Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Tilghman JM, Ling AY, Turner TN, et al. Molecular genetic anatomy and risk profile of Hirschsprung's disease. N Engl J Med 2019;380:1421-32. DOI: 10.1056/NEJMoa1706594

TABLE OF CONTENTS

Section **Page**

Supplementary Methods

Author contributions.. 2

Sample ascertainment.. 2

Genotyping.. 4

Exome sequencing, variant calling and annotation... 4

CNV analysis using exome sequence data... 5

Dataset analyzed in this study... 8

Statistical analyses... 8

Gene expression studies.. 12

Morpholino studies in zebrafish... 14

Supplementary Tables

Table S1: Genes with disease-associated variants (DAV) and pathogenic alleles (PA) reported in HSCR databases. 16

Table S2: Four common non-coding variants in Hirschsprung disease................. 17

Table S3: Exome sequence variation... 18

Table S4: Exome sequence data accuracy... 18

Table S5: Sequence similarity between cases and their relatives........................... 19

Table S6: Pathogenic allele distribution in cases versus controls........................... 19

Table S7: Distribution and effect of case-observed PAs by pathway....................... 22

Table S8: Identifying CNVs using exome sequence, SNP array and karyotype data 23

Table S9: Inferring the phenotypic consequences of karyotype variants and CNVs 24

Table S10: Comparison of genetic burden of classes of variation by sex................ 25

Table S11: Distribution of HSCR cases by mutation type and phenotype............... 26

Table S12: Functions of novel HSCR genes and their relevance to ENS development 27

Supplementary Figures

Figure S1: Allele frequency distribution of common variants among cases............ 28

Figure S2: Principal component analysis (PCA) of samples analyzed..................... 28

Figure S3: Sequence similarity between relatives.. 29

Figure S4: Assessment of genes significantly enriched for PAs.............................. 30

Figure S5: CNV burden in HSCR... 31

Figure S6: Gene expression of candidate HSCR genes in the embryonic human and mouse gut 32

Figure S7: Assessment of HSCR candidate genes in zebrafish................................. 33

References.. 34

SUPPLEMENTARY METHODS

Author contributions:

J.M.T, T.M.T., S.C., A.K., C.B. and A.C. designed experiments. J.M.T, C.B. and A.C. wrote and edited the initial draft of the manuscript and contributed to genotypic risk estimation. C.B. and A.C. selected HSCR families for inclusion in this study. J.M.T., A.Y.L., T.N.T., K.H.N. and C.B. contributed to calling and analysis of coding variation. N.K. performed exome-based CNV calling. A.Y.L. and T.N.T. contributed to CNV validation. M.X.S. performed zebrafish morpholino assays and analysis. S.C. performed gene expression assays and analysis. A.K. performed common risk polymorphism genotyping and analysis. B.C. performed control CNV genotyping. B.C. and E.E.E. contributed to functional classification and analysis of large CNVs. N.G. and S.G. performed genomics data generation. All authors had the opportunity to comment on and approve of the final manuscript.

Sample ascertainment:

Affected individuals were selected from our collection of 636 families comprising their phenotypes, medical, pathologic and family history, and a blood/cell line/DNA sample. Affected persons were classified by segment length of aganglionosis into three groups: short-segment (S-HSCR: aganglionosis up to the upper splenic flexure), long-segment (L-HSCR: aganglionosis beyond the splenic flexure) and total colonic aganglionosis (TCA). In addition, they were also classified by gender, familiality (positive family history) and occurrence of anomalies other than aganglionosis. We chose a sample of 304 HSCR cases for exome sequencing based on DNA availability and consent for genome studies. For sequence analyses, after data cleaning and quality control, we retained 190 independent, unrelated affecteds and their 47 affected relatives (data version 1.3); for this study, we did not use data on 35 individuals from a genetically isolated Old Order Mennonite population,[1] 5 samples with poor sequence quality, 24 admixed individuals and 3 individuals whose genetic relationships could not be verified against their pedigrees. The 190 independent unrelated individuals, whom we designate as ‘probands,’ were most often the actual proband but rarely an affected first degree relative with more complete data. The included individuals self-identified as being of European ancestry, which was checked for consistency with their genotype data (**Supplementary Figure 2**). The case sample was composed of: (1) 122 (64%) males and 68 (36%) females; (2) 82 (43%) S-HSCR, 67 (35%) L-HSCR/TCA and 41 (22%) unknown (unspecified) segment length; (3) 125 (66%) simplex and 65 (34%) multiplex families (24 sibs, 20 parent-child, 21 greater than first-degree); (4) 130 (68%) non-syndromic, 6 (3%) single gene syndromes (3 Central Congenital Hypoventilation syndrome (CCHS) and 1 each of Waardenburg (WS), L1CAM (L1CAMS) and Bardet-Biedl (BBS) syndromes), 17 (9%) chromosomal variants (11 with Down syndrome and 1 each with 16p11.2 del, 22q11.2 del, tetrasomy 22q, 47, XX, +der(15) t(4:15), 13q21.33-q31.1 del, 10q24.3-q26.13 inv) and 37 (20%) with multiple anomalies not recognized as a specific syndrome. This sample selection had features comparable to our total collection of 636 probands, except for an oversampling of known segment length cases, and comprised: 67%, 33% male/female; 39%, 29%, 32% S-/L-&TCA/unspecified; 70%, 30% simplex/multiplex; and, 63%/37% non-syndromic/ syndromic. Finally, the sampled sibship size was 1, 2, 3 or 4 for 155, 23, 7 and 1 individual, respectively. Subject ascertainment was conducted with written informed consent approved by the Institutional Review Board of Johns Hopkins University School of Medicine.

For European ancestry controls, we used publicly available exome sequence data from 370 NIMH controls (https://www.nimhgenetics.org/available_data/controls/) and 370 EUR 1000 Genome (1000G henceforth) samples (85 Toscani in Italy, 97 Utah residents of Northern or Western European ancestry, 96 Iberians in Spain, and 92 British in England and Scotland, but excluding 101 Finns owing to possible founder effects) (www.1000genomes.org) (ref.26 in main paper), generated using the same reagents and procedures by the Broad Institute. For assessing admixture, we included all 2,302 1000G individuals with diverse ancestries.

For common non-coding variant studies, we used a different set of controls genotyped in our laboratory because some of these genotypes were not publicly available: 404 EUR 1000G samples (excluding Finns) and an additional 223 pseudo-controls, generated from the chromosomes not transmitted to the affected from 254 HSCR parent-child trios (ref.13 in main paper). The differing numbers of EUR 1000G samples used depended on when the data were accessed and the numbers of samples available at that time.

For copy number variant (CNV) analyses, we used a third control set of 19,584 adult subjects of European ancestry, as described in reference 21 in the main text. Different European ancestry controls were required to accommodate risk factors of different frequencies and the assays available in control samples.

Genotyping:

Genotype data for the polymorphisms rs2435357, rs2506030 and rs7069590 at *RET* and rs11766001 at *SEMA3C/D* were previously generated using Taqman assays in our laboratory, and have been previously reported (refs.13 and 22 in main paper). In addition, HSCR cases with large copy number variants (CNVs; see below), together with their parents where available, were validated by genotyping using the Human Omni 2.5-4v1 BeadChip, using standard methods at the Broad Institute.

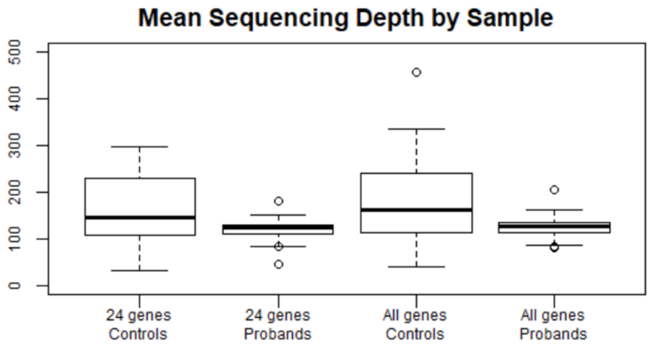
Exome sequencing, variant calling and annotation:

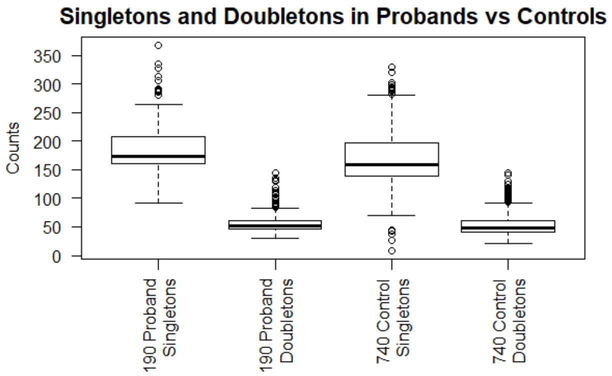
Genomic DNA was used to capture exomes using the Agilent 44Mb Sure-Select Human All Exon v2.0 capture, and sequenced using the 76 base paired-end method on an Illumina HiSeq2000 sequencer with >80% of bases at a coverage of 30X. The sequence data were aligned by the Picard (http://picard.sourceforge.net) pipeline using hg19 with the BWA algorithm and processed with the Genome Analysis Toolkit (GATK) to recalibrate base-quality scores and perform local realignment around known insertions and deletions (INDELs) (ref.17 in main paper).[2] BAM files were used to call single nucleotide variants (SNVs) and small (<50bp) insertions and deletions (INDELs) using the GATK's HaplotypeCaller algorithm. Variants were called simultaneously across all HSCR cohort members and controls (amounting to 3,176 samples) into a single VCF file. Initial filtering was done via the Variant Quality Score Recalibration (VQSR) method within GATK, which is based on detection of known variant sites. For SNVs, HapMap3.3 and Omni2.5 were used as training sites with HapMap3.3 used as the truth set. SNVs were filtered to obtain the highest confidence variant set achieving 99% truth sensitivity (1% false negative rate). For VQSR of INDELs, a set of curated INDELs obtained from the GATK resource bundle (Mills_and_1000G_gold_standard.indels.b37.vcf) were used as both a training and truth set. INDELs were filtered to obtain the highest confidence variant set achieving 91% truth sensitivity (9% false negative rate). Following initial filtering, an additional annotation was added for ancestral alleles using: ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/phase1/analysis_results/supporting/ancestral_alignments/human_ancestor_GRCh37_e59.README.

SNVs and INDELs meeting initial filtering criteria were further filtered using several hard quality filters. First, all multi-allelic sites were removed. Second, they were filtered on strand balance and homopolymer criteria (FS > 50 and HRun > 5.0 for SNVs, FS > 200 and

HRun > 10.0 for INDELs). Third, all individual genotypes with a depth < 10 were removed. Lastly, the dataset was filtered by deleting variants with >10% missing genotypes, separately for autosomes and sex chromosomes, and separately for males and females. ANNOVAR was used for annotation of variants (ref.19 in main paper). Average 46-way PhyloP conservation scores were added to each variant using internal lab scripts.

We assessed whether coding variant coverage and detection sensitivity were comparable between the 190 cases and 740 controls by summarizing sequencing coverage of coding genes targeted by our exome capture reagents (and the 24 HSCR genes specifically) and by counting singleton and doubleton variant sites per individual in each set, across all genes. As shown in the following plots, these metrics are comparable; there are no significant differences across case and control exomes, except that the variance in coverage is smaller in cases. Thus, the sensitivity of variation detection is identical between cases and controls excluding the possibility of false associations through differences in sensitivity.





CNV analysis using exome sequence data:

Mapping, CoNIFER and CNV Segmentation: Short reads from the exome sequencing experiment were split into 36bp chunks and mapped using the single-end mode of mrsFAST (up to two mismatches) to exons and 300bp flanking sequence extracted from the repeat-masked hg19 reference genome, using the target file for the Agilent SureSelect Target Enrichment capture platform. Next, we used CoNIFER and calculated RPKM values for 189,894 probes and exons derived from the target file. We set the --svd option to 12, and used default CoNIFER settings for all other options. Subsequently, the raw SVD-ZRPKM values were exported for downstream analysis. We used DNACopy and CGHCall to segment and assign probabilities to SVD-ZRPKM values. To prevent excessively strong SVD-ZRPKM signals from interfering with the models used by CGHCall to assign copy number, we clipped the signal at ±3 for each exon. Parameters for DNACopy were set to default and the alpha parameter was set to 0.01. Default options for CGHcall were used, and we allowed only “deletion” and “duplication” as called states. Using these parameters, we obtained 13,300 raw segments “deleted” or “duplicated”. These computational methods are not optimized to detect aneuploidies because the data are normalized by chromosome within each sample.

