**Copy Number Variants (CNVs) Affecting Cancer Predisposing Genes (CPGs) Detected As Incidental Findings In Routine Germline Diagnostic Chromosomal Micro-Array (CMA) Testing**

Josie Innes[1], Lisa Reali[2], Jill Clayton-Smith[1], Georgina Hall[1], Derek Lim[2], George Burghel[1], Kim French[1], Unzela Khan[1], Daniel Walker[1], Fiona Lalloo[1], D. Gareth R. Evans[1], Dom McMullan[2], Eamonn R. Maher[3], and Emma R. Woodward[2]

1. Manchester Centre for Genomic Medicine, St Mary's Hospital, Central Manchester University Hospitals NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, M13 9WL, UK.

2. West Midlands Regional Genetics Service, Birmingham Women's and Children's NHS Foundation Trust, Birmingham, B15 2TG, UK

3. Department of Medical Genetics, University of Cambridge and NIHR Cambridge Biomedical Research Centre and Cancer Research UK Cambridge Centre, CB2 OQQ, UK.

Corresponding author:

Dr Emma R. Woodward

Manchester Centre for Genomic Medicine, St Mary's Hospital, Central Manchester University Hospitals NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, M13 9WL, UK

Emma.Woodward@cmft.nhs.uk

Word count: 3,045

## Abstract

### Background

Identification of copy number variations (CNVs) through chromosomal microarray (CMA) testing is first line investigation in individuals with learning difficulties/congenital abnormalities. Although recognised that CMA testing may identify CNVs encompassing a cancer predisposition gene (CPG), limited information is available on the frequency and nature of such results.

### Methods

We investigated CNV gains and losses affecting 39 CPGs in 3,366 pilot index case individuals undergoing CMA testing, and then studied an extended cohort (n=10,454) for CNV losses at 105 CPGs and CNV gains at 9 proto-oncogenes implicated in inherited cancer susceptibility.

### Results

In the pilot cohort, 31/3,366 (0.92%) individuals had a CNV involving one or more of 16/39 CPGs. 30/31 CNVs involved a tumour suppressor gene (TSG), and 1/30 a proto-oncogene (gain of *MET*). *BMPR1A*, *TSC2* and *TMEM127* were affected in multiple cases. In the second stage analysis, 49/10,454 (0.47%) individuals in the extended cohort had 50 CNVs involving 24/105 CPGs. 43/50 CNVs involved a TSG and 7/50 a proto-oncogene (4 gains, 3 deletions). The most frequently involved genes, *FLCN* (n=10) and *SDHA* (n=7), map to the Smith-Magenis and cri-du-chat regions respectively.

### Conclusion

Incidental identification of a CNV involving a CPG is not rare and poses challenges for future cancer risk estimation. Prospective data collection from CPG-CNV cohorts ascertained incidentally and through syndromic presentations is required to determine the risks posed by specific CNVs. In particular, ascertainment and investigation of adults with CPG-CNVs and adults with learning disability and cancer, could provide important information to guide clinical management and surveillance.

## Introduction

The human genome contains marked structural variation and it is over 10 years since the first comprehensive copy number variant (CNV) map of the human genome was published (1). For children presenting with developmental delay/learning difficulties and/or congenital abnormalities, diagnostic germline chromosomal microarray (CMA) for causative CNVs is now a first line investigation and, together with advances in CMA technology leading to improving resolution (2), there are increasingly numbers of patients identified with CNVs of uncertain significance or for which the resulting phenotype is unclear. This is particularly pertinent where an identified CNV encompasses an inherited cancer (cancer predisposition gene/CPG) and there is no relevant personal or family history, a so-called incidental finding.

With the mainstreaming of modern genomic investigations, CMA testing is often ordered by non-genetics health care professionals (e.g. paediatricians) who may have limited familiarity with familial cancer syndromes and are unable to advise on the full significance of the CMA result. Previously, Pichert et al (3) described the frequency of CNVs affecting 47 CPGs in 4,805 CMA analyses. We report an independent replication study on a larger patient cohort (two fold increase in cases and a more extensive CPG list (n=105).

## Methods

Participants and samples

Samples were referred by paediatricians and clinical geneticists where constitutional diagnostic array comparative genomic hybridisation (aCGH) was requested to determine causes of developmental delay, learning difficulties, neurocognitive impairment and/or birth defects. Only arrays pertaining to the index case in a family were included. The following were excluded (i) patients with clinical features and/or a family history suggestive of the involvement of a known cancer predisposing gene (CPG) (ii) samples from prenatal diagnoses; and (iii) results involving whole chromosome or chromosome arm aneuploidy (iv) results where the CNV identified involving a CPG was present in mosaic form. Monozygotic twins were counted as one individual for the purposes of this study. A finding was considered positive where the involvement of a CPG was not suspected before testing (i.e. an incidental finding). CNVs included in the results were those where the CNV was considered to be causative of the index case' presenting features, and also CNVs of either benign or uncertain significance. Approval for the clinical audit study was provided by Birmingham Women's NHS Foundation Trust and Central Manchester University Hospitals NHS Foundation Trust.

Two cohorts of patients were analysed in two stages. Initially, a pilot cohort comprising 3,366 index case samples investigated between 1 Jan 2009 – 30 Sept 2013 at the West Midlands Regional Genetics Laboratory and then an extended cohort comprising 10,454 index case samples between 1 Jan 2011 and 31 Dec 2015 at the Manchester Centre for Genomic Medicine.

Gene Search Lists

The pilot cohort (n=3,366) was analysed for CNV gains and losses involving a core set of 39 genes (4 oncogenes, 35 tumour suppressor genes) associated with familial cancer predisposition syndromes (Table S1). In the 10k extended cohort (10,454 index patients), CNV losses were investigated in a panel of 105 known CPGs (including the 39 genes analysed in the pilot cohort) comprising, 94 genes on the Illumina Trusight Cancer Panel (4) and 11 further candidate CPGs [*CDKN1B*, *CDKN2B*, *ESR2*, *HIF2A*, *HOXB13*, *PDGFRA*, *POLD1*, *POLE*, *SMARCA4*, *SMARCE1*, *SDHA* (5-15)].

Additionally, in the 10k extended cohort, CNV gains at nine of the105 genes [*ALK, EGFR, HRAS, RHBDF2, CDK4, KIT, MET, PDGFRA, RET* (10, 16-23)] were investigated as activating alterations have been described in hereditary cancer predisposition. Partial or whole gene losses and gains were noted and counted as positive findings.

Laboratory methods and bioinformatics analysis

Testing was undertaken in CPA accredited laboratories. aCGH analysis was carried out using DNA extracted from peripheral blood or mouthwash samples using standard techniques.

For the pilot cohort, aCGH was carried out using either the BlueGnome CytoChip 1Mb BAC (Bacterial Artificial Chromosome) array or the Bluegnome 8x60k v2.0 (ISCA) design oligonucleotide array. aCGH data analysis was performed using BlueFuse Multi software. Copy number variant (CNV) detection using the Bluegnome 1Mb BAC array was carried out with a successful BAC inclusion threshold of >95%. Single clones were called as copy number variants using Log2 thresholds of +/-0.3. Copy number variant (CNV) detection using the Bluegnome 8x60k v2.0 (ISCA) design oligonucleotide array was carried out with a minimum 3 probe inclusion using Log2 ratio thresholds of +/-0.3. No minimum size threshold was applied for either platform.

For the 10k extended cohort, aCGH testing was carried out using Oxford Gene Technology (OGT) CytoSureTM ISCA v2 (8x60k) arrays for all cases with the exception of P102 which was tested using OGT CytoSureTM Constitutional v3 Array (8x60k). aCGH data analysis was performed using OGT CytoSureTM Interpret software. Copy number variant (CNV) detection was based on a minimum 4 probe inclusion using Log2 thresholds of $\geq 0.35$ for gains and $\leq -0.6$ for losses, and no minimum size threshold applied.

For both studies, automatically called CNVs were subject to manual assessment to exclude artefacts and a manual screen for mosaic aberrations was also performed. Inheritance studies were performed using karyotype analysis, targeted aCGH or *in situ* hybridisation studies, as appropriate where parental samples were available.

CNV co-ordinates described are based on the minimum affected region as per standard practice and all co-ordinates are GRCh37/hg19 except where otherwise stated.

Statistical Analysis

Data are displayed as mean ± SD. Continuous data were analysed using a two-tailed Student's t-test. A *P*-value $< 0.05$ was considered to be statistically significant.

## Results

Stage 1: Pilot Cohort Analysis

Within the 3,366 index patients there were 31 individuals (15 males, 16 females) harbouring 31 CNVs involving one or more of the 39 CPGs analysed (Table S2). The 'incidental finding' rate was 0.92% (31/3366). Mean age at CNV analysis in individuals with a positive finding was 51.9 months (SEM 14.2, range 0-312 months, median 8 months). In 16/31 cases the CPG-related CNV was considered to be relevant to the clinical phenotype and in 15 individuals the CNV identified was considered to be either unrelated or of uncertain clinical significance. In 10 cases the CNV encompassing the cancer gene was *de novo* and in 14 cases the CNV was inherited (including one where the child inherited the unbalanced form of a parental balanced translocation). The family history was known in 11 of the 14 cases where the CNV was inherited (excluding the case with the unbalanced form of the parental translocation), and there were no clinical features in keeping with a germline pathogenic alteration of the CPG.

Only one of the CNVs involved an oncogene, a gain encompassing *MET*. The remaining 30 CNVs (20 gains and 10 losses) involved a tumour suppressor gene and two CNVs affected multiple CPGs: a gain involving *MSH2* and *MSH6*; and a deletion encompassing *BMPR1A* and *PTEN* thought likely causative of the learning difficulties phenotype. In six cases the CNV arose as consequence of a complex chromosomal rearrangement and, in all six, resulted in the gain of a tumour suppressor gene (*TSC2* x3, *PMS2* x1, *VHL* x2). The 31 "incidental findings" CNVs involved 16/39 (41.0%) CPGs in the pilot stage gene list with *BMPR1A* (in 6 cases), *TSC2* (n=4) and *TMEM127* (n=3) affected in multiple cases (Fig. 1).

Stage 2: Extended Cohort Analysis

49 (17 females, 32 males) of 10,454 individuals (0.47%) had a total of 50 CNVs involving one or more CPGs on the Stage 2 gene list (see Table S1 and Table S3. The mean age at aCGH in individuals with a positive finding was 87.5 months (SEM 15.0, range 0-460 months, median 46.5 months). In 40 of the individuals the array finding involving the cancer gene was thought to be causative of the clinical phenotype in the index individual and in 9 cases the CNV identified was of uncertain clinical significance or unrelated to the *presenting* features. In 27 cases the CNV arose *de novo* and in 6 cases the CNV was inherited. In 3 of

these 6 cases the child had inherited the unbalanced form of a parental balanced translocation. In the remaining 3 cases, the family history was known in 2 cases and there were no clinical features in keeping with a germline pathogenic alteration of the CPG.

7 of the 50 CNVs involved an oncogene: 4 gains (*HRAS* x2, *MET* and *PDGFRA*) and 3 deletions (*MET*x2, one involving both *KIT/PDGFRA*). The remaining 43 CNVs were deletions involving TSGs. In one case there was a heterozygous contiguous deletion of two TSGs (*BLM* and *FANCI*). Four of the CNVs (in three individuals) arose as a consequence of complex chromosomal rearrangement resulting in gain of an oncogene in three cases *(HRAS* x2, *MET* x1) and loss of a tumour suppressor gene (*SDHA*) in one. These 50 CNVs affected 24 of the 105 genes on the search list (24/105 = 22.9%) with CNVs affecting *FLCN* accounting for 10/50 (20%) and of *SDHA* 7/50 (14%) (Fig. 2).

Joint Analysis of Stage 1 and Stage 2 data sets and CPG lists

Oncogene gains: 3 of 13,820 cases (0.02%) in the combined Stage 1 and Stage 2 cohorts had a CNV gain at one or more of the 4 oncogenes (*RET*, *PDGFRA*, *MET* and *KIT*) in the Stage 1 gene list with CNV gains occurring twice at the *MET* locus (including one individual with a complex rearrangement leading to gain) and once at the *PDGFRA* locus. Three individuals had a deletion of one or more of these oncogenes (including the one individual with the deletion of both *MET* and *PDGFRA* and two with a deletion of *MET*) (Table S4).

