0.000413	2
ZSCAN29	15	43653998	C	G	C	c.G1832C	p.C611S	exon5	NM_152455	NA	0.000000	0
DUOX2	15	45402687	C	A	C	c.G979T	p.E327X	exon9	NM_014080	NA	0.000000	5
VPS13C	15	62182571	G	A	C	c.C9005T	p.A3002V	exon65	NM_017684	NA	0.000000	3
TLN2	15	63055808	A	C	C	c.A5008C	p.I1670L	exon37	NM_015059	NA	0.000000	4
IGDCC4	15	65694752	C	T	C	c.G637A	p.V213M	exon4	NM_020962	rs143723846	0.000000	6
MAP2K1	15	66782894	C	A	C	c.C1123A	p.L375I	exon11	NM_002755	NA	0.000413	5
TLE3	15	70349848	C	T	C	c.G1006A	p.A336T	exon11	NM_001282982	NA	0.000000	5
THSD4	15	72057393	G	A	C	c.G1544A	p.R515K	exon10	NM_001286429	NA	0.000413	5
NR2E3	15	72104731	C	G	C	c.C627G	p.Y209X	exon5	NM_014249	NA	0.000000	1
SH2D7	15	78393619	T	G	C	c.T1024G	p.C342G	exon5	NM_001101404	NA	0.000000	0
ADAMTSL3	15	84566637	C	T	C	c.C1495T	p.R499W	exon14	NM_001301110	NA	0.000000	6
PIGQ	16	632973	C	T	C	c.C1622T	p.P541L	exon10	NM_148920	NA	0.000425	2
IGFALS	16	1840816	C	T	C	c.G1717A	p.G573S	exon2	NM_001146006	NA	0.000000	1
RPL3L	16	1997238	C	G	C	c.G645C	p.E215D	exon5	NM_005061	NA	0.000000	6
C16orf96	16	4606707	C	A	C	c.C217A	p.P73T	exon1	NM_001145011	NA	0.000827	0
C16orf96	16	4626216	G	A	C	c.G1735A	p.A579T	exon5	NM_001145011	rs368874235	0.000833	0
PPL	16	4934433	C	T	C	c.G4223A	p.R1408H	exon22	NM_002705	rs140117389	0.000417	4
PPL	16	4945680	C	T	C	c.G1010A	p.R337H	exon10	NM_002705	rs145641114	0.000000	0
ERCC4	16	14015973	A	G	C	c.A293G	p.Y98C	exon2	NM_005236	NA	0.000000	5
GGA2	16	23479011	C	T	C	c.G1742A	p.R581H	exon17	NM_015044	rs185467966	0.000826	7
ATXN2L	16	28843540	C	A	C	c.C1354A	p.P452T	exon11	NM_001308230	NA	0.000413	4
RABEP2	16	28925790	C	T	C	c.G661A	p.A221T	exon5	NM_024816	NA	0.000852	6
ITGAL	16	30522254	C	T	C	c.C2420T	p.S807L	exon21	NM_001114380	rs376221007	0.000000	0
SRCAP	16	30747554	C	T	C	c.C6763T	p.R2255C	exon32	NM_006662	NA	0.000000	6
ITFG1	16	47462720	G	T	C	c.C311A	p.T104K	exon6	NM_001305002	NA	0.000000	6
ADCY7	16	50327353	G	A	C	c.G776A	p.R259H	exon5	NM_001114	NA	0.000000	3
FAM65A	16	67579353	C	G	C	c.C3118G	p.L1040V	exon18	NM_001193522	NA	0.000841	0
WWP2	16	69970237	G	T	C	c.G682T	p.G228C	exon9	NM_199424	NA	0.000413	6
ZNF821	16	71898823	C	T	C	c.G169A	p.E57K	exon3	NM_017530	NA	0.000000	0
GLG1	16	74519713	T	C	C	c.A1519G	p.I507V	exon8	NM_001145666	NA	0.000000	4
ATP2C2	16	84493147	T	A	C	c.T2369A	p.V790E	exon24	NM_001286527	NA	0.000000	2
C16orf95	16	87339486	C	A	C	c.G356T	p.G119V	exon4	NM_001195125	NA	0.000000	1
JPH3	16	87730200	C	G	C	c.C2181G	p.I727M	exon5	NM_020655	rs200827482	0.000561	4
GEMIN4	17	648141	G	A	C	c.C3142T	p.R1048C	exon2	NM_015721	NA	0.000000	4
BHLHA9	17	1173958	C	T	C	c.C101T	p.S34L	exon1	NM_001164405	NA	0.000000	1
METTL16	17	2324092	C	T	C	c.G995A	p.R332H	exon9	NM_024086	rs142808274	0.000000	2
