medRxiv preprint doi: https://doi.org/10.1101/2020.11.02.20221481; this version posted November 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in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-NC-ND 4.0 International license .

### 1 Size matters: large copy number losses reveal novel Hirschsprung disease genes

- 2 Laura Kuil<sup>1</sup>, Katherine C. MacKenzie<sup>1</sup>, Clara S Tang<sup>3, 4</sup>, Jonathan D. Windster<sup>1, 2</sup>, Thuy Linh Le<sup>5</sup>, Anwarul
- 3 Karim<sup>3,</sup> Bianca M. de Graaf<sup>1</sup>, Robert van der Helm<sup>1</sup>, Yolande van Bever<sup>1</sup>, Cornelius E.J. Sloots<sup>2</sup>, Conny
- 4 Meeussen<sup>2</sup>, Dick Tibboel<sup>2</sup>, Annelies de Klein<sup>1</sup>, René M. H. Wijnen<sup>2</sup>, Jeanne Amiel<sup>5</sup>, Stanislas Lyonnet<sup>5</sup>,
- <sup>7</sup>, Maria-Mercè Garcia-Barcelo<sup>3</sup>, Paul K.H. Tam<sup>3, 4</sup>, Maria M. Alves<sup>1</sup>, Alice Brooks<sup>1</sup>, Robert M.W. Hofstra<sup>1,</sup>
- 6 <sup>6</sup>, Erwin Brosens<sup>1\*</sup>

7 <sup>1</sup>Department of Clinical Genetics, Erasmus University Medical Centre Rotterdam, The Netherlands.

- 8 <sup>2</sup> Department of Paediatric Surgery, Erasmus University Medical Centre, Rotterdam, The Netherlands.
- <sup>3</sup> Department of Surgery, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China.
- <sup>4</sup> Li Dak-Sum Research Centre, The University of Hong Kong Karolinska Institutet Collaboration in Regenerative Medicine, Hong
   Kong, China
- <sup>5</sup> Laboratory of embryology and genetics of malformations, Institut Imagine, Université de Paris, INSERM UMR1163, Necker
   Enfants malades University Hospital, APHP, Paris, France.
- <sup>6</sup> Stem Cells and Regenerative Medicine, UCL Great Ormond Street Institute of Child Health, London, UK
- 15 \* corresponding author: <u>e.brosens@erasmusmc.nl</u>
- 16

## 17 Abstract

- 18 Background
- 19 Hirschsprung disease (HSCR) is characterized by absence of ganglia in the intestine. Approximately
- 20 18% of patients have additional anatomical malformations or neurological symptoms (HSCR-AAM).
- 21 HSCR is a complex genetic disease in which the loss of enteric ganglia stems from a combination of
- 22 genetic alterations: rare coding variants, predisposing haplotypes and Copy Number Variation (CNV).
- 23 Pinpointing the responsible culprits within a large CNV is challenging as often many genes are
- 24 affected. We investigated if we could find deleterious CNVs and if we could identify the genes
- 25 responsible for the aganglionosis.

### 26 Results

27 Deleterious CNVs were detected in three groups of patients: HSCR-AAM, HSCR patients with a

28 confirmed causal genetic variant and HSCR-isolated patients without a known causal variant and

- 29 controls. Predisposing haplotypes were determined, confirming that every HSCR subgroup had
- 30 increased contributions of predisposing haplotypes, but their contribution was highest in isolated
- 31 HSCR patients without *RET* coding variants. CNV profiling proved that HSCR-AAM patients had larger
- 32 copy number losses. Gene enrichment strategies using mouse enteric nervous system transcriptomes
- 33 and constraint metrics were used to determine plausible candidate genes in Copy Number Losses.
- Validation in zebrafish using CRISPR/Cas9 targeting confirmed the contribution of UFD1L, TBX2,
   NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

- 35 *SLC8A1* and *MAPK8* to ENS development. In addition, we revealed epistasis between reduced Ret
- 36 and Gnl1 expression in vivo.
- 37 Conclusion
- 38 Rare large Copy Number losses often *de novo* contribute to the disease in HSCR-AAM patients
- 39 specifically. We proved the involvement of five genes in enteric nervous system development and
- 40 Hirschsprung disease.

#### Introduction 41

42 Hirschsprung disease (HSCR) is characterized by absence of the enteric nervous system (ENS) in the 43 intestine (aganglionosis). HSCR is a complex genetic disease where the phenotype stems from a 44 combination of genetic alterations: rare coding variants, predisposing haplotypes and Copy Number 45 Variation (CNV)[1]. Most patients have HSCR as an isolated congenital anomaly, while in others the 46 absence of enteric neurons is part of a spectrum of anomalies of a monogenetic syndrome[2]. 47 Approximately 18% of patients have additional anatomical malformations or neurological symptoms[3] 48 (HSCR-AAM). Deletions of 10q11[4, 5] led to the identification of the REarranged during Transfection 49 gene (RET), the major responsible gene for familial and sporadic isolated HSCR[6, 7]. HSCR 50 penetrance is influenced by predisposing RET risk haplotypes which increase disease risk 51 substantially, especially if homozygous, in specific combinations (rs2506030, rs7069590, 52 rs2435357)[8, 9], together with other risk loci near the semaphorin gene cluster (rs80227144)[9, 10], 53 or the neuregulin 1 gene (NRG1; rs7005606)[11-13]. 54 For many of HSCR-AAM patients the underlying cause is unknown. It is plausible that the associated 55 anomalies in these patients are the result of a pathogenic single nucleotide variant or 56 insertion/deletion. However, all different genomic variation types should be considered, since the 57 presence of unbalanced translocations and CNV has resulted in the identification of causal genes for several syndromes with HSCR as a key feature. For example, deletions of 13g[1, 14-18] resulted in 58 59 the identification of one of the genes responsible for Waardenburg-Shah syndrome type 4 60 (EDNRB)[19]. Deletions of 2q[20-23] and 4p[24] contributed to the discovery of genes responsible for 61 Mowatt-Wilson syndrome (ZEB2, formerly ZFHX1B)[25] and the gene responsible for Congenital 62 Central Hypoventilation Syndrome (PHOX2B)[26]. Furthermore, deletions (17g21[27] and 22g11[1, 28]) and duplications (17q23[29], dup22p[27, 30] and 22q11[31]) have been described in HSCR 63 patients. In the light of these findings, we hypothesized that rare CNVs significantly contribute to 64 65 aganglionosis in HSCR-AAM patients who lack a pathogenic coding variant in any of the known HSCR 66 genes.

67

#### 68 Results

#### 69 Rare deletions are enriched in HSCR-AAM patients

70 Patient and control CNVs were determined and classified as a "rare CNV" when absent from 71 large control cohorts (n=19,584). We also included known pathogenic or modifier CNVs[32, 33]. The 72 size, type and gene content characteristics of rare CNVs were compared to those of 326 controls 73 (Group 4, n=326). With this approach, 56 rare CNVs were detected in 34 HSCR patients (S3). We 74 determined segregation of the rare CNVs in nine patients, five of these were de novo: three gains 75 (22q11.21 - q11.22, 22q11.21 and 7q36.1) and two losses (17q23.1 - q23.2 and 6p22.1 - p21.33). 76 CNV regions did not have a lot of overlap within our cohort, previously published CNV studies or with 77 those described in the DECIPHER (https://decipher.sanger.ac.uk/) database (S3). The number of rare 78 CNVs per individual did not differ between subgroups or controls (S4). However, CNVs found in 79 HSCR-AAM patients (Group 1) were larger compared to controls (Group 4, p=7.297E-6, Figure 2A). 80 This difference is attributed to outliers, large size of losses in specific patients of group 1 (p=3.15E-81 08)(S5), strongly suggesting a role of these specific CNVs in HSCR-AAM patients.