TSG losses: 30 of 13,820 (0.22%) individuals in the combined Stage 1 and 2 cohorts had a partial or whole deletion involving one or more of 35 tumour suppressor genes (Table S4). In 22/30 cases the CNV identified was thought to be causative of the child's presenting features and was thought to be either unrelated to, or of uncertain significance, in the remaining eight. In 17 cases the CNV identified was *de novo* and was found to be inherited in 6 individuals (including one where the child had inherited the unbalanced form of the parental balanced translocation). 15 of the 39 (38.5%) genes on the common search list (12 tumour suppressor genes and 3 oncogenes) were affected by a deletion CNV with *SDHA* being involved in 9 CNVs and *BMPR1A* in 6 (Fig. 3) (Table S4).

CNVs encompassing CPGs residing within the known chromosomal microdeletion regions: 17p11.2 and 5p15.22

10 individuals (Stage 2 cohort) had a CNV encompassing *FLCN* (chr17:17,115,527-17,140,502) which resides within the Smith-Magenis Syndrome (SMS) region on 17p11.2 (Decipher chr17: 16,773,072 – 20,222,149) (24), accounting for 20% of the total CNVs identified (Fig. 2, Fig. 4). Five of the CNVs were known to be *de novo* and in five the inheritance was unknown. In four individuals (P113, P111, P116, P119) the majority of the CNV overlapped with but did not encompass the SMS region. In the remaining six, the SMS region was contained within the CNV.

Nine individuals (two from Stage 1 and seven from Stage 2) had a deletion encompassing *SDHA* (chr5:218,356-256,814) which resides toward the 5' end of the cri-du-chat Region (Decipher chr5:10,001 – 12,533,304) (24) accounting for 26.0% of the total number of CNVs identified in the combined cohorts (Fig. 3, Fig. 5). In five individuals the CNV was within the cri-du-chat region and, in the remaining four, the CNV identified extended 3' beyond the critical region (Fig. 5). The CNV was inherited in three individuals (including the individual with the unbalanced translocation leading to loss of *SDHA* and gain of *HRAS*), *de novo* in two and inheritance was unknown in four individuals.

## Discussion

CMA testing is now routinely ordered for individuals presenting with undiagnosed learning difficulties and/or developmental abnormalities and is often undertaken outwith the genetics clinic, for example in the paediatric mainstream setting.

Whilst these investigations provide the opportunity for diagnosis, the CNVs identified may encompass or involve genes where intragenic alteration or whole gene copy number losses are known to be associated with predisposition to other condition(s) unrelated to the presenting features and can be classed as incidental findings. Unlike other genome wide molecular genetic diagnostic strategies, such as whole exome and genome sequencing for which results can be filtered in a gene specific manner, identified CNVs are usually visible to the investigator. Incidental finding CNVs involving CPGs can present significant counselling challenges as (i) whilst the phenotype and cancer risks of intragenic mutations in a CPG may be well defined the risks associated with large CNVs are often unclear as deletion of additional *in cis* genes might modify cancer risks (25); (ii) the known cancer risks associated with CPGs are for individuals ascertained because of a family history and are likely to be lower for population-based ascertainment; and (iii) most CNVs involved CPGs associated with later onset cancers whereas CMA is more commonly performed in a paediatric setting (mean age at positive finding in our pilot and extended cohorts was 51.9m and 87.5m respectively). Nevertheless CPG-CNVs cannot be ignored - as exemplified by two infants (P136 and P137, ages at aCGH 0m and 3m respectively) with deletions of ~50Mb and 24.5Mb respectively encompassing *RB1* who subsequently developed clinical retinoblastoma after the CMA was requested. Whilst retinoblastoma is highly penetrant at a young age (mean age diagnosis of bilateral retinoblastoma 15m) (26), and the tumour penetrance for intragenic mutation of other CPGs is often more variable, this highlights that CNVs encompassing a CPG may be of clinical consequence and should be considered as a paradigm for the need to report such findings until more is known regarding their effects.

Indeed recent analysis of a range of CPGs showed that large deletions including whole gene deletions were associated with fairly typical cancer predisposition compared to point mutations (27). Deletions of CPGs with substantial childhood onset risks such as *SMARCB1* (malignant rhabdoid tumour) and *TP53* (brain and sarcoma) also appear to be not infrequent and there is no evidence these deletions are less penetrant than

point mutations (27). On the other hand, we also detected an inherited deletion encompassing *BMPR1A* (P005) where there was no family history of polyposis. Whilst *BMPR1A* mutations are of lower penetrance than *RB1* (28), it likely that other factors influencing penetrance/expression are also involved. Varying phenotypic consequences of large deletions encompassing disease-causing genes are a recognised challenge (29) and the mechanisms underlying such variable phenotypic effects may include combinations of underlying genomic architecture, long range regulatory effects and, more recently recognised, the influence of topology associated domains (30). In addition, for CNVs involving TSGs the somatic "second hit" might result in homozygous loss of many genes in the cancer cell and result in non-viability through loss of an essential gene or by producing, in combination with loss of the CPG, a synthetic lethal state (31).

In four individuals we identified partial deletions of a TSG. Whilst with CMA it is not possible to precisely characterise the breakpoints, we would expect multi-exon deletions to be pathogenic, particularly where they have been described in the corresponding familial cancer syndrome (32-35). However, in these partial deletion cases we cannot exclude the possibility of expression of an abnormally truncated gene product, although one patient, (P029), did subsequently develop features of tuberous sclerosis indicating pathogenicity.

For CNVs resulting in the gain of a TSG or proto-oncogene, the phenotypic consequences can also be very difficult to interpret. CMA gives no positional information (other than where there is also a cytogenetically characterised complex rearrangement, as occurred in 8 individuals) and a CMA-detected copy number gain might be caused by an intragenic duplication that inactivated a TSG or gain of a functional proto-oncogene – either of which might be associated with a cancer risk.

The CPGs most commonly involved in CNVs were *SDHA* and *FLCN* and both reside within the chromosomal micro-deletion regions for cri-du-chat (Decipher chr5:10,001–12,533,304) and Smith-Magenis (Decipher chr17: 16,773,072 – 20,222,149) respectively (24, 36-37). Although toward the 5' end, *SDHA* is within the cri-du-chat deleted region which has an incidence of 1:15,000 – 1:50,000 (38). Intragenic *SDHA* inactivating mutations may be associated with phaeochromocytoma, paraganglioma and gastrointestinal stromal tumours (GIST) (15, 39). Though the penetrance of familial *SDHA* mutations has been estimated at ~40% by age 40 years (40), other evidence suggests that the penetrance is much lower (39) and to date we

are not aware of any SDHA-related tumours reported in patients with cri-du-chat (41). Nevertheless, subject to appropriate ethical considerations, it would be of interest to investigate adults with cri-du-chat deletions involving *SDHA* for subclinical evidence of *SDHA*-related tumours.

*FLCN* lies within the Smith-Magenis syndrome (SMS) region, on 17p11.2 and accounted for 20% of the CNVs identified. Germline mutations in *FLCN* cause Birt-Hogg-Dube (BHD) syndrome which is characterised by the apprearance of fibrofolliculomas from the third decade and renal cell carcinoma in about 25-30% of cases (42). RCC has been described in patients with SMS (43) but the precise risk of RCC in SMS patients with *FLCN* loss is unclear and further information is required to determine whether surveillance for RCC should be offered routinely. Nevertheless, in the presence of lung cysts or fibrofolliculomas (which on average precede RCC in BHD syndrome) it would seem prudent to do so.

Our CNV detection rate encompassing a CPG was between 0.3% (Stage 1 and 2 combined, n=13,820) and 1% (Stage 1 pilot, n=3,366) in individuals undergoing diagnostic CMA. Pichert et al (2011) found CNVs affecting CPGs in 0.6% of 4805 diagnostic arrays and Boone et al (44) detected 0.9% in 9,005 arrays although this study involved a search list of 40 genes involved in adult onset disorders not specifically focussed toward cancer genes.

This study is the largest to date of CNVs affecting CPGs detected as incidental findings has demonstrated that optimal management of incidentally detected CPG-CNVs and will require systematic collection of long-term follow up data and international data sharing. In particular detailed studies of the clinical significance of SDHA and FLCN loss in patients with cri-du-chat and SMS would address the most frequently detected CPG-CNVs. Though CMAs are routinely performed in children with learning disability, significant numbers of adults with learning disability are likely not to have had high resolution CMA testing and routine reinvestigation of such patients could provide important information on cancer risks. In addition, we are compiling a database of adults with pathogenic CNVs and cancer and request that appropriate cases should be notified to Emma.Woodward@cmft.nhs.uk .

## Contributorship Statement

Josie Innes undertook the data extraction and analysis and contributed to the writing of the manuscript. Lisa Reali undertook the data extraction and analysis and contributed to the writing of the manuscript. Jill Clayton-Smith contributed to the study design. Georgina Hall contributed to the data extraction. Derek Lim undertook data extraction and analysis and contributed to the writing of the manuscript. George Burghel undertook data analysis. Kim French undertook data extraction. Unzela Khan undertook data extraction. Daniel Walker undertook data extraction. Fiona Lalloo contributed to the study design. D. Gareth R. Evans contributed to the writing of the manuscript. Dom McMullan contributed to the overall study design. Eamonn R. Maher contributed to the overall study design and writing of the manuscript. Emma R Woodward (ERW) was responsible for the overall design and content of the study and is overall guarantor for the study. ERW also wrote the manuscript.


## Funding

This study did not receive any specific grant funding.


## Competing Interests

There are no conflicts of interest to declare.


## Acknowledgements

Prof. Evans is an NIHR senior investigator and DGE and ERW are supported by the Manchester NIHR Biomedical Research Centre. ERM acknowledges support from NIHR, European Research Council Advanced Researcher Award and Cancer Research UK.

## References


1. Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, Cho EK. Global variation in copy number in the human genome. *Nature* 2006;444:444-54.

2. Miller DT, Adam MP, Aradhya S, Biesecker LG, Brothman AR, Carter NP, Church DM, Crolla JA, Eichler EE, Epstein CJ, Faucett WA, Feuk L, Friedman JM, Hamosh A, Jackson L, Kaminsky EB, Kok k, Krants ID, Kuhn RM, Lee C, Ostell JM, Rosenberg C, Scherer SW, Spinner NB, Stavropoulos DJ, Tepperberg JH, Thorland EC, Vermeesch JR, Waggoner DK, Watson MS, Martin CL, Ledbetter. DH. Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am J Hum Genet* 2010;86:749-64.

3. Pichert G, Mohammed SN, Ahn JW, Ogilvie CM, Izatt L. Unexpected findings in cancer predisposition genes detected by array comparative genomic hybridisation: what are the issues? J Med Genet 2011;48::535-9.

4. http://www.illumina.com/products/trusight_cancer.html. Accessed 21 June 2017.

5. Pellegata NS, Quintanilla-Martinez L, Siggelkow H, Samson E, Bink K, Höfler H, Fend F, Graw J, Atkinson MJ. Germ-line mutations in p27Kip1 cause a multiple endocrine neoplasia syndrome in rats and humans. *Proc Natl Acad Sci USA* 2006;103:15558-63.

6. Jafri M, Wake NC, Ascher DB, Pires DE, Gentle D, Morris MR, Rattenberry E, Simpson MA, Trembath RC, Weber A, Woodward ER, Donaldson A, Blundell TL, Latif F, Maher ER. Germline mutations in the CDKN2B tumor suppressor gene predispose to renal cell carcinoma. *Cancer Discov* 2015;5:723-9.

7. Smith J, Read ML, Hoffman J, Brown R, Bradshaw B, Campbell C, Cole T, Navas JD, Eatock F, Gundara JS, Lian E, McMullan D, Morgan NV, Mulligan L, Morrison PJ, Robledo M, Simpson MA,Smith VE,

Stewart S, Trembath RC, Sidhu S, Togneri FS, Wake NC, Wallis Y, Watkinson JC, Maher ER, McCabe CJ, Woodward ER. Germline ESR2 mutation predisposes to medullary thyroid carcinoma and causes up-regulation of RET expression. *Hum Mol Genet* 2016;25:1836-45.