OR1G1	17	3030757	G	A	C	c.C89T	p.S30F	exon1	NM_003555	NA	0.000413	1
CNTROB	17	7838405	A	G	C	c.A536G	p.N179S	exon4	NM_001037144	NA	0.000417	3
NCOR1	17	16004670	C	T	C	c.G2305A	p.V769I	exon17	NM_001190438	NA	0.000413	0
MMP28	17	34122353	C	T	C	c.G29A	p.R10H	exon1	NM_001032278	NA	0.000000	0
MYO19	17	34852184	G	A	C	c.C2224T	p.P742S	exon23	NM_025109	NA	0.000413	1
KLHL10	17	40003568	A	T	C	c.A1358T	p.Y453F	exon4	NM_152467	NA	0.000414	4
FAM134C	17	40733974	C	T	C	c.G1258A	p.G420R	exon9	NM_178126	NA	0.000826	0
WNK4	17	40948576	G	T	C	c.G3709T	p.A1237S	exon18	NM_032387	NA	0.000414	3
CDC27	17	45219686	A	C	C	c.T1104G	p.I368M	exon10	NM_001293091	rs78080061	0.000000	3
CDC27	17	45219688	T	C	C	c.A1102G	p.I368V	exon10	NM_001293091	rs200712699	0.000000	4
CDC27	17	45219702	T	C	C	c.A1088G	p.N363S	exon10	NM_001293091	rs199774308	0.000000	4
CDC27	17	45232087	T	C	C	c.A725G	p.N242S	exon7	NM_001293091	NA	0.000000	3
CDC27	17	45232102	A	C	C	c.T710G	p.L237X	exon7	NM_001293091	NA	0.000000	4
CDC27	17	45232109	A	G	C	c.T703C	p.S235P	exon7	NM_001293091	rs76383230	0.000000	4
MYCBPAP	17	48585968	G	A	C	c.G62A	p.G21D	exon1	NM_032133	NA	0.000000	2
ABCC3	17	48755543	C	T	C	c.C3668T	p.P1223L	exon25	NM_003786	NA	0.000000	8
SOX9	17	70120115	C	A	C	c.C1117A	p.P373T	exon3	NM_000346	NA	0.000000	1
DNAI2	17	72308160	C	T	C	c.C1477T	p.R493W	exon12	NM_001172810	NA	0.000000	5
GPRC5C	17	72443155	C	A	C	c.C1350A	p.Y450X	exon4	NM_018653	NA	0.000000	0
CD300LF	17	72699273	G	A	C	c.C392T	p.P131L	exon3	NM_001289084	NA	0.000000	0
DNAH17	17	76567792	G	C	C	c.C612G	p.I204M	exon4	NM_173628	NA	0.000867	0
CEP295NL	17	76887951	C	T	C	c.G635A	p.R212Q	exon3	NM_001243540	NA	0.000000	0
CCDC40	17	78061568	C	T	C	c.C2612T	p.S871L	exon15	NM_001243342	rs367596393	0.000494	1
CEP131	17	79173297	C	A	C	c.G1076T	p.R359L	exon10	NM_001009811	NA	0.000000	0
SLC38A10	17	79220401	G	A	C	c.C2315T	p.T772M	exon16	NM_001037984	rs374394310	0.000000	0
CHST9	18	24497189	A	C	C	c.T366G	p.S122R	exon6	NM_031422	NA	0.000000	0
RIT2	18	40323564	T	C	C	c.A548G	p.K183R	exon5	NM_002930	NA	0.000000	1
SALL3	18	76753088	C	T	C	c.C1097T	p.T366I	exon2	NM_171999	NA	0.000858	2
SBNO2	19	1123586	T	G	C	c.A404C	p.E135A	exon4	NM_001100122	NA	0.000000	2
GAMT	19	1399827	G	A	C	c.C292T	p.R98W	exon2	NM_000156	NA	0.000000	5
PTPRS	19	5214604	C	T	C	c.G3121A	p.V1041I	exon20	NM_130853	rs116515629	0.000827	2
ANGPTL4	19	8436337	G	A	C	c.G856A	p.V286I	exon5	NM_001039667	rs200918932	0.000000	1
MRPL4	19	10369185	C	G	C	c.C649G	p.L217V	exon7	NM_015956	NA	0.000000	6
C19orf44	19	16620790	T	C	C	c.T1630C	p.W544R	exon5	NM_001288834	NA	0.000000	4
FKBP8	19	18652762	G	A	C	c.C19T	p.P7S	exon2	NM_001308373	NA	0.000000	2

