

# SCIENTIFIC REPORTS



OPEN

## Integrating Functional Analysis in the Next-Generation Sequencing Diagnostic Pipeline of RASopathies

Gordon K. C. Leung<sup>1</sup>, H. M. Luk<sup>2</sup>, Vincent H. M. Tang<sup>3</sup>, W. W. Gao<sup>3</sup>, Christopher C. Y. Mak<sup>1</sup>, Mullin H. C. Yu<sup>1</sup>, W. L. Wong<sup>1</sup>, Yoyo W. Y. Chu<sup>1</sup>, W. L. Yang<sup>1</sup>, Wilfred H. S. Wong<sup>1</sup>, Alvin C. H. Ma<sup>4,5</sup>, Anskar Y. H. Leung<sup>4</sup>, D. Y. Jin<sup>3</sup>, Kelvin Y. K. Chan<sup>6</sup>, Judith Allanson<sup>7</sup>, Ivan F. M. Lo<sup>2</sup> & Brian H. Y. Chung<sup>1,6</sup>

RASopathies are a group of heterogeneous conditions caused by germline mutations in RAS/MAPK signalling pathway genes. With next-generation sequencing (NGS), sequencing capacity is no longer a limitation to molecular diagnosis. Instead, the rising number of variants of unknown significance (VUSs) poses challenges to clinical interpretation and genetic counselling. We investigated the potential of an integrated pipeline combining NGS and the functional assessment of variants for the diagnosis of RASopathies. We included 63 Chinese patients with RASopathies that had previously tested negative for *PTPN11* and *HRAS* mutations. In these patients, we performed a genetic analysis of genes associated with RASopathies using a multigene NGS panel and Sanger sequencing. For the VUSs, we evaluated evidence from genetic, bioinformatic and functional data. Twenty disease-causing mutations were identified in the 63 patients, providing a primary diagnostic yield of 31.7%. Four VUSs were identified in five patients. The functional assessment supported the pathogenicity of the *RAF1* and *RIT1* VUSs, while the significance of two VUSs in *A2ML1* remained unclear. In summary, functional analysis improved the diagnostic yield from 31.7% to 36.5%. Although technically demanding and time-consuming, a functional genetic diagnostic analysis can ease the clinical translation of these findings to aid bedside interpretation.

The application of next-generation sequencing (NGS) has rapidly expanded our knowledge of the mutational and phenotypic spectrum of genetic disorders. The massive parallel sequencing of millions of DNA fragments allows research and clinical laboratories to analyse multiple genes in the entire genome rapidly, and as a result, sequencing capacity is no longer a limitation. With the rapid decline in cost, various types of NGS, including multigene sequencing panels, whole-exome-sequencing (WES), or whole-genome sequencing (WGS), are now in routine clinical use to diagnose genetic disorders<sup>1–3</sup>. In general, gene panels are favoured for genetically heterogeneous conditions in which a clear clinical hypothesis is established before testing<sup>4–6</sup>. Studies have shown that the clinical sensitivity of disease-focused gene panels can actually be higher than WES/WGS<sup>7,8</sup>. In addition, gene panels usually have lower costs, and the data size is more manageable. Therefore, in some clinical laboratories, NGS gene panels are preferred over WES/WGS.

Despite the advantages of gene panels, the interpretation of genotype-phenotype relationships and the consequence of genetic variants are still major challenges to its clinical application. Although significant progress developing bioinformatic prediction tools and databases for international data sharing has been achieved, the interpretation of variants of unknown significance (VUSs) still impedes realizing the full application of genomic medicine. False pathogenicity assignments can have adverse consequences for patients, as genetic diagnosis is

<sup>1</sup>Department of Paediatrics and Adolescent Medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China. <sup>2</sup>Clinical Genetic Service, Department of Health, Hong Kong, China. <sup>3</sup>School of Biomedical Sciences, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China. <sup>4</sup>Department of Medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China. <sup>5</sup>Department of Health Technology and Informatics, The Hong Kong Polytechnic University, Hong Kong, China. <sup>6</sup>Department of Obstetrics and Gynaecology, Tsan Yuk Hospital, Hong Kong, China. <sup>7</sup>Department of Paediatrics, Faculty of Medicine, University of Ottawa, Ontario, Canada. Correspondence and requests for materials should be addressed to I.F.M.L. (email: [dr.ivanlo@gmail.com](mailto:dr.ivanlo@gmail.com)) or B.H.Y.C. (email: [bhychung@hku.hk](mailto:bhychung@hku.hk))

now used to manage patient care at different stages of life<sup>2</sup>. Extensive efforts have been made to elucidate the causality of VUSs in human diseases<sup>9–11</sup>. In summary, a two-step process is advised. First, evidence implicating a specific gene should be considered (gene level assessment). Second, a combined assessment of the genetic, computational and experimental support for an individual candidate genetic variant should then be performed (variant level assessment). Clinicians and clinical laboratories are usually more involved in variant assessment, and the usual strategy includes clinical correlation, family segregation, allelic frequency, and computational prediction, as well as some limited functional studies available in clinical settings, including biochemical analysis for metabolic diseases or mRNA studies. More recently, the ClinGen Clinical Validity Framework made a recommendation on experimental evidence, giving more weight to *in vivo* (model systems) compared to *in vitro* (function and functional alteration) experimental data<sup>12</sup>.

RASopathies are a group of developmental conditions caused by germline mutations in genes involved in the RAS/MAPK signalling pathway<sup>13,14</sup>. These conditions are clinically heterogeneous and include Noonan or Noonan-like syndromes, Costello syndrome, cardio-facio-cutaneous syndrome, neurofibromatosis type I and related disorders. They have both unique and overlapping clinical features recognizable by clinical geneticists who usually formulate a diagnostic hypothesis prior to genetic testing<sup>15</sup>. However, they are also genetically heterogeneous illnesses caused by gain-of-function mutations in multiple genes of the pathway, which regulate cell growth, proliferation and differentiation via a signalling cascade involving small GTPases and the phosphorylation of downstream molecules, such as MAPK3 (Legacy symbol: ERK1) and ELK1<sup>16</sup>. Various functional assays have been used for gene discovery, including *in vitro* luciferase assays and *in vivo* zebrafish (*Danio rerio*) modelling<sup>17–20</sup>. The aim of this study was to examine the feasibility of combining an NGS gene panel and functional assays for variant level assessment with an integrated pipeline for the diagnosis of RASopathies.

