

RESEARCH

Open Access



Sequence characterization of *RET* in 117 Chinese Hirschsprung disease families identifies a large burden of de novo and parental mosaic mutations

Qian Jiang^{1†}, Yang Wang^{2†}, Qi Li³, Zhen Zhang³, Ping Xiao⁴, Hui Wang¹, Na Liu⁵, Jian Wu⁵, Feng Zhang⁶, Aravinda Chakravarti⁷, Wei Cai^{2*†} and Long Li^{3*†}

Abstract

Background: Hirschsprung disease (HSCR) is an inherited congenital disorder characterized by the absence of enteric ganglia in the distal part of the gut. *RET* is the major causative gene and contains > 80% of all known disease-causing mutations.

Results: To determine the incidence of *RET* pathogenic variants, be they Mendelian inherited, mosaic in parents or true de novo variants (DNVs) in 117 Chinese families, we used high-coverage NGS and droplet digital polymerase chain reaction (ddPCR) to identify 15 (12.8%) unique *RET* coding variants (7 are novel); one was inherited from a heterozygous unaffected mother, 11 were DNVs (73.3%), and 3 full heterozygotes were inherited from parental mosaicism (2 paternal, 1 maternal): two clinically unaffected parents were identified by NGS and confirmed by ddPCR, with mutant allele frequency (13–27%) that was the highest in hair, lowest in urine and similar in blood and saliva. An extremely low-level paternal mosaicism (0.03%) was detected by ddPCR in blood. Six positive-controls were examined to compare the mosaicism detection limit and sensitivity of NGS, amplicon-based deep sequencing and ddPCR.

Conclusion: Our findings expand the clinical and molecular spectrum of *RET* variants in HSCR and reveal a high frequency of *RET* DNVs in the Chinese population.

Keywords: *RET*, Parental mosaicism, High coverage NGS, ddPCR, Chinese HSCR

Introduction

Pathogenic gene variation is a significant contributor to rare diseases, especially in children [1]. Thus, many genetic mutations of early development are inherited by children from their parents through the germline and are present in all cells of that individual, while others, mosaic or somatic mutations, may be acquired postzygotically

and are present in only a subset of an individual's cells [2]. It has long been known that cancer is a mosaic genetic disorder. However, a growing body of research suggests that analogous mosaicism may be a frequent feature in a diverse range of childhood disorders, including cerebral cortical malformations, autism spectrum disorder, epilepsies and other neuropsychiatric diseases [3–6]. In a previous study of Hirschsprung disease (HSCR) families, we identified mosaicism in 6 of 8 (75%) isolated cases [7]. This high frequency was surprising and prompted us to further investigate the frequency and nature of *RET* mosaic pathogenic variants.

HSCR or congenital aganglionosis, a heterogeneous genetic disorder, is characterized by the lack of ganglion cells along varying lengths of the intestine resulting in

* Correspondence: caiw204@sjtu.edu.cn; lilong23@126.com

[†]Qian Jiang, Yang Wang, Wei Cai and Long Li contributed equally to this work.

²Department of Pediatric Surgery, Xinhua Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai Key Laboratory of Pediatric Gastroenterology and Nutrition, Shanghai Institute for Pediatric Research, No. 1665 Kongjiang Rd., Yangpu District, Shanghai 200092, China

³Department of General Surgery, Capital Institute of Pediatrics Affiliated Children's Hospital, No. 2 Yabao Rd., Chaoyang District, Beijing 100020, China Full list of author information is available at the end of the article



the major cause of functional obstruction in children. According to the length of aganglionosis, the disorder is categorized into three types: short-segment (aganglionosis segment up to the upper sigmoid colon), long-segment (aganglionosis beyond the splenic flexure) and total colonic aganglionosis (TCA) [8]. The incidence of HSCR varies and is 15, 21 and 28 cases per 100,000 live births in infants with European, African and Asian ancestry, respectively. Genetic studies during the past 25 years have identified rare coding variants in 14 genes that together explain ~ 10% of HSCR cases [9–11]. Of these, the most frequent coding mutations occur in *RET*, which encodes a receptor tyrosine kinase that regulates the proliferation, differentiation and migration of the enteric neural crest cells to enteric neurons. However, family studies of these pathogenic variants demonstrate incomplete penetrance and variable expressivity, the causes of which remain largely unexplained [9, 12].

Numerous studies of *RET* pathogenic variants in HSCR show that they occur in 8.9–16.7% of cases with a contribution from de novo variants (DNVs) which occur in the parental germline [13, 14]. However, family studies of these variants are infrequent so that the distribution of Mendelian inherited versus DNVs is unknown, making risk prediction and genetic counselling of HSCR uncertain. Here, we set out to perform a prospective study of 117 HSCR parent-affected child trios to determine the frequency of *RET* Mendelian inherited, parental mosaic or true DNVs. Furthermore, we explored the mutant allele distribution patterns in multiple somatic tissues and gonadal tissue, and compared the detection accuracy of three commonly used molecular methods.

Subjects and methods

Subjects

One hundred and eighteen children diagnosed with isolated HSCR (85/33 male/female, 69/23/26 S-HSCR/L-HSCR/TCA; aged 2–18 months, mean = 16.1 months) from 117 pedigrees were recruited and studied here for the first time, together with their parents and siblings (357 individuals in total). Blood samples were collected from each child, their parents and siblings, and genomic DNA was isolated. Genomic DNA from multiple peripheral tissues, including saliva, urine, hair follicles and sperm, when available, was extracted using the TIANamp Micro DNA Kit (Tiagen Biotech, Beijing, China). Paternity testing was performed on a ProFlex PCR System (Applied Biosystems, USA) using the multiplex STR markers from the AmpFLSTR® Identifier Plus Amplification Kit (Applied Biosystems, USA).