Quality Control: We excluded samples with more than 200 calls after segmentation, as such sample have extremely high false discovery rates (FDR). Four samples (HSCR274, HSCR18, HSCR385 and HSCR178) with a total of 1,939 calls were excluded, leaving 11,361 calls.

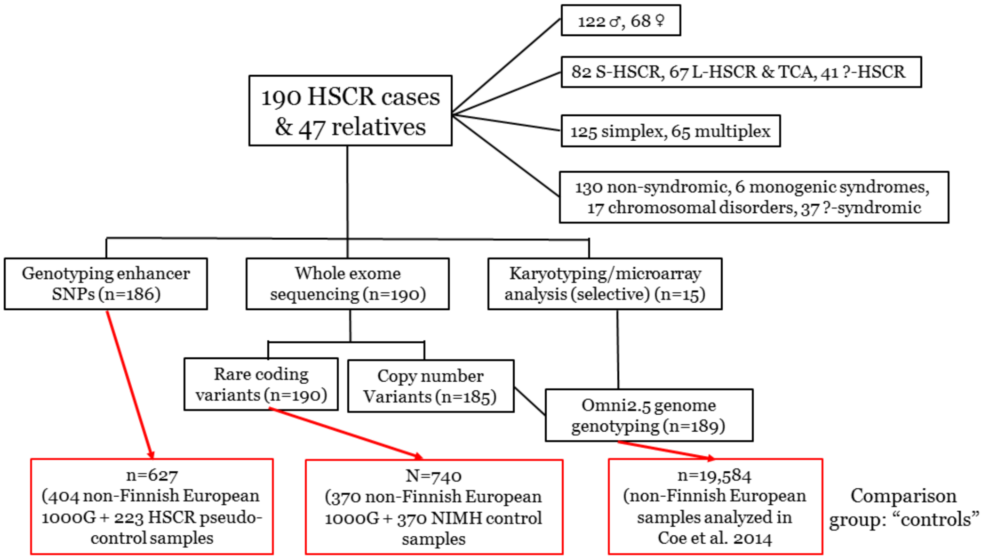
CNVR generation and call filtering: We clustered calls using a custom hierarchical clustering method which uses the pairwise reciprocal overlap (RO) between calls as a measure of distance. To prevent merging of large unique calls, the RO function was modified by a gamma tuning parameter, which weights the RO based on the total number of non-overlapping probes on each end. In this way, the function accounts for the uncertainty in breakpoints and RO for two small CNVs, while allowing two large overlapping CNVs to be counted as distinct entities. Calls were merged using hierarchical clustering (WPGMA, weighted pair group method with averaging), and we flattened the resulting trees to form CNVR clusters. Using this method, we generated 3,129 CNVRs.

Filtering segmentally duplicated regions and processed pseudogenes: We excluded CNVs which were found to have more than 50% of their probes within segmental duplications or duplicated regions of the genome (defined using previous methods from 1000 Genomes whole-genome depth-of-coverage analysis, where >80% of 34 unrelated genomes had a copy number three or greater in 500bp repeat-masked windows across the genome). Excluding calls which overlapped at least 50% with these regions resulted in the exclusion of 5,079 calls (45% of all calls), corresponding to 525 CNVRs. Next, we excluded calls which were likely to be due solely to the insertion of processed pseudogenes. CoNIFER, and most exome-based read-depth methods, are sensitive to copy number changes, specifically of exons, which can be the result of retro-insertion of processed mRNA transcripts (see reference 18 in main text for more details). We used two lists of commonly polymorphic processed pseudogenes generated using SPLIT-READ from 225 autism trios (data not reported here).[5] We excluded calls from our call list for which ≥90% of the probes corresponded to a gene which had been observed at least twice in 225 trios. This excluded a total of 1,063 of 10,927 calls. In sum, our processed pseudo-genes, segmental duplications and other duplicated portions of the genome accounted for 5,808 calls in 673 CNVRs.

Final filtering and call set generation: Our final set of calls and CNVRs was created by requiring at least one call in a CNVR with the following attributes: 1) an absolute median SVD-ZRPKM score (i.e., signal strength) of ≥ 1.5, 2) a CGHCall posterior probability of 0.95 or greater, and 3) passing additional filters for duplicated genes and regions as described above. Our final high-quality set of CNVRs contains 1,597 calls in 554 CNVRs. For the current study, we restricted attention to only 111 rare large CNVs, defined as deletions >500kb and duplications >1mb, and potentially with phenotypic impact (as assessed using external databases), but we also included all changes detected by karyotype or FISH (fluorescent *in situ* hybridization) for clinical diagnosis, loci for known genomic disorders and HSCR genes. All these cases were further validated by SNP genotyping using the Human Omni 2.5-4v1 BeadChip array (see **Genotyping**).

Validating CNV calls: Omni2.5 Beadchip genotype data were processed using a standard Illumina pipeline; 4 samples failed the QC process. We manually examined the data to confirm each CNV by plotting B allele frequency and the LogR Ratio (LRR) for each from the gtc file. For each chromosomal position, we ignored samples if either the B allele frequency or LogR Ratio was missing, if the GCScore ≤ 0.15, and if multiple discordant genotype calls were made. We also noted the following: annotated genes in the region, and exome sequencing coverage. We validated 31 cases, as shown in **Table 3** and **Supplementary Table S8**, 22 cases of 8 unique recurrent CNVs and 9 non-recurrent CNVs. Of the 14 CNVs not evident from karyotypes, 13 had the median SVD-ZRPKM cutoff >1.5, and 1 between 1 and 1.5. Of the 17 validated CNVs in 31 cases, we could determine parental origin in 8 cases: 4 were *de novo* (del13, +21, 1q21.1 del, inv 10), 2 were inherited from the father (17p11.2 (*CMT1A*) dup, dup7), 1 was inherited from the mother (t4;15); 1 was not inherited from the father but maternal origin could not be assessed since her sample failed QC (+21). We also identified an additional 4 kb *RET* deletion (chr10:42917793-42922026) in patient HSCR472 which was separately validated by qPCR analysis; note that the chromosome 13q21.33-q31.1 deletion included *EDNRB*.

Dataset analyzed in this study:


190 HSCR cases & 47 relatives
122 ♂, 68 ♀
82 S-HSCR, 67 L-HSCR & TCA, 41 ?-HSCR
125 simplex, 65 multiplex
130 non-syndromic, 6 monogenic syndromes, 17 chromosomal disorders, 37 ?-syndromic
Genotyping enhancer SNPs (n=186)
Whole exome sequencing (n=190)
Karyotyping/microarray analysis (selective) (n=15)
Rare coding variants (n=190)
Copy number Variants (n=185)
Omni2.5 genome genotyping (n=189)
n=627 (404 non-Finnish European 1000G + 223 HSCR pseudo-control samples
N=740 (370 non-Finnish European 1000G + 370 NIMH control samples
n=19,584 (non-Finnish European samples analyzed in Coe et al. 2014
Comparison group: "controls"

We present above an overview of all case and various comparison group samples analyzed, their sample sizes and the types of genetic analyses conducted on each.

Statistical Analyses:

Principal component analysis (PCA): To assess population structure and potential cases of admixture, we conducted standard PCA analysis using R package SNPRelate[4] on all 301 HSCR cases and 2,672 controls (370 NIMH samples and 2,302 1000G) to identify 29 highly-admixed HSCR individuals (potentially African- and Asian- Americans). We used genotypes for 7,536 autosomal SNPs from exome sequencing that had allele frequencies ≥10%, missingness <5%, LD trimmed using an r^2 threshold of 0.02. In **Supplementary Figure 2** we show European ancestry HSCR probands and all 2,672 PCA controls plotted along the first three PCs, followed by the first three PC's of a Europeans-only PCA, showing that the only European ancestry group from whom HSCR probands can be discriminated is Finns.

Sequence similarity between relatives: The exome sequence data were used to assess the overall genetic relationship between each case and his/her relative. Our sample included 42 relatives of 190 probands yielding 53 relative pairs from 32 families. We used the exome sequences of each pair to compute a similarity statistic S:

$$S = \frac{n_{xy}\left(1/n_x + 1/n_y\right)}{2}$$

where, n_x, n_y and n_{xy} refer to the number of distinct alleles at variant sites in individual x, in individual y, and shared by x and y, respectively, at a variant site and is summed across all variant sites.[6] S = 0 whenever n_x, n_y or n_{xy} is zero. S is the proportion of shared sites relative to the harmonic mean of the number of variants in the pair compared. The coefficient of relationship (r) is then estimated as:

$$r = (\bar{S} - U)/(1 - U),$$

where $\bar{S}$ is the average S across all variants, estimated by summing the numerator and denominator in the above formula, and U is the similarity statistic from unrelated individuals, estimated from all possible pairings of the 190 cases.[6]

To assess whether susceptibility variants were enriched in affected relatives of probands, we estimated S for the 24 HSCR genes, based on pathogenic alleles. We estimated the mean S for all such relative pairs and compared it to its expectation by obtaining its empirical distribution from 5,000 means based on 24 randomly selected genes for each relative pair, restricting analysis to only those gene sets with at least one pathogenic allele in the proband. This distribution was used to calculate a one-sided test of excess sharing in these relatives. Across all relative pairs, $\bar{S}$ = 0.75 and was significantly greater than the mean permuted value of 0.45 (P = 0.0054).

Discovery of genes enriched for rare coding single nucleotide variants (SNV): We compared rare pathogenic SNVs, defined as coding alleles that are nonsense, highly conserved missense (PhyloP score ≥ 4) or changes that alter the canonical ±2bp splice junctions, with frequency ≤ 5% in cases and controls for genes in which at least one pathogenic SNV was observed in controls. For analysis, we used the observed number (d) of distinct pathogenic SNVs among the 190 cases (d_o), motivated by the known distribution of allele multiplicity for alleles of a defined selection coefficient.[7] To assess whether the observed value was higher than expected we used 740 European ancestry NIMH and 1000G controls to randomly sample 190 individuals and calculate d for each replicate. Repeating this sampling

10,000 times, with replacement, provided an estimate of the distribution of the random variable d with mean $\bar{d}$. We estimated the significance value (α) of the hypothesis of no gene effect as α = Prob $\{d \geq d_o | \bar{d}\}$ and by assuming d is Poisson distributed, an assumption that was tested from the empirical distribution of d across the replicates. This assumption was conservative since the observed variance of the distribution was smaller than the average (**Supplementary Figure 4**). Note that these are gene-specific estimates and so no corrections for gene size or sequence features are necessary, although the statistical power of detecting departures in individual genes decreases with a gene's increasing intrinsic rate of pathogenic variation in controls. This empirical probability distribution, contrasted to the expected distribution, for all human genes was used for testing whether there is an excess of pathogenic variants in specific genes (see QQ plot in **Supplementary Figure 4**).

Discovery of copy number variants (CNV): CNV burden was compared between cases and controls for rare CNVs (prevalence <1%) using CNV length, excluding gaps and regions annotated as segmental duplications (hg18). The 19,584 controls, described in reference 21 in main text, were obtained by combining 8,329 controls from Cooper *et al.* (dbVar study accession nsdt54) with 11,255 new controls profiled on Affymetrix SNP6 arrays from the Wellcome Trust Case Control Consortium 2 (WTCCC2) 58C cohort, as well as the ARIC (Atherosclerosis Risk in Communities) Cohort (database of Genotypes and Phenotypes, dbGaP accession phs000090.v1.p1) (Supplementary Table 1 in main text reference 21). The details of CNV calling in controls are described there. CNV calls that falsely extended across centromeric gaps due to small polymorphisms on both arms were trimmed. These CNVs are shown in Supplementary Figure 1 of reference 21 in main text. Burden was defined using only the largest CNV to account for the large number of bases encompassed by small CNVs and the difference in resolution between cases (exome sequence) and controls (SNP arrays). The overall incidence of rare deletions and duplications among these 19,584 controls was 0.020 (391 instances) and 0.014 (282 instances), respectively. These controls did not include individuals with intellectual disability and so this estimate was supplemented by the prevalence of Down syndrome in the population (8.27/10,000 individuals),[8] and its value imputed for controls of equivalent size (i.e., 19,584).