## Results

**Mutation analysis.** In this study, 20 pathogenic or likely pathogenic mutations were identified in the 63 patients recruited. Of these, 18 were identified in the NGS panel and two were from *RIT1* sequencing, and all were missense mutations. Six (9.5%) were found in *SOS1*; two (3.1%), in *RAF1*; three (4.7%), in *KRAS*; three (4.7%), in *MAP2K1*; two (3.1%), in *RIT1*; two (3.1%), in *BRAF*; and two (3.1%), in *SHOC2*. These mutations have been previously reported in medical literature and various mutation databases (Table 1), and the clinical features of these patients were compatible with RASopathies (Fig. 1 and suppl. Table 1). Combining the NGS panel and *RIT1* sequencing, a molecular diagnosis was established in 31.7% (20/63) of the patients.

Four VUSs were identified in five patients (7.9%) according to the ACMG guideline<sup>10</sup>. Three were missense mutations, and one was a nonsense mutation. *RAF1*:c.1173 G > C:p.(R391S) was identified in a mother and son, both of whom exhibited features of Noonan syndrome. In addition, one *RIT1* (c.259 G > C:p.(D87H)) and two *A2ML1* (c.256 C > T:p.(P86S), c.4261 C > T:p.(Q1421X)) variants were observed in three other patients. From the bioinformatics analysis, an evaluation of these variants showed stronger evidence of pathogenicity for the *RIT1* and *RAF1* variants than the two *A2ML1* variants (Table 2). The clinical features of the patients carrying these variants are shown in Fig. 1 and in suppl. Table 2.

**Functional characterization of the VUSs.** To assess the pathogenicity of the four VUSs found in this study, we introduced mutations in the human coding transcript. Previous studies suggest that a majority of mutations causing RASopathies lead to the activation of the RAS/MAPK signalling pathway<sup>17,21,22</sup>. Increased signalling is measured by the level of downstream substrate phosphorylation, such as phosphorylated ELK1 (pELK1). To evaluate whether the RAS/MAPK pathway was activated by the VUSs in our study, we introduced mutant reporter plasmids into HEK-293T cells to assess phosphorylation levels using a dual luciferase reporter assay. Both *RAF1*:p.R391S and *RIT1*:p.D87H significantly increased pELK1 expression compared with the WT transcripts (Fig. 2). In contrast, the two *A2ML1* mutations did not produce a significant fold change in pELK1 levels *in vitro*.

The phenotypic impact of the VUSs on embryonic development was examined by transiently expressing WT or mutant transcripts in zebrafish embryos at the single-cell stage. We scored the embryos according to their morphological appearance (Fig. 3), the width-to-length ratio of the head (Fig. 4), and *in situ* hybridization of the cardiac marker *cmlc1* (Fig. 5). These features correspond to the typical phenotypes of Noonan syndrome (cranio-facial dysmorphism and cardiac abnormalities) in human patients. The morphometric analysis agreed with the results from the luciferase assay, and the activation of the RAS signals in the zebrafish embryos was also compatible with the phenotypes (data not shown). A significant portion of zebrafish embryos injected with human mRNA transcripts with novel mutations in *RAF1* and *RIT1* developed atypical craniofacial and heart structures, whereas the expression of the *A2ML1* mutations did not affect zebrafish development.

In summary, both *in vitro* and *in vivo* analyses provided further supportive evidence for the functional significance of the *RAF1* and *RIT1* variants, whereas the significance of the two VUSs in *A2ML1* remained unclear. A combined assessment of the genetic, bioinformatic and experimental evidence from the *RAF1* and *RIT1* variants led to a re-classification of the VUSs as likely pathogenic mutations, which increased the overall diagnostic yield from 31.7% to 36.5%.

## Discussion

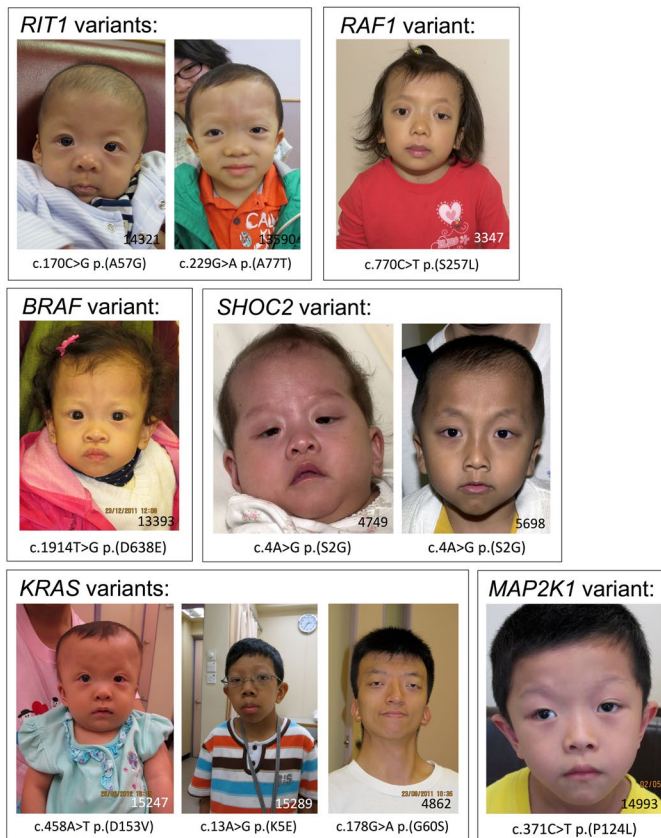
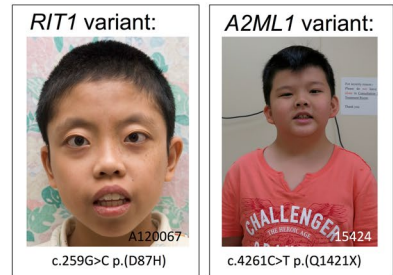
This study aimed to establish a diagnostic pipeline for RASopathies that integrates an NGS panel with functional genetics. We adopted the experimental approach used for gene discovery applied to *RAF1*, *RIT1* and *A2ML1*<sup>17–19</sup> on variant assessment for VUSs in these genes such that the same criteria can be used to judge their pathogenicity. Our initial diagnostic yield was 31.7%, which was lower than that of other RASopathy studies (diagnostic yield of 60–70%)<sup>7,15</sup>, because the patients recruited had already tested negative for *PTPN11* and *HRAS* mutations. The functional genetic studies contributed to resolving two of four VUSs, increasing the diagnostic yield from 31.7% to 36.5%.