Genetic analysis

The coding region of *RET* (RefSeq NM_020975.5) and its annotated functional noncoding elements (putative enhancers, promoters, untranslated regions, exon-intron

boundaries ranging from – 50 to + 50 bp, etc.) were enriched from genomic DNA using a GenCap Custom Enrichment Kit (MyGenostics, Beijing, China) [15] as previously described. After sequencing, low-quality reads were filtered out, and adaptor sequences were removed using cutadapt software (<http://code.google.com/p/cutadapt/>, v1.9.1). Next, we used BWA to align reads to the human reference genome (hg19). After removing duplicates with Picard (v2.2.3), single-nucleotide variants (SNV) and small insertions/deletions (INDEL) were identified using the GATK HaplotypeCaller program (v3.7) and VarScan (v2.3.7). We annotated the identified SNVs and INDELS using ANNOVAR (<http://annovar.openbioinformatics.org/en/latest/>). Short read alignment and candidate SNP and INDEL validation were performed using IGV. To select putative DNVs, the following criteria were used: 1) minimal 10X coverage in patients and parents; 2) a minimal genotype quality score of 10 for both patients and parents; 3) at least 10% of the reads showing the alternative allele in patients; and 4) not more than 10% of the reads showing the alternative allele in parents. To predict whether a missense change is damaging to the resultant protein function or structure, the following criteria were used: the evolutionary conservation of an amino acid with GERP, the location and context within the protein sequence with InterPro, and the biochemical consequence of the amino acid substitution using SIFT, PolyPhen and MutationTaster.

Quantification of mosaicism

To validate and quantify putative mosaic events, ~ 12 ng of DNA was used per ddPCR reaction, using previously described methods [16, 17]. Analysis was performed using QuantaSoft software with wells < 8000 total droplets excluded from analysis. Mutant (FAM) and wild-type (HEX) droplet fluorescence were read on the QX200™ Droplet Digital™ PCR System. Alternate allele frequency was calculated as the percentage of mutant-positive droplets divided by the total number of DNA-containing droplets. Multiple wells were merged for analysis, and Poisson confidence intervals were defined using QuantaSoft software (Bio-Rad, Hercules, CA). Samples were deemed “positive” when the 95% Poisson confidence intervals did not overlap the wild-type negative control. Although some samples showed a few positive droplets, they were still deemed negative when their 95% confidence intervals overlapped with wild-type results. Additional statistical analysis was performed in R-Studio (Boston, MA).

Results

Novel *RET* coding-region variants detected in 117 families with HSCR

On average, 823.3 million cleaned reads of 100-bp length were generated per sample, except for XHYY019, a male

Table 1 *RET* sequence variants in 15 Chinese HSCR probands with three mosaic events highlighted in red

Family #	Subject ID	Sex	Phenotype ^a	Coordinates (hg19)	Protein change ^b	Type of variant ^c	Previously reported ^d	Frequency in CMDB ^e	Frequency in gnomAD ^f	Inheritance	ddPCR on parents (Alt/Ref) ^g
1	HSCRFM197 (proband)	Female	TCA	chr10: 43595928 C>T	p.(Ser32Leu)	Missense (LP)	Yes	0	0	Paternal mosaicism	F: 1.2/3740 M: 0/3580
	HSCRFM198 (twin sister)	Female	TCA	chr10: 43595928 C>T	p.(Ser32Leu)	Missense (LP)	Yes	0	0	Paternal mosaicism	
2	XHY057 (proband)	Female	L	chr10: 43596110 G>A	p.(Gly93Ser)	Missense (LP)	Yes	0	0	de novo	F: 0/3380 M: 0/3780
3	HSCRFM191 (proband)	Female	S	chr10: 43597990 C>T	p.(Arg180*)	Nonsense (LP)	Yes	0	0	de novo	F: 0/4420 M: 0/4460
4	XHY022 (proband)	Male	S	chr10: 43600465 C>T	p.(Arg231Cys)	Missense (LP)	No	0	0	de novo	F: 0/3760 M: 0/3800
5	XHY051 (proband)	Female	L	chr10: 43600618 dupT	p.(Val282Valfs*71)	Frameshift (LP)	No	0	0	Maternal mosaicism	F: 0/4000 M: 854/2760
6	XHY087 (proband)	Male	S	chr10: 43604677 A>C	p.(Gln421Pro)	Missense (LP)	No	0	0	de novo	F: 0/3420 M: 0/3260
7	HSCRFM233 (proband)	Male	L	chr10: 43606856 G>A	p.(Asp489Asn)	Missense (LP)	Yes	0.022	0.00209	Mother	NA
8	HSCRFM181 (proband)	Female	TCA	chr10: 43609058 G>A	p.(Gly605Asp)	Missense (LP)	No	0	0	de novo	F: 0/241 M: 0/235
9	XHY093 (proband)	Female	TCA	chr10: 43612087 G>A	p.(Gly731Glu)	Missense (LP)	No	0	0	de novo	F: 0/3840 M: 0/3500
10	HSCRFM072 (proband)	Male	TCA	chr10: 43613844 C>T	p.(Arg770*)	Nonsense (LP)	Yes	0	0	Paternal mosaicism	F: 338/2300 M: 0/3400
11	HSCRFM230 (proband)	Male	TCA	chr10: 43615611 G>A	p.(Arg897Gln)	Missense (LP)	Yes	0	0	de novo	F: 0/4140 M: 0/4320
12	HSCRFM075 (proband)	Male	TCA	chr10: 43615611 G>A	p.(Arg897Gln)	Missense (LP)	Yes	0	0	de novo	F: 0/4440 M: 0/4280
13	HSCRFM024 (proband)	Male	S	chr10: 43622168 A>G	p.(Tyr1062Cys)	Missense (LP)	Yes	0	1.45e-5	de novo	F: 0/3900 M: 0/3400
14	HSCRFM007 (proband)	Female	TCA	chr10: 43623639 A>G	p.(Arg1089Arg)	Synonymous	No	0	0	de novo	F: 0/4100 M: 0/4000
15	HSCRFM156 (proband)	Female	TCA	chr10: 43615526 C>G	c. 2608-3C>G	Splicing	No	0	0	de novo	F: 0/4120 M: 0/3140

^a S, L, TCA, short-segment, long-segment and total colonic aganglionosis.

^b RefSeq NM_020975.5, NP_066124.1.

^c LP, Likely pathogenic according to the ACMG guideline.

^d No, novel mutations.