medRxiv preprint doi: <https://doi.org/10.1101/2023.08.24.23294455>; this version posted August 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Gene	Chromosome	Position (kb)	Reference Allele	Variant Allele	Variant Type	Gene	Transcript	Exon	RefSeq ID	Variant ID	Frequency
GRAMD1A	19	35814392	A	G	C	c.A454C	p.S64G	exon17	NM_01136199	rs190168534	0.000415
FFAR2	19	35941491	G	T	C	c.G4362T	p.E1454D	exon32	NM_001083961	NA	0.000000
WDR62	19	36595720	G	T	C	c.G379T	p.V127L	exon4	NM_001243116	NA	0.000000
NFKB1B	19	39397878	G	T	C	c.G379T	p.V127L	exon4	NM_001243116	NA	0.000000
SARS2	19	39406709	T	A	C	c.A1315T	p.T439S	exon14	NM_017827	rs200202461	0.000416
PAK4	19	39664500	C	A	C	c.C489A	p.D163E	exon3	NM_001014834	NA	0.000000
CEACAM5	19	42219668	A	G	C	c.A803G	p.Q268R	exon4	NM_001291484	NA	0.000413
SIX5	19	46269328	C	T	C	c.G1651A	p.V551M	exon3	NM_175875	NA	0.000000
CCDC8	19	46915205	C	T	C	c.G863A	p.G288E	exon1	NM_032040	rs201827579	0.000000
CRX	19	48342884	C	T	C	c.C560T	p.T187I	exon4	NM_000554	NA	0.000414
HAS1	19	52217119	C	G	C	c.G1295C	p.W432S	exon5	NM_001297436	NA	0.000855
ZNF468	19	53348534	A	G	C	c.T158C	p.I53T	exon4	NM_001277120	rs73069429	0.000000
LAIR1	19	54866948	G	A	C	c.C739T	p.R247W	exon9	NM_001289023	rs200218727	0.000000
RDH13	19	55568152	C	T	C	c.G209A	p.R70Q	exon3	NM_001145971	NA	0.000000
ZNF628	19	55994880	A	G	C	c.G2320A	p.V774I	exon3	NM_033113	NA	0.000426
ZNF154	19	58213953	C	T	C	c.G364A	p.G122S	exon3	NM_001085384	rs202082426	0.000000
STK35	20	2082672	G	A	C	c.G145A	p.A49T	exon1	NM_080836	NA	0.000618
SIGLEC1	20	3687141	C	A	C	c.G262T	p.E88X	exon2	NM_023068	rs150358287	0.000826
HSPA12B	20	3729935	G	A	C	c.G904A	p.V302I	exon9	NM_001197327	NA	0.000000
SLX4IP	20	10579389	C	G	C	c.C305G	p.T102R	exon5	NM_001009608	rs199963485	0.000568
RRBP1	20	17640746	G	T	C	c.C407A	p.P136H	exon2	NM_004587	rs200124781	0.000000
BCL2L1	20	30309628	G	C	C	c.C394G	p.R132G	exon2	NM_138578	NA	0.000000
NECAB3	20	32247379	C	T	C	c.G803A	p.G268D	exon8	NM_031232	NA	0.000836
MYH7B	20	33583163	G	A	C	c.G2851A	p.E951K	exon28	NM_020884	rs370896124	0.000535