Patient number	Gene	Mutation	ClinVar*	HGMD#	NSEuroNet^	Publication with functional analysis
5132	SOS1	c.512T > C:p.(V171A)	likely pathogenic	—	—	
15547	SOS1	c.1297G > A:p.(E433K)	pathogenic	disease-causing mutation	Mutation (count = 19)	
14626	SOS1	c.1644T > G:p.(S548R)	pathogenic	disease-causing mutation	Mutation (count = 1)	Smith <i>et al.</i> (2013)
5608	SOS1	c.1644T > A:p.(S548R)	pathogenic	—	—	
6381	SOS1	c.1644T > A:p.(S548R)	pathogenic	—	—	
9233	SOS1	c.1654A > G:p.(R552G)	pathogenic	disease-causing mutation	Mutation (count = 52)	
4748	RAF1	c.770C > T:p.(S257L)	likely pathogenic	disease-causing mutation	Mutation (count = 73)	
3347	RAF1	c.770C > T:p.(S257L)	likely pathogenic	disease-causing mutation	Mutation (count = 73)	
15289	KRAS	c.13A > G:p.(K5E)	likely pathogenic	disease-causing mutation	Mutation (count = 3)	
4862	KRAS	c.178G > A:p.(G60S)	pathogenic	disease-causing mutation	Mutation (count = 1)	
15247	KRAS	c.458A > T:p.(D153V)	pathogenic	disease-causing mutation	Mutation (count = 13)	
6575	MAP2K1	c.199G > A:p.(D67N)	likely pathogenic	—	Mutation (count = 6)	
14993	MAP2K1	c.371C > T:p.(P124L)	pathogenic	disease-causing mutation	Mutation (count = 1)	Emery <i>et al.</i> (2009)
4012	MAP2K1	c.389A > G:p.(Y130C)	pathogenic	disease-causing mutation	Mutation (count = 35)	Cheng <i>et al.</i> (2012)
14321	RIT1	c.170C > G:p.(A57G)	pathogenic	disease-causing mutation	Mutation (count = 23)	Chen <i>et al.</i> (2014)
13590	RIT1	c.229G > A:p.(A77T)	likely pathogenic	disease-causing mutation	Mutation (count = 6)	
5153	BRAF	c.1785T > G:p.(F595L)	likely pathogenic	disease-causing mutation	Mutation (count = 5)	Cheng <i>et al.</i> (2012)
13393	BRAF	c.1914T > G:p.(D638E)	pathogenic	disease-causing mutation	Mutation (count = 5)	Cheng <i>et al.</i> (2012)
4749	SHOC2	c.4A > G:p.(S2G)	likely pathogenic	disease-causing mutation	Mutation (count = 169)	
5698	SHOC2	c.4A > G:p.(S2G)	likely pathogenic	disease-causing mutation	Mutation (count = 169)	

**Table 1.** A summary of 20 patients identified with pathogenic or likely pathogenic mutations from our cohort using the multigene panel sequencing approach. \*Clinical significance was labelled according to the ClinVar database from NCBI (<https://www.ncbi.nlm.nih.gov/clinvar/>). #Variant class was determined according to the HGMD (Human Gene Mutation Database [http://www.hgmd.cf.ac.uk/docs/new\\_help.html](http://www.hgmd.cf.ac.uk/docs/new_help.html)). ^Variant classification was determined according to the NSEuroNet (European Network on Noonan Syndrome and Related Disorders; <https://nseuro.net/php/about.php>).

For the four VUSs we identified, *in silico* predictions suggested that the variants were either disease-causing or damaging in nature (Table 2). For the *RAF1*:p.R391S variant, it was located at an evolutionarily conserved genomic locus at conserved region 3 (CR3) of *RAF1* gene. CR3 encodes multiple protein kinase catalytic domains for ATP binding<sup>23,24</sup>; however, p.R391S does not directly interfere with the functional motif of CR3. The phenotype-genotype correlation between Noonan syndrome and the *RAF1* mutation were previously reported<sup>19,21</sup>, and the two patients in our cohort with the p.R391S mutation had typical features of Noonan syndrome. *RIT1*:p.D87H was reported once previously as a VUS in the ClinVar database. Our patient with the *RIT1* mutation presented with typical facial features of Noonan syndrome (Fig. 1), as well as hypertrophic cardiomyopathy (HCM) and sub-aortic stenosis. Additional features included congenital chylothorax, multifocal bronchomalacia, and bilateral lower-limb lymphangiectasia of the calves. The variant p.D87H is not located in one of the five functional domains of *RIT1*. Still, pathogenic mutations for Noonan syndrome have been reported in an amino acid position adjacent to p.D87H<sup>17</sup>. In patients with *RIT1* mutations, HCM (56%), atrial defects, and pulmonary stenosis are the more frequently reported clinical manifestations<sup>25</sup>. The clinical presentation of our patient was consistent with these published phenotypes and included cardiac anomalies and lymphatic complications<sup>26,27</sup>. Therefore, combining all of the evidence and functional analysis, we classified the two genetic variants *RAF1*:p.R391S and *RIT1*:p.D87H as likely pathogenic mutations, according to the latest guidelines<sup>9,10,12</sup>.

The genetic variants *A2ML1*:p.P86S and *A2ML1*:p.Q1421X did not appear functionally significant after *in vitro* or *in vivo* assessment. To date, the association of *A2ML1* with Noonan-related syndrome has only been reported in one study<sup>18</sup>. Vissers *et al.* identified a *de novo* variant of *A2ML1* in an individual with Noonan syndrome by WES, and additional analyses of 155 individuals revealed two more families with Noonan syndrome carrying *A2ML1* variants. From the current understanding, the majority of reported pathogenic mutations in RASopathies are gain-of-function mutations, whereas one of the two *A2ML1* variants in our study is a null mutation. Constraint analysis from ExAC suggested that heterozygous loss-of-function variants in *A2ML1* are common across the gene in unaffected populations<sup>28</sup>, and intragenic deletions have also been reported in normal individuals<sup>29</sup>. Therefore, it is unlikely that *A2ML1* mutations act via haploinsufficiency. An independent assay, similar to the methodology used in the first publication on *A2ML1*<sup>18</sup> was used for functional analysis. We illustrated that the two *A2ML1* variants in our cohort did not activate the RAS/MAPK pathway, and they did not affect the development of the zebrafish embryos. Our findings suggested that the functional significance of the two *A2ML1* variants might not affect the RAS/MAPK pathway via gain-of-function nor dominant negative mechanism, and the clinical significance of the two variants remains unclear. This observation is actually in line with Aoki *et al.*<sup>30</sup>, stating that the functional properties of *A2ML1* and the mechanisms by which *A2ML1* regulates the RAS/MAPK pathway is largely unknown. Furthermore, the phenotypes of the two patients were reviewed by a highly experienced clinical geneticist who is an internationally recognized expert in RASopathies, blinded to the functional analysis and computational prediction. It was concluded that both patients in our cohort did not have the classical phenotype of Noonan syndrome (clinical photograph of one patient is shown in Fig. 1).