Quantifying pathogenic allele (PA) enrichment: We tested for enrichment of PAs by class in individuals of European ancestry. (1) Common variants were allele, haplotype and genotype counted in 186 cases and 627 controls with tests conducted using contingency chi-square methods with significance calculated using a 2-sided Fisher's exact probability. Frequency differences were represented as corrected odds ratios (ref. 13 in main paper) with variances and tests of significance as estimated using the Haldane bias correction.[9] (2) Rare coding variants were compared using exome sequence data in 190 cases and 740 1000G and NIMH controls. For overall enrichment, the PA definition for rare coding variants was extended to include all INDELs overlapping the coding sequence and restricted to only PAs identified in cases. For these alleles, we report their allele frequencies in ancestry-matched, non-neuropsychiatric samples from ExAC,[10] comprising a much larger sample size of 21,071 subjects. Frequency differences were represented as corrected (owing to small numbers in some cells) odds ratios (ref. 13 in main paper) with variances and tests of significance as estimated using the Haldane bias correction.[9] (3) Each recurrent and non-recurrent CNV was compared against its frequency in a sample of 19,584 adult subjects of European ancestry. Tests of enrichment used a 2-sided Fisher's exact probability. Note that these controls were all ascertained as adults and therefore were depleted for high penetrance CNVs such as trisomy 21, which we corrected for individually. The other CNVs detected are not *a priori* known to lead to high penetrance phenotypes.

Estimating disease penetrance: Phenotype penetrance refers to the probability of phenotypic expression for specific genotypes and, as such, are marginal effects averaging across the phenotypic effects (if any) across all other genes (genetic background). As such, this penetrance is also the disease incidence given that genotype. The incidence is the frequency (rate) of new cases which may manifest or be recognized at different ages. However, for HSCR these are nearly identical because most cases are recognized and treated in the first years of their life. Consequently, genotype-dependent penetrance can be calculated as follows:

$$P\{D|G\} = P\{G|D\}P\{D\}/P\{G\},$$

where G, D and D^c are genotype, phenotype (disease) and the phenotype complement, respectively, and P {.} is probability. If we examined n cases and m controls, and P{G} was the population frequency of a specific genotype class (either at one locus or at many) then it could be estimated from the control data, while P{G|D} could be estimated from the genotype distribution among cases. P{D} is the disease incidence and for HSCR is set to 15/100,000 live births. The penetrance estimated from the above equation has the expected standard deviation of:

$$P\{D|G\}\sqrt{\frac{1-P\{G|D\}}{nP\{G|D\}}+\frac{1}{mP\{G\}}},$$

which is estimated by replacing all expected values by their observed quantities.

Estimating population attributable risk: Population attributable risk (PAR) is the proportion of disease in a population involving a given exposure, which is useful in this study for comparing the relative contributions of each risk factor to the development of HSCR. In order to estimate PAR, one requires an estimate of the relative risk (RR) of each risk factor as well as the proportion of the general population exposed. When disease incidence is low, as is the case with HSCR, RR ~ OR. Therefore, PAR can be estimated for HSCR risk factors as follows:

$$PAR = \frac{P_e\ (OR-1)}{P_e\ (OR-1)+1}$$

where P_e is the proportion of the general population exposed to that risk factor derived from one of our three control populations.

Gene expression studies:

Taqman gene expression assays of human and mouse gut tissue: We studied 8 human fetal guts at Carnegie Stage 22 (CS22) obtained from the Human Developmental Biology Resource (www.hdbr.org; grant 099175/Z/12/Z). All HDBR samples were collected according to local Research Ethics Committee review by the NHS Health Research Authority National Research Ethics Service and in line with the ethical guidelines laid out in the Polkinghorne Report (Review of the Guidance on the Research Use of Fetuses and Fetal Material, 1989). The HDBR is also licensed as a tissue bank by the Human Tissue Authority. The samples were approved for use in this study by the Institutional Review Board of Johns Hopkins University School of Medicine. All mouse guts used were from three E10.5 wild type C57BL/6J male mice purchased from The Jackson Laboratory. Total RNA was extracted from these tissues using TRIzol (Life Technologies, USA) and cleaned on RNeasy columns (Qiagen, USA). 500ng of total RNA was converted to cDNA using SuperScriptIII reverse transcriptase (Life Technologies, USA) and Oligo-dT primers. The diluted (1/5) total cDNA was subjected to Taqman gene expression (Life Technologies, USA) using transcript-specific probes and primers. Human or mouse β-actin was used as an internal loading control to normalize data. Each sample was assayed 3 times and the data presented are means with their standard errors. The relative fold-change was calculated based on the $2^{\Delta\Delta Ct}$ (threshold cycle) method, with the highest expressing transcript (lowest Ct value) set to unity. Any gene with Ct value >30 was considered not expressed. Only one potential gene- *MMAA* had a Ct value >30 in both mouse and human. The following Taqman probes were used from Applied Biosystems: For human: *RET* (Hs01120032_m1), *EDNRB (*Hs00240747_m1*), ADAMTS17 (*Hs00330236_m1*), ACSS2* (Hs01120914_m1*), SLC27A4 (*Hs00192700_m1*), SH3PXD2A (*Hs01046313_m1*), MMAA (*Hs00604098_m1*), ENO3 (*Hs01093275_m1*), FAM213A (*Hs00800009_s1*)* and *UBR4 (*Hs00390223_m1*)*. For mouse: *Ret* (Mm00436305_m1), *Ednrb* (Mm00432989_m1), *Adamts17* (Mm01318914_m1), *Acss2* (Mm00480101_m1), *Slc27a4* (Mm01327405_m1), *Sh3pxd2a* (Mm01205065_m1), *Mmaa* (Mm04209905_m1), *Eno3* (Mm00468267_m1), *Fam213a* (Mm00510430_m1) and *Ubr4* (Mm01348737_m1).

RNA-seq gene expression assays of mouse gut tissue: Total RNA was extracted from 3 male mouse guts at E10.5. cDNA was prepared by oligo dT beads to select mRNA from the total RNA sample followed by heat fragmentation and cDNA synthesis from the RNA template as part of the Illumina Tru Seq™ RNA Sample Preparation protocol. The resultant cDNA was used for library preparation (end repair, base 'A' addition, adapter ligation, and enrichment) using standard Illumina protocols. Libraries were sequenced on a HiSeq 2000 using manufacturer's protocols to a depth of 15 million reads per samples (75 base pair, paired end). The primary data were analyzed using the Broad Institute's Picard pipeline, which includes de-multiplexing, and data aggregation. The resultant BAM files were mapped to the mouse genome (assembly mm10/GRCm38) using *TopHat* with its setting for paired end, non-strand specific library.[10] Successfully mapped reads were used to assemble transcripts and estimate their abundances using *Cufflinks*.[11] The resulting data assigned Fragments Per Kilobase of Transcript per Million mapped reads (FPKM) values for each transcript and gene. To further assign which genes were "expressed" in the gut, we did qPCR analysis of multiple genes with FPKM ranging from 1-10. Since we did not always detect expression of genes with FPKM < 5, we set FPKM of 5 as the threshold for genes to be considered gut expressed. All data have been deposited in NCBI's GEO and are accessible at accession number GSE99232**.**

Morpholino studies in zebrafish:

Zebrafish Maintenance and embryo collection: Zebrafish (AB strain) were raised and maintained under standard conditions. All animal research was approved by the Institutional Animal Care and Use Committee at Johns Hopkins University. Embryos were collected and staged as described previously.[12,13]

Morpholino microinjections: Translation blocking morpholinos (MO) were designed against each zebrafish ortholog to the human gene and ordered from Gene Tools, LLC along with a standard negative control morpholino; the sequences are provided below. All genes had a single zebrafish ortholog except *EDNRB* for which both zebrafish orthologs were tested.

Gene	***Transcript id***	***Morpholino sequence***
Control	-	CCTCTTACCTCAGTTACAATTTATA
ret	NM_181662	ACACGATTCCCCGCGTACTTCCCAT
ednrba	NM_131197	GGAAACGCATGACTATTTAACAGTC
ednrbb	XM_683473.5	GCAGCAGAATGACCGATGATGCCAT
ubr4	XM_005162190	CTCCATCTCCTCCACTCGACGCCAT
eno3	NM_214723	GCGTGAATCTTACTAATGGACATCC
mmaa	NM_001105112	AAAACTCTAGATGGACGCATCTTTC
sh3pxd2aa	NM_001160022	TTGGGAACTTGTCGAGTATCTGCAT
slc27a4	NM_001017737	TGGCACACGCCAACCGCAACATCCT
acss2	NM_001002641	CAATCAGAGAGTGCCAACACATATC
fam213aa	NM_001193525	CAAGGCCAAGTGACCACATGCCCAT

Injections were performed on 1-2-cell zebrafish embryos (n=50) independently on at least 2 different days. Survival of uninjected, negative control and transcript-specific morpholino-injected embryos were recorded to assess the effect of the transcript-specific morpholinos on survival. Different concentrations were injected for each MO to determine the optimal concentration at which a phenotype was detected. Only two concentrations are being reported for each MO for simplicity; the lowest concentration at which an effect, if any, is seen and the highest concentration before the morpholino is lethal to the embryo.

Immunostaining and visualization: Injected zebrafish embryos were fixed at 6 dpf (days post-fertilization) with 4% paraformaldehyde (PFA). Monoclonal anti-HuC antibody (Invitrogen #A-21271.) followed by Alexa Fluor 568 F (ab')$_2$ fragment of goat anti-mouse IgG secondary antibody (Invitrogen #A11019) were used for fluorescent labeling of enteric neurons as previously described, with a mild modification (see Reference 12 in main text).[14] The embryos were bleached after fixing in 4% PFA by incubating in 3% H_2O_2/0.5% KOH medium until there was a complete loss of epidermal pigmentation (~30-45 min), followed by a 5 min wash with PBS to stop the bleaching reaction. Stained embryos were visualized using a Nikon SMZ 1500 fluorescent microscope using a DS red filter to assess the colonization of enteric neurons in the gut of each embryo.

Cell counting: Stained neurons were counted using the Image-based Tool for Counting Nuclei (ITCN) plugin in ImageJ visualization software,[15] with the following parameters: width 9 pixels, minimum distance 4.5 pixels, threshold of 1 and using a selected Region of Interest (ROI). Since the enteric neurons are mostly lost caudally in the gut in well-established HSCR models in zebrafish, we chose our region of interest as 8 somites starting at the caudal end of the gut and going rostral. 15 embryos were used for cell counting for each concentration of morpholino for each gene; 20 embryos were counted for controls.

SUPPLEMENTARY TABLES

Supplementary Table S1: *Genes with disease-associated variants (DAV) and pathogenic alleles (PA) reported in HSCR mutation databases.*

Gene	*Locus*	*Syndrome*	*# DAVs*[a]	*# PAs*[b]	*# (%) of null PAs*[c]
PHOX2B[1,5]	4p13	Central Congenital Hypoventilation (CCHS)	29	26	22 (84.6%)
SOX10[1,5]	22q13.1	Waardenburg, type 4 (WS4)	38	36	33 (91.7%)
TCF4[1,5]	18q21.2	Pitt Hopkins (PHS)	49	49	32 (65.3%)
ZEB2[1,5]	2q22.3	Mowat Wilson (MWS)	150	144	135 (93.8%)
GDNF[2]	5p13.2	-	5	0	0
NRTN[2]	19p13.3	-	2	0	0
GFRA1[2]	10q25.3	-	2	1	0
RET[2]	10q11.21	-	132	77	29 (37.7%)
ECE1[3]	1p36.12	-	1	0	0
EDN3[3,5]	20q13.32	Waardenburg, type 4 (WS4)	15	6	4 (66.7%)
EDNRB[3,5]	13q22.3	Waardenburg, type 4 (WS4)	42	36	12 (33.3%)
SEMA3C[4]	7q21.11	-	2	2	0
SEMA3D[4]	7q21.11	-	3	2	0
KIF1BP[5] *(KIAA1279)*	10q22.1	Goldberg Shprintzen (GOSHS)	4	4	3 (75.0%)
L1CAM[5]	Xq28	L1CAM (L1S)	10	8	1 (12.5%)
IKBKAP[5]	9q31.3	Riley-Day (RDS)	2	2	0
NRG1	8p12	-	3	2	1 (50.0%)
Total	-	-	489	395 (80.8%)	263 (66.6%)
Mean allele frequency	-	-	5.53×10^{-4}	3.61×10^{-5}	9.05×10^{-8}