BPI	20	36938936	A	G	C	c.A430G	p.I144V	exon4	NM_001725	rs144469725	0.000826
SNX21	20	44462562	C	T	C	c.C4T	p.H2Y	exon1	NM_001042632	NA	0.000000
SULF2	20	46305861	T	A	C	c.A1211T	p.K404M	exon9	NM_001161841	NA	0.000000
SPATA2	20	48524990	T	C	C	c.A38G	p.D13G	exon2	NM_001135773	NA	0.000413
NRIP1	21	16338385	C	T	C	c.G2129A	p.R710H	exon4	NM_003489	NA	0.000000
PFKL	21	45725785	C	T	C	c.C202T	p.P68S	exon4	NM_001002021	rs138883373	0.000000
TRPM2	21	45821691	C	T	C	c.C2449T	p.L817F	exon16	NM_003307	NA	0.000413
C21orf58	21	47734824	C	T	C	c.G97A	p.V33M	exon4	NM_001286463	rs75798864	0.000000
IL17RA	22	17584415	A	G	C	c.A794G	p.N265S	exon8	NM_001289905	rs201250724	0.000000
CECR2	22	17983890	C	G	C	c.C223G	p.P75A	exon7	NM_001290046	NA	0.000413
SLC25A18	22	18070827	G	A	C	c.G712A	p.A238T	exon8	NM_001303484	NA	0.000000
CABIN1	22	24561553	A	T	C	c.A4816T	p.T1606S	exon30	NM_001201429	NA	0.000000
TFIP11	22	26894893	C	G	C	c.G1378C	p.E460Q	exon10	NM_012143	NA	0.000000
RHBDD3	22	29656346	A	G	C	c.T952C	p.W318R	exon6	NM_012265	NA	0.000000
SLC5A4	22	32650193	A	G	C	c.T143C	p.L48P	exon2	NM_014227	NA	0.000000
MYH9	22	36688118	G	C	C	c.C4258G	p.Q1420E	exon31	NM_002473	rs200510675	0.000000
IL2RB	22	37524507	G	A	C	c.C1285T	p.L429F	exon10	NM_000878	NA	0.000434
SLC16A8	22	38477353	T	C	C	c.A692G	p.E231G	exon4	NM_013356	NA	0.000480
BAIAP2L2	22	38484968	G	A	C	c.C905T	p.S302L	exon10	NM_025045	NA	0.000430
GTSE1	22	46712073	G	A	C	c.G1196A	p.R399Q	exon7	NM_016426	NA	0.000000
CELSR1	22	46931599	T	C	C	c.A1469G	p.Q490R	exon1	NM_014246	rs200630932	0.000000
OFD1	X	13779315	C	T	C	c.C2372T	p.P791L	exon17	NM_003611	NA	0.000826
DMD	X	31697509	T	C	C	c.A475G	p.K159E	exon10	NM_004013	NA	0.000000
CXorf67	X	51151176	C	G	C	c.C1308G	p.N436K	exon1	NM_203407	NA	0.000000
WDR44	X	117532396	C	A	C	c.C1162A	p.P388T	exon7	NM_001184966	NA	0.000827
DCAF12L1	X	125685729	A	T	C	c.T863A	p.L288Q	exon1	NM_178470	NA	0.000000
UBL4A	X	153713967	G	C	C	c.C385G	p.R129G	exon4	NM_014235	NA	0.000852
UBL4A	X	153713969	C	A	C	c.G383T	p.S128I	exon4	NM_014235	NA	0.000000