8. Lorenzo FR, Yang C, Fui MN, Vankayalapati H, Zhuang Z, Huynh T, Grossmann M, Pacak K, Prchal JT. A novel EPAS1/HIF2A germline mutation in a congenital polycythemia with paraganglioma. *J Mol Med (Berl)* 2013; 91:507-12.

9. Ewing CM, Ray AM, Lange EM, Zuhlke KA, Robbins CM, Tembe WD, Wiley KE, Isaacs SD, Johng D, Wang Y, Bizon C, Yan G, Gelzak M, Partin AW, Shanmugam V, Izatt T, Sinari S, Craig DW, Zheng SL, Walsh PC, Montie JE, Xu J, Carpten JD, Isaacs WB, Cooney KA. Germline mutations in HOXB13 and prostate-cancer risk. *N Engl J Med* 2012;366:141-9.

10. Chompret A, Kannengiesser C, Barrois M, Terrier P, Dahan P, Tursz T, Lenoir GM, Bressac-De Paillerets B. PDGFRA germline mutation in a family with multiple cases of gastrointestinal stromal tumor. *Gastroenterology* 2004;126:318-21.

11. Palles C, Cazier JB, Howarth KM, Domingo E, Jones AM, Broderick P, Kemp Z, Spain SL, Guarino E, Salguero I, Sherborne A, Chubb D, Carvajal-Carmona LG, Ma Y, Kaur K, Dobbins S, Barclay E, Gorman M, Martin L, Kovac MB, Humphray S; CORGI Consortium; WGS500 Consortium, Lucassen A, Holmes CC, Bentley D, Donnelly P, Taylor J, Petridis C, Roylance R, Sawyer EJ, Kerr DJ, Clark S, Grimes J, Kearsey SE, Thomas HJ, McVean G, Houlston RS, Tomlinson I. Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas. *Nat Genet* 2013;45:136-44.

12. Schneppenheim R, Frühwald MC, Gesk S, Hasselblatt M, Jeibmann A, Kordes U, Kreuz M, Leuschner I, Subero JI, Obser T, Oyen F, Vater I, Siebert R. Germline nonsense mutation and somatic inactivation of

SMARCA4/BRG1 in a family with rhabdoid tumor predisposition syndrome. *Am J Hum Genet* 2010;86:279-84.

13. Witkowski L, Carrot-Zhang J, Albrecht S, Fahiminiya S, Hamel N, Tomiak E, Grynspan D, Saloustros E, Nadaf J, Rivera B, Gilpin C, Castellsagué E, Silva-Smith R, Plourde F, Wu M, Saskin A, Arseneault M, Karabakhtsian RG, Reilly EA, Ueland FR, Margiolaki A, Pavlakis K, Castellino SM, Lamovec J, Mackay HJ, Roth LM, Ulbright TM, Bender TA, Georgoulias V, Longy M, Berchuck A, Tischkowitz M, Nagel I, Siebert R, Stewart CJ, Arseneau J, McCluggage WG, Clarke BA, Riazalhosseini Y, Hasselblatt M, Majewski J, Foulkes WD. Germline and somatic SMARCA4 mutations characterize small cell carcinoma of the ovary, hypercalcemic type. *Nat Genet* 2014;46:438-43.

14. Smith MJ, O'Sullivan J, Bhaskar SS, Hadfield KD, Poke G, Caird J, Sharif S, Eccles D, Fitzpatrick D, Rawluk D, du Plessis D, Newman WG, Evans DG. Loss-of-function mutations in SMARCE1 cause an inherited disorder of multiple spinal meningiomas. *Nat Genet* 2013;45:295-8.

15. Burnichon N, Brière JJ, Libé R, Vescovo L, Rivière J, Tissier F, Jouanno E, Jeunemaitre X, Bénit P, Tzagoloff A, Rustin P, Bertherat J, Favier J, Gimenez-Roqueplo AP. SDHA is a tumor suppressor gene causing paraganglioma. *Hum Mol Genet* 2010;19:3011-20.

16. Mossé YP, Laudenslager M, Longo L, Cole KA, Wood A, Attiyeh EF, Laquaglia MJ, Sennett R, Lynch JE, Perri P, Laureys G, Speleman F, Kim C, Hou C, Hakonarson H, Torkamani A, Schork NJ, Brodeur GM, Tonin GP, Rappaport E, Devoto M, Maris JM. Identification of ALK as a major familial neuroblastoma predisposition gene. *Nature* 2008 16;455:930-5.

17. Bell DW, Gore I, Okimoto RA, Godin-Heymann N, Sordella R, Mulloy R, Sharma SV, Brannigan BW, Mohapatra G, Settleman J, Haber DA. Inherited susceptibility to lung cancer may be associated with the T790M drug resistance mutation in EGFR. *Nat Genet* 2005;37:1315-6.

18. Aoki Y, Niihori T, Kawame H, Kurosawa K, Ohashi H, Tanaka Y, Filocamo M, Kato K, Suzuki Y, Kure S, Matsubara Y. Germline mutations in HRAS proto-oncogene cause Costello syndrome. *Nat Genet* 2005;37:1038-40..

19. Blaydon DC, Etheridge SL, Risk JM, Hennies HC, Gay LJ, Carroll R, Plagnol V, McRonald FE, Stevens HP, Spurr NK, Bishop DT, Ellis A, Jankowski J, Field J K, Leigh IM, South AP, Kelsell DP. RHBDF2 mutations are associated with tylosis, a familial esophageal cancer syndrome. Am J Hum Genet 2012;90:340-6.

20. Zuo L, Weger J, Yang Q, Goldstein AM, Tucker MA, Walker GJ, Hayward N, Dracopoli NC. Germline mutations in the p16INK4a binding domain of CDK4 in familial melanoma. *Nat Genet* 1996;12:97-9.

21. Nishida T, Hirota S, Taniguchi M, Hashimoto K, Isozaki K, Nakamura H, Kanakura Y, Tanaka T, Takabayashi A, Matsuda H, Kitamura Y. Familial gastrointestinal stromal tumours with germline mutation of the KIT gene. *Nat Genet* 1998;19:323-4.

22. Schmidt L, Duh FM, Chen F, Kishida T, Glenn G, Choyke P, Scherer SW, Zhuang Z, Lubensky I, Dean M, Allikmets R, Chidambaram A, Bergerheim UR, Feltis JT, Casadevall C, Zamarron A, Bernues M, Richard S, Lips CJ, Walther MM, Tsui LC, Geil L, Orcutt ML, Stackhouse T, Lipan J, Slife L, Brauch H, Decker J, Niehans G, Hughson MD, Moch H, Storkel S, Lerman MI, Linehan WM, Zbar B. Germline and somatic mutations in the tyrosine kinase domain of the MET proto-oncogene in papillary renal carcinomas. *Nat Genet* 1997;16:68-73.

23. Mulligan LM, Kwok JB, Healey CS, Elsdon MJ, Eng C, Gardner E, Love DR, Mole SE, Moore JK, Papi L. Germ-line mutations of the RET proto-oncogene in multiple endocrine neoplasia type 2A. *Nature* 1993;363:458-60.

24. https://decipher.sanger.ac.uk/browser. Accessed 21 June 2017

25. McNeill A, Rattenberry E, Barber R, Killick P, MacDonald F, Maher ER. Genotype–phenotype correlations in VHL exon deletions. *Am J Med Genet* 2009;149A:2147-51.

26. Lohmann DR, Brandt B, Höpping W, Passarge E, Horsthemke B. The spectrum of RB1 germ-line mutations in hereditary retinoblastoma. *Am J Hum Genet* 1996;58:940-9.

27. Smith MJ, Urquhart JE, Harkness EF, Miles EK, Bowers NL, Byers HJ, Bulman M, Gokhale C, Wallace AJ, Newman WG, Evans DG. The contribution of whole Gene deletions and large rearrangements to the mutation spectrum in inherited tumor predisposing syndromes. *Hum Mutat* 2016;37:250-6.

28. Aytac E, Sulu B, Heald B, O'malley M, LaGuardia L, Remzi FH, Kalady MF, Burke CA, Church JM. Genotype-defined cancer risk in juvenile polyposis syndrome. *Br J Surg* 2015;102:114-8.

29. Girirajan S, Eichler EE. Phenotypic variability and genetic susceptibility to genomic disorders. *Hum Mol Genet* 2010;19:R176-87.

30. Ordulu Z, Kammin T, Brand H, Pillalamarri V, Redin CE, Collins RL, Blumenthal I, Hanscom C, Pereira S, Crandall BF, Gerrol P, Hayden MA, Hussain N, Kanengisser-Pines B, Kantarci S, Levy B, Macera MJ, Quintero-Rivera F, Spiegel E, Stevens B, Ulm JE, Warburton D, Wilkins-Haug LE, Yachelevich N, Gusella JF, Talkowski ME, Morton CC. Structural Chromosomal Rearrangements Require Nucleotide-Level Resolution: Lessons from Next-Generation Sequencing in Prenatal Diagnosis. *Am J Hum Genet* 2016;99:1015-33.

31. Nijman S. Synthetic lethality: general principles, utility and detection using genetic screens in human cells. *FEBS lett* 2011;585:1-6.

32. Rendtorff ND, Bjerregaard B, Frödin M, Kjaergaard S, Hove H, Skovby F, Brøndum-Nielsen K, Schwartz M, Danish Tuberous Sclerosis Group. Analysis of 65 tuberous sclerosis complex (TSC) patients by TSC2 DGGE, TSC1/TSC2 MLPA, and TSC1 long-range PCR sequencing, and report of 28 novel mutations. *Hum Mutat* 2005;26:374-83.

33. Kozlowski P, Roberts P, Dabora S, Franz D, Bissler J, Northrup H, Au KS, Lazarus R, Domanska-Pakiela D, Kotulska K, Jozwiak S, Kwiatowski DJ. Identification of 54 large deletions/duplications in TSC1 and TSC2 using MLPA, and genotype-phenotype correlations. *Hum Genet* 2007;121:389-400.

34. Aretz S, Stienen D, Uhlhaas S, Stolte M, Entius MM, Loff S, Back W, Kaufmann A, Keller KM, Blaas SH, Siebert R, Vogt S, Spranger S, Holinski-Feder E, Sunde L, Propping P, Friedl W. High proportion of large genomic deletions and a genotype–phenotype update in 80 unrelated families with juvenile polyposis syndrome. *J Med Genet* 2007;44:702-9.

35. Cheah PY, Wong YH, Chau YP, Loi C, Lim KH, Lim JF, Koh PK, Eu KW. Germline bone morphogenesis protein receptor 1A mutation causes colorectal tumorigenesis in hereditary mixed polyposis syndrome. *Am J Gastroenterol* 2009;104:3027-33.

36. Lejeune J, Lafourcade J, Berger R, Vialatte J, Boeswillwald M, Seringe P, Turpin R. 3 Cases of Partial Deletion of the Short Arm of a 5 Chromosome. C R Hebd Seances Acad Sci 1963;257:3098-102.

37. Smith A, McGavran L, Robinson J, Waldstein G, Macfarlane J, Zonona J, Reiss J, Lahr M, Allen L, Magenis E, Opitz JM, Reynolds JF. Interstitial deletion of (17)(p11. 2p11. 2) in nine patients. *Am J Med Genet* 1986;24:393-414.

38. Mainardi PC, Perfumo C, Cali A, Coucourde G, Pastore G, Cavani S, Zara F, Overhauser J, Pierluigi M, Bricarelli FD. Clinical and molecular characterisation of 80 patients with 5p deletion: genotype-phenotype correlation. *J Med Genet* 2001;38:151-8.

39. Casey RT, Ascher DB, Rattenberry E, Izatt L, Andrews KA, Simpson HL, Challis B, Park SM, Bulusu VR, Lalloo F, Pires DEV, West H, Clark GR, Smith PS, Whitworth J, Papathomas TG, Taniere P, Savisaar R, Hurst LD, Woodward ER, Maher ER. SDHA related tumorigenesis: a new case series and literature review for variant interpretation and pathogenicity. *Mol Genet Genomic Med* 2017;5:237-50.