[1] Transcription factors: *PHOX2B, SOX10, TCF4, ZEB2*; [2] RET pathway: *GDNF, NRTN, GFRA1, RET*; [3] EDNRB pathway: *ECE1, EDN3, EDNRB*; [4] SEMA3 pathway: *SEMA3C, SEMA3D*; [5]Single gene syndromes: *PHOX2B, SOX10, TCF4, ZEB2, EDN3, EDNRB, KIF1BP, L1CAM, IKBKAP;* [a] DAV: disease-associated variant as reported in *HGMD* (ref. 16) and *ClinVar* (ref. 17); [b] PA: pathogenic alleles as defined in Supplementary Methods; [c] Null: nonsense alleles and frame-shifting INDELs. Note that alleles with multiple functional classifications were classified with the following order of priority: nonsense, splice junction, coding INDEL and conserved missense. The mean allele frequency was estimated from non-Finnish European ancestry subjects from the ExAC database; only individuals without a neuro-psychiatric disorder were included.[10]

HGMD and ClinVar reported 489 DAVs for HSCR, but our criteria for identifying a PA would have identified a smaller set of 395 (80.8%) alleles. These databases do not specify why most alleles are considered pathogenic. Note that the average allele frequency of these PAs is ~15X smaller than the corresponding DAVs suggesting a greater deleterious effect (penetrance). However, these PAs are a biased set since 66.6% of them are null alleles which are easier to recognize as pathogenic and are, therefore, preferentially reported in the literature: 96 (24.3%) missense, 95 (24.1%) nonsense, 27 (6.8%) splice junction, (42.5%) frame-shifting INDELs, and (2.3%) non-frame-shifting INDELs. As expected, the null alleles are extremely rare and at ~400X lower frequency than all PAs, demonstrating an even greater deleterious effect or higher penetrance. Determining causality for missense variants is much more difficult and requires statistical analysis of enrichment using controls or functional studies or both. Note the wide variation in reported DAVs and PAs across the HSCR genes, including that of null alleles, indicating differential allelic effects across genes. Consequently, the reliance on null alleles only for gene discovery and reporting creates an extreme bias in HSCR, and other genetic studies, for gene identification. Unbiased studies of PAs in different genes require appropriate control data on those same alleles. The overall PA detection rate is not possible to estimate from these data since the total numbers of patients screened were not reported. In contrast, we can estimate the maximum ‘false’ positive rate of PA detection *under our criteria* at ~13.2% since 52 such PAs were identified in 98 of 740 NIMH and 1000G controls (**Table S6**), in whom knowledge regarding HSCR family history is absent. This is an upper estimate since true causal variants have low penetrance and are expected to appear at low rates in controls.

Supplementary Table S2: *Four common non-coding variants in Hirschsprung disease.*

190 cases and 627 (404 Non-Finnish EUR 1000G + 223 HSCR pseudo-controls from 254 trios) controls were genotyped for rs2435357, rs2506030 and rs11766001; rs7069590 had genotypes for 186 HSCR samples. The disease associations of these variants have been previously published (references 12, 13, 22 in main text); their properties in the sample studied here are shown below. Unsurprisingly, the odds ratios have the same magnitudes as reported earlier in larger samples, providing confidence that the sampled cases studied here are representative of HSCR.

Table S2.1: *Common variant associations in HSCR.*

Gene	**SNP (risk/non-risk alleles)**	**Case-control samples***		
		Risk allele (case/control frequency)	**Odds ratio (95% CI)**	***P***
RET	rs2435357 (T/C)	0.59/0.23	4.8 (3.8-6.1)	6.0×10^{-40}
RET	rs7069590 (T/C)	0.84/0.74	1.8 (1.3-2.5)	9.7×10^{-5}
RET	rs2506030 (G/A)	0.54/0.40	1.7 (1.4-2.2)	2.2×10^{-6}
SEMA3	rs11766001 (C/A)	0.21/0.16	1.4 (1.1-1.9)	0.02

The genotypes of 186 HSCR cases available for all four markers were next used to count the total number of risk alleles *per individual*, a summary measure of susceptibility arising from common variants in cases and controls, as shown below.

Table S2.2: *Common variant susceptibility distribution in HSCR.*

# risk alleles	Number (%) of cases (n = 186)	Number (%) of controls (n = 627)
0	4 (2.2)	16 (2.6)
1	5 (2.7)	71 (11.3)
2	13 (7.0)	137 (21.9)
3	40 (21.5)	172 (27.4)
4	34 (18.3)	124 (19.8)
5	35 (18.8)	84 (13.4)
6	41 (22.0)	18 (2.9)
7	12 (6.5)	5 (0.8)
8	2 (1.0)	0 (0.0)

Supplementary Table S3: *Exome sequence variation.*

The data used (v1.3) represents joint calling and analysis of 301 HSCR cases and 740 controls all sequenced at the Broad Institute, Cambridge, MA. There were a total of 306,910 SNVs, 3,646 insertions and 7,900 deletions for a total of 318,456 variants. Considering coding sequences only, there were 112,489 SNVs, 844 insertions and 2,753 deletions for a total of 116,086 variants. The properties of these variants passing quality control (**Supplementary Methods**) among the 190 European ancestry cases by type, genomic location and frequency (common defined as a minor allele frequency (MAF) ≥ 10%) are as follows:

Variant type	*Coding*			*Total*		
	SNV	*INS/DEL*	*Common*	*SNV*	*INS/DEL*	*Common*
Autosomal	41,295	267/872	6,012	126,089	1,588/3,058	28,167
X-linked	743	2/9	117	2,553	16/38	579
Total	42,043	269/881	6,129	128,642	1,604/3,096	28,746

In summary, we identified 49,322 coding variants in the 190 independent, European ancestry HSCR probands of which 8,506 were pathogenic (616 nonsense, 478 splice junction variants, 924 coding INDELS and 6,488 conserved (PhyloP ≥ 4) missense). Pathogenic variants were distributed across 5,271 genes.

Supplementary Table S4: *Exome sequence data accuracy.*

For quality control (QC) of these data we compared the sequences of six duplicate HSCR samples and assessed their concordance to those of 33 duplicate pairs in the 1000G data. The case samples were compared at between 777,945 and 896,652 sites with discordance varying between 7.4 x 10^{-5} and 3.24 x 10^{-4} in contrast to a discordance of 1.07 x 10^{-4} ± 4 x 10^{-5} in controls. Therefore, the average discordancy rate is 1.57x 10^{-4}.

Sample	*# sites*	*# missing*	*# concordant*	*# discordant*	*discordance rate*
HSCR2564	883,181	16,422	883,110	71	0.000080
HSCR18	777,945	121,658	777,794	151	0.000194
HSCR218	783,669	115,934	783,415	254	0.000324
HSCR1572	896,652	2,951	896,565	87	0.000097
HSCR430	892,302	7,301	892,236	66	0.000074
HSCR3685	890,554	9,049	890,403	151	0.000170

Table S5: Sequence similarity between cases and their relatives.

As a further check on data quality, we used the exome sequence data to assess the expected versus estimated genetic relationship between each affected and his/her sequenced relative. For these analyses, we used genotype data on 27,411 common autosomal exome variants for which allele frequencies were available from external controls (a subset of the data reported in **Supplementary Figure S1**). These estimates demonstrate the linear fit of observations to theoretical expectations.[6] From these results we estimated the coefficient of relationship as shown below (see **Supplementary Methods**):

Table S5.1: *Similarity measures using common variants.*

Expected coefficient of relationship (r)	*S*					
	Min.	*1st Quartile*	*Median*	*Mean*	*3rd Quartile*	*Max.*
0.5 (n = 41)	0.8910	0.9066	0.9079	0.9073	0.9092	0.9257
0.25 (n = 7)	0.8587	0.8627	0.8667	0.8661	0.8683	0.8750
0.125 (n = 4)	0.8338	0.8355	0.8367	0.8389	0.8400	0.8484
0.0625 (n = 1)	0.8248					
0 (n = 17, 955)	0.8025	0.8125	0.8146	0.8145	0.8166	0.8270

Table S5.2: *Coefficient of relationship corresponding to similarity in Table S5.1.*

Expected coefficient of relationship (r)	*R*					
	Min.	*1st Quartile*	*Median*	*Mean*	*3rd Quartile*	*Max.*
0.5 (n = 41)	0.4123	0.4962	0.5032	0.5002	0.5103	0.5995
0.25 (n = 7)	0.2381	0.2597	0.2810	0.2778	0.2899	0.3261
0.125 (n = 4)	0.1036	0.1132	0.1193	0.1312	0.1374	0.1827
0.0625 (n = 1)	0.0551					

Supplementary Table S6: *Pathogenic allele distribution in cases versus controls.*

Exome sequence analyses of the 190 HSCR cases identified 10 genes showing enrichment in the number of *distinct* SNVs, including *RET* and *EDNRB*, which serve as positive controls (**Table 2**). Based on gene expression studies in the human and mouse embryonic gut, all genes except *MMAA* were considered to be HSCR-relevant. The 7 novel genes identified had a total of 39 PAs (1 nonsense, 36 missense, 1 intronic and 1 exonic splicing change) which occurred in 40 of the 190 (21.1%) subjects. However, these cases also had additional PAs in the 17 previously identified HSCR genes (**Table S1**). For completeness, we list below by gene (column 1), the numbers of PAs (column 2) and the numbers of affected individuals with these PAs (column 3) in all 24 HSCR genes among the 190 cases (**Table S6**). The allele frequencies of these PAs as estimated from non-Finnish European ancestry subjects without a neuro-psychiatric disorder from ExAC[10] are shown in column 4 (**Table S6**). These data from HSCR patients are compared to two types of controls. In the first, we compared the numbers of PAs (column 5) and the numbers of individuals (column 7) with these alleles, defined

Table S6: *Distribution of pathogenic alleles in independent HSCR cases and controls.* ([a]: number of distinct PAs in controls; [b]: same as in [a] but restricted to alleles observed in cases).

Gene	***190 cases***			***740 controls***				
	# unique PAs	***# cases with PAs***	***average ExAC allele frequency by gene***	***# unique PAs***		***# controls with PAs***		***average ExAC allele frequency***
SOX10	1	1	0	0[a]	0[b]	0[a]	0[b]	-
PHOX2B	1	1	0	0	0	0	0	-
ZEB2	2	2	0	2	0	2	0	3.57 x 10^{-5}
TCF4	0	0	-	0	0	0	0	-
GDNF	0	0	-	0	0	0	0	-
NRTN	0	0	-	0	0	0	0	-
GFRA1	1	1	7.16 x 10^{-5}	1	1	1	0	2.35 x 10^{-5}
RET	9	12	3.33 x 10^{-4}	3	0	5	3	1.18 x 10^{-3}
ECE1	0	0	-	2	0	2	0	9.79 x 10^{-5}
EDN3	1	1	0	1	0	4	0	2.10 x 10^{-3}
EDNRB	7	7	0	1	0	1	0	2.37 x 10^{-5}
KIF1BP	1	1	7.18 x 10^{-5}	3	0	5	0	1.40 x 10^{-4}
L1CAM	1	1	0	0	0	0	0	-
IKBKAP	1	1	1.00 x 10^{-3}	5	1	6	1	2.69 x 10^{-4}
SEMA3C	3	3	2.10 x 10^{-3}	3	1	25	10	4.28 x 10^{-3}
SEMA3D	6	8	9.25 x 10^{-4}	4	2	9	7	1.83 x 10^{-3}
NRG1	2	3	2.05 x 10^{-3}	4	1	11	8	1.04 x 10^{-3}
ADAMTS17	5	5	2.95 x 10^{-5}	1	0	1	0	4.00 x 10^{-4}
ACSS2	6	6	2.37 x 10^{-4}	2	1	2	1	6.00 x 10^{-4}
SLC27A4	4	4	1.55 x 10^{-4}	1	0	1	0	2.00 x 10^{-4}
SH3PXD2A	4	4	4.62 x 10^{-4}	2	1	4	1	2.50 x 10^{-4}
ENO3	5	5	1.50 x 10^{-4}	2	0	2	0	3.60 x 10^{-5}
FAM213A	4	6	7.30 x 10^{-4}	1	1	2	2	2.80 x 10^{-3}
UBR4	11	15	3.50 x 10^{-4}	14	3	17	4	4.10 x 10^{-4}
All Genes	**75**	**66**	**4.22 x 10^{-4}**	**52**	**12**	**98**	**37**	**8.26 x 10^{-4}**

identically as in HSCR cases, among 740 non-Finnish European ancestry 1000G and NIMH controls (**Table S6**). In the second, we counted the numbers of PAs and cases only for alleles observed in cases (columns 6 and 8). Estimating these numbers from ExAC is not possible because we do not have access to the genotypes of individuals.