82 Candidate genes within a loss are constraint and expressed during ENS development

83 HSCR disease genes (S6) are often constraint and expressed in the affected tissue, genes 84 within a CNV region associated with HSCR should share these characteristics (Figure 1A). In Group 1, 85 rare CNVs are enriched for ENS genes (p=4.565E-6, S7). Moreover, losses in HSCR-AAM patients 86 (Group 1) contained more ENS genes that are constraint[34, 35], compared to losses found in 87 controls (Group 4, Figure 2B, p=0,0004). Therefore, constraint genes within CN losses identified in 88 group 1, and expressed in the ENS were prioritized as candidate genes for HSCR: AKT3, GNL1, GABBR1, SLC8A1, MAPK8, UFD1L, TBX2, USP32 and TUBB (Table 1). 89

90 Zebrafish validation confirms the impact of candidate gene disruption on ENS development

91 To validate the effect of losing one copy of these constraint genes in ENS development and 92 HSCR, we disrupted their orthologues in zebrafish by CRISPR/Cas9 targeting. Zebrafish are highly 93 suited for reverse-genetic screening due to their rapid, ex-utero development and transparency and 94 have proven instrumental in the identification of HSCR disease genes[36]. Premature termination of 95 enteric neuron migration was observed in a subgroup of larvae injected with guide RNAs targeting 96 ufd1l (p=0,0208), tbx2 (p=0,0373), slc8a1 (p=0,0073), and mapk8 (p=0,0022, Figure 2D).

#### 97 HSCR disease coding genetics is heterogeneous

98 More evidence for their involvement could be provided by finding additional patients with 99 putative deleterious alterations in genes affected by rare CNV. However, there was no significant 100 enrichment for nonsense, splice site or missense variants in genes impacted by rare gains nor losses 101 (S9) when comparing 443 short segment HSCR patients and 493 controls[37], underlining the genetic 102 heterogeneity between HSCR patients. Nonsense or splice-site variants were detected in eight genes 103 (Table 2). Only one variant, detected in DNA derived from blood and isolated ENS cells, affected a 104 constraint gene overexpressed in the mouse ENS and present in a CN loss: a frameshift (TUBB; 105 NM 001293212.1:c.1330 1331delCAinsA, p.Gln444Serfs\*35) in a patient with isolated HSCR (table 106 1, S8).

#### 107 Copy number loss alone is likely not sufficient to result in HSCR

108 Considering that aganglionosis has a low prevalence in patients with 22q11[32, 33, 38] and 109 17q23[39, 40] deletions, CNVs alone are probably not sufficient to cause HSCR. HSCR penetrance is 110 influenced by predisposing risk haplotypes [8-13]. Five non-coding risk haplotypes were tested in our 111 three patient groups and compared to an *in-house* population dataset (group 5) (n=727, Risk noncoding (RSnc) = 2.54). As expected, HSCR groups 1, 2 and 3 all had a higher RSnc compared to 112 113 group 5 (p= 4.59825E-14, p= 0.000960603, p= 1.21875E-23, respectively) (Figure 2C), indicating a contribution of the risk haplotypes in all HSCR subgroups. (S10; Figure 1B). Associated common risk 114 115 haplotypes have epistatic interactions, not only with each other, but also with known HSCR genes -116 such as RET and NRG1[10, 12, 41] - and modify HSCR penetrance in monogenetic syndromes and 117 Down syndrome[42-46]. The effect of epistatic interactions is smaller in high penetrant HSCR 118 monogenetic disorders and larger in disorders in which HSCR penetrance is lower[42, 43]. In line with 119 this, contribution of these risk haplotypes was the lowest in group 2 (RSnc=3.70), followed by group 1 120 (RSnc=4.71), and the highest in group 3 (RSnc= 5.66). Thus, risk haplotypes have indeed a higher 121 impact in HSCR patients without known coding variants, compared to those that contain known coding variants (S11). To model reduced RET expression in patients with predisposing haplotypes, we 122 123 injected zebrafish with an antisense ATG blocking morpholino specific for ret. Disruption of the 124 candidate genes mapk8 (p=0,011, Figure 2E) and gnl1 (p=0,0405, Figure 2F) in combination with the 125 ret morpholino resulted in less neurons. Considering that disruption of gnl1 alone did not induce a

5

HSCR phenotype, these results suggest that only in a sensitized background loss of *gnl1* can increaseHSCR penetrance.

128 Candidate genes are associated to ENS development.

129 CRISPR/CAS9 targeted knockdown of the candidate genes SLC8A1, GNL1, MAPK8, UFD1L 130 and TBX2 connects these genes to ENS development. These genes are overexpressed in mouse 131 ENS and have associations to ENS development. For instance, Mitogen-Activated Protein Kinase 8 132 gene (MAPK8) is part of the Mitogen activated protein (MAP) kinase pathway and required for ENCCs 133 to migrate properly[47]. UFD1L, as this encodes for a downstream target of HAND2[48]. Hand2<sup>-/-</sup> mice 134 have decreased numbers of enteric neurons, neuronal differentiation defects and a disorganized 135 ENS[49, 50]. T-Box 2 (TBX2) codes for a transcription factor already known to be involved in the regulation of neural crest derived melanocytes[51], a process that is also hampered in specific HSCR 136 137 patients carrying pathogenic variants in EDN3, SOX10 and EDNRB[52, 53]. Candidate genes are 138 discussed further in S8.

### 139 Discussion

140 The results presented here highlight the importance of Copy Number profiling, leading us to 141 propose a HSCR risk model that consists of risk scores for: rare coding variants, non-coding variants, 142 rare CNVs and currently unknown risk factors (e.g. epigenetic modifications and perhaps 143 environmental factors). Considering HSCR-AAM (group 1), 36% of patients have contained a CNV 144 harboring genes overexpressed in the mouse ENS (Figure 3A). Our risk model (Figure 3C) indicates 145 major contributing factors in each group: in group 1 CNVs, in group 2 coding variants and in group 3 146 noncoding variants (Figure 3A, C). The predisposing RET haplotypes reduce RET expression[54] and 147 in combination with the other risk haplotypes increase HSCR risk substantially. Translating this 148 population risk to a threshold for HSCR development in individual patients is challenging as it is not 149 precisely known how many of (and if) these known predisposing haplotypes are sufficient to result in 150 aganglionosis. If we set a threshold at a relatively high level (RSnc>6), this would explain HSCR in 151 30% of patients in group 3 (S11, Figure 2C, 3A). Most HSCR-AAM patients have a lower RSnc, which makes sense, as this group contains more high impact CNVs and whilst they lack RET coding 152 153 variants, they potentially have variants in other genes (Figure 2C, 3A, S11). Translating our findings 154 back to our full cohort including all patients seen by a clinical geneticist and screened for RET

155 mutations, HSCR coding variants explain 21% of HSCR cases (Figure 3B). If we consider CNVs 156 containing genes expressed in the ENS, we could explain an additional 10% (Figure 3B). Importantly, 157 this is an underestimation, since for 55% of patients without a RET coding variant no copy number 158 profiling is performed (Figure 3B) and we focussed exclusively on rare CNVs. Since more common 159 CNVs can also modify HSCR penetrance[55, 56], our results emphasize the need for a more 160 widespread genomic analyses in all subgroups. Noncoding risk scores could be one of the 161 determinants for a more elaborate genetic evaluation. A high noncoding risk score in a HSCR patient 162 without associated anomalies is less likely to have a deleterious RET coding variant, patients with a low score are more likely to benefit from exome sequencing. Similarly, patients with large de novo 163 164 losses often have a more complex phenotype and more intensive clinical investigations might be 165 indicated.