^e CMDB, The Chinese Millionome Database: <https://db.cngb.org/cmdb/>

^f gnomAD, genome Aggregation Database: <http://gnomad.broadinstitute.org/>

^g F, M, father and mother; NA, not available; all listed are ddPCR test results on blood samples.

patient with short segment HSCR, which had 74.3 million cleaned reads of 100-bp length. We achieved a minimum of 20-fold coverage per base on average for 99.7% of the target region at a mean coverage of 2962 reads (Basic QC metrics are shown in Additional file 1: Table S1, Table S2). Altogether 16 patients (from 15 families) were discovered to carry *RET* coding-region variants, out of 118 cases (13.6%) but two of these were full siblings. Thus, the variant detection frequency is 15/117 or 12.8%. Note that, two independent probands had the same variant (p. Arg897Gln) and the 14 unique variants consisted of 2 nonsense (p. Arg180*, p. Arg770*), 1 frameshift (p. Val282Valfs*71), 1 splicing (c. 2608-3C>G), 9 missense (p. Ser32Leu, p. Gly93Ser, p. Arg231Cys, p. Gln421Pro, p. Asp489Asn, p. Gly605Asp, p. Gly731Glu, p. Arg897Gln, p. Tyr1062Cys) and 1 synonymous (p. Arg1089Arg) variant. 12 of these changes are absent in both CMDB and gnomAD databases, one (p. Tyr1062Cys) is absent in CMDB and has a very low frequency (1.45×10^{-5}) in gnomAD, while the

last (p. Asp489Asn) may be common (~ 2% in CMDB and 0.2% in gnomAD). Half of these variants (p. Val282Valfs*71, c. 2608-3C>G, p. Arg231Cys, p. Gln421Pro, p. Gly605Asp, p. Gly731Glu, p. Arg1089Arg) have never been reported in HSCR patients before. Variant annotation suggests that 11 of 14 (78.6%) variants in this sample of HSCR cases are likely pathogenic according to the 2015 ACMG Standards and Guidelines (3 null variants that are absent from controls, 3 previously reported pathogenic de novo missense variants that are absent from controls, 1 de novo missense variant that is absent from controls and affects the amino acid known to be pathogenic, 4 de novo missense variants that are absent from controls and predicted to be deleterious by multiple bioinformatic programs) (Table 1) [18].

Large burden of de novo and parental mosaic pathogenic mutations in HSCR

We next studied the inheritance pattern of each variant using blood DNA of the patients and their parents. One

missense variant in family 7 (HSCRFM233) was confirmed to be inherited from the boy's heterozygous mother. However, surprisingly, the other fourteen families were suspected to have de novo mutations or were parental mosaics. When examined by NGS, the probands showed a mutant: wildtype allelic ratio of $48.0 \pm 3.6\%$ (range: 40.5–53.1%) and ddPCR a ratio of $50.0 \pm 1.1\%$ (range: 48.4 and 52.1%). The identical twin females in family 1 (HSCRFM197 and HSCRFM198) were first suspected to be post-zygotic mosaics with a mutant allele frequency of 40.5% (alternative allele reads/total coverage: 194/479, similarly hereinafter) and 42.2% (564/1336) according to NGS, but based on ddPCR results of 48.4% (1388/2870) and 50.1% (1836/3666), respectively, were confirmed as true heterozygotes (see Additional file 1: Table S3). The mutant ratios of the other patients are shown as follows, in the order of NGS and ddPCR respectively, with the number of alternative allele reads and total coverage in parenthesis: XHYY057: 46.7% (436/933) vs. 49.4% (1994/4038), HSCRFM191: 48.7% (1110/2281) vs. 52.1% (2000/3840), XHYY022: 50.5% (650/1287) vs. 49.8% (1582/3176), XHYY051: 51.2% (463/905) vs. 49.8% (1824/3660), XHYY087: 50.1% (610/1218) vs. 50.2% (1956/3894), HSCRFM181: 49.7% (441/887) vs. 49.1% (108/220), XHYY093: 50.8% (705/1388) vs. 49.2% (1692/3442), HSCRFM072: 46.7% (307/657) vs. 52.1% (1654/3178), HSCRFM230: 53.1% (129/243) vs. 50.0% (2376/4756), HSCRFM075: 47.0% (379/806) vs. 52.0% (2342/4502), HSCRFM024: 50.4% (1149/2280) vs. 49.6% (1682/3394), HSCRFM007: 48.6% (688/1416) vs. 49.3% (2102/4262), HSCRFM156: 43.4% (162/373) vs. 49.3% (2072/4200).