40. Bausch B, Schiavi F, Ni Y, Welander J, Patocs A, Ngeow J, Wellner U, Malinoc A, Taschin E, Barbon G, Lanza V, Söderkvist P, Stenman A, Larsson C, Svahn F, Chen JL, Marquard J, Fraenkel M, Walter MA, Peczkowska M, Prejbisz A, Jarzab B, Hasse-Lazar K, Petersenn S, Moeller LC, Meyer A, Reisch N, Trupka A, Brase C, Galiano M, Preuss SF, Kwok P, Lendvai N, Berisha G, Makay Ö, Boedeker CC, Weryha G, Racz K, Januszewicz A, Walz MK, Gimm O, Opocher G, Eng C, Neumann HP; European-American-Asian Pheochromocytoma-Paraganglioma Registry Study Group. Clinical characterization of the pheochromocytoma and paraganglioma susceptibility genes sdha, tmem127, max, and sdhaf2 for gene-informed prevention. *JAMA oncology* 2017 Apr 6. [Epub ahead of print]

41. Guala A, Spunton M, Kalantari S, Kennerknecht I, Danesino C. Neoplasia in Cri du Chat Syndrome from Italian and German Databases. Case Rep Genet 2017; Apr 24;.

42. Menko FH, Van Steensel MA, Giraud S, Friis-Hansen L, Richard S, Ungari S, Nordenskjöld M, vO Hansen T, Solly J, Maher ER, European BHD Consortium. Birt-Hogg-Dubé syndrome: diagnosis and management. *Lancet Oncol* 2009;10:1199-206.

43. Dardour L, Verleyen P, Lesage K, Holvoet M, Devriendt K. Bilateral renal tumors in an adult man with Smith-Magenis syndrome: The role of the FLCN gene. *Eur J Med Genet* 2016;59:499-501.

44. Boone PM, Soens ZT, Campbell IM, Stankiewicz P, Cheung SW, Patel A, Beaudet AL, Plon SE, Shaw CA, McGuire AL, Lupski JR. Incidental copy-number variants identified by routine genome testing in a clinical population. *Genet Med* 2013;15:45-54.

## Figure Legends

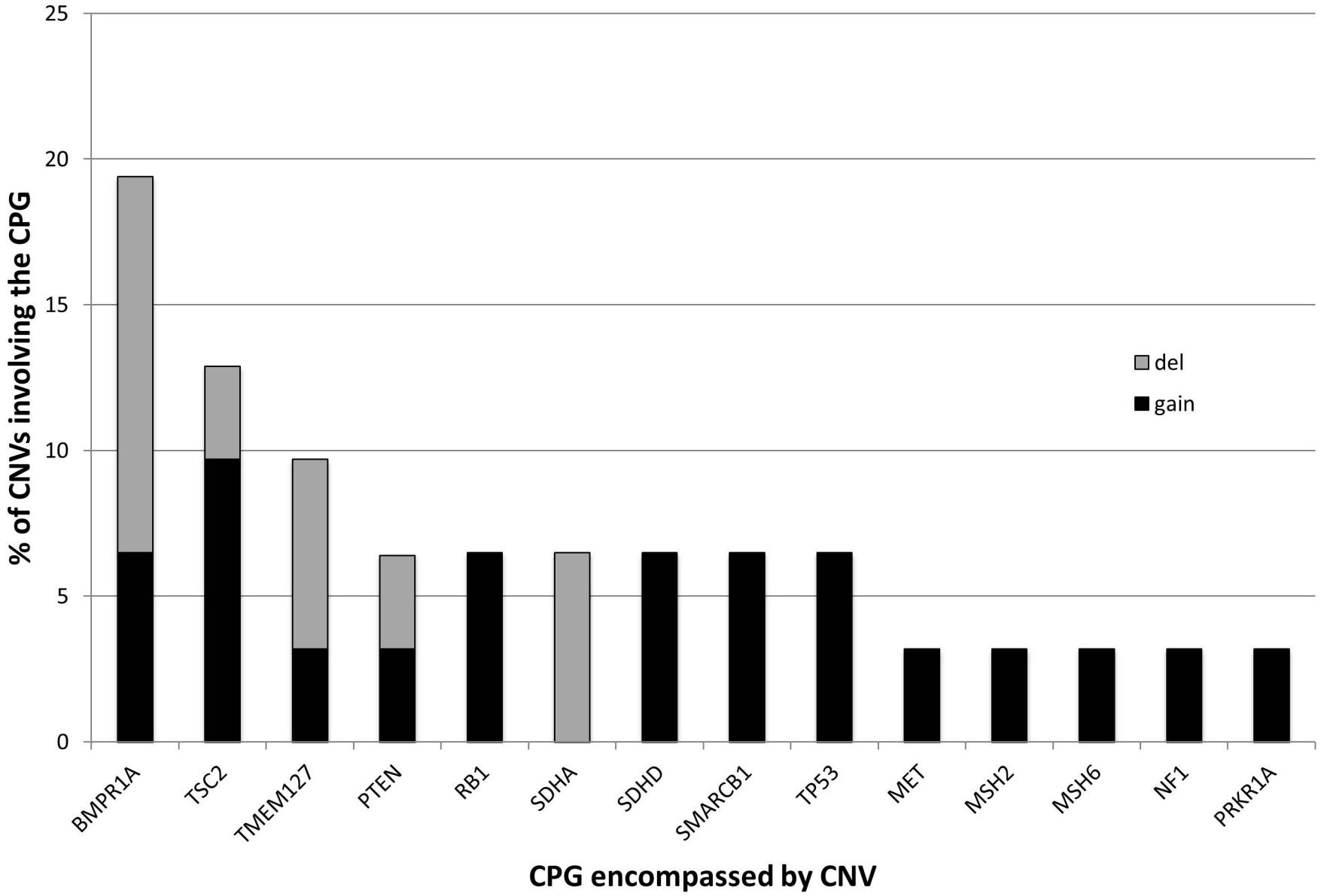
**Figure 1.** Percentage of CNVs detected in the Stage 1 pilot cohort affecting the CPG shown and whether the CNVs detected were gains or deletions. Where a CPG gene present in Table S1 is not shown then no CNV involving it was detected.

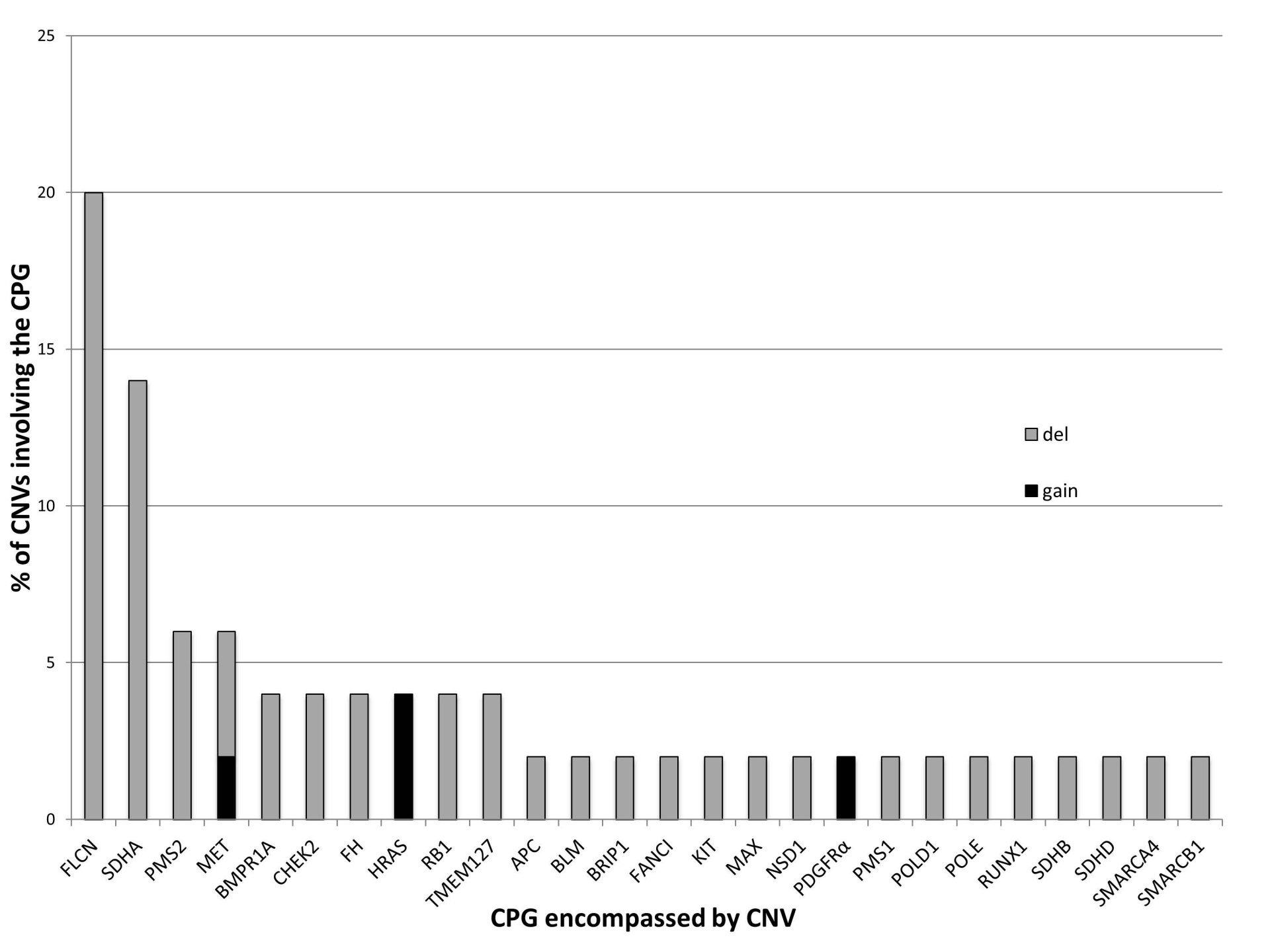
**Figure 2.** Percentage of CNVs detected in the Stage 2 extended cohort affecting the CPGs shown and whether the CNVs detected were gains or deletions. Where a CPG is not shown then no CNV affecting it was detected.

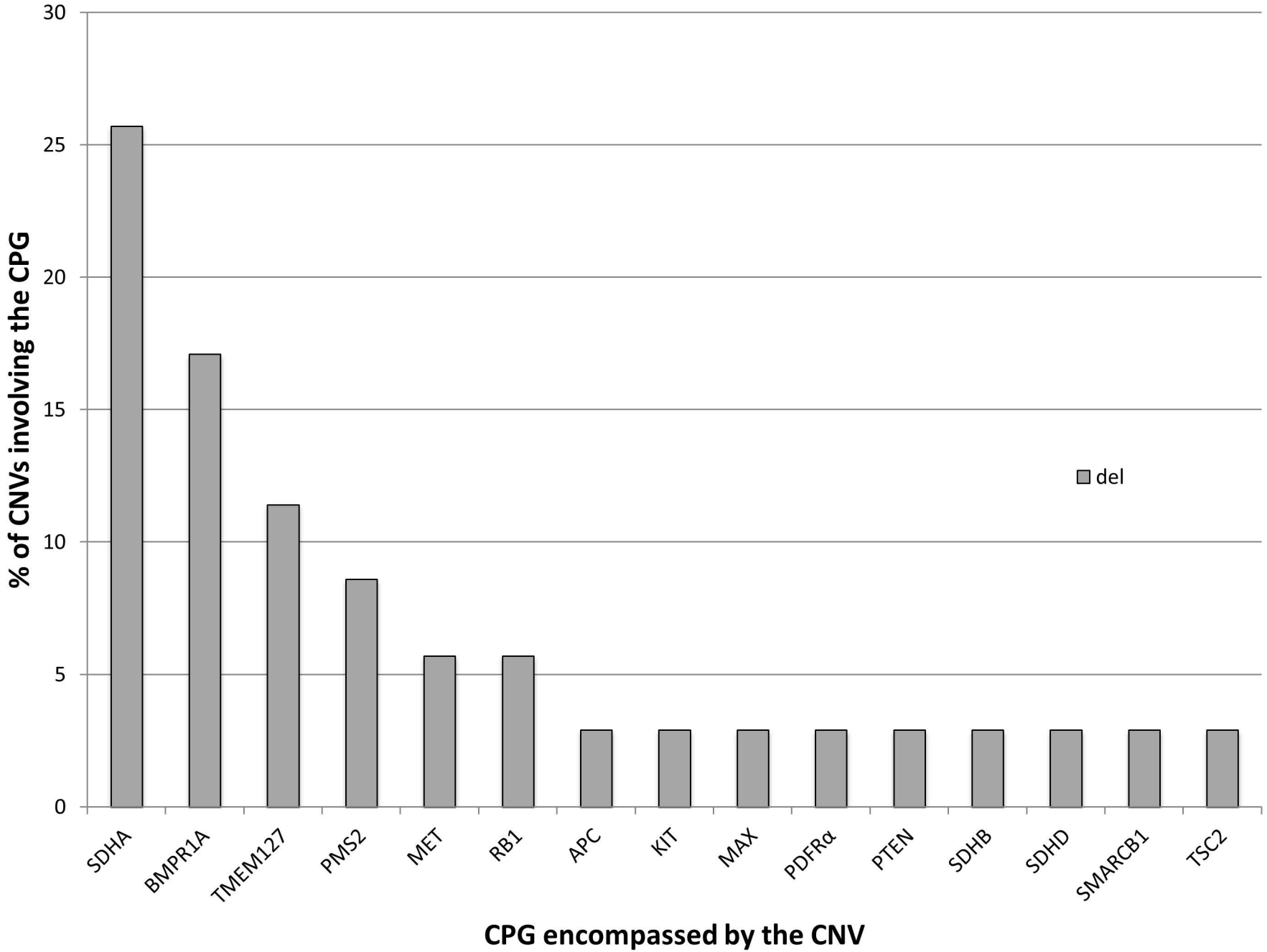
**Figure 3.** Percentage of deletion CNVs detected in the combined Stage 1 pilot and Stage 2 extended cohorts affecting the CPGs shown. Where a CPG is not shown then no deletion CNV affecting it was detected.