**a. Patients with pathogenic / likely pathogenic variant****b. Patients with VUS**

**Figure 1.** Clinical photographs of patients with variants in genes of the RAS/MAPK signalling pathway. Patients with (a) pathogenic or likely pathogenic variants and (b) VUSs from our study.

Herein, we used a multigene sequencing panel to identify the disease-causing mutations in patients with RASopathies with a primary diagnostic yield of 31.7%. We also performed a functional characterisation of VUSs and confirmed the pathogenic nature of two. The inclusion of functional data improved the diagnostic yield by 4.8%. As suggested by ACMG guidelines, well-established functional studies can serve as a strong evidence of pathogenicity (i.e. PS3). Integrating robust functional analysis into routine clinical laboratory tests is challenging yet feasible. Some of these challenges include the requirement of specialised expertise and collaboration with research laboratories (e.g., zebrafish facilities). The knowledge of disease-causing mechanisms is also essential and experiments can only be designed for known disease pathways or known structural changes (e.g. ciliopathies). It may seem practically difficult to apply this one-by-one approach to the large number of genetic variants discovered by WES and WGS. However, a recent report showed that a multiplexed functional assay can be used to map sequence-function relationships with base-pair resolution for hundreds to thousands of variants in both coding and non-coding regions of the genome in a single experiment<sup>31</sup>. This innovation can potentially overcome the current limitations of functional genetics and can be implemented successfully into the clinical diagnostic pipeline in the near future.

Molecular diagnosis using NGS can facilitate the clinical management for patients with genetic diseases. The concept of predictive genomics has been reviewed in clinical oncology, in which the genomic data of tumor was used to predict clinical phenotypes and tumor progression to better design patient treatment<sup>32</sup>. Based on a systematic approach, we could predict personalized drug targets and outcome for cancer patients<sup>33</sup>. For patients with other genetic syndromes, timely diagnosis and predictable clinical outcomes are always beneficial to the patient<sup>34</sup>. Therefore, the establishment of functional analysis for genetic syndromes would help clinical decision-making in the long term and facilitate comprehensive interpretation of new variants by clinicians.

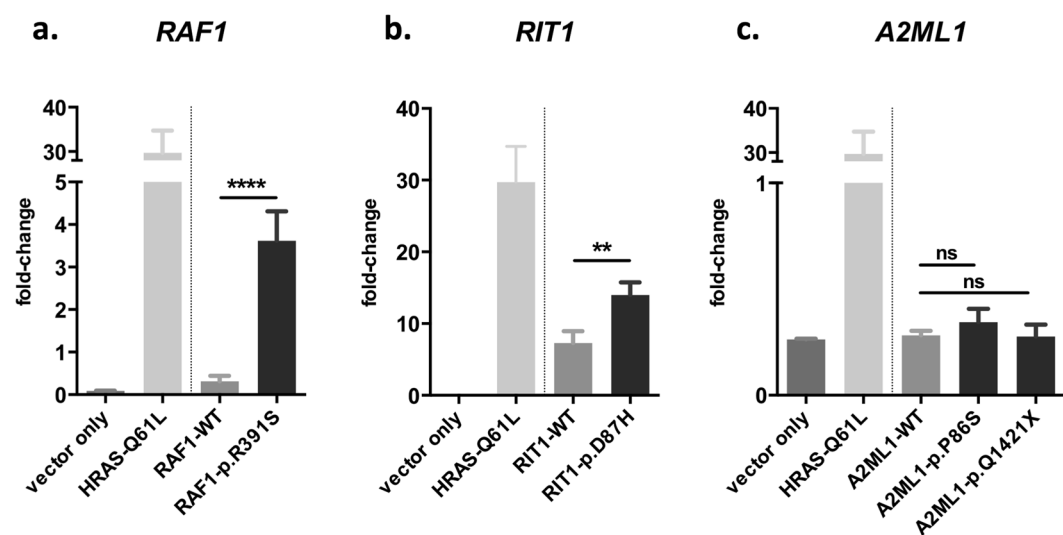
## Materials and Methods

**Patient recruitment.** This study was approved by the Institutional Review Board of the University of Hong Kong/ Hospital Authority Hong Kong West Cluster (HKU/HA HKW IRB). Sixty-three Chinese patients presenting with clinical suspicion of RASopathies were recruited from the Clinical Genetics Service, Department of Health, Hong Kong, and Department of Paediatrics and Adolescent Medicine, the University of Hong Kong, Hong Kong. These subjects were evaluated by clinical geneticists from the two centres. Because *PTPN11* and *HRAS* sequencing were available at these clinics, all 63 subjects were tested and were negative for mutations in these two genes. Informed consent was obtained from the parents and/or patients for the participation, and publication of identifying images in this study.