### 166 Conclusions

167 To summarize, HSCR genetics is complex with contributions of predisposing haplotypes in all 168 HSCR subgroups. Rare large CNVs - often de novo (S5) - contribute substantially to the disease in 169 HSCR-AAM patients. These CNVs are enriched for CN losses and for genes intolerant to variation 170 that are overexpressed in the developing mouse ENS. Disruption of some of these genes in zebrafish confirmed that reduced expression of these genes increased the occurrence of HSCR, alone, or in 171 172 epistasis with Ret. These genes have functional overlap with known HSCR disease genes: e.g. 173 UFD1L is involved signalling receptor binding and MAPK8 in axon guidance. Our "ENS expressed gene" based approach lead to the identification of new HSCR disease genes (Figure 3; Table 2): 174 175 UFD1L, TBX2, SLC8A1, GNL1 and MAPK8.

### 176 Materials and methods

177 All authors had access to the study data and reviewed and approved the final manuscript

### 178 Patient inclusion

- 179 This project was approved by the Medical ethics committee of the Erasmus Medical Centre
- 180 (Hirschsprung disease: no 2012-582, addendum No. 1 and no.193.948/2000/159, addendum No. 1
- and 2, MEC-20122387). We selected 58 out of 197 patients (Figure 3) for which DNA and informed
- 182 consent were available, and in whom *RET* was screened. Three subgroups of HSCR patients were

- 183 included in the CNV detection study: patients with HSCR and additional anatomical malformations or 184 neurological defects, but without a RET pathogenic variant, or other causal genetic defect (group 1, 185 n=23, S1), patients with HSCR and a known variant in RET or another causal gene (group 2, n=15, 186 S2), and patients with only HSCR, without a deleterious RET coding variant or other causal genetic 187 defect (group 3, n=20). Additionally, we included unaffected control individuals (group 4, n=326). 188 Genotypes of noncoding predisposing loci were compared with those of other unaffected controls 189 (group 5, n=727).

### 190 Determination of Copy Number Variation (Figure 1a)

- CNV profiles were determined with either the HumanCytoSNP-12 v2.1 or the Infinium Global 191
- 192 Screening Array-24 v1.0 (Illumina Inc., San Diego, CA, USA), using methods previously described[57].
- 193 CNV profiles were inspected visually in Biodiscovery Nexus CN8.0 (Biodiscovery Inc., Hawthorne, CA,
- 194 USA). CNVs with an overlap of at least 75% with similar state CN changes, were either classified as
- 195 rare, when absent from large control cohorts (n=19,584), or as a known modifier[32, 33]. CNV
- 196 number, size, type and gene content of rare CNVs were determined in HSCR patients (n=58) and
- 197 unaffected controls (n=326) and compared between the control groups and previously described
- 198 HSCR subgroups. All rare CNVs were uploaded to the ClinVar database
- 199 (https://www.ncbi.nlm.nih.gov/clinvar/) and are depicted in S3.

#### 200 Exclusion of the involvement of known disease genes (Figure 1a)

201 The presence of RET coding mutations and those in intron-exon boundaries, in all patients in this 202 study are determined. Furthermore, if a specific monogenetic syndrome (S6) was suspected based on 203 the phenotypic spectrum observed, the suspected gene(s) were evaluated using a targeted NGS 204 panel or whole exome sequencing. In four HSCR patients with associated anomalies (Group 1) and nine HSCR patients without associated anomalies (Group 3), the involvement of other known disease 205 206 genes was excluded[2, 36, 37, 58] using whole exome sequencing (WES) with previously described 207 pipelines[59, 60] and variant prioritization methods[61].

#### 208 Evaluation of candidate gene expression (Figure 1a)

209 We prioritized candidate genes based on gene characteristics: genes that are intolerant to variation

- 210 and/or dosage sensitive (see variant filter criteria), and overexpressed in the developing mouse ENS
- 211 compared to whole gut, between embryonic day E11.5 and E15.5 [62, 63]. Since data from human

- 212 colon was only available for embryonic week 12, 14 and 16, we evaluated gene expression in these
- 213 time points[64]. Data was downloaded from the gene expression omnibus (GSE34208 and
- 214 GSE100130). Genes that were dosage sensitive, and either overexpressed in the mouse ENS or
- 215 highly expressed in human embryonic colon, were selected.

#### 216 Variant prioritization in loss of function genes (Figure 1a)

- 217 To determine whether a gene affected by a rare putative deleterious CNV was constraint[34, 35], we
- 218 allowed some tolerance to account for reduced penetrance (S9). Variants from NGS data previously
- 219 generated were prioritized according to settings described in S9: These data included a WES cohort of
- 220 sporadic HSCR (n=76, 149 controls) and (2) a Whole Genome Sequencing (WGS) cohort of 443 short
- 221 segment HSCR patients and 493 unaffected controls[37]. Using RVTESTS[65], a variant burden test
- 222 was done comparing the variant burden in 443 short segment HSCR patients and 493 controls (S9).
- 223 All rare putative deleterious loss of function variants unique to the HSCR cohort in constraint genes
- 224 are described in table 1.

#### 225 Genotyping of HSCR associated SNPs

- 226 Sanger sequencing was used to genotype all patients for SNPs known to be associated to HSCR[8-
- 227 10, 55]. Primer sequences can be found in S11. We used (proxy) SNPs present on the GSAMD-v1
- 228 chip to determine the Rotterdam population background for these risk haplotypes (S10).

#### 229 **Experimental animals**

- Zebrafish were kept on a 14h/10h light–dark cycle at 28°C, during development and adulthood. 230
- 231 Tg(phox2bb:GFP) animals were used for all experiments[66]. Larvae were kept in HEPES-buffered E3
- 232 medium and 0.003% 1-phenyl 2-thiourea (PTU) was added 24 hours post fertilization (hpf), to prevent
- 233 pigmentation.

#### 234 **CRISPR/Cas9** gene disruption

- 235 Gene targeting using CRISPR/Cas9 was performed as described previously[67]. The Alt-R CRISPR-
- 236 Cas9 System from Intergrated DNA Technologies (IDT) was used[68] (Figure 2A). crRNAs were
- 237 designed using the IDT tool and selected on high target efficiency and low off target effect. gRNA
- 238 sequences are listed in S13, primers used are listed in supplementary S14. Efficiency of indel
- 239 generation was determined as described previously[67] (S15).