Cases in this study harbored 36 distinct PAs in 17 previously known, 39 PAs in 7 novel genes or a total of 75 distinct PAs in all 24 genes. These PAs occur in 41 (21.6%), 40 (21.1%) and 66 (34.7%) individuals for the known, novel and all HSCR genes. The mean allele frequencies of these PAs in our sample of HSCR cases for known, novel and all HSCR genes are 5.58 x 10^{-4}, 2.96 x 10^{-4} and 4.22 x 10^{-4}, respectively, showing relatively little difference between these three categories, but being ~12 times larger than identically defined PAs reported for known genes in databases (i.e., 3.61 x 10^{-5}) (**Table S1**). We suspect that this is owing to the selective reporting of more severe and rarer alleles in public databases and missing true disease variants of lower penetrance which are expected to have higher allele frequencies.

(i) We first tested whether our definition of PAs enriches for causal alleles among HSCR cases in the identified genes as compared to the 740 1000G and NIMH controls. In controls, we identified 29 distinct PAs in the 17 previously known genes, 23 PAs in 7 novel genes and 52 PAs in all 24 genes, and these occurred in 71 (9.6%), 28 (3.8%) and 98 (13.2%) controls, respectively. Overall, identically defined PAs were identified in 34.7% (66/190) of cases in contrast to 13.2% (98/740) of controls, demonstrating a 2.63-fold enrichment (2-sided: P = 4.08 x 10^{-12}). This enrichment was evident for both known genes (41/190 in cases versus 71/740 in controls: 2.25-fold, P = 5.97 x 10^{-6}) and, necessarily (identified using this criterion), novel genes (40/190 in cases versus 28/740 in controls: 5.56-fold, P = 3.46 x 10^{-16}). Observe also that in controls, these PAs had average ExAC allele frequencies of 1.11 x 10^{-3}, 4.74 x 10^{-4}, 8.26 x 10^{-4}, in known, novel and all HSCR genes respectively, and were ~2-fold *higher* than the corresponding allele frequencies in HSCR patients (2-sided: P = 2.14 x 10^{-5}, 0.066 and 1.62 x 10^{-5}, respectively, for known, novel and all HSCR genes). Thus, our definition of PAs leads to enrichment of causal variants because these selected variants exist in significantly greater numbers in cases than in controls and they are significantly rarer in the population than similarly defined variant alleles in controls. Note that 98 of 740 controls or 13.2% of controls have PAs: these represent both non-causal alleles and low penetrance disease alleles unobserved in our cases. Thus, we have a maximum false positive rate of 13.2% in identification of causal alleles in cases. The true proportion of falsely identified PAs in cases is, however, much lower because causal alleles are enriched in cases. In any case, we have significant statistical evidence of an enrichment of HSCR causal alleles across all 24 genes.

(ii) Given the effects of the 24 genes in HSCR, we assessed the impact of observed variants at these genes by performing direct association tests of variant frequencies in cases and controls by gene, i.e., we restricted attention to only PAs observed in cases. We observed 12 case-specific PAs (6 each for known and novel HSCR genes) among 37 (5.0%) of 740 controls (29 and 8 individuals for known and novel HSCR genes). Across all 24 genes, this number is significantly smaller than the 75 among HSCR cases (P = 9.15 x 10^{-55}) and they occur in 37 controls which is also considerably smaller than that in the 66 HSCR cases (P = 2.27 x 10^{-31}). The number of PAs identified in HSCR cases is 24-fold higher than in controls, and the number of individuals with such alleles is 7-fold greater than in controls. These significant differences are true for both known and novel genes as a group. We do not have statistical power to assess these effects for individual genes but the results can be accumulated by pathways (see **Table S1**), as in the following **Table S7**, so that the *relative contributions* of different gene classes to HSCR risk can be estimated.

Table S7: *Distribution and effect of case-observed PAs by pathway.*
These are data in **Table S6** rearranged by gene class (defined in **Table S1**) with statistically significant odds ratios in bold.

Pathway	***Genes***	***# cases (n = 190)***		***# controls (n = 740)***		**Pathway odds ratio (95% CI)**
RET	*GDNF*	0	13	0	3	**16.03** (5.21-49.28)
	NRTN	0		0		
	GFRA1	1		0		
	RET	12		3		
EDNRB	*ECE1*	0	8	0	0	**68.98** (8.68-547.92)
	EDN3	1		0		
	EDNRB	7		0		
SEMA3	*SEMA3C*	3	11	10	17	**2.65** (1.25-5.60)
	SEMA3D	8		7		
TFs	*SOX10*	1	4	0	0	**35.73** (4.15-307.72)
	ZEB2	2		0		
	PHOX2B	1		0		
	TCF4	0		0		
remaining genes	*KIF1BP*	1	7	0	9	**3.15** (1.22-8.09)
	L1CAM	2		0		
	IKBKAP	1		1		
	NRG1	3		8		
17 known genes	*all the above*	*41*		*29*		**6.70** (4.06 – 11.04)
novel genes	*ADAMTS17*	5	40	0	8	**23.19** (11.04-48.72)
	ACSS2	6		1		
	SLC27A4	4		0		
	SH3PXD2A	4		1		
	ENO3	5		0		
	FAM213A	6		2		
	UBR4	15		4		
All 24 genes	*all the above*	*66*		*37*		**10.02** (6.45 – 15.58)

[1] Odds ratios were calculated using the Haldane bias correction and by comparing 190 cases with 740 controls based on coding PAs observed in cases only.

Table S8: *Identifying CNVs using exome sequence, SNP array and karyotype data.*
Details of each CNV detected and validated, based on multiple data types, are shown with CNV location, type, size, chromosomal locus and observed numbers in 185 cases and 19,584 controls. We separately validated a 4 kb *RET* deletion (chr10:42917793-42922026) in patient HSCR472 from a low-quality CNV.

Sample ID	***Karyotype, CNV***		***CNV size (kb)***	***CNV Location***[1]	***Karyotype, Microarray results***	***SNP Validation***[2]	***Observed #***		***P value***[4]
	State	***Chr.***					***Case***	***Control***[3]	
many	dup	21	47,710	1-46,709,983	47 XX & XY, +21	+	11	17[4]	6.68 x 10^{-16*}
HSCR2970	del	16	985	28299106-29283882	16p11.2 del	+	3	12	3.38 x 10^{-4*}
HSCR4220	del	16	906	29372452-30278662	-	+			
HSCR71	del	16	740	29372452-30112616	-	+			
HSCR4522	dup	1	509	144126136-144634799	-	+	3	27	2.72 x 10^{-3}
HSCR4584	dup	1	1,185	143565872-144750520	-	+			
HSCR46	dup	1	971	143663355-144634799	-	+			
HSCR3886	del	1	1,425	144751127-145931774	-	+	1	6	6.37 x 10^{-2*}
HSCR491	del	22	8,000	16280000-24230000	22q11.2 del	+	1	0	9.36 x 10^{-3*}
HSCR3186	dup	22	1,447	15826987-17273998	Tetrasomy 22q	+	1	0	9.36 x 10^{-3*}
HSCR522	dup	17	1,835	13340236-15175628	-	+	1	5	5.49 x 10^{-2*}
HSCR11	dup	4	7,768	183482167-191247414	47, XX, +der(15)t(4:15)	+	1	0[5]	9.36 x 10^{-3*}
HSCR11	dup	15	3,800	61487053-65287054					
HSCR73	del	1	582	46916717-47498799	-	+	1	0	9.36 x 10^{-3}
HSCR208	del	12	554	8102597-8656675	-	+	1	0	9.36 x 10^{-3}
HSCR4368	del	13	14,356	70912755-85268645	13q21.33-q31.1 del	+	1	0	9.36 x 10^{-3*}
HSCR3305	del	2	8,847	133437762-142284441	-	+	1	0	9.36 x 10^{-3*}
HSCR423	del	8	579	1501255-2080313	-	+	1	0	9.36 x 10^{-3}
HSCR241	dup	2	1,377	31566-1397283	-	+	1	1	1.86 x 10^{-2}
HSCR500	dup	7	1,498	88227224-89725429	-	+	1	11	1.06 x 10^{-1}
HSCR4178	inv	10	25,600	101900000-127500000	10q24.3-q26.13inv	-	1	0[5]	-

[1]hg18 genome coordinates; [2] Human Omni 2.5-4 v1 BeadChip data; [3]observed numbers (50% reciprocal overlap of each CNV) in 19,584 controls from reference 21 in main text; [4] controls used were ascertained as adults and not expected to include trisomy 21, the rate of which in 19,584 births was estimated from population studies (ref. 8). [5]Note that the array studies in controls could not detect aneuploidies, translocations and inversions. The control counts for 47, XX, +der(15) t(4:15) are for the two duplications at the translocation site; for the 10q24.3-q26.13 inversion, control counts were not available and not expected. P-values with an asterisk indicate pathogenic CNVs as designated in Table S9.

Table S9: *Inferring the phenotypic consequences of karyotype variants and CNVs.*

Karyotype/ copy number variant[1]	*Syndrome*	*P*[2]	*Assessment of Causality*[3,4]
Free & mosaic trisomy 21	Y	**6.68 x 10^{-16}**	Pathogenic – known association (HSCR)
16p11.2 del	Y/2N	**3.38 x 10^{-4}**	Pathogenic – known association (DD)
1q21.1 dup	N	**2.72 x 10^{-3}**	Likely benign
1q21.1 del	Y	6.37 x 10^{-2}	Pathogenic – known association (DD)
22q11.2 del	Y	9.36 x 10^{-3}	Pathogenic – known association (DD)
Tetrasomy 22q	Y	9.36 x 10^{-3}	Pathogenic – known association (cat eye)
17p11.2 dup	N	5.49 x 10^{-2}	Pathogenic - known association (CMT1A)
47, XX, +der(15) t(4:15)	Y	9.36 x 10^{-3}	Pathogenic – large duplication with 4q partial trisomy
1p33 del	N	9.36 x 10^{-3}	VOUS
12p13.31 del	Y	9.36 x 10^{-3}	VOUS (large segmental duplication content)
13q21.33-q31.1 del	Y	9.36 x 10^{-3}	Pathogenic – known association (DD)
2q21.2-q22.2 del	Y	9.36 x 10^{-3}	Pathogenic – known association (DD)
8p23.3 del	Y	9.36 x 10^{-3}	VOUS (genes in interval have deletions in controls)
2p25.3 dup	N	1.86 x 10^{-2}	Likely benign
7q21.12 dup	N	1.06 x 10^{-1}	Benign
10q24.3-q26.13 inv	Y	-	VOUS

[1] CNVs of interest were defined as deletions >500kb and duplications >1 mb with a control frequency of <1%.[18] Considering all of these CNVs of interest (listed in Table S8) except for the 10q24.3-q26.13 inversion (because a control frequency could not be determined), we observed a total of 29 cases (of 185) having a CNV of interest compared to an expected control frequency of 700 (of 19,584) corresponding to an odds ratio of 5.10 (95% CI: 3.43 – 7.57, P = 4.27 x 10^{-11}). However, most of these changes in controls do not have a phenotypic effect and were assessed against primarily known causal changes, which is why we decided to use only a smaller set of known pathogenic variants for risk estimation. [2]Entries in bold are statistically significant after multiple (15) test corrections with overall significance level of 5%. [3] The presence of a CNV in a HSCR patient can be a causal event or an incidental finding. We assessed known CNV-HSCR associations, statistical evidence of a new CNV association (column 3) and previous CNV association with a developmental phenotype from a set of 29,085 cases of developmental disorders (DD) described in reference 21 in the main text and reference 18, for assessing CNV pathogenicity. [4] VOUS is variant of unknown significance. We ultimately classified variants as "pathogenic" based on a known association with a developmental disorder; these pathogenic CNVs include Free and

mosaic trisomy 21, 16p11.2 del, 1q21.1 del, 22q11.2 del, tetrasomy 22q, 17p11.2 dup, 47, XX, +der(15) t(4:15), 13q21.33-q31.1 del and 2q21.2-q22.2 del.

Table S10: *Comparison of genetic burden of classes of variation by sex.*

Disease-associated risk allele class		% frequency		Male odds ratio (95% CI)	Female odds ratio (95% CI)
		cases (M/F)	controls		
Enhancers, common variants[1]	known[4]	54.2/38.2	17.1	5.78 (4.01-8.33)	3.05 (1.86-5.00)
Coding genes, rare variants[2]	known & novel[4]	35.2/27.9	5.0	10.32 (6.41-16.63)	7.46 (4.09-13.60)
	known[4]	23.0/16.2	3.9	7.31 (4.22-12.65)	4.86 (2.36-10.04)
Copy number alterations, rare variants[3]	known & novel[4]	14.2/6.2	0.20	81.99 (45.51-147.70)	35.60 (13.05-97.13)
	known[4]	8.3/1.5	0.09	100.95 (46.36-219.83)	24.79 (4.61-133.21)

[1] Five or more common disease variants (Table 1) were observed in 90 of 186 cases and 107 of 627 controls; [2] rare coding sequence variants (Table 2) were identified in 66 of 190 cases with an expected rate of 37 in 740 controls; [3] copy number variants (Table 3) were identified in 21 of 185 cases with an expected rate of 40 in 19,584 controls. [4] The data relevant to 24 known and novel loci, and the 18 known loci, respectively, are shown subdivided by sex and are the same data as in Table 4 of the main paper.