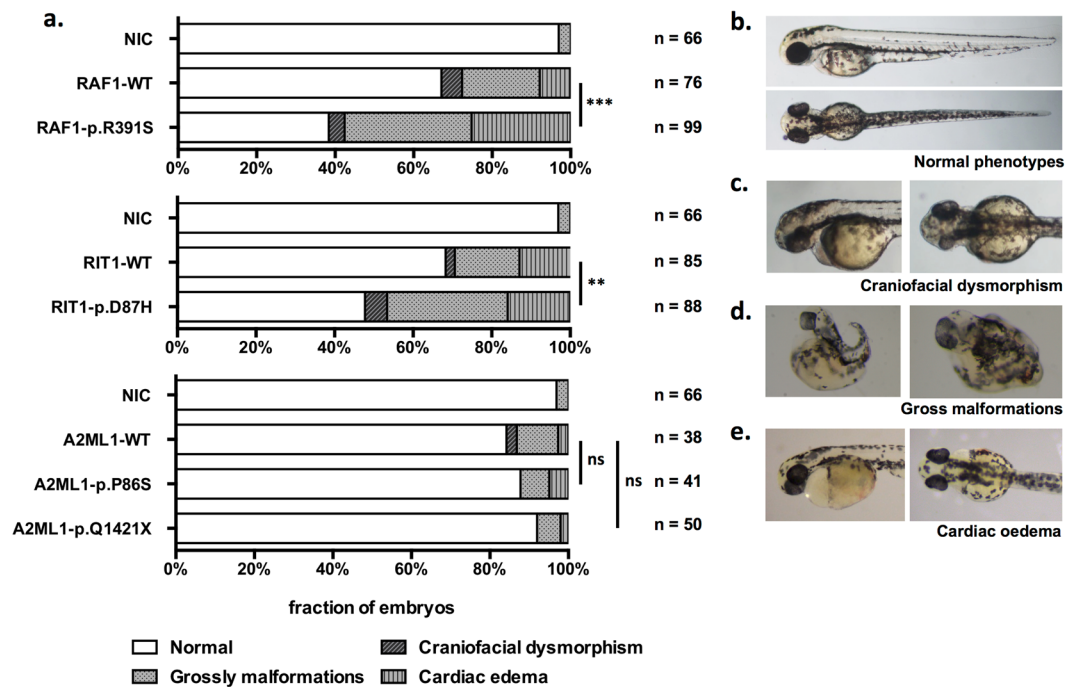
Patient number	Gene	Mutation	Clinvar <sup>a</sup>	HGMD <sup>b</sup>	NseuroNet <sup>c</sup>	Allelic frequency in gnomAD	Allelic frequency in ExAC	exceeded maximum credible population allelic frequency <sup>%</sup>	GERP score	DANN score	Mutation Taster	PROVEAN	SIFT
4868	<i>RAF1</i>	c.1173 G > C.p. (R391S)	—	—	—	Not reported	Not reported	—	4.82	0.9977	Disease-causing	Damaging	Damaging
4869	<i>RAF1</i>	c.1173 G > C.p. (R391S)	—	—	—	Not reported	Not reported	—	4.82	0.9977	Disease-causing	Damaging	Damaging
A120067	<i>RIT1</i>	c.259 G > C.p. (D87H)	uncertain significance	—	—	Not reported	Not reported	—	5.76	0.9949	Disease-causing	Damaging	Damaging
10235	<i>A2ML1</i>	c.256 C > T.p. (P86S)	—	—	—	15 in 245,782 alleles $6.103 \times 10^{-5}$	9 in 119,026 alleles $7.561 \times 10^{-5}$	No	3.99	0.9984	Polymorphism	Damaging	Tolerated
15424	<i>A2ML1</i>	c.4261 C > T.p. (Q1421X)	—	—	—	58 in 246,126 alleles $2.357 \times 10^{-4}$	27 in 120,712 alleles $2.237 \times 10^{-4}$	Yes	3.32	0.9952	Disease-causing (automatic)	—	—

**Table 2.** A summary of five patients identified with VUSs from our cohort using the multigene panel sequencing approach. <sup>a</sup>Clinical significance was labelled according to the ClinVar database from NCBI (<https://www.ncbi.nlm.nih.gov/clinvar/>). <sup>b</sup>Variant class was determined according to the HGMD (Human Gene Mutation Database [http://www.hgmd.cf.ac.uk/docs/new\\_help.html](http://www.hgmd.cf.ac.uk/docs/new_help.html)). <sup>c</sup>Variant classification was determined according to the NSEuroNet (European Network on Noonan Syndrome and Related Disorders; <https://nseuro.net.com/php/about.php>). <sup>%</sup>The maximum credible population allelic frequency for Noonan syndrome ( $1.0 \times 10^{-4}$ ) was calculated based on disease prevalence, maximum allelic contribution, maximum genetic contribution and penetrance<sup>37</sup>. Allelic frequency data was extracted from ExAC (Exome Aggregation Consortium; <http://exac.broadinstitute.org>).

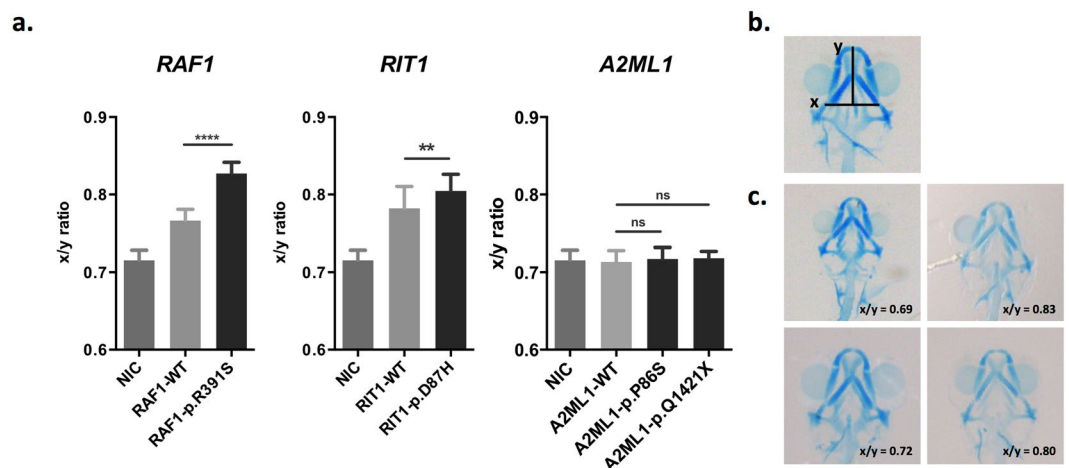


**Figure 2.** Dual luciferase assay of the phosphorylation activity changes to ELK1 from 293 T cells transfected with the corresponding expression and reporter plasmids. Statistical significance was derived using an unpaired t-test comparing the mutants with the corresponding WT. \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ; \*\*\*\* $p < 0.0001$ ; ns: not significant. Mutated human transcript with VUSs from (a) *RAF1* (b) *RIT1* and (c) *A2ML1* from this study were compared with the wild-type transcript. The data are the mean  $\pm$  SD from 3 determinations.