### 240 Morpholino injections

- 241 Fertilized zebrafish eggs were injected with the CRISPR/Cas9 complex. Subsequently, 0,5 or 1 ng of
- translation blocking morpholino against *ret* (5'- ACACGATTCCCCGCGTACTTCCCAT -3'), and 1:10
- 243 Phenol Red (Sigma-Aldrich), was injected between the 1- and 4-cell stage (see figure 2D)[69]. To
- 244 minimize variability, the same needle was used to inject the controls, that were not injected with
- 245 CRISPR/Cas9 complex.

### 246 Imaging

- 247 Images of 5 dpf tg(*phox2bb:*GFP) larvae were taken using a Leica M165 fluorescent dissection
- 248 microscope with the Leica LASX software. Larvae were anesthetized with 0.016% MS-222 in HEPES
- buffered E3 medium and positioned on their lateral side on a 1.5 2% agarose coated petridish, to
- 250 enable visualization of the enteric neurons.

### 251 Scoring HSCR phenotypes in zebrafish

- 252 To determine whether the larvae present with a HSCR phenotype five categories were made: normal,
- 253 hypoganglionic (less neurons), ultra-short segment HSCR (only most distal end of the gut lack
- 254 neurons), short segment HSCR (neurons colonize until the midgut), and total colonic HSCR (neurons
- are absent from the total intestine) (see figure 2C).

### 256 Statistical analysis

- 257 The number and size of rare CNVs, the number of rare losses and gains, the number of genes
- intolerant to variation (SNVs and CNVs), the number of genes overexpressed in mouse ENS per rare
- 259 CNV, and the relative weighted risk score, were determined and compared for the different groups
- 260 with a single ANOVA test. If group differences existed (P<0.05), we determined which subgroups were
- significantly different, using a two-tailed T-test. For statistical analysis of the zebrafish experiments the
- 262 online tool MedCalc (MedCalc software ltd., Ostend, Belgium), was used together with the "N-1" Chi-
- squared test, number of larvae used and p values can be found in supplementary S16.

### 264 Supplemental Data description

- 265 CNV\_HSCR\_MedRxiv\_Kuil\_etal\_supplement contains 16 elements including detailed patients
- 266 descriptions, additional results, extended methods, primers and gRNA sequences.
- 267 Author Contributions

- 268 Guarantor of the article [EB]; Conception or design of the work [EB, AB, RH], the acquisition [RW, JW,
- 269 TL, RH, YB, CS, CM, DT, LK], analysis [LK, KM, CT, BM, AK], or interpretation [EB, LK, CT] of data;
- 270 drafted the work [EB, KM, LK] or substantively revised it [MA, RH, AB, JA, SL, MB, PT].

### 271 Acknowledgements

- 272 This study makes use of data generated by the DECIPHER community[70]. A full list of centers who
- 273 contributed to the generation of the data is available from http://decipher.sanger.ac.uk and via email
- from decipher@sanger.ac.uk.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.02.20221481; this version posted November 4, 2020. 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-NC-ND 4.0 International license .

#### 275 References

276 Tilghman JM, Ling AY, Turner TN, Sosa MX, Krumm N, Chatterjee S, et al. Molecular Genetic 1. 277 Anatomy and Risk Profile of Hirschsprung's Disease. N Engl J Med. 2019;380(15):1421-32. PubMed 278 PMID: 30970187.

279 2. Alves MM, Sribudiani Y, Brouwer RW, Amiel J, Antinolo G, Borrego S, et al. Contribution of 280 rare and common variants determine complex diseases-Hirschsprung disease as a model. Dev Biol. 281 2013;382(1):320-9. Epub 2013/05/28. doi: 10.1016/j.ydbio.2013.05.019. PubMed PMID: 23707863. 282 Amiel J, Sproat-Emison E, Garcia-Barcelo M, Lantieri F, Burzynski G, Borrego S, et al. 3.

283 Hirschsprung disease, associated syndromes and genetics: a review. J Med Genet. 2008;45(1):1-14. 284 PubMed PMID: 17965226.

285 Fewtrell MS, Tam PK, Thomson AH, Fitchett M, Currie J, Huson SM, et al. Hirschsprung's 4. 286 disease associated with a deletion of chromosome 10 (q11.2q21.2): a further link with the neurocristopathies? J Med Genet. 1994;31(4):325-7. PubMed PMID: 7915329. 287

Martucciello G, Bicocchi MP, Dodero P, Lerone M, Silengo Cirillo M, Puliti A, et al. Total 288 5. 289 colonic aganglionosis associated with interstitial deletion of the long arm of chromosome 10. Pediatric 290 Surgery International. 1992;7(4):308-10. doi: 10.1007/bf00183991.

291 Edery P, Lyonnet S, Mulligan LM, Pelet A, Dow E, Abel L, et al. Mutations of the RET proto-6. 292 oncogene in Hirschsprung's disease. Nature. 1994;367(6461):378-80. PubMed PMID: 8114939.

293 Romeo G, Ronchetto P, Luo Y, Barone V, Seri M, Ceccherini I, et al. Point mutations affecting 7. 294 the tyrosine kinase domain of the RET proto-oncogene in Hirschsprung's disease. Nature. 295 1994;367(6461):377-8. PubMed PMID: 8114938.

296 Chatterjee S, Kapoor A, Akiyama JA, Auer DR, Lee D, Gabriel S, et al. Enhancer Variants 8. 297 Synergistically Drive Dysfunction of a Gene Regulatory Network In Hirschsprung Disease. Cell. 298 2016;167(2):355-68.e10. Epub 2016/10/04. doi: 10.1016/j.cell.2016.09.005. PubMed PMID: 299 27693352; PubMed Central PMCID: PMCPMC5113733.

300 9. Kapoor A, Jiang Q, Chatterjee S, Chakraborty P, Sosa MX, Berrios C, et al. Population 301 variation in total genetic risk of Hirschsprung disease from common RET, SEMA3 and NRG1 302 susceptibility polymorphisms. Human molecular genetics. 2015;24(10):2997-3003. Epub 2015/02/11. 303 doi: 10.1093/hmg/ddv051. PubMed PMID: 25666438; PubMed Central PMCID: PMCPMC4406299.

304 10. Jiang Q, Arnold S, Heanue T, Kilambi KP, Doan B, Kapoor A, et al. Functional loss of 305 semaphorin 3C and/or semaphorin 3D and their epistatic interaction with ret are critical to 306 Hirschsprung disease liability. Am J Hum Genet. 2015;96(4):581-96. Epub 2015/04/04. doi:

307 10.1016/j.ajhg.2015.02.014. PubMed PMID: 25839327; PubMed Central PMCID: PMCPMC4385176. 308 11. Garcia-Barcelo MM, Tang CS, Ngan ES, Lui VC, Chen Y, So MT, et al. Genome-wide 309 association study identifies NRG1 as a susceptibility locus for Hirschsprung's disease. Proc Natl Acad 310 Sci U S A. 2009;106(8):2694-9. Epub 2009/02/07. doi: 10.1073/pnas.0809630105. PubMed PMID: 311 19196962; PubMed Central PMCID: PMCPMC2650328.