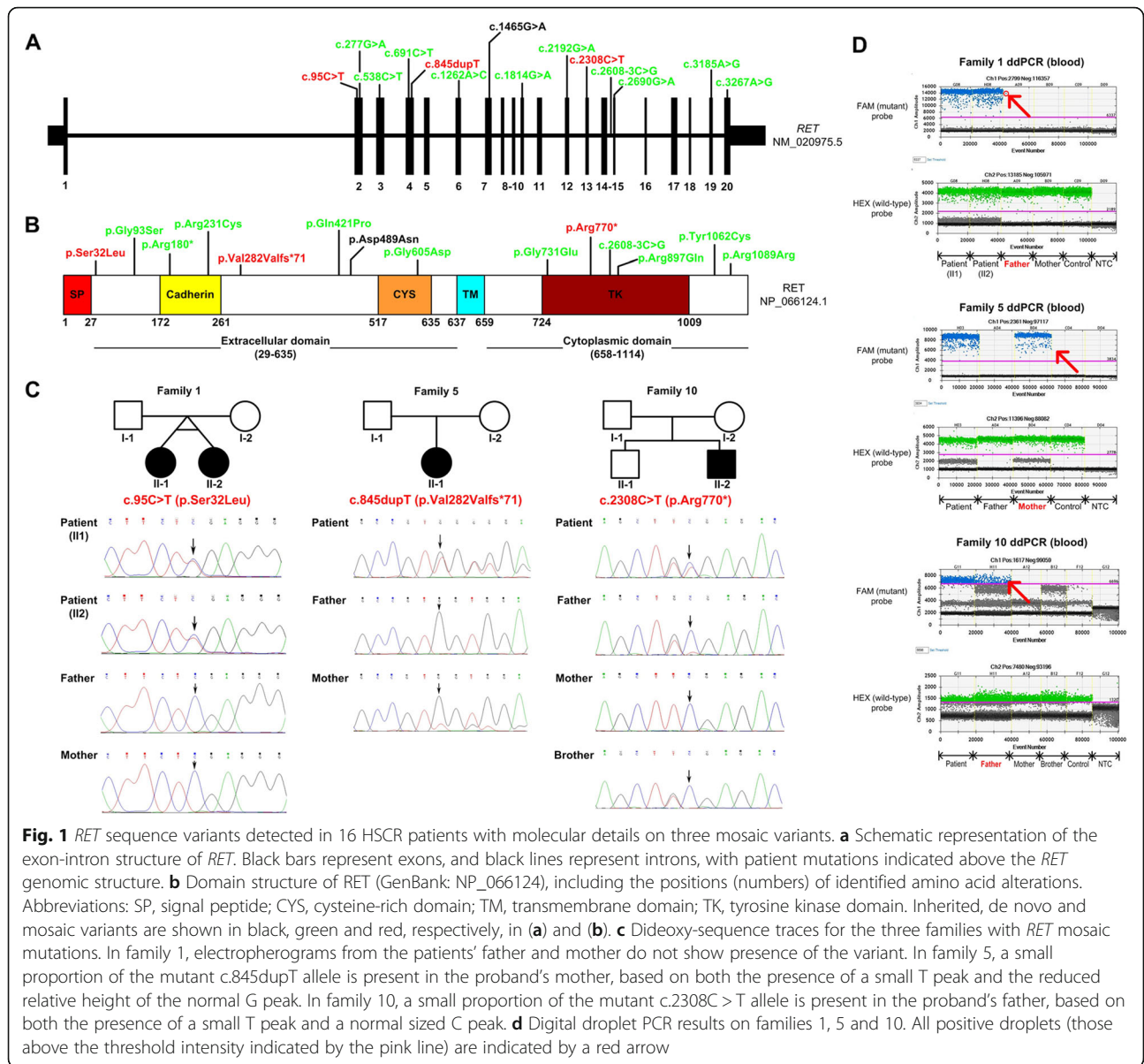
Nevertheless, true mosaicism was identified in two clinically unaffected parents by NGS at a sequencing depth of 192X and 703X in families 5 (XHYY051) and 10 (HSCRFM072), respectively. Sanger sequencing detected a small mutant allele peak in the dideoxy-sequence traces for each of them. ddPCR revealed a similar pattern of the mutant allele frequency distribution among multiple tissues: p. Val282Valfs*71 in family 5: 26.9% in hair (mutant-positive droplets/DNA-containing droplets: 728/2708, similarly hereinafter), 18.4% in urine (746/4046), 23.6% in blood (854/3614) and 22.6% in saliva (690/3050); and p. Arg770* in family 10: 16.9% in hair (374/2214), 12.6% in urine (438/3478), 12.8% in blood (338/2638) and 14.1% in saliva (394/2794). An extremely low-level of paternal mosaicism was missed by NGS at a depth of 674X in family 1 but detected by ddPCR with a very low mutant allele frequency in blood at 0.03% (1/3741) (Figs. 1 and 2). Thus, in the 15 HSCR cases, we identified 1 Mendelian inherited, 3 parental germline mosaics (2 paternal, 1 maternal) and 11 DNVs. Functional annotation of these variants shows that 0, 3 and 8, respectively, are likely pathogenic.

Detection limit and sensitivity of high-coverage NGS, ADS and ddPCR

To determine the detection limit and sensitivity of the three different mutation analysis methods, we examined six positive-control samples, previously demonstrated to carry pathogenic mosaic mutations in *RET*, using amplicon-based deep sequencing (ADS), NGS and ddPCR. Overall, NGS showed a mosaicism detection performance comparable to that of ADS and ddPCR, while ADS displayed a much more reliable detection accuracy and good sensitivity down to a lower limit of $\sim 1\%$: (1) p. Trp85*: 28.0, 41.9 and 28.3% in blood, hair and saliva by ADS; 26.9% in blood by NGS; 29.3, 39.3 and 29.2% in blood, hair and saliva by ddPCR; (2) p. Gln860*: 2.1 and 2.0% in blood and saliva by ADS; and 1.8% in blood by NGS; (3) p. Arg77Cys: 1.3, 2.9 and 4.0% in blood, saliva and sperm by ADS; 0.9% in blood by NGS; 1.3, 3.4 and 4.8% in blood, saliva and sperm by ddPCR (Fig. 2, Table 2).

Discussion

Several lines of evidence suggest that the mechanism of *RET* involvement in HSCR is the result of partial or total loss of *RET* function, with mutant penetrance depending on the degree of functional loss. We presume that the threshold is $> 50\%$ loss because heterozygotes for a *RET* nonsense mutation do not have 100% penetrance in humans [11] but homozygotes for a *Ret* null mutation do have 100% penetrance in mice [19]. One missense variant (p. Asp489Asn) was confirmed to be inherited in a male patient's unaffected mother in our study. Similarly, multiple putative *RET* mutations were inherited from one of the unaffected parents in a previous study. The underlying mechanism, as stated, is that although a substitution is not thought to be causative of disease in and of itself, it may influence the phenotype, especially given the multigenic nature of HSCR [20, 21]. Here, we identify 2 patients with *RET* nonsense mutations and 1 with frameshift mutation, all resulting in a premature stop codon that is expected to produce non-functional *RET*. In addition, most of the *RET* HSCR missense mutations involved amino-acids conserved in multiple species and were scattered in the functional domain of *RET*, which is consistent with the diversity of events predicted to be associated with gene inactivation [21–24]. In brief, those lying within the extracellular domain are proposed to interfere with *RET* maturation and its translocation to the plasma membrane. Variants residing within the TK domain are likely to reduce the catalytic activity of the receptor, and mutations sitting in the region around Y1062 may compromise the efficiency with which *RET* binds to its effector molecules. Finally, we also discovered 1 synonymous and 1 splicing variant in families 14 (HSCRFM007) and 15 (HSCRFM156). At face value, these variants are likely benign; however, their absence in large databases suggests that