**Next generation sequencing.** Genomic DNA was extracted from patient peripheral blood using standard protocols. We designed a gene enrichment panel for 13 genes known to cause RASopathies in 2013, which included *A2ML1* (NM\_144670.4), *BRAF* (NM\_004333.4), *CBL* (NM\_005188.3), *HRAS* (NM\_005343.3), *KRAS* (NM\_004985.3), *MAP2K1* (NM\_002755.3), *MAP2K2* (NM\_030662.3), *NRAS* (NM\_002524.4), *PTPN11* (NM\_002834.3), *RAF1* (NM\_002880.3), *SHOC2* (NM\_007373.3), *SOS1* (NM\_005633.3), and *SPRED1* (NM\_152594.2). The panel covered the coding exons (flanked by  $\pm 25$  bp) of these 13 selected genes (suppl. Figure 1). The size of the final target region was 101.38 kbp with 97.99% coverage of the target region. Target sequencing was performed using the Haloplex custom target enrichment system (Agilent Technologies, CA, USA) on the MiSeq sequencing platform with the MiSeq reagent kit V2. The average read depth in the regions of interest was  $>500\times$ , and over 98% of the covered region had a sequencing depth of at least  $30\times$ . The data were analysed using NextGENe software v.2.3.4.1 (Softgenetics, PA, USA) with the default settings. Regions with low coverage ( $<20\times$  read depth) were re-sequenced by Sanger sequencing. Genetic variant pathogenicity was assessed according to the guidelines for the interpretation of sequence variants suggested by the American College of Medical Genetics and Genomics (ACMG)<sup>10</sup>. Disease-causing mutations were Sanger validated. During the study period, *RIT1* (NM\_006912.4) was identified as a new causative gene for Noonan syndrome<sup>17</sup>. We subsequently designed primers and sequenced the five coding exons of *RIT1*.

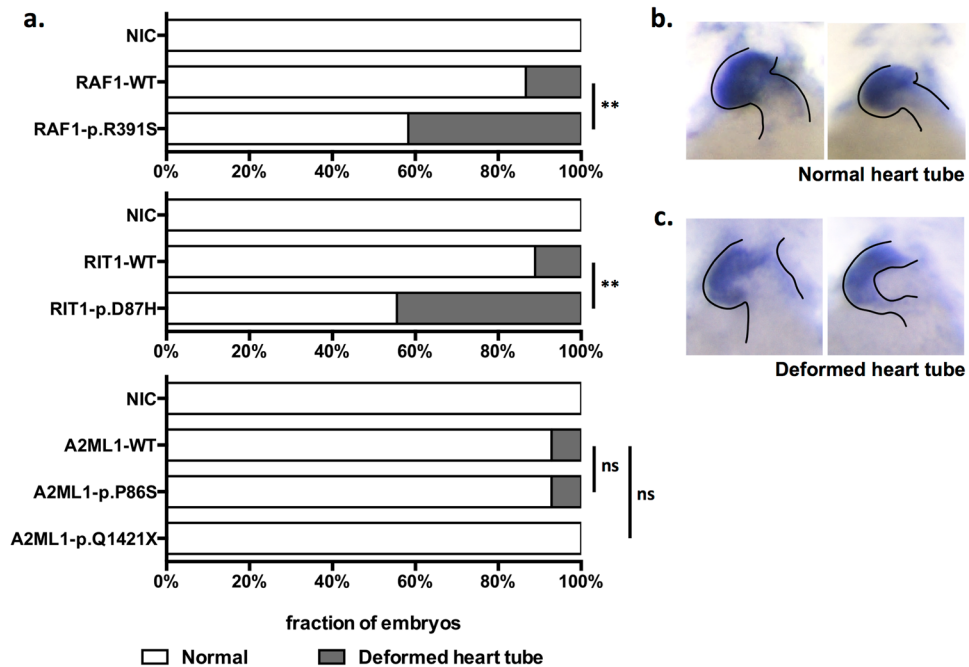


**Figure 3.** Transient expression of RNA transcripts in zebrafish embryos. The injection dosage of each RNA transcript was optimised for comparison (*RAF1*: 50 pg/embryo; *RIT1*: 400 pg/embryo; *A2ML1*: 200 pg/embryo). A morphometric analysis was performed to compare the effect of the VUSs compared to the wild-type transcript at three dpf. Statistical significance was derived using a two-sided Fisher's exact test to compare mutants with the corresponding WT. \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ; ns: not significant. NIC: No injection control; (a) Proportion of zebrafish embryos with normal or diseased phenotypes. (b) Representative zebrafish embryos with a normal phenotype (c) craniofacial dysmorphism (d) gross malformations and (e) cardiac oedema.



**Figure 4.** Craniofacial assessment of zebrafish embryos. The zebrafish embryos were treated with Alcian blue and washed with acidic alcohol 3 dpf. The ratio of the width of the ceratohyal (x) to the tip of Meckel's cartilage (y) was used as a measure of craniofacial defects. (a) Comparison of the x-to-y ratio between mutant and wild-type zebrafish embryos. The data are the mean  $\pm$  SD. Statistical significance was derived using an unpaired t-test comparing the mutants with the corresponding WT. \*\* $p < 0.01$ ; \*\*\* $p < 0.0001$ ; ns: not significant. NIC: No injection control. The number in bold represents the total count of zebrafish embryos in the experimental group. (b) Alcian-blue stained cartilage in the zebrafish. The x-axis shows the width of the ceratohyal, and the y-axis shows the length of the tip of Meckel's cartilage. (c) Representative embryos with different x-to-y ratios.

**Functional analysis.** Genetic variants classified as VUSs were selected for functional characterization. The VUSs were *RAF1*:NM\_002880.2:c.1173 G > C:p.(R391S), *A2ML1*:NM\_144670.4:c.256 C > T:p.(P86S), *A2ML1*:NM\_144670.4:c.4261 C > T:p.(Q1421X) and *RIT1*:NM\_006912.4:c.259 G > C:p.(D87H). Expression plasmids for *RAF1* (#RC201983), *A2ML1* (#RC219615) and *RIT1* (#RC220552) were ordered from Origene (MD, USA). We introduced a



**Figure 5.** Structural assessment of cardiac tubes by *in situ* hybridization of *cmlc1*. Statistical significance was derived using a two-sided Fisher's exact test to compare mutants with the corresponding WT. \*\* $p < 0.01$ ; ns: not significant. **(a)** The proportion of zebrafish embryos with a normal or deformed heart tube structure. **(b)** Representative embryos with a normal and **(c)** deformed heart tube.

single-base substitution into the corresponding expression vector using Q5<sup>®</sup> Site-Directed Mutagenesis Kit (NEB, UK). The *RAF1*, *A2ML1* and *RIT1* coding sequences were validated using Sanger sequencing before subsequent analysis.