312 Gui H, Tang WK, So MT, Proitsi P, Sham PC, Tam PK, et al. RET and NRG1 interplay in 12. 313 Hirschsprung disease. Hum Genet. 2013;132(5):591-600. Epub 2013/02/13. doi: 10.1007/s00439-013-314 1272-9. PubMed PMID: 23400839.

315 Tang CS, Tang WK, So MT, Miao XP, Leung BM, Yip BH, et al. Fine mapping of the NRG1 13. 316 Hirschsprung's disease locus. PLoS One. 2011;6(1):e16181. Epub 2011/02/02. doi:

10.1371/journal.pone.0016181. PubMed PMID: 21283760; PubMed Central PMCID:

317 318 PMCPMC3024406.

319 14. Bottani A, Xie YG, Binkert F, Schinzel A. A case of Hirschsprung disease with a chromosome 320 13 microdeletion, del(13)(q32.3q33.2): potential mapping of one disease locus. Hum Genet.

321 1991;87(6):748-50. PubMed PMID: 1937482.

322 Kiss P, Osztovics M. Association of 13q deletion and Hirschsprung's disease. J Med Genet. 15. 323 1989;26(12):793-4. PubMed PMID: 2614805.

324 16. Lamont MA, Fitchett M, Dennis NR. Interstitial deletion of distal 13g associated with

325 Hirschsprung's disease. J Med Genet. 1989;26(2):100-4. PubMed PMID: 2918536.

326 Shanske A, Ferreira JC, Leonard JC, Fuller P, Marion RW. Hirschsprung disease in an infant 17. with a contiguous gene syndrome of chromosome 13. Am J Med Genet. 2001;102(3):231-6. PubMed 327 328 PMID: 11484199.

329 Sparkes RS, Sparkes MC, Kalina RE, Pagon RA, Salk DJ, Disteche CM. Separation of 18.

retinoblastoma and esterase D loci in a patient with sporadic retinoblastoma and del(13)(q14.1q22.3). 330

331 Hum Genet. 1984;68(3):258-9. PubMed PMID: 6500578. medRxiv preprint doi: https://doi.org/10.1101/2020.11.02.20221481; this version posted November 4, 2020. 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

332 19. Puffenberger EG, Hosoda K, Washington SS, Nakao K, deWit D, Yanagisawa M, et al. A missense mutation of the endothelin-B receptor gene in multigenic Hirschsprung's disease. Cell. 333 334 1994;79(7):1257-66. Epub 1994/12/30. PubMed PMID: 8001158. Amiel J, Espinosa-Parrilla Y, Steffann J, Gosset P, Pelet A, Prieur M, et al. Large-scale 335 20. 336 deletions and SMADIP1 truncating mutations in syndromic Hirschsprung disease with involvement of 337 midline structures. Am J Hum Genet. 2001;69(6):1370-7. PubMed PMID: 11595972. 338 21. Lurie IW, Supovitz KR, Rosenblum-Vos LS, Wulfsberg EA. Phenotypic variability of del(2) 339 (q22-q23): report of a case with a review of the literature. Genet Couns. 1994;5(1):11-4. PubMed 340 PMID: 8031530. 341 McMilin KD, Reiss JA, Brown MG, Black MH, Buckmaster DA, Durum CT, et al. Clinical 22. 342 outcomes of four patients with microdeletion in the long arm of chromosome 2. Am J Med Genet. 343 1998;78(1):36-43. PubMed PMID: 9637421. 344 23. Mowat DR, Croaker GD, Cass DT, Kerr BA, Chaitow J, Ades LC, et al. Hirschsprung disease, 345 microcephaly, mental retardation, and characteristic facial features: delineation of a new syndrome and identification of a locus at chromosome 2q22-q23. J Med Genet. 1998;35(8):617-23. PubMed 346 PMID: 9719364. 347 348 Benailly HK, Lapierre JM, Laudier B, Amiel J, Attie T, De Blois MC, et al. PMX2B, a new 24. 349 candidate gene for Hirschsprung's disease. Clin Genet, 2003:64(3):204-9. PubMed PMID: 12919134. 350 Wakamatsu N, Yamada Y, Yamada K, Ono T, Nomura N, Taniguchi H, et al. Mutations in 25. 351 SIP1, encoding Smad interacting protein-1, cause a form of Hirschsprung disease. Nat Genet. 352 2001;27(4):369-70. Epub 2001/03/30. doi: 10.1038/86860. PubMed PMID: 11279515. Amiel J, Laudier B, Attie-Bitach T, Trang H, de Pontual L, Gener B, et al. Polyalanine 353 26. 354 expansion and frameshift mutations of the paired-like homeobox gene PHOX2B in congenital central 355 hypoventilation syndrome. Nat Genet. 2003;33(4):459-61. Epub 2003/03/18. doi: 10.1038/ng1130. 356 PubMed PMID: 12640453. 357 Amiel J, Lyonnet S. Hirschsprung disease, associated syndromes, and genetics: a review. J 27. 358 Med Genet. 2001;38(11):729-39. PubMed PMID: 11694544. Brewer C, Holloway S, Zawalnyski P, Schinzel A, FitzPatrick D. A chromosomal deletion map 359 28. 360 of human malformations. Am J Hum Genet. 1998;63(4):1153-9. PubMed PMID: 9758599. Brewer C, Holloway S, Zawalnyski P, Schinzel A, FitzPatrick D. A chromosomal duplication 361 29. map of malformations: regions of suspected haplo- and triplolethality--and tolerance of segmental 362 363 aneuploidy--in humans. Am J Hum Genet. 1999;64(6):1702-8. PubMed PMID: 10330358. Mahboubi S, Templeton JM, Jr. Association of Hirschsprung's disease and imperforate anus in 364 30. 365 a patient with "cat-eye" syndrome. A report of one case and review of the literature. Pediatr Radiol. 366 1984;14(6):441-2. PubMed PMID: 6504608. 367 31. Kerstjens-Frederikse WS, Hofstra RM, van Essen AJ, Meijers JH, Buys CH. A Hirschsprung 368 disease locus at 22g11? J Med Genet. 1999;36(3):221-4. PubMed PMID: 10204849. 369 32. Coe BP, Witherspoon K, Rosenfeld JA, van Bon BW, Vulto-van Silfhout AT, Bosco P, et al. 370 Refining analyses of copy number variation identifies specific genes associated with developmental 371 delay. Nat Genet. 2014;46(10):1063-71. PubMed PMID: 25217958. 372 Cooper GM, Coe BP, Girirajan S, Rosenfeld JA, Vu TH, Baker C, et al. A copy number 33. 373 variation morbidity map of developmental delay. Nat Genet. 2011;43(9):838-46. PubMed PMID: 374 21841781. 375 Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of 34. 376 protein-coding genetic variation in 60,706 humans. Nature. 2016;536(7616):285-91. PubMed PMID: 377 27535533. 378 35. Ruderfer DM, Hamamsy T, Lek M, Karczewski KJ, Kavanagh D, Samocha KE, et al. Patterns 379 of genic intolerance of rare copy number variation in 59,898 human exomes. Nat Genet. 380 2016;48(10):1107-11. Epub 2016/08/18. doi: 10.1038/ng.3638. PubMed PMID: 27533299; PubMed Central PMCID: PMCPMC5042837. 381 Gui H, Schriemer D, Cheng WW, Chauhan RK, Antinolo G, Berrios C, et al. Whole exome 382 36. 383 sequencing coupled with unbiased functional analysis reveals new Hirschsprung disease genes. 384 Genome Biol. 2017;18(1):48. Epub 2017/03/10. doi: 10.1186/s13059-017-1174-6 385 10.1186/s13059-017-1174-6 [pii]. PubMed PMID: 28274275; PubMed Central PMCID: PMC5343413. 386 37. Tang CS, Li P, Lai FP, Fu AX, Lau ST, So MT, et al. Identification of Genes Associated With Hirschsprung Disease, Based on Whole-Genome Sequence Analysis, and Potential Effects on Enteric 387 388 Nervous System Development. Gastroenterology. 2018;155(6):1908-22.e5. Epub 2018/09/16. doi: 389 10.1053/j.gastro.2018.09.012. PubMed PMID: 30217742.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.02.20221481; this version posted November 4, 2020. 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

perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

390 38. Rosenfeld JA, Coe BP, Eichler EE, Cuckle H, Shaffer LG. Estimates of penetrance for recurrent pathogenic copy-number variations. Genet Med. 2013;15(6):478-81. Epub 2012/12/22. doi: 391 392 10.1038/gim.2012.164. PubMed PMID: 23258348; PubMed Central PMCID: PMCPMC3664238. Ballif BC, Theisen A, Rosenfeld JA, Traylor RN, Gastier-Foster J, Thrush DL, et al. 393 39. Identification of a recurrent microdeletion at 17q23.1q23.2 flanked by segmental duplications 394 395 associated with heart defects and limb abnormalities. Am J Hum Genet. 2010;86(3):454-61. PubMed 396 PMID: 20206336. 397 40. Laurell T, Lundin J, Anderlid BM, Gorski JL, Grigelioniene G, Knight SJ, et al. Molecular and 398 clinical delineation of the 17q22 microdeletion phenotype. Eur J Hum Genet. 2013;21(10):1085-92. 399 PubMed PMID: 23361222. 400 Tang CS, Gui H, Kapoor A, Kim JH, Luzon-Toro B, Pelet A, et al. Trans-ethnic meta-analysis 41. 401 of genome-wide association studies for Hirschsprung disease. Human molecular genetics. 402 2016;25(23):5265-75. Epub 2016/10/06. doi: 10.1093/hmg/ddw333. PubMed PMID: 27702942. de Pontual L, Pelet A, Trochet D, Jaubert F, Espinosa-Parrilla Y, Munnich A, et al. Mutations 403 42. 404 of the RET gene in isolated and syndromic Hirschsprung's disease in human disclose major and 405 modifier alleles at a single locus. J Med Genet. 2006;43(5):419-23. Epub 2006/01/31. doi: 406 10.1136/jmg.2005.040113. PubMed PMID: 16443855; PubMed Central PMCID: PMCPMC2649010. 407 de Pontual L. Pelet A. Clement-Ziza M. Trochet D. Antonarakis SE. Attie-Bitach T. et al. 43. 408 Epistatic interactions with a common hypomorphic RET allele in syndromic Hirschsprung disease. 409 Hum Mutat. 2007;28(8):790-6. Epub 2007/04/03. doi: 10.1002/humu.20517. PubMed PMID: 410 17397038. 44. 411 de Pontual L, Zaghloul NA, Thomas S, Davis EE, McGaughey DM, Dollfus H, et al. Epistasis 412 between RET and BBS mutations modulates enteric innervation and causes syndromic Hirschsprung 413 disease. Proc Natl Acad Sci U S A. 2009;106(33):13921-6. Epub 2009/08/12. doi: 10.1073/pnas.0901219106. PubMed PMID: 19666486; PubMed Central PMCID: PMCPMC2728996. 414 415 Arnold S, Pelet A, Amiel J, Borrego S, Hofstra R, Tam P, et al. Interaction between a 45. 416 chromosome 10 RET enhancer and chromosome 21 in the Down syndrome-Hirschsprung disease 417 association. Hum Mutat. 2009;30(5):771-5. Epub 2009/03/24. doi: 10.1002/humu.20944. PubMed 418 PMID: 19306335; PubMed Central PMCID: PMCPMC2779545. 419 46. Jannot AS, Pelet A, Henrion-Caude A, Chaoui A, Masse-Morel M, Arnold S, et al. 420 Chromosome 21 scan in Down syndrome reveals DSCAM as a predisposing locus in Hirschsprung 421 disease. PLoS One. 2013;8(5):e62519. Epub 2013/05/15. doi: 10.1371/journal.pone.0062519. PubMed PMID: 23671607; PubMed Central PMCID: PMCPMC3646051. 422 423 Natarajan D, Marcos-Gutierrez C, Pachnis V, de Graaff E. Requirement of signalling by 47. 424 receptor tyrosine kinase RET for the directed migration of enteric nervous system progenitor cells 425 during mammalian embryogenesis. Development. 2002;129(22):5151-60. Epub 2002/10/26. PubMed 426 PMID: 12399307. 427 48. Yamagishi H, Garg V, Matsuoka R, Thomas T, Srivastava D. A molecular pathway revealing a 428 genetic basis for human cardiac and craniofacial defects. Science. 1999;283(5405):1158-61. PubMed 429 PMID: 10024240. 430 Hendershot TJ, Liu H, Sarkar AA, Giovannucci DR, Clouthier DE, Abe M, et al. Expression of 49. 431 Hand2 is sufficient for neurogenesis and cell type-specific gene expression in the enteric nervous 432 system. Dev Dyn. 2007;236(1):93-105. PubMed PMID: 17075884. 433 D'Autreaux F, Margolis KG, Roberts J, Stevanovic K, Mawe G, Li Z, et al. Expression level of 50. Hand2 affects specification of enteric neurons and gastrointestinal function in mice. Gastroenterology. 434 435 2011;141(2):576-87, 87 e1-6. PubMed PMID: 21669203. 436 Pan L, Ma X, Wen B, Su Z, Zheng X, Liu Y, et al. Microphthalmia-associated transcription 51. 437 factor/T-box factor-2 axis acts through Cyclin D1 to regulate melanocyte proliferation. Cell Prolif. 438 2015;48(6):631-42. Epub 2015/10/22. doi: 10.1111/cpr.12227. PubMed PMID: 26486273. 439 Stanchina L, Baral V, Robert F, Pingault V, Lemort N, Pachnis V, et al. Interactions between 52. Sox10, Edn3 and Ednrb during enteric nervous system and melanocyte development. Dev Biol. 440 441 2006;295(1):232-49. Epub 2006/05/03. doi: 10.1016/j.ydbio.2006.03.031. PubMed PMID: 16650841. 442 53. Watanabe Y, Stanchina L, Lecerf L, Gacem N, Conidi A, Baral V, et al. Differentiation of 443 Mouse Enteric Nervous System Progenitor Cells Is Controlled by Endothelin 3 and Requires 444 Regulation of Ednrb by SOX10 and ZEB2. Gastroenterology. 2017;152(5):1139-50.e4. Epub 2017/01/09. doi: 10.1053/j.gastro.2016.12.034. PubMed PMID: 28063956. 445 Miao X, Leon TY, Ngan ES, So MT, Yuan ZW, Lui VC, et al. Reduced RET expression in gut 446 54. 447 tissue of individuals carrying risk alleles of Hirschsprung's disease. Human molecular genetics.