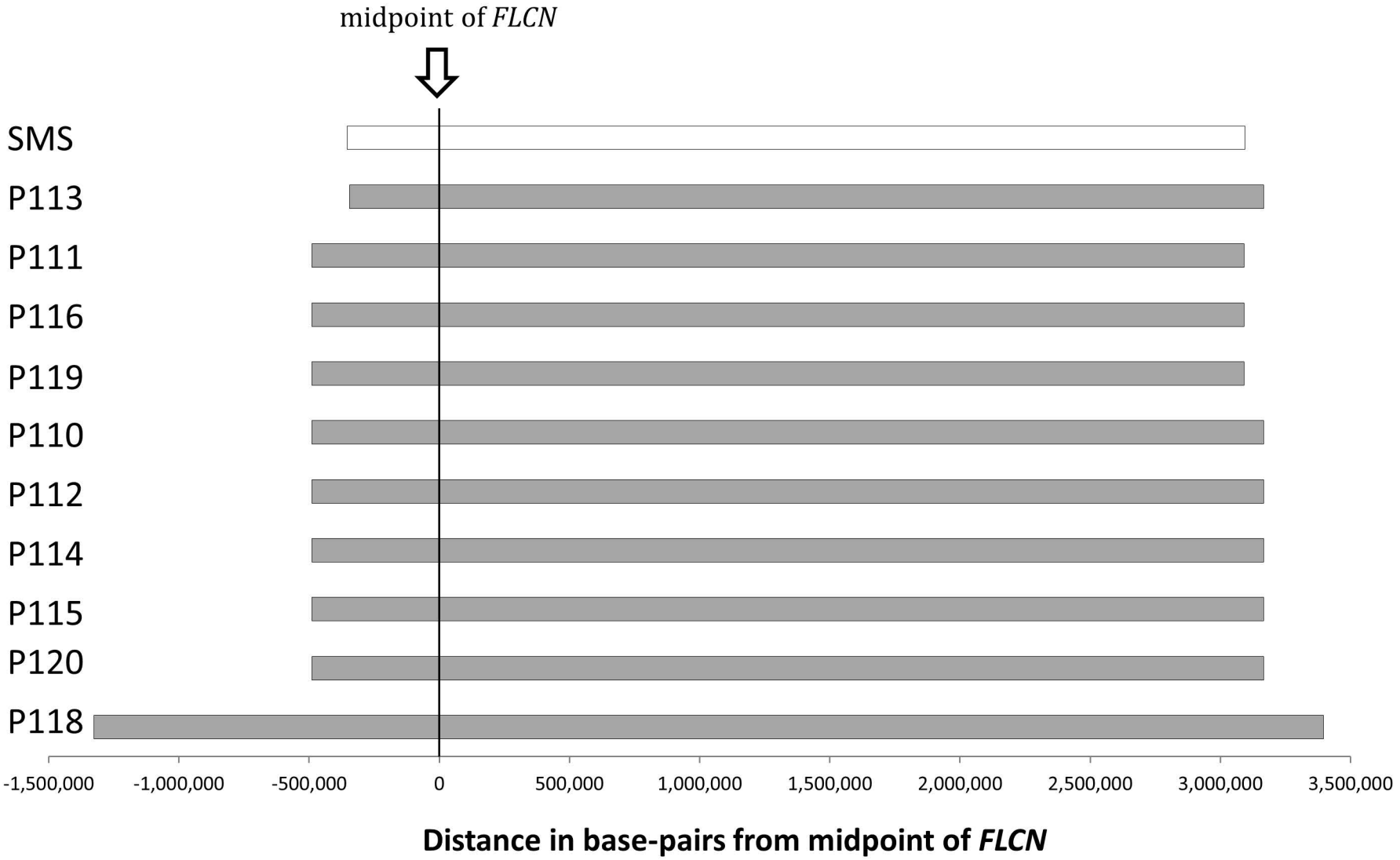
**Figure 4.** Schematic diagram of the 10 CNVs encompassing *FLCN* relative to the mid-point of *FLCN*. Each bar represents the CNV identified in the patient shown. The central axis represents the mid-point of *FLCN*, and the distance in base-pairs from this mid-point is shown on the horizontal X-axis.

**Figure 5.** Schematic diagram of the 9 CNVs encompassing *SDHA* relative to the mid-point of *SDHA*. Each bar represents the CNV identified in the patient shown. The central axis represents the mid-point of *SDHA*, and the distance in base-pairs from this mid-point is shown on the horizontal X-axis.


25
20
15
10
5
0
% of CNVs involving the CPG
del
gain
BMPR1A
TSC2
TMEM127
PTEN
RB1
SDHA
SDHD
SMARCB1
TP53
MET
MSH2
MSH6
NF1
PRKR1A
CPG encompassed by CNV


25
20
15
10
5
0
% of CNVs involving the CPG
del
gain
FLCN
SDHA
PMS2
MET
BMPR1A
CHEK2
FH
HRAS
RB1
TMEM127
APC
BLM
BRIP1
FANCI
KIT
MAX
NSD1
PDGFRα
PMS1
POLD1
POLE
RUNX1
SDHB
SDHD
SMARCA4
SMARCB1
CPG encompassed by CNV


30
25
20
15
10
5
0
% of CNVs involving the CPG
del
SDHA
BMPR1A
TMEM127
PMS2
MET
RB1
APC
KIT
MAX
PDFRα
PTEN
SDHB
SDHD
SMARCB1
TSC2
CPG encompassed by the CNV


midpoint of *FLCN*
SMS
P113
P111
P116
P119
P110
P112
P114
P115
P120
P118
-1,500,000
-1,000,000
-500,000
0
500,000
1,000,000
1,500,000
2,000,000
2,500,000
3,000,000
3,500,000
Distance in base-pairs from midpoint of *FLCN*

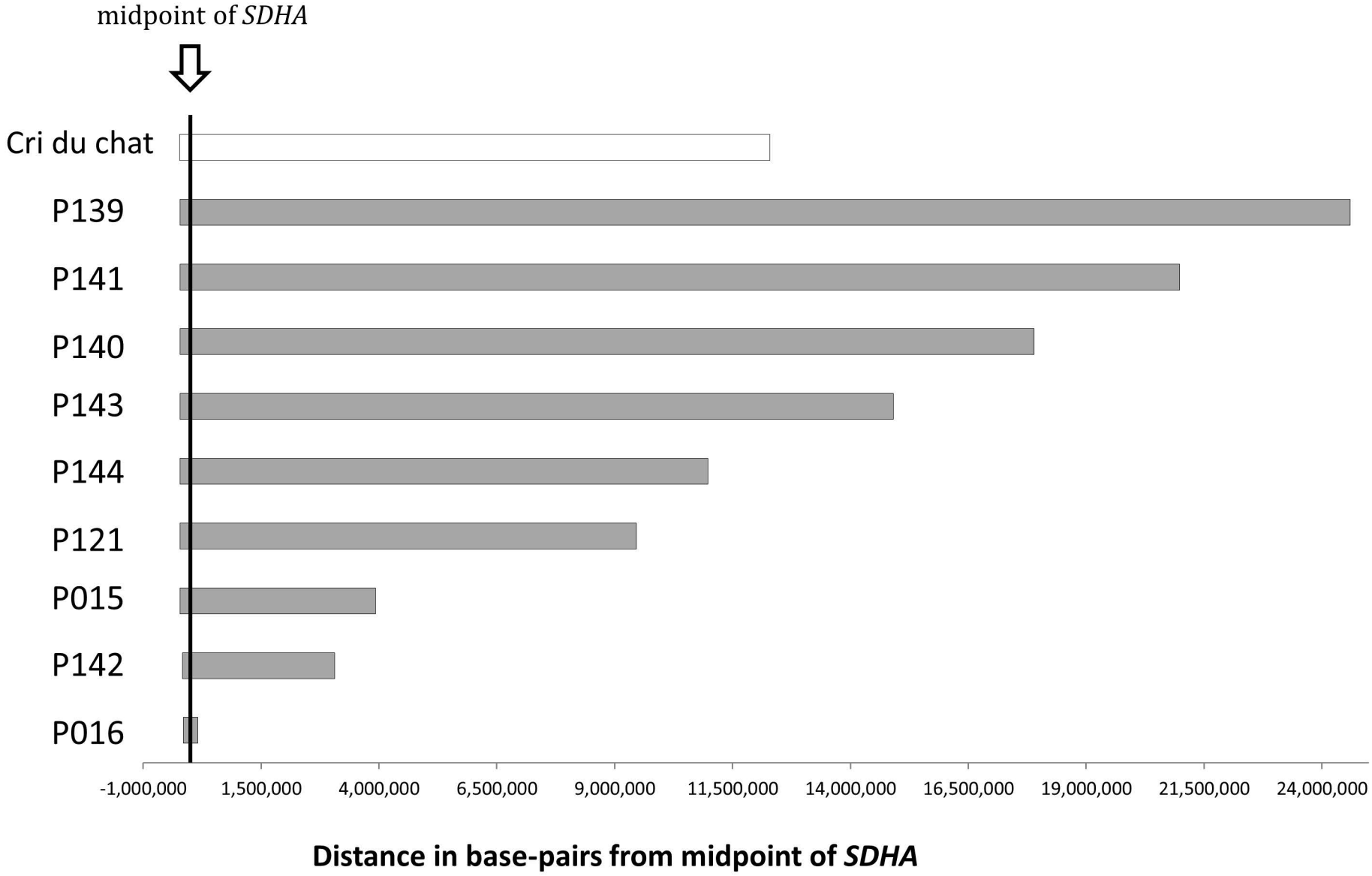

midpoint of *SDHA*
Cri du chat
P139
P141
P140
P143
P144
P121
P015
P142
P016
-1,000,000
1,500,000
4,000,000
6,500,000
9,000,000
11,500,000
14,000,000
16,500,000
19,000,000
21,500,000
24,000,000
Distance in base-pairs from midpoint of *SDHA*