**In vitro dual luciferase reporter assay.** To evaluate the molecular effect of the VUSs identified in the cohort, mutant transcripts were expressed in HEK-293T cells. The phosphorylation level of ELK1, which is a downstream substrate of the RAS/MAPK pathway, was measured as a reporter of pathway activity<sup>17,35</sup>. HEK-293T cells were cultured in DMEM and transfected with expression plasmids using Genejuice transfection reagent (EMD Chemicals, NJ, USA). We introduced pGal-LUC (Clontech), pGal-ELK1 and the expression plasmids into HEK-293T cells. The expression plasmids included the mutant constructs of *RAF1*, *A2ML1* and *RIT1*, or the corresponding wild-type gene vectors for comparison. In each experimental setup, an empty vector was used as the negative control, while a HRAS expression plasmid with p.Q61L mutation was used as a positive control<sup>36</sup>. The cells were harvested 48 hours post-transfection. The luciferase activity was measured using a Dual-Luciferase Reporter assay system (Promega, WI, USA). Transfection efficiencies were normalised to the control plasmid, which was pSV-RLuc (Promega, WI, USA) expressing *Renilla* luciferase. The expression of the DDK-tagged transcript was validated by Western blot using anti-DDK antibodies.

**In vivo transient expression in zebrafish embryos.** To assess the effect of the VUSs on organismal phenotypes, zebrafish were used according to the methods previously described in experiments assessing gene level evidence<sup>9–11</sup>. Briefly, the in-frame coding sequence of *RAF1*, *A2ML1* and *RIT1* were subcloned into the pCMV6-AC-GFP vector such that the GFP coding sequence was fused in-frame to the 3' end of the inserted gene. GFP-tagged RNA transcripts were synthesised using the mMACHINE T7 Transcription Kit (Thermo Fisher Scientific, MA, USA). RNA transcripts were injected into zebrafish embryos at the single-cell stage. GFP signals in the zebrafish embryos indicated synthetic RNA expression and served as a positive selection marker for morphometric assessment. Phenotypic observation was performed 3 days post-fertilization (dpf), as previously described<sup>18</sup>. The embryos were anaesthetised with tricaine methanesulfonate and fixed in 4% paraformaldehyde (PFA), and their general body structure and the development of craniofacial and cardiac structure were assessed. Cartilage was stained with 0.1% Alcian Blue in acidic ethanol. The width and length of the cranio-cartilage were measured using ImageJ software. The ratio of the width of the ceratohyal to the tip of Meckel's cartilage was determined as a measure of craniofacial defects. For cardiac defects, PFA fixed embryos were probed with digoxigenin-labelled *cmlc1*, which is a cardiac development marker for zebrafish, by *in situ* hybridization. Total protein from zebrafish embryos was also extracted to examine the correlation between the zebrafish phenotypes and the activation of RAS signals. The study was approved by the Committee of the Use of Live Animals for Teaching and Research in The University of Hong Kong.