448 2010;19(8):1461-7. Epub 2010/01/22. doi: 10.1093/hmg/ddq020. PubMed PMID: 20089534.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.02.20221481; this version posted November 4, 2020. 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.

449 55. Jiang Q, Ho YY, Hao L, Nichols Berrios C, Chakravarti A. Copy number variants in candidate 450 genes are genetic modifiers of Hirschsprung disease. PLoS One. 2011;6(6):e21219. Epub 451 2011/06/30. doi: 10.1371/journal.pone.0021219. PubMed PMID: 21712996; PubMed Central PMCID: 452 PMCPMC3119685. Tang CS, Cheng G, So MT, Yip BH, Miao XP, Wong EH, et al. Genome-wide copy number 453 56. 454 analysis uncovers a new HSCR gene: NRG3. PLoS Genet. 2012;8(5):e1002687. Epub 2012/05/17. 455 doi: 10.1371/journal.pgen.1002687. PubMed PMID: 22589734; PubMed Central PMCID: 456 PMCPMC3349728. 457 57. Brosens E, Marsch F, de Jong EM, Zaveri HP, Hilger AC, Choinitzki VG, et al. Copy number 458 variations in 375 patients with oesophageal atresia and/or tracheoesophageal fistula. Eur J Hum 459 Genet. 2016;24(12):1715-23. PubMed PMID: 27436264. 460 Lai FP, Lau ST, Wong JK, Gui H, Wang RX, Zhou T, et al. Correction of Hirschsprung-58. Associated Mutations in Human Induced Pluripotent Stem Cells Via Clustered Regularly Interspaced 461 Short Palindromic Repeats/Cas9, Restores Neural Crest Cell Function. Gastroenterology. 462 463 2017;153(1):139-53 e8. doi: 10.1053/j.gastro.2017.03.014. PubMed PMID: 28342760. Brouwer RWW, van den Hout M, Kockx CEM, Brosens E, Eussen B, de Klein A, et al. 464 59. 465 Nimbus: A design-driven analyses suite for amplicon based NGS data. Bioinformatics. 2018. Epub 466 2018/03/15. doi: 10.1093/bioinformatics/btv145. PubMed PMID: 29538618. 467 Brouwer RW, van den Hout MC, Grosveld FG, van licken WF. NARWHAL, a primary analysis 60. 468 pipeline for NGS data. Bioinformatics. 2012;28(2):284-5. Epub 2011/11/11. doi: 469 10.1093/bioinformatics/btr613. PubMed PMID: 22072383. 470 61. Halim D, Brosens E, Muller F, Wangler MF, Beaudet AL, Lupski JR, et al. Loss-of-Function 471 Variants in MYLK Cause Recessive Megacystis Microcolon Intestinal Hypoperistalsis Syndrome. Am J 472 Hum Genet. 2017;101(1):123-9. Epub 2017/06/13. doi: 10.1016/j.ajhg.2017.05.011. PubMed PMID: 473 28602422; PubMed Central PMCID: PMCPMC5501771. 474 Schriemer D, Sribudiani Y, A IJ, Natarajan D, MacKenzie KC, Metzger M, et al. Regulators of 62. 475 gene expression in Enteric Neural Crest Cells are putative Hirschsprung disease genes. Dev Biol. 476 2016;416(1):255-65. PubMed PMID: 27266404. 477 Memic F, Knoflach V, Morarach K, Sadler R, Laranjeira C, Hjerling-Leffler J, et al. 63. 478 Transcription and Signaling Regulators in Developing Neuronal Subtypes of Mouse and Human 479 Enteric Nervous System. Gastroenterology. 2018;154(3):624-36. PubMed PMID: 29031500. 480 McCann CJ, Alves MM, Brosens E, Natarajan D, Perin S, Chapman C, et al. Neuronal 64. 481 Development and Onset of Electrical Activity in the Human Enteric Nervous System. 482 Gastroenterology. 2019. Epub 2019/01/06. doi: 10.1053/j.gastro.2018.12.020. PubMed PMID: 30610864. 483 484 65. Zhan X, Hu Y, Li B, Abecasis GR, Liu DJ. RVTESTS: an efficient and comprehensive tool for 485 rare variant association analysis using sequence data. Bioinformatics. 2016;32(9):1423-6. PubMed 486 PMID: 27153000. 487 Nechiporuk A, Linbo T, Poss KD, Raible DW. Specification of epibranchial placodes in 66. 488 zebrafish. Development. 2007;134(3):611-23. PubMed PMID: 17215310. 489 Kuil LE, Oosterhof N, Geurts SN, van der Linde HC, Meijering E, van Ham TJ. Reverse 67. 490 genetic screen reveals that II34 facilitates yolk sac macrophage distribution and seeding of the brain. Dis Model Mech. 2019;12(3). PubMed PMID: 30765415. 491 Hoshijima K, Jurynec MJ, Klatt Shaw D, Jacobi AM, Behlke MA, Grunwald DJ. Highly Efficient 492 68. 493 CRISPR-Cas9-Based Methods for Generating Deletion Mutations and F0 Embryos that Lack Gene 494 Function in Zebrafish. Dev Cell. 2019;51(5):645-57 e4. PubMed PMID: 31708433. 495 69. Heanue TA, Pachnis V. Ret isoform function and marker gene expression in the enteric 496 nervous system is conserved across diverse vertebrate species. Mech Dev. 2008;125(8):687-99. 497 PubMed PMID: 18565740. 498 Firth HV, Richards SM, Bevan AP, Clayton S, Corpas M, Rajan D, et al. DECIPHER: 70. 499 Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. Am J

- 500 Hum Genet. 2009;84(4):524-33. PubMed PMID: 19344873.
- 501

#### 502 **Figure legends**

503 Figure 1. Schematic overview of our overall study design and methods used. a) In brief: We 504 determined the Copy Number profiles of different subgroups of HSCR patients (1-3) and controls (4). 505 We determined RET and / or known disease gene coding mutations and ranked the genes affected by 506 a CNV according to frequency in controls, expression in the developing mouse ENS and gene tolerance 507 to variation. Next, we determined whether disruption of the main candidate genes resulted in a reduction 508 of enteric neurons in zebrafish. b) In parallel, we evaluated the contribution of different genetic risk 509 factors, by comparing the contribution of known non-coding predisposing haplotypes across groups. o/e 510 = overexpressed, VI = intolerant to variation.

511 Figure 2. ENS overexpressed genes intolerant to variation were enriched in HSCR-complex 512 patients and their disruption in zebrafish caused HSCR phenotypes a) Graph showing CNV size, 513 each dot represents a CNV. b) graph showing the number of variant intolerant (VI) genes overexpressed 514 in the mouse ENS per patient, each dot represents one patient. c) Graph showing the RSnc, each dot 515 represents one patient. Error bars represent standard deviation, statistical analysis used: students t-test 516 (a-c). d) CRISPR/Cas9 complex injections in zebrafish fertilized eggs induced HSCR phenotypes upon 517 disruption of four genes (accumulated data from multiple experiments). e) Graph showing accumulated 518 data of the percentage of fish with HSCR phenotypes upon injection of a morpholino targeting Ret 519 translation at a concentration that induced HSCR phenotypes in approximately 50% of the fish. Ret 520 morpholino injections in combination with disruption of mapk8 shows epistasis. Disruption of gnl1 521 showed a trend towards higher penetrance of HSCR phenotypes. f) A second injection round targeting 522 gnl1, using a lower dose of ret morpholino, confirmed gnl1 epistasis with Ret.