Table S11. *Distribution of HSCR by mutation type and phenotype.*

Common Variant[a]	*Rare Variant*[b]	*CNV*[c]	*# (%) Cases*[d]	*# (%) Male / Female*	*# (%) Short / Long & TCA*[e]	*# (%) Simplex / Multiplex*	*# (%) non-syndromic / MA*[f]
-	-	-	50 (28)	26 (52) / 24 (48)	17 (46) / 20 (54)	28 (56) / 22 (44)	38 (76) / 12 (24)
+	-	-	53 (30)	35 (66) / 18 (34)	24 (57) / 18 (43)	36 (68) / 17 (32)	46 (87) / 7 (13)
-	+	-	27 (15)	14 (52) / 13 (48)	9 (41) / 13 (59)	14 (52) / 13 (48)	17 (63) / 10 (37)
-	-	+	13 (7)	11 (85) / 2 (15)	9 (82) / 2 (18)	12 (92) / 1 (8)	2 (15) / 11 (85)
+	+	-	29 (16)	24 (83) / 5 (17)	14 (58) / 10 (42)	24 (83) / 5 (17)	20 (69) / 9 (31)
+	-	+	1 (1)	1 (100) / 0 (0)	0 (0) / 1 (100)	1 (100) / 0 (0)	1 (100) / 0 (0)
-	+	+	3 (2)	3 (100) / 0 (0)	2 (67) / 1 (33)	2 (67) / 1 (33)	0 (0) / 3 (100)
+	+	+	3 (2)	1 (33) / 2 (67)	2 (67) / 1 (33)	3 (100) / 0 (0)	0 (0) / 3 (100)
Totals			179 (100)	115 (64) / 64 (36)	77 (54) / 66 (46)	120 (67) / 59 (33)	124 (69) / 55 (31)

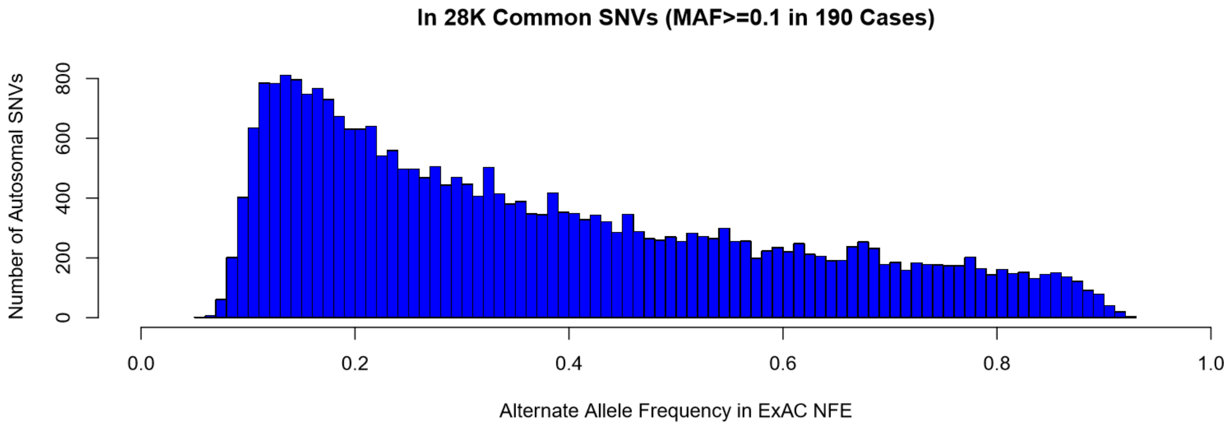
[a] Common variant: 5 or more risk alleles at *RET* (rs2435357, rs2506030, rs7069590) and *SEMA3D* (rs11766001); [b] Rare Variant: 1 or more rare, deleterious variants in any of 17 known and 7 new susceptibility genes identified in this study; [c] CNV (copy number variant): a clinically identified alteration (trisomy 21, 22q deletion, etc.), recurrent CNV or unique rare deletion >500kb or duplication >1000kb identified as pathogenic in Table S9; [d] 179 affected individuals with complete data for all three mutation classes; [e] Cases where segment length was uncertain have been excluded here; [f] Non-syndromic cases have no clinical diagnosis of recognized syndromes or multiple anomalies (MA) in addition to HSCR.

Table S12: *Functions of novel HSCR genes and their relevance to ENS development.*

Relevance to ENS development.	***Gene***	***Encoded functions***
Regulation of axonal guidance	*ADAMTS17*	A plasma membrane protein whose knockdown induces breast cancer cell apoptosis; acts as a versicanase in development and is dysregulated by epigenetic alterations[19,20]
	SH3PXD2A	A lipid-binding cytoskeletal protein resident in the embryonic mesenchyme, binds many ADAM proteins and functions to locally degrade extracellular matrix during axon guidance through tissues. Analysis of zebrafish embryos and neural crest cells *in vitro* have indicated that Src-activated Tks5 (protein encoded by *SH3PXD2A*) is necessary for proper neural crest cell migration.[21]
Cell growth & proliferation	*ACSS2*	Acetyl-Coenzyme A synthetase 2 is both cytoplasmic and nuclear. Despite having many functions in lipid synthesis and energy generation, it can affect transcriptional control and gene expression through p300-catalyzed control of histone acetylation versus crotonylation.[22]
	SLC27A4	A fatty acid transport protein localized to the endoplasmic reticulum and the plasma membrane which has acyl-CoA ligase activity and, therefore, could have functions that interact with *ACSS2*, since increased fatty acid synthesis is required to meet the demand for membrane expansion of rapidly growing cells.
	UBR4	A ubiquitin E3 protein ligase (component N-Recognin 4) localized to the cytoskeleton and the nucleus. Despite having a function required for the termination of RET signaling (performed by CBL[23]), UBR4 may also be involved in regulating acetylation versus ubiquitylation by competing for the same lysine residues in the regulation of fatty acid synthesis and cell growth.[24]
	ENO3	Encodes a muscle-specific enolase active during development.
Local inflammation	*FAM213A*	A cytoplasmic and mitochondrial redox-regulatory protein. Recently, sulfhydryl-mediated redox signaling in inflammation has been shown to have a significant role in neuro-degenerative diseases using RET target proteins.[25]

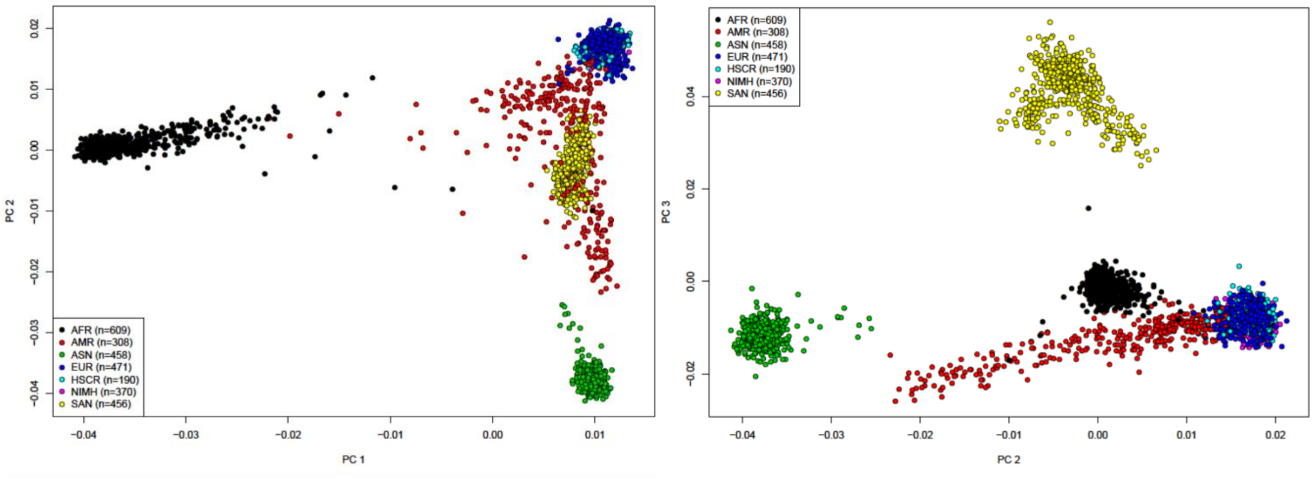
SUPPLEMENTARY FIGURES

Supplementary Figure S1: *Allele frequency distribution of 28,746 common autosomal variants among the 190 HSCR cases* (see Table S3).

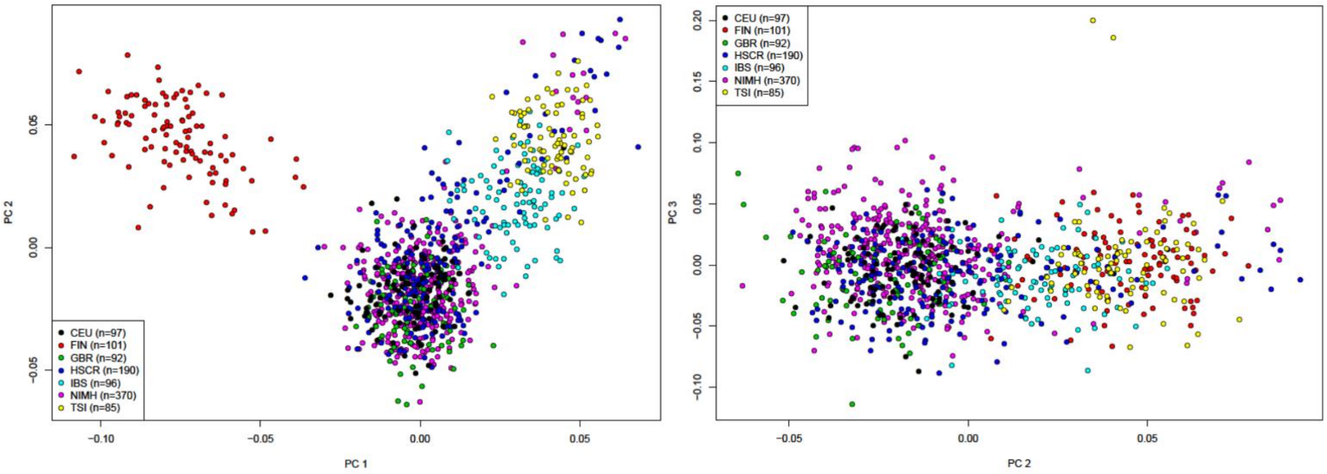


Supplementary Figure S2: *Principal component analysis (PCA) of HSCR samples.*

The first three PCs are plotted below for PCA of 190 HSCR non-Mennonite independent cases (HSCR NI); 370 European American samples from NIMH (NIMH); 458 East Asian samples from 1000G (ASN); 471 European samples from 1000G (EUR); 609 African samples from 1000G (AFR); 308 American samples from 1000G (AMR); 456 South Asian samples from 1000G (SAN). The results show clear overlap for all 190 HSCR cases with reference individuals of European ancestry.