**Table S1. Cancer Predisposing Genes (CPGs).**

| CPG | CPG in Pilot Cohort Analysis ? | CPG in Extended Cohort Analysis? | CPG MIM number | Chromosomal Location | Bi-allelic or Mono-allelic (with respect to cancer predisposion) | Condition associated with tumour predisposition | Phenotype(s) MIM number(s) |
|---|---|---|---|---|---|---|---|
| AIP | No | Yes | 605555 | 11q13.3 | Monoallelic | Pituitary adenoma predisposition | 600634 / 102200 |
| ALK* | No | Yes | 105590 | 2p23.2-23.1 | Monoallelic | Neuroblastoma susceptibility | 613014 |
| APC[#] | Yes | Yes | 611731 | 5q21-q22 | Monoallelic | Familial Adenomatous Polyposis Coli | 175100 |
| ATM | No | Yes | 607585 | 11q22.3 | Monoallelic | Breast cancer susceptibility | 114480 |
| | | | | | Biallelic | Ataxia Telangiectasia | 208900 |
| BAP1 | No | Yes | 603089 | 3p21.1 | Monoallelic | Tumour Predisposition Syndrome | 614327 |
| BLM | No | Yes | 604610 | 11q26.1 | Biallelic | Bloom Syndrome | 210900 |
| | | | | | Monoallelic | Colorectal cancer susceptibility | 114500 |
| BMPR1A | Yes | Yes | 174900 | 10q23.2 | Monoallelic | Juvenile Polyposis Syndrome | 174900 |
| BRCA1 | Yes | Yes | 113705 | 17q21.31 | Monoallelic | Breast and ovarian cancer susceptibility | PS604370 |
| BRCA2 | Yes | Yes | 600185 | 13q13.1 | Monoallelic | Breast and ovarian cancer susceptibilty | PS604370 |
| | | | | | Biallelic | Fanconi anaemia (D1) | 605724 |
| BRIP1 | No | Yes | 605882 | 17q23.2 | Monoallelic | Breast cancer susceptibility | 114480 |
| | | | | | Biallelic | Fanconi anaemia (J) | 609054 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| BUB1B | No | Yes | 602860 | 15q15.1 | Biallelic | Mosaic Variegated Aneuploidy Syndrome 1 | 257300 |
| CDC73 | Yes | Yes | 607393 | 1q31.2 | Monoallelic | Hyperparathyroidism-jaw tumour syndrome | 145001 |
| CDH1 | Yes | Yes | 192090 | 16q22.1 | Monoallelic | Hereditary diffuse gastric cancer | 137215 |
| CDK4* | No | Yes | 123829 | 9p21.3 | Monoallelic | Melanoma susceptibility | 609048 |
| CDKN1B | No | Yes | 600778 | 12p13.1 | Monoallelic | Multiple endocrine neoplasia type IV | 610755 |
| CDKN1C | No | Yes | 600856 | 11p15.4 | Monoallelic | Beckwith-Wiedemann Syndrome | 130650 |
| CDKN2A | Yes | Yes | 600160 | 9p21.3 | Monoallelic | Melanoma susceptibility | 606719 / 155755 / 155601 |
| CDKN2B | No | Yes | 600431 | 9p21.3 | Monoallelic | Renal cancer susceptibility | 14700 |
| CEBPA | No | Yes | 116897 | 19q13.11 | Monoallelic | Acute myeloid leukaemia susceptibility | 601626 |
| CEP57 | No | Yes | 607951 | 11q21 | Bialleleic | Mosaic variegated aneuploidy syndrome 2 | 614114 |
| CHEK2 | No | Yes | 604373 | 22q12.1 | Monoallelic | Breast cancer susceptibility | 114480 |
| CYLD | No | Yes | 605018 | 16q12.1 | Monoallelic | Familial cylindromatosis | 132700 / 601606 / 605041 |
| DDB2 | No | Yes | 60811 | 11p11.2 | Biallelic | Xeroderma pigmentosum (E) | 278740 |
| DICER1[#] | No | Yes | 606241 | 14q32.13 | Monoallelic | Familial pleuropulmonary blastoma tumour predisposition syndrome | 601200 / 138800 / 180295 |
| DIS3L2 | No | Yes | 614184 | 2q37.1 | Biallelic | Perlman syndrome | 267000 |
| EGFR* | No | Yes | 131550 | 7p11.2 | Monoallelic | Non small cell lung cancer susceptibility | 211980 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| EPCAM | No | Yes | 185535 | 2p21 | Monoallelic | Lynch syndrome | 613244 |
| ERCC2 | No | Yes | 126340 | 19q13.32 | Biallelic | Xeroderma pigmentosum (D) | 278730 |
| ERCC3 | No | Yes | 133510 | 2q14.3 | Biallelic | Xeroderma pigmentosum (B) | 610651 |
| ERCC4 | No | Yes | 133520 | 16p13.12 | Biallelic | Xeroderma pigmentosum (F) | 278760 |
| ERCC4 | No | Yes | 133520 | 16p13.12 | Biallelic | Fanconi Anaemia (Q) | 615272 |
| ERCC5 | No | Yes | 133530 | 13q33.1 | Biallelic | Xeroderma pigmentosum (G) | 278780 |
| ESR2 | No | Yes | 601163 | 14q23.2-q23.3 | Monoallelic | Familial Medullary Thyroid Carcinoma | 155240 |
| EXT1 | No | Yes | 608177 | 8q24.11 | Monoallelic | Multiple Exostoses, type 1 | 133700 |
| EXT2 | No | Yes | 608210 | 11p11.2 | Monoallelic | Multiple Exostoses, type 2 | 133701 |
| EZH2 | No | Yes | 6011573 | 7q36.1 | Biallelic | Weaver syndrome | 277590 |
| FANCA | No | Yes | 607139 | 16q24.3 | Biallelic | Fanconi anaemia (A) | 227650 |
| FANCB | No | Yes | 300515 | Xp22.2 | Biallelic | Fanconi anaemia (B) | 300514 |
| FANCC | No | Yes | 613899 | 9q22.32 | Biallelic | Fanconi anaemia (C | 227645 |
| FANCD2 | No | Yes | 613984 | 3p25.3 | Biallelic | Fanconi anaemia (D2) | 227646 |
| FANCE | No | Yes | 613976 | 6p21.31 | Biallelic | Fanconi anaemia (E) | 600901 |
| FANCF | No | Yes | 613897 | 11p14.3 | Biallelic | Fanconi anaemia (F) | 603467 |
| FANCG | No | Yes | 602956 | 9p13.3 | Biallelic | Fanconi anaemia (G) | 614082 |
| FANCI | No | Yes | 611360 | 15q26.1 | Biallelic | Fanconi anaemia (I) | 609053 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| FANCL | No | Yes | 608111 | 2p16.1 | Biallelic | Fanconi anaemia (L) | 614083 |
| FANCM | No | Yes | 609644 | 14q21.2 | Biallelic | Fanconi anaemia? | PS227650 |
| FH | No | Yes | 136850 | 1q43 | Monoallelic | Hereditary leiomyomatosis and renal cell cancer | 150800 |
| | | | | | Biallelic | Fumarase deficiency (no known tumour predisposition) | 606812 |
| FLCN | No | Yes | 607273 | 17p11.2 | Monoallelic | Birt-Hogg-Dube syndrome | 135150 |
| GATA2 | No | Yes | 137295 | 3q21.3 | Monoallelic | Acute myeloid leukemia susceptibility | 614038 / 614172 / 601626 / 614286 |
| GPC3 | No | Yes | 300037 | Xq26.2 | Monoallelic (XLR) | Simpson-Golabi-Behmel Syndrome | 312870 |
| HIF2A | No | Yes | 603349 | 2p21 | Monoallelic | Paraganglioma susceptibility | PS168000 |
| HNF1A | No | Yes | 142410 | 12q24.31 | Monoallelic | Renal cancer susceptibility | 144700 |
| HOXB13 | No | Yes | 604607 | 17q21.31 | Monoallelic | Prostate cancer susceptibility | 176807 |
| HRAS* | No | Yes | 190020 | 11p15.5 | Monoallelic | Costello Syndrome | 218040 |
| KIT* | Yes | Yes | 164920 | 4q12 | Monoallelic | Familial gastrointestinal stromal tumours | 606764 |
| MAX[§] | Yes | Yes | 154950 | 14q23.3 | Monoallelic | Familial Paraganglioma-Pheochromocytoma Syndrome | 171300 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| MEN1 | Yes | Yes | 613733 | 11q13.1 | Monoallelic | Multiple endocrine neoplasia type I | 131100 |
| MET* | Yes | Yes | 164860 | 7q31.2 | Monoallelic | Familial papillary renal cell cancer type 1 | 605074 |
| MLH1 | Yes | Yes | 120436 | 3p22.2 | Monoallelic | Lynch Syndrome | PS120435 |
| MSH2 | Yes | Yes | 609309 | 2p21 | Biallelic<br>Monoallelic | Congenital mismatch repair deficiency<br>Lynch Syndrome | 276300<br>PS120435 |
| MSH6 | Yes | Yes | 600678 | 2p16.3 | Biallelic<br>Monoallelic | Congenital mismatch repair deficiency<br>Lynch Syndrome | 276300<br>PS120435 |
| | | | | | Biallelic | Congenital mismatch repair deficiency | 276300 |
| MUTYH | Yes | Yes | 604933 | 1p34.1 | Biallelic | MUTYH associated polyposis | 608456 |
| NBN | No | Yes | 602667 | 8q21.3 | Biallelic | Nijmegen Breakage Syndrome | 251260 |
| NF1[#] | Yes | Yes | 613113 | 17q11.2 | Monoallelic | Neurofibromatosis type I | 16220 |
| NF2[#] | Yes | Yes | 607379 | 22q12.2 | Monoallelic | Neurofibromatosis type II | 101000 |
| NSD1 | No | Yes | 606681 | 5q35.3 | Monoallelic | Sotos Syndrome | 117550 |
| PALB2 | No | Yes | 610355 | 16p12.2 | Mononallelic | Breast cancer susceptibility | 114480 |
| | | | | | Biallelic | Fanconi anaemia (N) | 610832 |
| PDGFRA* | Yes | Yes | 173490 | 4q12 | Monoallelic | Familial gastrointestinal stromal tumor | 606764 |
| PHOX2B | No | Yes | 603851 | 4p13 | Monoallelic | Neuroblastoma susceptibility | 613013 |
| PMS1 | No | Yes | 600258 | 2q32.2 | Monoallelic | Lynch Syndrome | PS120435 |
| PMS2 | Yes | Yes | #614337 | 7p22.1 | Monoallelic | Lynch Syndrome | PS120435 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| POLD1 | No | Yes | 174761 | 19q13.33 | Biallelic<br>Monoallelic | Congenital mismatch repair deficiency<br>Colorectal cancer susceptibility | 276300<br>612591 |
| POLE | No | Yes | 174762 | 12q24.33 | Monoallelic | Colorectal cancer susceptibility | 615083 |
| PRF1 | No | Yes | 170280 | 10q22.1 | Biallelic | Familial Hemophagocytic lymphohistiocytosis,(associated with tumour predisposition?) | 267700 |
| PRKAR1A | Yes | Yes | 188830 | 17q24.2 | Mono-allelic | Carney complex | 160980 |
| PTCH1# | Yes | Yes | 601309 | 9q22.32 | Mono-allelic | Gorlin syndrome | 109400 |
| PTEN | Yes | Yes | 601728 | 10q23.31 | Mono-allelic | Cowden Syndrome | 158350 |
| RAD51C | No | Yes | 602774 | 17q22 | Mono-allelic | Ovarian cancer susceptibility | 16700 |
| | | | | | Biallelic | Fanconi Anaemia (O) | 613390 |
| RAD51D | No | Yes | 602954 | 17q12 | Mono-allelic | Breast - ovarian cancer susceptibility | PS604370 |
| RB1# | Yes | Yes | 614041 | 13q14.2 | Monoallelic | Hereditary retinoblastoma | 180200 |
| RECQL4 | No | Yes | 603780 | 8q24.3 | Biallelic | Rothmund-Thompson Syndrome | 268400 |
| RET* | Yes | Yes | 164761 | 10q11.21 | Monoallelic | Multiple endocrine neoplasia, Familial medullary thyroid cancer | 171400 / 162300 / 155240 |
| RHBDF2* | No | Yes | 614404 | 17q25.1 | Monoallelic | Tylosis with oesephaeal cancer | 148500 |
| RUNX1 | No | Yes | 151385 | 21q22.12 | Monoallelic | Acute myeloid leukaemia, Familial platelet disorder with associated myeloid malignancy | 601626 / 601399 |
| SBDS | No | Yes | 607444 | 7q11.21 | Biallelic | Schwachman-Diamond Syndrome | 260400 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| SDHAF2 | Yes | Yes | 613019 | 11q12.2 | Monoallelic | Familial Paraganglioma-Phaeochromocytoma Syndrome | PS168000 |
| SDHA | Yes | Yes | 600857 | 5p15.33 | Monoallelic | Familial Paraganglioma-Pheochromocytoma Syndrome | PS168000 |
| SDHB | Yes | Yes | 185470 | 1p36.13 | Monoallelic | Familial Paraganglioma-Pheochromocytoma Syndrome | PS168000 |
| SDHC | Yes | Yes | 602413 | 1q23.3 | Monoallelic | Familial Paraganglioma-Pheochromocytoma Syndrome | PS168000 |
| SDHD[§] | Yes | Yes | 602690 | 11q23.1 | Monoallelic | Familial Paraganglioma-Pheochromocytoma Syndrome | PS168000 |
| SLX4 | No | Yes | 613278 | 16p13.3 | Biallelic | Fanconi anaemia (P) | 613951 |
| SMAD4 | Yes | Yes | 600993 | 18q21.2 | Monoallelic | Juvenile Polyposis Syndrome | 174900 |
| SMARCA4 | No | Yes | 603254 | 19p13.2 | Monoallelic | Rhabdoid tumour predisposition syndrome-2<br>Small cell Ca ovary hypercalcaemic type | PS609322<br>N/A |
| SMARCB1[#] | Yes | Yes | 601607 | 22q11.23 | Monoallelic | Rhabdoid Predisposition Syndrome | PS609322 |
| SMARCE1 | Yes | Yes | #603111 | 17q21.2 | Monoallelic | Meningioma susceptibility | 607174 |
| STK11 | Yes | Yes | 602216 | 19p13.3 | Monoallelic | Peutz-Jeghers Syndrome | 175200 |
| SUFU[#] | No | Yes | 607035 | 10q24.32 | Monoallelic | Familial medulloblastoma<br>Basal cell naevus syndrome | 607174 / 109400 |
| TMEM127 | Yes | Yes | 613403 | 2q11.2 | Monoallelic | Familial Paraganglioma-Pheochromocytoma Syndrome | PS168000 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| TP53 | Yes | Yes | 191170 | 17p13.1 | Monoallelic | Li Fraumeni Syndrome | 151623 |
| TSC1 | Yes | Yes | #191100 | 9q34.13 | Monoallelic | Tuberous sclerosis | PS191100 |
| TSC2 | Yes | Yes | 191092 | 16p13.3 | Monoallelic | Tuberous sclerosis | PS191100 |
| VHL[#] | Yes | Yes | 608537 | 3p25.3 | Monoallelic | Von Hippel Lindau Syndrome | 193300 |
| WRN | No | Yes | 604611 | 8p12 | Biallelic | Werner Syndrome | 277700 |
| WT1[#] | Yes | Yes | 607102 | 11p13 | Monoallelic | WAGR<br>Familial Wilms tumour<br>Denys-Drash<br>Frasier Syndrome | 194072 / 194070 /<br>194080 / 136680 |
| XPA | No | Yes | 611153 | 9q22.33 | Biallelic | Xeroderma pigmentosum (A) | 278700 |
| XPC | No | Yes | 613208 | 3p25.1 | Biallelic | Xeroderma pigmentosum (C) | 27820 |