**Data availability.** The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

## References

- Sun, Y. *et al.* Next-generation diagnostics: gene panel, exome, or whole genome? *Hum Mutat* **36**, 648–655, <https://doi.org/10.1002/humu.22783> (2015).
- Yang, Y. *et al.* Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *The New England journal of medicine* **369**, 1502–1511, <https://doi.org/10.1056/NEJMoa1306555> (2013).
- Biesecker, L. G. & Green, R. C. Diagnostic clinical genome and exome sequencing. *The New England journal of medicine* **371**, 1170, <https://doi.org/10.1056/NEJMc1408914> (2014).
- Glockle, N. *et al.* Panel-based next generation sequencing as a reliable and efficient technique to detect mutations in unselected patients with retinal dystrophies. *European Journal of Human Genetics* **22**, 99–104, <https://doi.org/10.1038/ejhg.2013.72> (2014).
- Consugar, M. B. *et al.* Panel-based genetic diagnostic testing for inherited eye diseases is highly accurate and reproducible, and more sensitive for variant detection, than exome sequencing. *Genetics in Medicine* **17**, 253–261, <https://doi.org/10.1038/gim.2014.172> (2015).
- Pajusalu, S. *et al.* Large Gene Panel Sequencing in Clinical Diagnostics—Results from 501 Consecutive Cases. *Clinical genetics*. <https://doi.org/10.1111/cge.13031> (2017).
- Rehm, H. L. Evolving health care through personal genomics. *Nature Reviews Genetics* **18**, 259–267 (2017).
- Xue, Y., Ankala, A., Wilcox, W. R. & Hegde, M. R. Solving the molecular diagnostic testing conundrum for Mendelian disorders in the era of next-generation sequencing: single-gene, gene panel, or exome/genome sequencing. *Genetics in medicine: official journal of the American College of Medical Genetics* **17**, 444–451, <https://doi.org/10.1038/gim.2014.122> (2015).
- Wallis, Y. *et al.* Practice guidelines for the evaluation of pathogenicity and the reporting of sequence variants in clinical molecular genetics. *Association for Clinical Genetic Science and the Dutch Society of Clinical Genetic Laboratory Specialists* (2013).
- Richards, S. *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. *Genetics in medicine: official journal of the American College of Medical Genetics* **17**, 405–424, <https://doi.org/10.1038/gim.2015.30> (2015).
- MacArthur, D. G. *et al.* Guidelines for investigating causality of sequence variants in human disease. *Nature* **508**, 469–476, <https://doi.org/10.1038/nature13127> (2014).
- Strande, N. T. *et al.* Evaluating the Clinical Validity of Gene-Disease Associations: An Evidence-Based Framework Developed by the Clinical Genome Resource. *Am J Hum Genet* **100**, 895–906, <https://doi.org/10.1016/j.ajhg.2017.04.015> (2017).
- Aoki, Y., Niihori, T., Narumi, Y., Kure, S. & Matsubara, Y. The RAS/MAPK syndromes: novel roles of the RAS pathway in human genetic disorders. *Hum Mutat* **29**, 992–1006, <https://doi.org/10.1002/humu.20748> (2008).
- Kruszka, P. *et al.* Noonan syndrome in diverse populations. *American journal of medical genetics. Part A* **173**, 2323–2334, <https://doi.org/10.1002/ajmg.a.38362> (2017).
- Cizmarova, M. *et al.* New Mutations Associated with Rasopathies in a Central European Population and Genotype-Phenotype Correlations. *Annals of human genetics* **80**, 50–62, <https://doi.org/10.1111/ahg.12140> (2016).
- Kim, E. K. & Choi, E. J. Pathological roles of MAPK signaling pathways in human diseases. *Biochimica et biophysica acta* **1802**, 396–405, <https://doi.org/10.1016/j.bbadis.2009.12.009> (2010).
- Aoki, Y. *et al.* Gain-of-function mutations in RIT1 cause Noonan syndrome, a RAS/MAPK pathway syndrome. *Am J Hum Genet* **93**, 173–180, <https://doi.org/10.1016/j.ajhg.2013.05.021> (2013).
- Visser, L. E. *et al.* Heterozygous germline mutations in A2ML1 are associated with a disorder clinically related to Noonan syndrome. *European journal of human genetics: EJHG* **23**, 317–324, <https://doi.org/10.1038/ejhg.2014.115> (2015).
- Pandit, B. *et al.* Gain-of-function RAF1 mutations cause Noonan and LEOPARD syndromes with hypertrophic cardiomyopathy. *Nature genetics* **39**, 1007–1012, <https://doi.org/10.1038/ng2073> (2007).
- Bonetti, M. *et al.* Distinct and overlapping functions of ptpn11 genes in Zebrafish development. *Plos One* **9**, e94884. <https://doi.org/10.1371/journal.pone.0094884> (2014).
- Kobayashi, T. *et al.* Molecular and Clinical Analysis of RAF1 in Noonan Syndrome and Related Disorders: Dephosphorylation of Serine 259 as the Essential Mechanism for Mutant Activation. *Human Mutation* **31**, 284–294, <https://doi.org/10.1002/humu.21187> (2010).
- Aoki, Y. *et al.* Germline mutations in HRAS proto-oncogene cause Costello syndrome. *Nature genetics* **37**, 1038–1040, <https://doi.org/10.1038/ng1641> (2005).
- Chan, E. Y. W., Stang, S. L., Bottorff, D. A. & Stone, J. C. Mutations in conserved regions 1, 2, and 3 of Raf-1 that activate transforming activity. *Mol Carcinogen* **33**, 189–197, <https://doi.org/10.1002/mc.10031> (2002).
- Hanks, S. K., Quinn, A. M. & Hunter, T. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science* **241**, 42–52 (1988).
- Yaoita, M. *et al.* Spectrum of mutations and genotype-phenotype analysis in Noonan syndrome patients with RIT1 mutations. *Human genetics* **135**, 209–222, <https://doi.org/10.1007/s00439-015-1627-5> (2016).
- Calcagni, G. *et al.* Congenital heart defects in Noonan syndrome and RIT1 mutation. *Genetics in medicine: official journal of the American College of Medical Genetics* **18**, 1320, <https://doi.org/10.1038/gim.2016.137> (2016).
- Kouz, K. *et al.* Genotype and phenotype in patients with Noonan syndrome and a RIT1 mutation. *Genetics in medicine: official journal of the American College of Medical Genetics* **18**, 1226–1234, <https://doi.org/10.1038/gim.2016.32> (2016).
- Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291, <https://doi.org/10.1038/nature19057> (2016).
- MacDonald, J. R., Ziman, R., Yuen, R. K., Feuk, L. & Scherer, S. W. The Database of Genomic Variants: a curated collection of structural variation in the human genome. *Nucleic Acids Res* **42**, D986–992, <https://doi.org/10.1093/nar/gkt958> (2014).
- Aoki, Y., Niihori, T., Inoue, S. & Matsubara, Y. Recent advances in Rasopathies. *Journal of human genetics* **61**, 33–39, <https://doi.org/10.1038/jhg.2015.114> (2016).
- Gasperini, M., Starita, L. & Shendure, J. The power of multiplexed functional analysis of genetic variants. *Nature protocols* **11**, 1782–1787, <https://doi.org/10.1038/nprot.2016.135> (2016).
- Wang, E. *et al.* Predictive genomics: A cancer hallmark network framework for predicting tumor clinical phenotypes using genome sequencing data. *Semin Cancer Biol* **30**, 4–12, <https://doi.org/10.1016/j.semcancer.2014.04.002> (2015).
- Zaman, N. *et al.* Signaling Network Assessment of Mutations and Copy Number Variations Predict Breast Cancer Subtype-Specific Drug Targets. *Cell reports* **5**, 216–223, <https://doi.org/10.1016/j.celrep.2013.08.028> (2013).
- Katsanis, S. H. & Katsanis, N. Molecular genetic testing and the future of clinical genomics. *Nature Reviews Genetics* **14**, 415–426, <https://doi.org/10.1038/nrg3493> (2013).
- Molzan, M. *et al.* Impaired binding of 14-3-3 to C-RAF in Noonan syndrome suggests new approaches in diseases with increased Ras signaling. *Mol Cell Biol* **30**, 4698–4711, <https://doi.org/10.1128/MCB.01636-09> (2010).
- Ng, M. H., Ng, R. K., Kong, C. T., Jin, D. Y. & Chan, L. C. Activation of Ras-dependent Elk-1 activity by MLL-AF4 family fusion oncoproteins. *Experimental hematology* **38**, 481–488, <https://doi.org/10.1016/j.exphem.2010.03.014> (2010).
- Whiffin, N. *et al.* Using high-resolution variant frequencies to empower clinical genome interpretation. *bioRxiv*, 073114 (2016).



## Acknowledgements

The project was funded by The Society for the Relief of Disabled Children, Hong Kong. We would like to acknowledge the patients and their family members for their support in the project.

## Author Contributions

G.K.C.L. wrote the main manuscript and performed the experiments. H.M.L. provided clinical consultation and informed consent to the patients in the study. V.H.M.T., W.W.G. and D.Y.J. provided assistance and analysis for the dual luciferase assay experiment. C.C.Y.M., M.H.C.Y. and W.H.S.W. aided in the analysis and preparation of the manuscript. W.L.W. and Y.W.Y.C. helped with the sample preparation and analysis. W.L.W. and K.Y.K.C. provided comments on the bioinformatic analysis and genetic variant prioritisation. A.C.H.M. and A.Y.H.L. provided assistance and analysis for transient RNA expression in zebrafish embryos. J.A. offered secondary advice on the patients with VUSs in our cohort. I.F.M.L and B.H.Y.C. were the principal investigators of the project. All of the authors reviewed the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-20894-0>.

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018