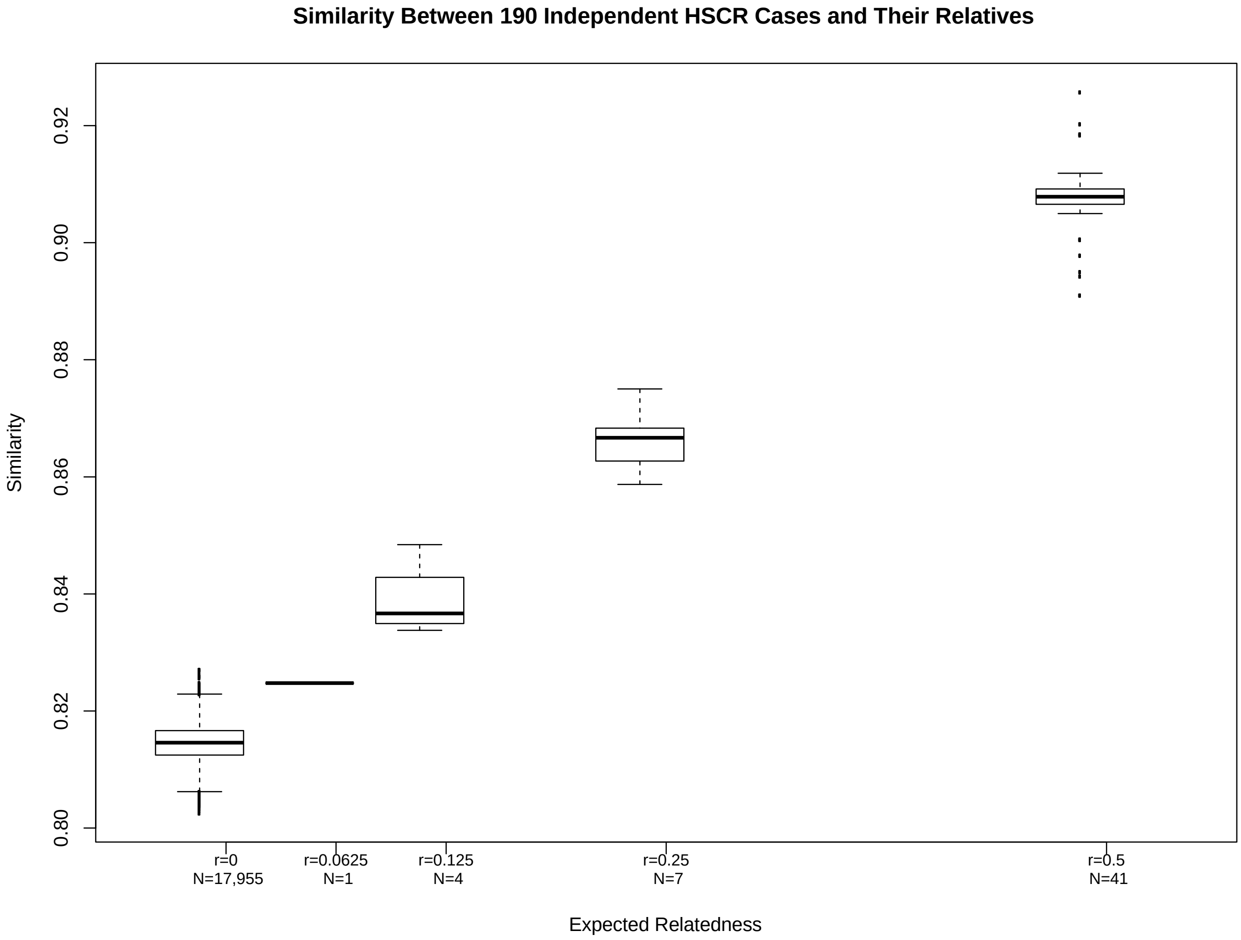


PCA of Europeans only showed that the HSCR cases cannot be distinguished from any European ancestry group except the Finns. The first three PCs of this analysis are plotted below.

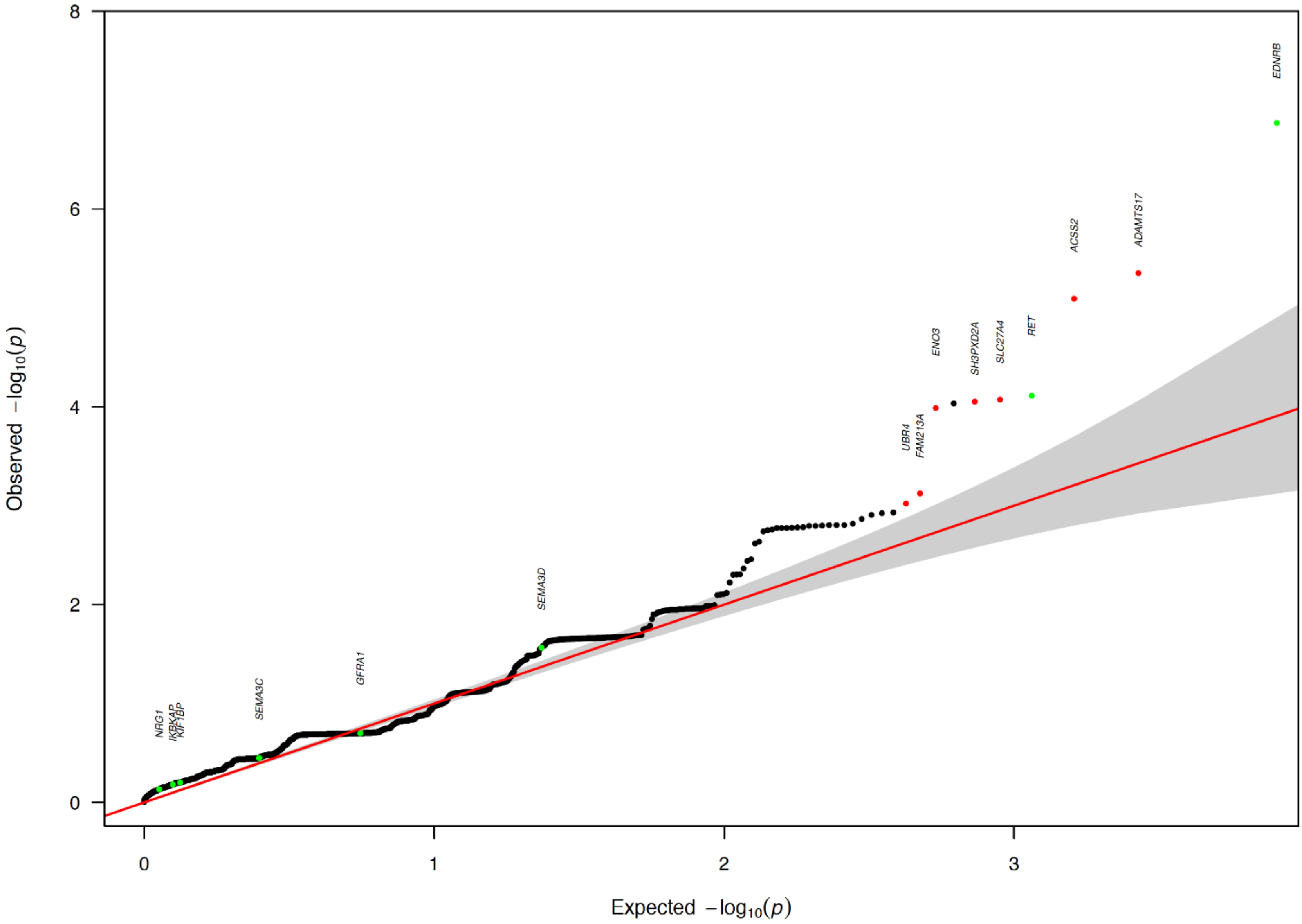


Supplementary Figure S3: *Sequence similarity between relatives.*

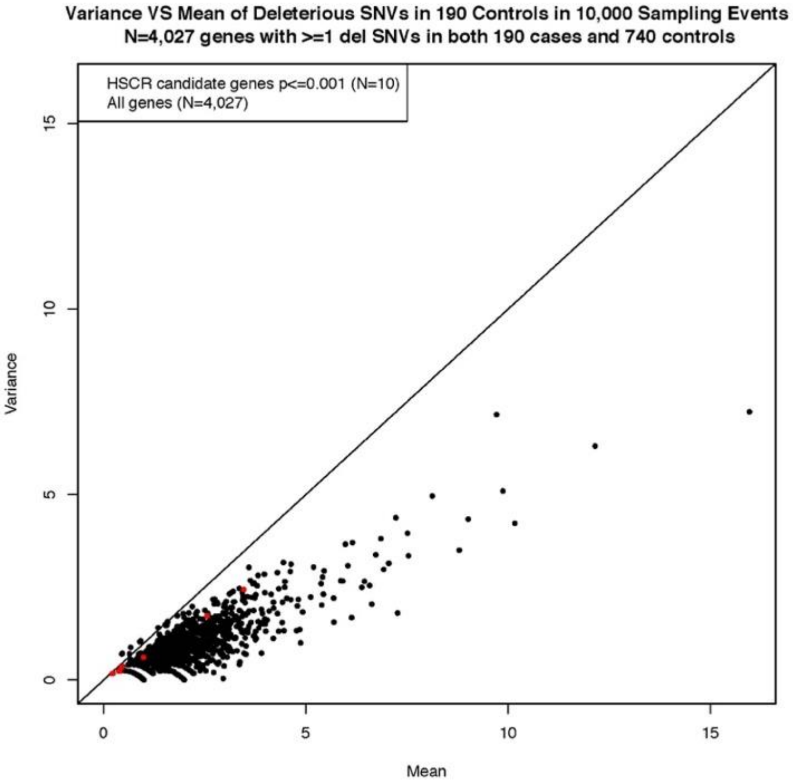
The distribution of similarity scores (S) for the expected (pedigree-based) degree of relationship is summarized below (see data in **Tables S5**). S is linearly related to the coefficient of relationship, as expected, verifying the putative relationships with genetic data.



Supplementary Figure S4: *Assessment of genes significantly enriched for PAs.*



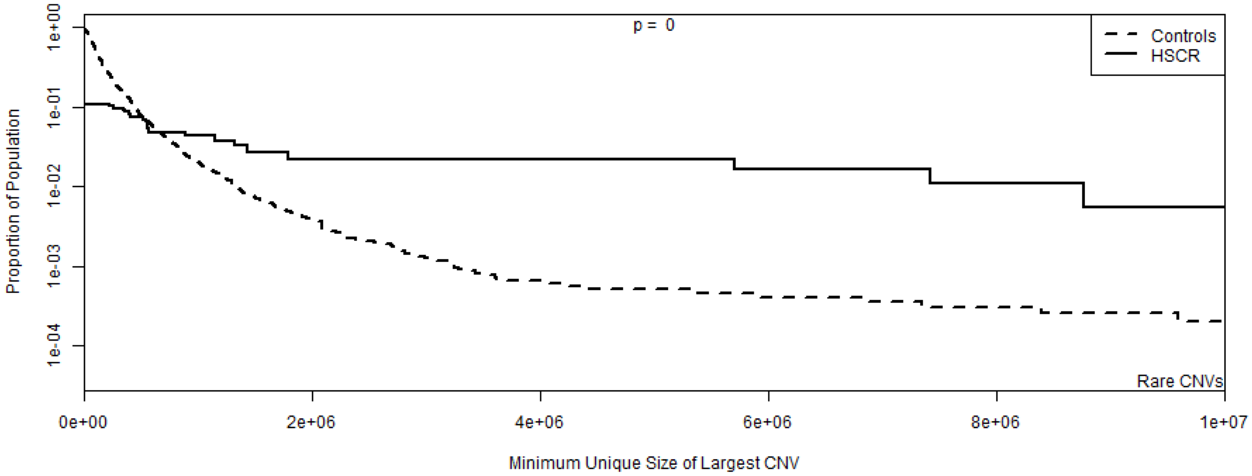
We used computer simulations, using the control exome sequence data, to compare the observed to expected distribution of distinct pathogenic alleles (PA) for each of 4,027 genes with at least one such variant in cases and controls. These were compared to their observed numbers in cases and are compared in the QQ plot below with a 95% confidence interval at each point. As explained in the main text, the top 10 genes were enriched as a group (P<0.001). Genes marked in green were previously identified HSCR genes and those marked in red are novel genes identified in this study.



The statistical test for comparing observed to expected numbers of distinct PAs assumed a Poisson distribution of the number of distinct PAs in a sample. This is a conservative assumption because comparisons of the variance to the mean of the number of distinct PAs in 190 samples, as assessed from replicate sampling from controls, shows considerably less-dispersion (see plot on the right). The same statistical method was used to identify candidate HSCR genes from small INDELs. The test was applied to rare (MAF ≤ 0.05 in 190 cases or 740 controls) and common (MAF>0.05 in cases or controls) alleles for small insertions and deletions separately. There were 551 genes with rare small INDELs in both cases and controls but only one gene, *FAN1*, had a P value below 0.01. None of the 132 genes with common small INDELs showed any statistical significance. This is unsurprising given that most genes have very few (at most 3 rare and 2 common) INDELs.