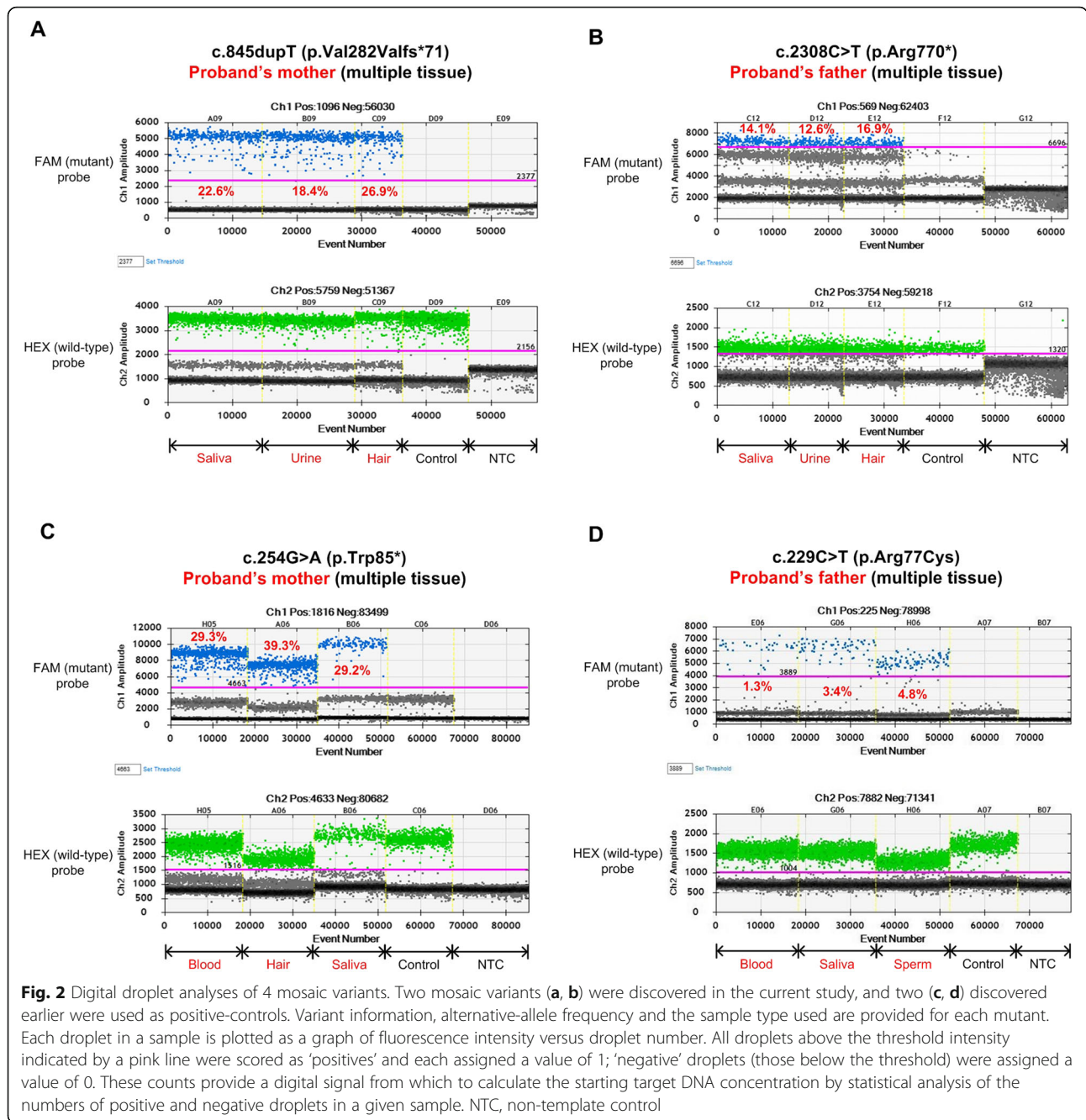


they may have a functional effect acting through activating or abrogating cryptic splice sites or their enhancers [25].

A second intriguing part of this study is the discovery of only one full heterozygote inherited from constitutional heterozygous parent (6.7%) and three heterozygotes inherited from parental mosaics (20%). Genomic mosaicism results from postzygotic events occurring predominantly in early embryogenesis but can arise throughout life and result in genetically distinct cell lines within one individual. Human gastrulation, the process by which the three germ layers are established, is thought to occur at approximately day 16. Primordial germ cells are thought to arise from the primary ectoderm during the second week of development. Therefore, the presence of a somatic variant in blood, saliva (mesodermal tissues), urine (endodermal origin) and hair root

bulbs (ectodermal tissue) indicates that the variant arose early enough to potentially also be present in germ cells and is therefore transmissible to the next generation. This high rate of mosaicism suggests that in some families with apparent DNVs, the pathogenic variant is actually mosaic in the parents, and indeed inherited, and that the risk of HSCR in subsequent children is not infinitesimal. This distinction between non-mosaic inherited DNV (heterozygous in proband and variant not detected in parent) and mosaic inherited DNV (heterozygous in proband, and variant detected mosaic in parent) is important for genetic prognosis and counseling. However, it is very difficult to distinguish true DNV from low allele fraction mosaic mutations in reality.

Here, we surveyed 14 families with both NGS and ddPCR on blood DNA. The degree of allelic ratio bias in



our NGS results is larger than that in most previous studies, the source of which is still unknown. Among those four where deviation from the expected ~ 50/50 allele ratio of true heterozygosity was observed in NGS, three individuals (HSCRFM197, HSCRFM230 and HSCRFM156) were covered by less than 500X. One exception was HSCRFM198, which had a mutant ratio of 42.2% at a whole coverage of 1336X. In contrast, one sample (HSCRFM181) was covered by less than 500X but ddPCR correctly recognized the mutant status (allele ratio 49.1%), which is not surprising given the nature of

the method. NGS can serve as an effective and less expensive technique for screening and quantifying variants; however, it should be noted that many factors may interfere with the results (quality) of the reads/coverage/biallelic ratio by NGS, such as DNA quality (affect baits affinity), biased PCR amplification, sequence context of the variant, pooled DNA isolated from multiple cells as template, the short-read length, sequencing errors and bioinformatic workflow which may filter out biased allele calls. In ddPCR assays, by contrast, template DNA is partitioned into tens of thousands of individual droplets

Table 2 Comparison of amplicon-based deep sequencing (ADS), targeted next-generation sequencing (NGS) and droplet digital polymerase chain reaction (ddPCR) for detection of mosaicism