[§], paternal transmission of mutant allele associated with tumour risk; *, gain of function; [#], heterozygous deletion mutant allele associated with substantial childhood onset cancer risk

**Table S2. CNVs detected involving a CPG in the Stage 1 Pilot Cohort Analysis.**

| Patient | CPG | Del / gain of CPG | Co-ordinates of CNV affecting CPG | CNV: benign / pathological / unrelated to clinical features | CNV dn or inh | Other array / cytogenetic findings | Other array / cytogenetic findings: benign / pathological / unrelated | Additional notes | Age at array (m) | Sex F/M |
|---|---|---|---|---|---|---|---|---|---|---|
| P001 | BMPR1A | gain | arr[GRCh37] 10q22.3q23.2(81,489,155-89,047,672)x3 | unrelated | inh | Nil | N/A | No relevant FHx | 31 | F |
| P002 | BMPR1A | gain | arr[NCBI36] 10q22.3q23.2(81,618,854-88,977,697)X3 | pathogenic | unknown | Nil | N/A | Nil | 60 | M |
| P003 | BMPR1A | del | arr[GRCh37] 10q22.3q23.2(81,489,155-88,847,906)x1 | pathogenic | dn | Nil | N/A | Nil | 13 | M |
| P004 | BMPR1A | del | arr[GRCh37] 10q22.3q23.2(81,489,155-89,189,038)x1 | pathogenic | dn | Nil | N/A | Nil | 1 | M |
| P005 | BMPR1A | del | arr[GRCh37] 10q22.3q23.2(81,628,948-89,189,038)x1 | pathogenic | inh | Nil | N/A | No relevant FHx | 0 | M |
| P006 | BMPR1A & PTEN | del | arr[GRCh37] 10q23.1q23.31(87,785,872-90,301,307)x1 | pathogenic | dn | Nil | N/A | Nil | 0 | F |
| P007 | MET* | gain | arr[GRCh37] 7q31.2(114,687,978-116,355,918)x3 | unrelated | inh | Nil | N/A | No relevant FHx | 0 | F |
| P008 | MSH2 & MSH6 | gain | arr[NCBI36] 2p22.2p16.2(37,472,079-54,403,333)x3 | pathogenic | dn | Nil | n/a | Child has inherited unbalanced form of balanced deletion / insertion event on chr2 from a carrier parent | 312 | F |
| P009 | NF1 | gain | arr[NCBI36] 17q11.2(25,997,842-27,385,919.5)x3 | unrelated | inh | Nil | N/A | No relevant FHx | 0 | M |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| P010 | PMS2 | gain[#] | arr[GRCh37] 7p22.3p21.3(54,215-7,690,132)x3 | pathogenic | unknown | 46,XY,der(22)t(7;22)(p21.3;p11) | pathogenic | Likely child inherited unbalanced form of parental translocation. Both parental samples not available. | 50 | M |
| P011 | PRKAR1A | gain | arr[GRCh37] 17q24.2q24.3(61,652,704-66,662,580)x3 | pathogenic | unknown | Nil | N/A | Nil | 55 | F |
| P012 | PTEN | gain | arr[GRCh37] 10q23.31(89,635,555-89,665,005)x3 | unrelated | inh | arr[GRCh37] 15q13.2q13.3(30,491,443-32,509,897)x3 | unrelated | No relevant FHx. Child has in inherited one duplication from each parent. Partial gain of PTEN. | 1 | F |
| P013 | RB1 | gain | 13q14.11q21.1(RP11-125A7->RP11-200F15)x3 | uncertain significance | inh | Nil | N/A | No relevant FHx. Analysis by BAC array | 180 | M |
| P014 | RB1 | gain | arr[GRCh37] 13q14.11q21.2(43,517,174-60,107,850)x3 | pathogenic | inh | Nil | N/A | No relevant FHx | 192 | F |
| P015 | SDHA | del | arr[GRCh37] 5p15.33(22,179-4,163,877)x1 | pathogenic | dn | Nil | N/A | Nil | 8 | F |
| P016 | SDHA | del | arr[GRCh37] 5p15.33(86,236-401,759)x1 | uncertain significance | inh | arr[GRCh37] 7q35(146,175,061-146,376,898)x1 46,XY,t(4;18)(q23;q23) | uncertain significance | FHx unknown. Child has inherited two deletions inherited from one parent and a balanced translocation from the other parent. | 2 | M |
| P017 | SDHD | gain | arr[GRCh37] 11q23.1(111,883,595-112,272,219)x3 | unrelated | inh | arr[GRCh37] 2p13.1p12(74,433,160-75,699,530)x3 | unrelated | No relevant FHx. Likely two separate duplications inherited from one parent (no FISH undertaken). | 175 | F |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| P018 | SDHD | gain | arr[GRCh37] 11q22.3q23.1(110,281,343-112,218,364)x3 | uncertain significance | unknown | Nil | N/A | FHx unknown | 8 | F |
| P019 | SMARCB1 | gain | arr[GRCh37] 22q11.23(23,720,201-24,959,798)x3 | unrelated | unknown | Nil | N/A | FHx unknown | 0 | M |
| P020 | SMARCB1 | gain | arr[GRCh37] 22q11.23(23,822,957-24,178,146)x3 | uncertain significance | unknown | Nil | N/A | No relevant FHx | 124 | F |
| P021 | TMEM127 | gain | arr[GRCh37] 2q11.1q11.2(96,545,380-98,013,837)x3 | unrelated | inh | Nil | N/A | No relevant FHx | 1 | F |
| P022 | TMEM127 | del | arr[GRCh37] 2q11.2(95,910,348-97,570,502)x1 | uncertain significance | unknown | Nil | N/A | Nil | 35 | M |
| P023 | TMEM127 | del | arr[GRCh37] 2q11.1q11.2(96,766,590-98,206,184)x1 | unrelated | inh | Nil | N/A | No relevant FHx | 6 | M |
| P024 | TP53 | gain | arr[GRCh37] 17p13.1(7,394,419-7,772,158)x3 | uncertain significance | inh | arr[GRCh37] 10p15.3(226,113-332,378)x3, arr[GRCh37] 10p15.3(536,734-816,599)x3 | uncertain significance | FHx unknown. All 3 duplications inherited from one parent. | 5 | F |
| P025 | TP53 | gain | arr[GRCh37] 17p13.1(7,572,205-7,583,268)x3 | unrelated | inh | Nil | N/A | No relevant FHx. Partial gain of TP53. | 179 | M |
| P026 | TSC2 | gain[#] | arr[NCBI36] 16p13.3p13.13(12,798-10,656,496)x3 | pathogenic | inh | 46,XY.ish der(10)t(10;16)(q26.2;p13.13)(10ptel-,16ptel+) | pathogenic | Child has inherited unbalanced rearrangement from balanced carrier parent. | 0 | F |
| P027 | TSC2 | gain[#] | arr[NCBI36] 16p13.3p12.3(12,798-17,285,210)x3 | pathogenic | dn | 46,XX,der(1)t(1;16)(q44;p12.3)dn | pathogenic | Nil | 9 | F |
| P028 | TSC2 | gain[#] | arr[GRCh37] 16p13.3(93,748-3,702,950)x3 | pathogenic | dn | 46,XY,der(3)t(3;7)(p26.1;p21)dn, der(7)t(7;16)(p21;p16.3)dn, arr[GRCh37] 3p26.3p26.1(93,979-8,325,145)x1 dn | pathogenic | Nil | 50 | M |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| P029 | TSC2 | del | arr[GRCh37] 16p13.3(2,097,019-2,099,039)x1dn | pathogenic | dn | Nil | N/A | TS clinically diagnosed subsequently. Deletion exons 1-2 TSC2 on minimum co-ordinates. Deletion exons 1-2 and part of exon 3 on maximum co-ordinates, arr[GRCh37] 16p13.3(2,047,554-2,100,464)x1. Pathogenic deletions of exon 3 and 5' UTR to exon 3 been described in TS (32-33). | 7 | M |
| P030 | VHL | gain[#] | arr[GRCh37] 3p26.1p24.2(7,543,632-24,901,101)x3 | pathogenic | dn | 45,XY,psu dic(3;21)dup(3)(3qter->3p26.?2::3p26.2->3p25.1::21p11.1)dn | pathogenic | Nil | 104 | M |
| P031 | VHL | gain[#] | 3p26.1 to 3p24.3 (RP11-277D17->RP11-208G16)x3 | pathogenic | dn | 46,XX,der(3)t(3;6)(p26.1;q25.2)dup(3)(p26.1p24.3)dn , arr cgh 3p26.3p26.1(RP11-86C13->RP11-324K11)x1 dn, 6q25.2q27(RP1-66H9->RP11-226I21)x3 dn | pathogenic | Nil | 0 | F |

del, deletion; FHx, family history; dn, de novo; inh, inherited; * oncogene; [#] presence of a complex chromosomal rearrangement; F, female; M, male