^aPhysical position in UCSC Genome Browser in Human Feb. 2009 Human GRCh37/hg19 Assembly.

^bReference allele in UCSC Genome Browser.

^cAllele observed in this study.

^dAllele frequency in Japanese in Human Genetic Variation Database (<https://www.hgvd.genome.med.kyoto-u.ac.jp/>).

^eNumber of methods for predicting variant as deleterious.

Supplementary Table3. *De novo* variants in each family

Gene symbol	Family	No. of predictions ^a	Chr	Position ^b	Ref ^c	Alt ^d	Substitution			SNV	Frequency		Gene description	Gene expression ^g	
							cDNA	Amino acid	Exon		Transcript ID	Japanese ^e			European ^f (non-Finnish)
<i>MYO18B</i>	P5	4	22	26177721	C	G	c.C2232G	c.C2232G	exon10	NM_032608	rs3718444616	0.000507	0.000000	Myosin XVIIB	Skeletal muscle, heart
<i>BTAF1</i>	P7	4	10	93748956	C	G	c.C2473G	p.P825A	exon20	NM_003972	rs774232595	0.000000	0.000000	BTAF1 RNA polymerase II, B-TFIIID transcription factor-associated	Lymphocyte, nerve
<i>EED</i>	P11	1	11	85981134	G	A	c.G976A	p.A326T	exon10	NM_001308007	rs1023860960	0.000000	0.000000	Embryonic ectoderm development	Lymphocyte, brain
<i>MUC3A</i>	P15	0	7	100550974	A	G	c.A1555G	p.T519A	exon2	NM_005960	NA	0.000000	0.000000	Mucin-3A	Small intestine, colon

Abbreviations: NA, not applicable

^aNumber of methods for predicting deleterious variant.

^bPhysical position in UCSC Genome Browser in Human Feb. 2009 Human GRCh37/hg19 Assembly.

^cReference allele in UCSC Genome Browser.

^dAllele observed in this study.

^eAllele frequency in Japanese in Human Genetic Variation Database (<https://www.hgvdb.genome.med.kyoto-u.ac.jp/>).

^fAllele frequency in European gnomAD (<https://gnomad.broadinstitute.org/>).

^gTop two tissues showing highest expression level in GTEx (release V8, 54 non-diseased tissue sites across nearly 1000 individuals)

Supplementary Table4. Splice site variants in each family

Gene symbol	Family	Chr	Position ^a	Ref ^b	Alt ^c	Splice region	Substitution	SNV	Frequency		Gene description	Gene expression ^f
									Japanese ^d	European ^e (non-Finnish)		
<i>ATG7</i>	P4	3	11468277	G	C	splice acceptor	AG→AC	rs756690200	0.000000	0.000008	Autophagy related 7	Fibroblast, esophagus
<i>WDR45B</i>		17	80579677	T	G	splice acceptor	AG→CG	NA	0.000000	0.000000	WD repeat domain 45B	Testis, fallopian tube
<i>HBD</i>	P11	11	5255444	C	T	splice acceptor	AG→AA	rs35408031	0.000000	0.000026	Hemoglobin, delta	Whole blood, spleen
<i>LILRB3</i>		19	54721187	C	T	splice donor	GT→AT	rs147121453	0.000000	0.000134	Leukocyte immunoglobulin-like receptor, subfamily B, member 3	Whole blood, spleen
<i>FAM216B</i>	P15	13	43361020	G	T	splice donor	GT→TT	NA	0.000413	0.000000	Family with sequence similarity 216, member B	Heart, fallopian tube

Abbreviations: NA, not applicable

^aPhysical position in UCSC Genome Browser in Human Feb. 2009 Human GRCh37/hg19 Assembly.

^bReference allele in UCSC Genome Browser.

^cAllele observed in this study.

^dAllele frequency in Japanese in Human Genetic Variation Database (<https://www.hgvd.genome.med.kyoto-u.ac.jp/>).

^eAllele frequency in European gnomAD (<https://gnomad.broadinstitute.org/>).

^fTop two tissues showing highest expression level in GTEx (release V8, 54 non-diseased tissue sites across nearly 1000 individuals)