#	Variant (hg19)	Protein change ^a	Sample type	Method [#]						
				ADS		NGS		ddPCR		
				Total coverage	Alternate-allele reads (%)	Total coverage	Alternate-allele reads (%)	Total coverage	Alternate-allele reads (%)	
1	chr10: 43600528: G>T	p.(Glu252*)	Blood	12023	4674 (38.9%)	225	109 (48.4%)	NA	NA	
			Hair	20801	8440 (40.6%)					
			Saliva	2961	1150 (38.8%)	NA				
			Colon (FFPE)	893	396 (44.3%)					
2	chr10: 43596029: insTCC	p.(Arg67insLeu)	Blood	5931	2081 (35%)	159	67 (42.1%)	NA	NA	
			Hair	3247	1219 (38%)					
			Saliva	3838	1480 (39%)	NA				
			Colon (FFPE)	1881	720 (38%)					
3	chr10: 43600563: C>G	p.(Tyr263*)	Blood	15127	2102 (13.9%)	29	9 (31.0%)	NA	NA	
			Hair	575	139 (24.2%)					
			Saliva	4901	769 (15.7%)					
4	chr10: 43596087: G>A	p.(Trp85*)	Blood	5958	1671 (28.0%)	566	152 (26.9%)	3960	1160 (29.3%)	
			Hair	7098	2977 (41.9%)					
			Saliva	2794	791 (28.3%)	NA	754			220 (29.2%)
5	chr10: 43615164: C>T	p.(Gln860*)	Blood	15246	318 (2.1%)	870	16 (1.8%)	NA	NA	
			Hair	1178	0 (0%)					
			Saliva	305	6 (2.0%)					
6	chr10: 43596062: C>T	p.(Arg77Cys)	Blood	11060	140 (1.3%)	797	7 (0.9%)	4052	52 (1.3%)	
			Hair	894	0 (0%)					
			Saliva	2846	83 (2.9%)	NA	2774			94 (3.4%)
			Sperm	987	39 (4.0%)	3382	162 (4.8%)			

^a RefSeq NM_020975.5, NP_066124.1.

[#] Notice the different performance of NGS for detecting mosaicism at depths over (red) and under (blue) 500X, as compared to ADS and ddPCR.

^a RefSeq NM_020975.5, NP_066124.1

[#] Notice the different performance of NGS for detecting mosaicism at depths over (red) and under (blue) 500X, as compared to ADS and ddPCR

so that at low DNA concentrations the vast majority of droplets contain no more than one copy of template DNA. PCR within each droplet produces a fluorescent readout to indicate the presence or absence of the target of interest, allowing for the accurate “counting” of the number of copies present in a sample [16]. The number of partitions is large enough to assay somatic mosaic events with frequencies down to less than 1%. This excellent accuracy is credited with increased signal-to-noise ratio and removal of PCR bias. As we have shown here, by examining 6 positive-control samples carrying different levels of mosaicism, both ddPCR and ADS surpass the performance of the prevailing NGS and Sanger sequencing.

Interestingly, 11 families (out of 15, 73.3%) were determined to carry non-mosaic inherited DNVs in *RET*, at a significantly higher rate than in any previously reported study: 42.9% in Indonesia, 43.8% in France, and 58.3% in Hong Kong, China [14, 26, 27]. Of these, 72.7% are likely pathogenic. These data raise two issues. First, the pathogenic nature of the DNV needs to be established since *RET* is a commonly mutable gene [28], or rather, its mutants in sperm have a survival advantage [29]. Second, why is the DNV mutation frequency so high? Although our finding may be a chance event it is unlikely because we have observed this before in our studies [7]. A possible and

intriguing reason is that many *RET* DNVs may not be disease-causing or be penetrant on their own but can be in a specific *RET* genetic background that is more permissive in infants with Chinese (Asian) than European ancestry; note that the frequency is also high in the Hong Kong Chinese sample but not the Indonesian one. A candidate for this difference is the *RET* enhancer polymorphism rs2435357 (MCS + 9.7 or RET+ 3) at which a hypomorphic allele that significantly reduces *RET* transcription, has a background allele frequency of 24% (homozygotes ~ 6%) in Europe but 45% (homozygotes ~ 20%) across Asia, a ~ 4-fold difference [30, 31]. MCS + 9.7 does not act on *RET* transcription alone but in concert with at least two other enhancers that also contribute to this genetic background difference [31]. Thus, we hypothesize that this increased widespread susceptibility in China allows a greater number of milder *RET* variants to be HSCR-associated, including DNVs, accounting for the higher frequency of DNVs in Chinese HSCR patients. Regardless, both paternal age and the sequencing sensitivity of different technologies should be taken into account when making the final statement.

Every human gene is subject to random mutation multiple times within each individual. However, most variants are either benign or never reach a fraction high enough to cause disease. Thus, whether a pathogenic variant is disease

penetrant or not depends on the physiological function of the encoded molecule and the fraction of cells possessing the mutation in a given tissue. Somatic mutations that lead to a gain of function or growth advantage might cause disease if they are present in even one cell, as in cancer. On the other hand, somatic mutations that lead to a loss of function might need to occur in a larger clonal fraction in order to cause a clinical phenotype. Therefore, for every deleterious somatic mutation there likely exists a threshold mosaic fraction above which the mutation causes disease but below which it does not and so remains undetected [32]. Of course, for de novo changes the penetrance is likely dependent on the number of cells affected, as well as the specific mutation, the disease involved, and the genetic background of the individual. Thus, distinguishing non-mosaic inherited DNV (germline DNV) from true postzygotic DNVs is important, as is the threshold mosaic fraction. These analyses need to be quantitative because in some cases, in clinically significant cortical malformations, the disorder can result from somatic mutations in as few as 1% of cells [33]. The threshold mosaic fraction for HSCR is important to investigate because it is likely a critical determinant of HSCR penetrance and expressivity.

Conclusion

Together with previously reported cases, our study broadened the clinical and molecular spectrum of HSCR and revealed a large burden of de novo and parental mosaic pathogenic mutations in *RET* in the Chinese population. All the observations indicated that distinguishing non-mosaic inherited DNV from mosaic inherited DNV is important for both genetic prognosis and accurate counseling.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s13023-019-1194-2>.

Additional file 1: Table S1. Summary of the quality of targeted next-generation sequencing data on *RET*. **Table S2.** Quality of targeted next-generation sequencing data on *RET* of all the samples in the current study. **Table S3.** Comparison of mutant ratio in 15 HSCR patients carrying 13 different kinds of *RET* variants examined by targeted next-generation sequencing (NGS) and droplet digital polymerase chain reaction (ddPCR). (DOCX 91 kb)

Acknowledgements

We are grateful to the many affected individuals and their families whose cooperation made this study possible. We also thank Mr. Chao Li, Ms. Xiaotong Li, Mr. Jianyong Gao, and Mr. Geng Tian from Geneis (Beijing) Co., Ltd. for assistance with ddPCR mosaic validation.