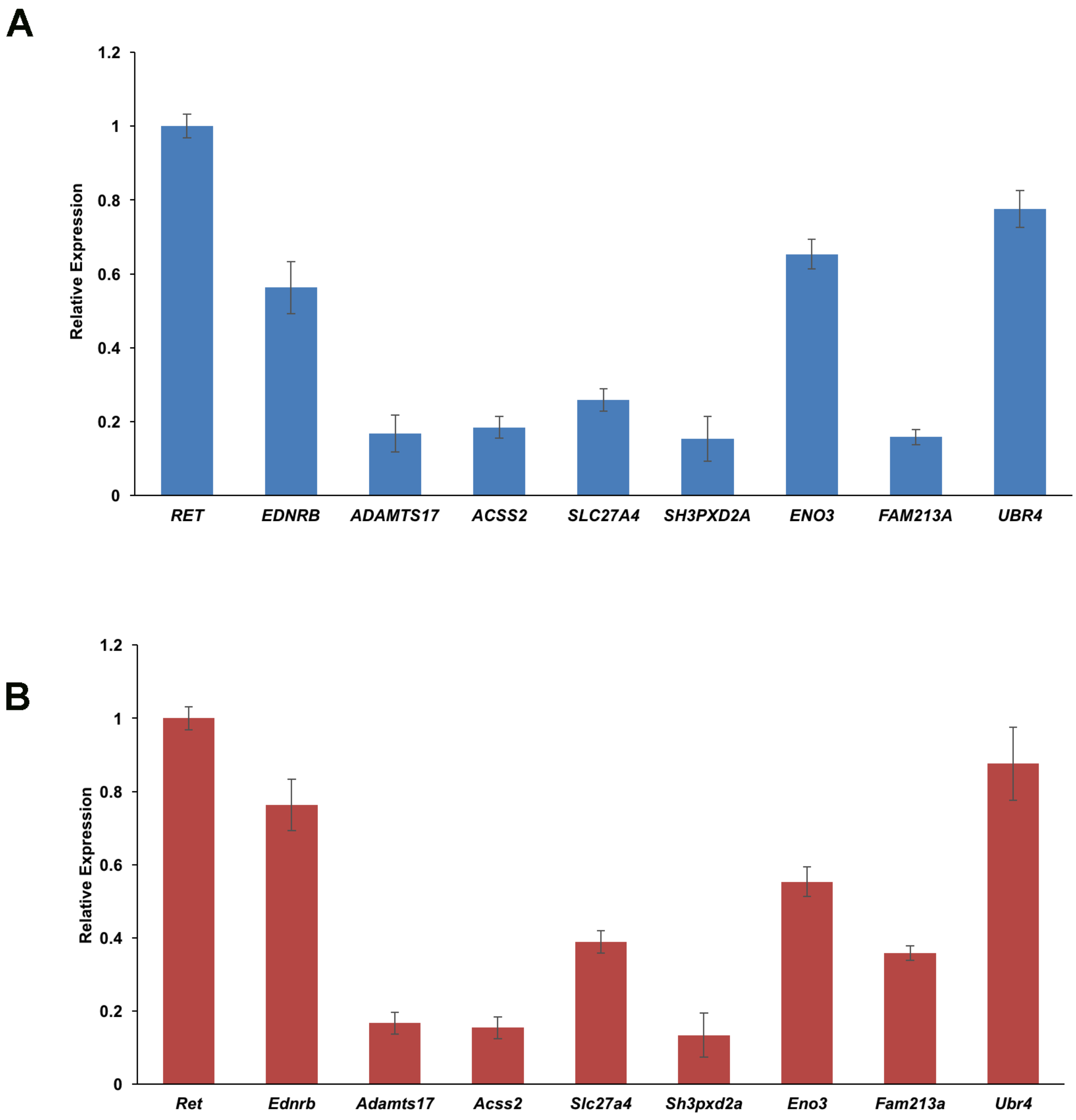
Supplementary Figure S5: *CNV burden in HSCR.*

The proportion of samples with any CNV, in either HSCR or controls, is plotted against the minimum unique size of the largest CNV. The data shows that the distribution of CNVs in HSCR is significantly greater ($P<2.2x10^{-16}$) than in controls by both the log-rank test and the Peto and Peto modification of the Gehan-Wilcoxon tests (https://stat.ethz.ch/R-manual/R-devel/library/survival/html/survdiff.html).[26] The lines cross at 500 kb. Note that CNV size in this analysis is corrected for segmental duplication content.



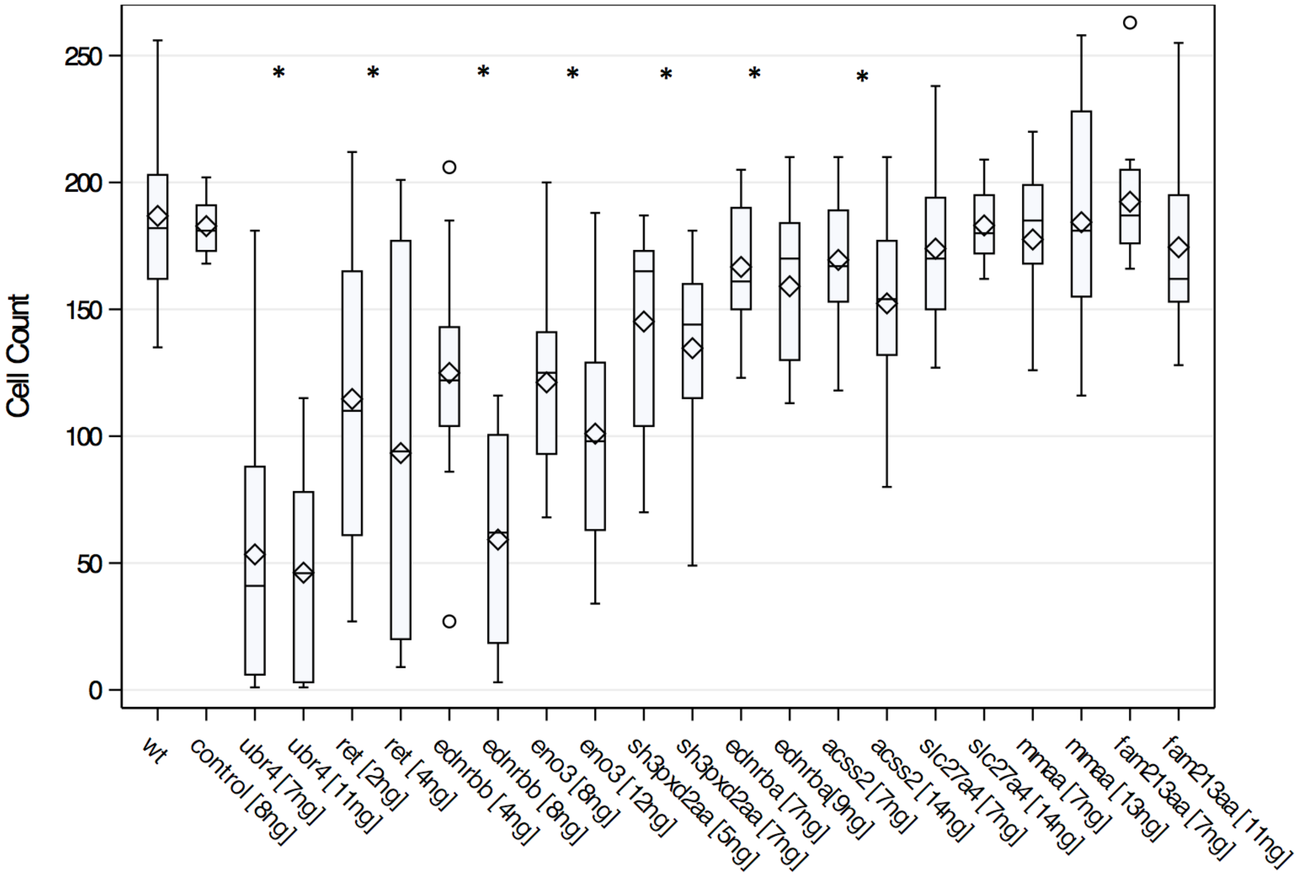
Supplementary Figure S6: *Gene expression of candidate HSCR genes in the embryonic human and mouse gut.*

Taqman gene expression profiles in human fetal gut tissue at Carnegie stage 22 shows all genes except *MMAA* are expressed at the relevant time point in development (A), with similar data from mouse gut tissues at E10.5 (B). The transcript with the highest expression was set to unity to compare the relative expression of other genes. The error bars represent standard errors of the mean from multiple measurements.



Supplementary Figure S7: *Assessment of HSCR candidate genes in zebrafish.*

Distribution of HuC positive migratory enteric neuronal precursors in 6 dpf zebrafish embryos from controls and knockdown of HSCR candidate gene orthologs. Genes with a statistically significant reduction in cell numbers are indicated by an asterisk. Note that there are two *ednrb* zebrafish orthologs but only *ednrbb* was significant in these assays; further *acss2* was significant only at the higher concentration.



References:

1) Puffenberger EG, Hosoda K, Washington SS, et al. A missense mutation of the Endothelin-B Receptor Gene in Multigenic Hirschsprung's Disease. *Cell* 1994; 79:1257-1266.

2) Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatic*s 2009; 25:1754-1760.

3) Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A. Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res* 2010; 20:110–121.

4) Zheng X, Levine D, Shen J, et al. A High-performance Computing Toolset for Relatedness and Principal Component Analysis of SNP Data. *Bioinformatics* 2012; 28:3326-3328.

5) Karakoc E, Alkan C, O'Roak BJ, et al. Detection of structural variants and indels within exome data. *Nat Methods* 2011;9:176-178.

6) Li CC, DE Weeks, Chakravarti A. Similarity of DNA fingerprints due to chance and relatedness. *Human Heredity* 1993; 43:45-52.

7) Hartl DL, Campbell RP. Allele multiplicity in simple Mendelian disorders. *Am J Hum Genet* 1982; 34:866-873.

8) Presson AP, Partyka G, Jensen KM, et al. Current estimate of Down Syndrome population prevalence in the United States. *J Pediatrics* 2013; 163:1163-1168.

9) Haldane JBS. The estimation and significance of the logarithm of a ratio of frequencies. *Ann Hum Genet* 1956; 20:309-311.

10) Lek M, Karczewski KJ, Minikel EV et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 2016; 536:285-291.

11) Trapnell C, Roberts A, Goff L, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols* 2012; 7:562-578.

12) Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF. Stages of embryonic development of the zebrafish. *Dev Dynamics* 1995; 203:253-310.

13) Westerfield, M. The Zebrafish Book. A Guide for the Laboratory Use of Zebrafish (*Danio rerio*). 4th ed. Eugene, Oregon: University of Oregon Press, 1991.

14) Kuhlman J, Eisen JS. Genetic screen for mutations affecting development and function of the enteric nervous system. *Dev Dynamics* 2007; 236:118-127.

15) Abramoff MD, Magalhaes PJ, Ram SJ. Image Processing with ImageJ. *Biophotonics International* 2004; 11:36-42.

16) Stenson PD, Mort M, Ball EV et al. The Human Gene Mutation Database: building a comprehensive mutation repository for clinical and molecular genetics, diagnostic testing and personalized genomic medicine. *Hum Genet* 2014;133:1-9.

17) Landrum MJ, Lee JM, Benson M et al. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res* 2016; 44:D862-868.

18) Kaminsky EB, Kaul V, Paschall J et al. An evidence-based approach to establish the functional and clinical significance of copy number variants in intellectual and developmental disabilities. *Genet Med* 2011; 13:777-784.

19) Kelwick R, Desanlis I, Wheeler GN, Edwards DR. The ADAMTS (A Disintegrin and Metalloproteinase with Thrombospondin motifs) family. *Genome Biol* 2015; 16:113.

20) Jia Z, Gao S, M'Rabet N, et al. Sp1 is necessary for gene activation of Adamts17 by estrogen. *J Cell Biochem* 2014; 115:1829-1839.

21) Murphy DA, Diaz B, Bromann PA, et al. A Src-Tks5 pathway is required for neural crest cell migration during embryonic development. *PLoS One* 2011; 6:e22499.

22) Sabari BR, Tang Z, Huang H, et al. Intracellular crotonyl-CoA stimulates transcription through p300-catalyzed histone crotonylation. *Mol Cell* 2015;58:203-215.

23) Mulligan LM. RET revisited: expanding the oncogenic portfolio. *Nature Rev Cancer* 2014; 14:173-186.

24) Lin R, Tao R, Gao X, et al. Acetylation Stabilizes ATP-Citrate Lyase to Promote Lipid Biosynthesis and Tumor Growth. *Mol Cell* 2013; 51:506-518.

25) Miller G, Mieyal JJ. Sulfhydryl-mediated redox signaling in inflammation: role in neurodegenerative diseases. *Arch Toxicol* 2015;89:1439-1467.

26) Harrington DP, Fleming TR. A Class of Rank Test Procedures for Censored Survival Data *Biometrika* 1982; 69:553-566.