523 Figure 3. Complex HSCR genetics: genetic predispositions of HSCR patient groups a) Visual 524 representation of the distribution of genetic predispositions over HSCR patient groups. In total 197 525 patients born between 1973 and 2018 were evaluated by a clinical geneticist in the department of Clinical Genetics, Erasmus Medical Center, Rotterdam. Of these, 114 did not have associated 526 527 anomalies nor a known syndrome. 29 patients had a known HSCR related genetic syndrome, including Down syndrome (n=18). 153 out of 197 patients were genetically evaluated for RET gene involvement 528 529 and 21 had a pathogenic RET variant. b) Pie charts showing the incidence of rare CNVs containing 530 genes overexpressed in the developing mouse ENS and coding variants in HSCR patients. c) Graphical

- 531 representation of a hypothetical model explaining the relative contributions of the risk scores in our 3
- 532 patient groups. Error bars represent standard deviation.

### 533 Table 1. Genes in rare CN losses that are overexpressed in mouse ENS and intolerant to

- 534 genetic variation Genes marked with an # also have a loss of function variant in an independent
- 535 HSCR cohort (see table 2). Depicted are the RSnc (see S11) and the CNV and variant intolerance
- 536 scores derived from http://gnomad.broadinstitute.org/ and http://exac.broadinstitute.org/about.
- 537

## 538 Table 2: WES: Nonsense and splice site variants in rare CNV genes

- 539 Rare putative deleterious nonsense variants and variants predicted to affect splicing in a whole exome
- 540 sequencing cohort of HSCR (n=76, 149 controls) [36] and whole genome sequencing cohort of 443
- short segment HSCR patients and 493 unaffected controls[37]. Variants in genes intolerant to
- 542 variation that were also impacted by the de novo 17q23.1 q23.2 loss (INTS2, MED13), the de novo
- 543 6p22.1 p21.33 loss (PRRC2A, TUBB), the 9p21 loss (LINGO2), the maternal inherited 10q11.22 -
- 544 q11.23 loss (SGMS1), de novo 7q36.1 gain (KMT2C) and the 3q24 gain (SLC6A6).

			Ex	AC/ Gn	omAD intolerance s		
Patient	RS <sub>NC</sub>	Gene	Missense Z	pLl	Deletion single	Deletion CNV	CNV region
P_000479	4,16	SLC8A1	2.23	1.00	-0.02	0.31	chr2:40,624,267-40,646,501
P_000512	8,31	TUBB <sup>#</sup>	5.71	0.98	-2.85	-1.72	chr6:28,005,012-31,683,185
		GNL1	2.52	1.00	1.03	0.70	
		GABBR1	4.98	1.00	1.36	1.23	
P_000537	6,92	MAPK8	2.92	1.00	0.84	-2.25	chr10: 49,033,586 -52,431,193
P_000561	6,08	UFD1L	2.77	1.00	1.06	-2.53	chr22:18,861,209-21,630,630
P_000567	7,98	TBX2	1.75	0.99	-0.01	0.53	chr17:58,076,721-60,362,868
		USP32	3.55	1.00	-2.85	-0.93	
P_002431	9,37	AKT3	4.03	1.00	-2.61	-1.20	chr1:243,963,527-244,016,804

# Table 1. Genes in rare CN losses that are overexpressed in mouse ENS and intolerant to genetic variation

Sample	Chr	Start	Stop	Ref	Alt	Exon	Gene	Туре	location	Effect	HGVS cDNA-level	CADD	gnomADv2.1	gnomADV2.1
													Exomes	Genomes
SE14-0527	17	59946466	59946466	•	Т	23	INTS2	insertion	exonic	frameshift	NM_001330417.1:c.3172dupA	·	0	0
HK19-0006	17	60072727	60072727	С	Т	10	MED13	snv	intronic	splicing	NM_005121:c.1968-1G>A	22.4	0	0
SE16-3114	17	59469360	59469360	С	т	26	BCAS3	snv	exonic	stopgain	NM_001320470.1:c.2773C>T	18.4	0.0002	0.0007
SE14-0656	6	31604286	31604286	G	т	27	PRRC2A	snv	splicing	splicing	NM_004638.3:c.5836-1G>T	23.6	0.0001	3.24E-05
SE16-3123	6	30692109	30692110	CA	A	4	TUBB	substitution	exonic	frameshift	NM_001293212.1:c.1330_133 1delCAinsA		0	0
SE17-3220	7	151962296	151962296	Т	С	8	KMT2C	snv	splicing	splicing	NM_170606.2:c.1013-2A>G	22	8.21E-06	0
HK19-0002	9	28476025	28476025	Т	G		LINGO2	snv	intronic	splicing	NM_001258282:c395-2A>C	23.2	0	0
HK19-0003	10	52103343	52103344	TA	Т	7	SGMS1	deletion	exonic	frameshift	NM_147156:c.T529delT- :p.F177del	33	0	0
HK19-0004	10	52104106	52104108	CTG	С		SGMS1	deletion	intronic	splicing	NM_147156:c313-2CAG>G	24.5	0	0
HK19-0005	10	52349911	52349911	А	G		SGMS1	snv	intronic	splicing	NM_147156:c683+2T>C	23	0	0
SE16-3109	22	21065731	21065731	G	А	51	PI4KA	snv	exonic	stopgain	NM_058004.3:c.5821C>T	51	0.0002	0.0002
HK19-0001	3	14485130	14485130	А	G		SLC6A6	snv	intronic	splicing	NM_001134367:c.297-6A>G	15.92	0	0

# Table 2: WES: Nonsense and splice site variants in rare CNV genes in HSCR patients



а



medRxiv preprint doi: https://doi.org/10.1101/2020.11.02.20221481; this version posted November 4, 2020. 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. re 2 It is made available under a CC-BY-NC-ND 4.0 International license . Figure 2 b С а # of VI genes o/e in ENS in CN loss 4 10 \*\*\* 6,000 8 3 RS<sub>NC</sub> 6 2 . . . . 4 1 2 0 0 0 groups oroup 3 dioup 5 oroup2 group 2 oroup group2 groups group group group group d ribonucleoprotein complex normal 5 dpf microinjection (crRNA:tracrRNA:Cas9) hypogangliosis

ultra-short segment HSCR

short segment HCSR

total colonic HSCR

g,

Bercentage of fish

100

е

\*\* \*\*







Rare CNV containing ENS o/e gene(s)

Rare CNV containing VI ENS o/e gene(s)



- HSCR coding variant
- Rare CNV containing ENS o/e gene(s)
- Rare CNV containing VI ENS o/e gene(s)
- No coding variants in RET or rare CNVs
- No coding variants in RET, no CN prolifing done