Authors' contributions

QJ and YW prepared the manuscript. QJ and YW performed the high-coverage NGS and droplet digital PCR. QL, ZZ, PX and HW conducted sample acquisition and collected clinical information. NL, JW, FZ and AC conducted genetic analysis. WC and LL designed the study and edited the manuscript. All authors read and approved the final manuscript.

Funding

This work was made possible by a grant from the Chinese Academy of Medical Sciences Initiative for Innovative Medicine (CAMS-I2M) and the CAMS Central Public Welfare Scientific Research Institute Basal Research Expenses (2017PT31004) and is supported by the National Natural Science Foundation of China (81771620), the Beijing Nova Program (Z171100001117125) and the Basic Foundation of the Capital Institute of Pediatrics (FX-2016-02) to Q.J. Y. W. was supported by the National Natural Science Foundation of China (81670469) and Shanghai Municipal Commission of Health and Family Planning (201840028). Z.Z. was supported by the National Natural Science Foundation of China (81700451). F.Z. was supported by the National Natural Science Foundation of China (31625015). A.C. was supported by grants from the US National Institutes of Health (HD28088). W.C. was supported by the National Natural Science Foundation of China (81630039) and Shanghai Key Laboratory of Pediatric Gastroenterology and Nutrition (17DZ2272000). L.L. was supported by grants from the Public Welfare Industry Research Special Foundation of China (201402007) and a Clinical Medicine Development Project of the Beijing Hospital Administration Bureau (ZYLX201306).

Availability of data and materials

All data generated or analysed during this study are included in this published article and its supplementary information files.

Ethics approval and consent to participate

The patient sampling was conducted through the Hirschsprung Disease Consortium of China (HDCC), spanning two clinical centres across north and south China. This study was approved by the medical ethics committee of the Capital Institute of Pediatrics (SHERLL 2013039) and Xinhua Hospital (XHEC-C-2016-263) and was carried out according to the Declaration of Helsinki. All caregivers and participants (when able) provided written informed consent.

Consent for publication

Written informed consents for publication were obtained from all caregivers.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Medical Genetics, Beijing Municipal Key Laboratory of Child Development and Nutriomics, Capital Institute of Pediatrics, Beijing 100020, China. ²Department of Pediatric Surgery, Xinhua Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai Key Laboratory of Pediatric Gastroenterology and Nutrition, Shanghai Institute for Pediatric Research, No. 1665 Kongjiang Rd., Yangpu District, Shanghai 200092, China. ³Department of General Surgery, Capital Institute of Pediatrics Affiliated Children's Hospital, No. 2 Yabao Rd., Chaoyang District, Beijing 100020, China. ⁴Department of Pathology, Capital Institute of Pediatrics Affiliated Children's Hospital, Beijing 100020, China. ⁵MyGenosics Inc, Beijing 101318, China. ⁶Obstetrics and Gynecology Hospital, Collaborative Innovation Center of Genetics and Development, Fudan University, Shanghai 200011, China. ⁷Center for Human Genetics and Genomics, New York University School of Medicine, New York, NY 10016, USA.

Received: 28 April 2019 Accepted: 4 September 2019

Published online: 30 October 2019

References

1. Wright CF, FitzPatrick DR, Firth HV. Paediatric genomics: diagnosing rare disease in children. *Nat Rev Genet.* 2018;19(5):253–68.
2. Biesecker LG, Spinner NB. A genomic view of mosaicism and human disease. *Nat Rev Genet.* 2013;14(5):307–20.
3. Jamuar SS, Lam AT, Kircher M, D'Gama AM, Wang J, Barry BJ, Zhang X, Hill RS, Partlow JN, Rozzo A, et al. Somatic mutations in cerebral cortical malformations. *N Engl J Med.* 2014;371(8):733–43.
4. D'Gama AM, Pochareddy S, Li M, Jamuar SS, Reiff RE, Lam AN, Sestan N, Walsh CA. Targeted DNA sequencing from autism Spectrum disorder brains implicates multiple genetic mechanisms. *Neuron.* 2015;88(5):910–7.

5. Myers CT, Hollingsworth G, Muir AM, Schneider AL, Thuesmann Z, Knupp A, King C, Lacroix A, Mehaffey MG, Berkovic SF, et al. Parental mosaicism in "De novo" epileptic encephalopathies. *N Engl J Med*. 2018;378(17):1646–8.
6. King DA, Jones WD, Crow YJ, Dominiczak AF, Foster NA, Gaunt TR, Harris J, Hellens SW, Homfray T, Innes J, et al. Mosaic structural variation in children with developmental disorders. *Hum Mol Genet*. 2015;24(10):2733–45.
7. Jiang Q, Liu F, Miao C, Li Q, Zhang Z, Xiao P, Su L, Yu K, Chen X, Zhang F, et al. RET somatic mutations are underrecognized in Hirschsprung disease. *Genet Med*. 2018;20(7):770–7.
8. Amiel J, Sproat-Emison E, Garcia-Barcelo M, Lantieri F, Burzynski G, Borrego S, Pelet A, Arnold S, Miao X, Griseri P, et al. Hirschsprung disease, associated syndromes and genetics: a review. *J Med Genet*. 2008;45(1):1–14.
9. Alves MM, Sribudiani Y, Brouwer RW, Amiel J, Antinolo G, Borrego S, Ceccherini I, Chakravarti A, Fernandez RM, Garcia-Barcelo MM, et al. Contribution of rare and common variants determine complex diseases-Hirschsprung disease as a model. *Dev Biol*. 2013;382(1):320–9.
10. Chakravarti AMAS, Lyonnet S. Hirschsprung disease. In: BV DV, Kinzler KW, Antonarakis SE, Ballabio A, Gibson K, Mitchell G, editors. *Metabolic and molecular bases of inherited disease*: McGraw-Hill; 2001.
11. Jiang Q, Arnold S, Heaneu T, Kilambi KP, Doan B, Kapoor A, Ling AY, Sosa MX, Guy M, Burzynski G, et al. Functional loss of semaphorin 3C and/or semaphorin 3D and their epistatic interaction with ret are critical to Hirschsprung disease liability. *Am J Hum Genet*. 2015;96(4):581–96.
12. Bahrami A, Joodi M, Moetamani-Ahmadi M, Maftouh M, Hassanian SM, Ferns GA, Avan A. Genetic background of Hirschsprung disease: a bridge between basic science and clinical application. *J Cell Biochem*. 2018;119(1):28–33.
13. Gui H, Schriemer D, Cheng WW, Chauhan RK, Antinolo G, Berrios C, Bleda M, Brooks AS, Brouwer RW, Burns AJ, et al. Whole exome sequencing coupled with unbiased functional analysis reveals new Hirschsprung disease genes. *Genome Biol*. 2017;18(1):48.
14. Attie T, Pelet A, Ederly P, Eng C, Mulligan LM, Amiel J, Boutrand L, Beldjord C, Nihoul-Fekete C, Munnich A, et al. Diversity of RET proto-oncogene mutations in familial and sporadic Hirschsprung disease. *Hum Mol Genet*. 1995;4(8):1381–6.
15. Zhang Z, Li Q, Diao M, Liu N, Cheng W, Xiao P, Zou J, Su L, Yu K, Wu J, et al. Sporadic Hirschsprung disease: mutational Spectrum and novel candidate genes revealed by next-generation sequencing. *Sci Rep*. 2017;7(1):14796.
16. Hindson BJ, Ness KD, Masquelier DA, Belgrader P, Heredia NJ, Makarewicz AJ, Bright JJ, Lucero MY, Hiddessen AL, Legler TC, et al. High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Anal Chem*. 2011;83(22):8604–10.
17. Luks VL, Kamitaki N, Vivero MP, Uller W, Rab R, Bovee JV, Rialon KL, Guevara CJ, Alomari AI, Greene AK, et al. Lymphatic and other vascular malformative/overgrowth disorders are caused by somatic mutations in PIK3CA. *J Pediatr*. 2015;166(4):1048–54.
18. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17(5):405–24.
19. Schuchardt A, D'Agati V, Larsson-Blomberg L, Costantini F, Pachnis V. Defects in the kidney and enteric nervous system of mice lacking the tyrosine kinase receptor ret. *Nature*. 1994;367(6461):380–3.
20. Ederly P, Lyonnet S, Mulligan LM, Pelet A, Dow E, Abel L, Holder S, Nihoul-Fekete C, Ponder BA, Munnich A. Mutations of the RET proto-oncogene in Hirschsprung's disease. *Nature*. 1994;367(6461):378–80.
21. Angrist M, Bolk S, Thiel B, Puffenberger EG, Hofstra RM, Buys CH, Cass DT, Chakravarti A. Mutation analysis of the RET receptor tyrosine kinase in Hirschsprung disease. *Hum Mol Genet*. 1995;4(5):821–30.
22. Romeo G, Ronchetto P, Luo Y, Barone V, Seri M, Ceccherini I, Pasini B, Boccardi R, Lerone M, Kaariainen H, et al. Point mutations affecting the tyrosine kinase domain of the RET proto-oncogene in Hirschsprung's disease. *Nature*. 1994;367(6461):377–8.
23. Geneste O, Bidaud C, De Vita G, Hofstra RM, Tartare-Deckert S, Buys CH, Lenoir GM, Santoro M, Billaud M. Two distinct mutations of the RET receptor causing Hirschsprung's disease impair the binding of signalling effectors to a multifunctional docking site. *Hum Mol Genet*. 1999;8(11):1989–99.
24. Nagar B, Overduin M, Ikura M, Rini JM. Structural basis of calcium-induced E-cadherin rigidification and dimerization. *Nature*. 1996;380(6572):360–4.
25. Bolk S, Pelet A, Hofstra RM, Angrist M, Salomon R, Croaker D, Buys CH, Lyonnet S, Chakravarti A. A human model for multigenic inheritance: phenotypic expression in Hirschsprung disease requires both the RET gene and a new 9q31 locus. *Proc Natl Acad Sci U S A*. 2000;97(1):268–73.
26. Widowati T, Melhem S, Patria SY, de Graaf BM, Sinke RJ, Viel M, Dijkhuis J, Sadewa AH, Purwahardjono R, Soenarto Y, et al. RET and EDNRB mutation screening in patients with Hirschsprung disease: functional studies and its implications for genetic counseling. *Eur J Hum Genet*. 2016;24(6):823–9.
27. Garcia-Barcelo M, Sham MH, Lee WS, Lui VC, Chen BL, Wong KK, Wong JS, Tam PK. Highly recurrent RET mutations and novel mutations in genes of the receptor tyrosine kinase and endothelin receptor B pathways in Chinese patients with sporadic Hirschsprung disease. *Clin Chem*. 2004;50(1):93–100.
28. Lim J, Maher GJ, Turner GD, Dudka-Ruszkowska W, Taylor S, Rajpert-De Meys E, Goriely A, Wilkie AO. Selfish spermatogonial selection: evidence from an immunohistochemical screen in testes of elderly men. *PLoS One*. 2012;7(8):e42382.
29. Shinde DN, Elmer DP, Calabrese P, Boulanger J, Arnheim N, Tiemann-Boege I. New evidence for positive selection helps explain the paternal age effect observed in achondroplasia. *Hum Mol Genet*. 2013;22(20):4117–26.
30. Emison ES, Garcia-Barcelo M, Grice EA, Lantieri F, Amiel J, Burzynski G, Fernandez RM, Hao L, Kashuk C, West K, et al. Differential contributions of rare and common, coding and noncoding ret mutations to multifactorial Hirschsprung disease liability. *Am J Hum Genet*. 2010;87(1):60–74.
31. Chatterjee S, Kapoor A, Akiyama JA, Auer DR, Lee D, Gabriel S, Berrios C, Pennacchio LA, Chakravarti A. Enhancer variants synergistically drive dysfunction of a gene regulatory network in Hirschsprung disease. *Cell*. 2016;167(2):355–68.
32. Rodin RE, Walsh CA. Somatic mutation in pediatric neurological diseases. *Pediatr Neurol*. 2018;87:20–2.
33. Nakashima M, Saito H, Takei N, Tohyama J, Kato M, Kitaura H, Shiina M, Shirozu H, Masuda H, Watanabe K, et al. Somatic mutations in the MTOR gene cause focal cortical dysplasia type IIb. *Ann Neurol*. 2015;78(3):375–86.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