**Table S3. CNVs detected involving a CPG in the Stage 2 Extended Cohort Analysis.**

| Patient | CPG | Del / gain of CPG | Co-ordinates of CNV affecting CPG | CNV: benign / pathological / unrelated to clinical features | CNV dn or inh | Other array / cytogenetic findings | Other array / cytogenetic findings: benign / pathological / unrelated | additional notes | age at array (m) | Sex F/M |
|---|---|---|---|---|---|---|---|---|---|---|
| P101 | APC | deletion | arr[GRCh37] 5q15-q23.1(95,864,493-115,563,383)x1 | pathogenic | dn | t(5;8)(q15;q13) dn | N/A | nil | 205 | female |
| P102 | BLM & FANCI | deletion | arr[GRCh37] 15q25.3q26.1(87,891,894-92,197,207)x1 | pathogenic | unknown | nil | N/A | nil | 12 | male |
| P103 | BMPR1A | deletion | arr[GRCh37] 10q23.2(88,197,182-88,680,492)x1 | unrelated | inh | t(1;10)(p32;q23.2) inh | unrelated | No FHx relevant to cancer gene. Deletion exons 1-10 BMPR1A on minimum co-ordinates. Maximum co-ordinates, arr[GRCh37] 10q23.2(88,150,393-88,828,085)x1 encompasses BMPR1A entirely. Pathogenic deletions of exons 1-3 and 2-11 been described (34-35). | 26 | male |
| P104 | BMPR1A | deletion | arr[GRCh36] 10q22.3q23.2(81,631,898-88,930,398)x1 | pathogenic | dn | nil | N/A | nil | 175 | male |
| P105 | BRIP1 | deletion | arr[GRCh37] 17q23.1q23.2(58,172,677-60,395,815)x1 | pathogenic | unknown | nil | N/A | nil | 102 | male |
| P106 | CHEK2 | deletion | arr[GRCh36] 22q12.1(26,802,174-27,739,571)x1 | uncertain significance | inh | nil | N/A | No FHx relevant to cancer gene | 177 | male |
| P107 | CHEK2 | deletion | arr[GRCh37] 22q12.1q12.2(27,988,300-29,622,742)x1 | uncertain significance | dn | arr[GRCh37] 5p15.2(12,260,190-13,141,797)x1, arr[GRCh37] 15q13.3(31,972,643-32,509,932)x3 | both uncertain significance | one additional CNV inherited from each parent | 195 | female |
| P108 | FH | deletion | arr[GRCh37] 1q43q44(240,871,091-247,124,412)x1 | pathogenic | dn | nil | N/A | nil | 2 | male |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| P109 | FH | deletion | arr[GRCh37] 1q43(239,940,368-243,128,788)x1 | pathogenic | dn | nil | N/A | nil | 47 | male |
| P110 | FLCN | deletion | arr[GRCh37] 17p11.2(16,637,872-20,294,010)x1 | pathogenic | dn | nil | N/A | nil | 1 | male |
| P111 | FLCN | deletion | arr[GRCh37] 17p11.2(16,637,872-20,219,455)x1 | pathogenic | dn | nil | N/A | nil | 12 | female |
| P112 | FLCN | deletion | arr[GRCh37] 17p11.2(16,637,872-20,294,010)x1 | pathogenic | unknown | nil | N/A | nil | 33 | male |
| P113 | FLCN | deletion | arr[GRCh37] 17p11.2(16,782,547-20,294,010)x1 | pathogenic | dn | nil | N/A | nil | 40 | female |
| P114 | FLCN | deletion | arr[GRCh37] 17p11.2(16,637,872-20,294,010)x1 | pathogenic | dn | nil | N/A | nil | 45 | male |
| P115 | FLCN | deletion | arr[GRCh37] 17p11.2(16,637,872-20,294,010)x1 | pathogenic | unknown | nil | N/A | nil | 63 | female |
| P116 | FLCN | deletion | arr[GRCh37] 17p11.2(16,637,872-20,219,455)x1 | pathogenic | unknown | nil | N/A | nil | 95 | male |
| P118 | FLCN | deletion | arr[GRCh37] 17p12p11.2(15,801,294-20,524,000)x1 | pathogenic | dn | nil | N/A | nil | 106 | male |
| P119 | FLCN | deletion | arr[GRCh37] 17p11.2(16,637,872-20,219,455)x1 | pathogenic | unknown | nil | N/A | nil | 136 | female |
| P120 | FLCN | deletion | arr[GRCh37] 17p11.2(16,637,872-20,294,010)x1 | pathogenic | unknown | nil | N/A | nil | 444 | female |
| P121 | HRAS* | gain# | arr[GRCh37] 11p15.5p15.2(113,082-14,430,533)x3. | pathogenic | inh | 46,XX,del(5)(p14.3p15.1)inh, der(5)t(5;11)(p15.31;p15.2)inh. arr[GRCh37] 5p15.33p15.31(22,149-9,700,223)x1,5p15.1p14.3(16,168,384-22,925,861)x1 | pathogenic | Child has inherited from each parent (i) unbalanced form of familial t(5;11)(p15.31;p15.2) (ii) familial del(5)(p14.3p15.1) associated with learning difficulties. Same | 0 | female |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | individual as below. | | |
| P121 | SDHA | deletion# | arr[GRCh37] 5p15.33p15.31(22,149-9,700,223)x1 | pathogenic | inh | 46,XX,del(5)(p14.3p15.1)inh, der(5)t(5;11)(p15.31;p15.2)inh. arr[GRCh37] 5p15.1p14.3(16,168,384-22,925,861)x1,11p15.5p15.2(113,082,14,430,533)x3 | pathogenic | Child has inherited from each parent (i) unbalanced form of familial t(5;11)(p15.31;p15.2) (ii) familial del(5)(p14.3p15.1) associated with learning difficulties. Same individual as above. | 0 | female |
| P122 | HRAS* | gain# | arr[GRCh37] 11p15.5(113,082-2,550,805)x3 | pathogenic | inh | arr[GRCh37] 2q37.3(237,463,855-243,087,748)x1 | pathogenic | Child has inherited unbalanced form of parental ish t(2;11)(q37.3;p15.5) | 6 | female |
| P123 | KIT* and PDGFRA* | deletion | arr[GRCh37] 4q12q13.1(53,850,058-65,512,919)x1 | pathogenic | dn | nil | N/A | nil | 292 | female |
| P124 | MAX | deletion | arr[GRCh37] 14q23.3(65,538,709-67,351,711)x1 | uncertain significance | dn | nil | N/A | Not known if deletion on pat or mat chr. | 100 | male |
| P125 | MET* | deletion | arr[GRCh37] 7q22.3q31.2(106,871,009-122,986,270)x1 | pathogenic | dn | nil | N/A | nil | 162 | female |
| P126 | MET* | deletion | arr[GRCh37] 7q31.1q31.31(109,958,035-117,946,370)x1 | pathogenic | dn | nil | N/A | nil | 272 | female |
| P127 | MET* | gain# | arr[GRCh37] 7q21.3q36.3(94,257,592-159,124,141)x3 | pathogenic | dn | 46,X,der(X)t(X;7)(p21.3;q21.3). arr[GRCh37] Xp22.33p21.3(310,953-28,515,804)x1 | pathogenic | nil | 29 | female |
| P128 | NSD1 | deletion | arr[GRCh37] 5q35.2q35.3(174,897,894-177,013,956)x1 | pathogenic | dn | nil | N/A | nil | 0 | male |
| P129 | PDGFRα * | gain | min: arr[GRCh37] 4q12(54,873,967-55,081,970)x3 max: arr[GRCh37] 4q12(54,830,498-55,133,838)x3 | uncertain significance | unknown | nil | N/A | Maximum region of CNV contains all of exons 1-6 and part of exon 7 of transcript NM_006206.4 of PDGFRα | 33 | male |
| P130 | PMS1 | deletion | arr[GRCh36] 2q32.2q33.1(189,294,117-200,158,719)x1 | pathogenic | dn | nil | N/A | nil | 57 | male |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| P131 | PMS2 | deletion | arr[GRCh37] 7p22.1(5,151,574-6,745,570)x1 | pathogenic | dn | nil | N/A | nil | 6 | male |
| P132 | PMS2 | deletion | arr[GRCh36] 7p22.1(5,337,139-6,263,352)x1 | pathogenic | dn | nil | N/A | nil | 30 | male |
| P133 | PMS2 | deletion | arr[GRCh37] 7p22.1(5,617,810-6,350,273)x1 | uncertain significance | unknown | nil | N/A | nil | 80 | male |
| P134 | POLD1 | deletion | arr[GRCh37] 19q13.33(50,126,448-51,026,193)x1 | pathogenic | unknown | nil | N/A | nil | 46 | male |
| P135 | POLE | deletion | arr[GRCh37] 12q24.33(133,237,077-133,773,393)x1 | uncertain significance | unknown | arr[GRCh37] 2p21(44,518,437-44,542,918)x3 | uncertain significance | Deletion exons 1-25 POLE on minimum co-ordinates. Maximum co-ordinates, arr[GRCh37] 12q24.33(133,153,306-133,851,895)x1, encompasses POLE. | 122 | male |
| P136 | RB1 | deletion | arr[GRCh37] 13q13.1q31.1(33,786,863-83,585,442)x1 | pathogenic | dn | nil | N/A | subsequently developed RB | 0 | female |
| P137 | RB1 | deletion | arr[GRCh37] 13q14.11q21.32(41,717,885-66,044,521)x1 | pathogenic | dn | nil | N/A | subsequently developed RB as did MZ twin | 3 | male |
| P138 | RUNX1 | deletion | arr[GRCh36] 21q22.11q22.12(33,833,276-35,618,531)x1 | pathogenic | dn | arr[GRCh36] 5q33.2(155,149,896-155,496,459)x3 | inh. uncertain signficance | child previously noted to have unexplained low platelet count | 137 | male |
| P139 | SDHA | deletion | arr[GRCh37] 5p15.33p14.1(22,149-24,835,505)x1 | pathogenic | unknown | nil | N/A | nil | 0 | male |
| P140 | SDHA | deletion | arr[GRCh37] 5p15.33p15.1(22,149-18,133,322)x1 | pathogenic | unknown | arr[GRCh37] 5p15.1p14.1(18,215,614-26,972,861)x3 | uncertain significance | nil | 3 | female |
| P141 | SDHA | deletion | arr[GRCh37] 5p15.33p14.3(22,149-21,217,291)x1 | pathogenic | dn | nil | N/A | nil | 6 | female |
| P142 | SDHA | deletion | arr[GRCh36] 5p15.33(75,149-3,292,621)x1 | uncertain significance | inh | nil | N/A | nil | 16 | male |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| P143 | SDHA | deletion | arr[GRCh37] 5p15.33p15.1(22,149-15,151,743)x1 | pathogenic | dn | nil | N/A | nil | 26 | male |
| P144 | SDHA | deletion | arr[GRCh37] 5p15.33p15.2(22,149-11,213,706)x1 | pathogenic | unknown | nil | N/A | nil | 460 | male |
| P145 | SDHB | deletion | arr[GRCh36] 1p36.21-p36.12(15,422,198-21,235,203)x1 | pathogenic | dn | nil | N/A | nil | 74 | male |
| P146 | SDHD | deletion | arr[GRCh37] 11q22.3q23.3(105,289,128-116,644,381)x1 | pathogenic | dn | nil | N/A | del on pat chr | 13 | female |
| P147 | SMARCA4 | deletion | min: arr[GRCh37] 19p13.2(10,466,712-11,053,974)x1 max: arr[GRCh37] 19p13.2(10,418,713-11,117,391)x1 | uncertain significance | unknown | nil | N/A | Maximum region of CNV deletes exons 1-14 of SMARCA4. Minimum region not involve SMARCA4. | 50 | male |
| P148 | SMARCB1 | deletion | arr[GRCh36] arr 22q11.21q11.23(20,052,254-22,923,820)x1 | pathogenic | unknown | nil | N/A | nil | 148 | male |
| P149 | TMEM127 | deletion | arr[GRCh36] 2q11.1q11.2(94,892,764-100,915,841)x1 | pathogenic | dn | nil | N/A | nil | 97 | male |
| P150 | TMEM127 | deletion | arr[GRCh37] 2q11.1q11.2(96,766,564-97,039,740)x1 | uncertain significance | unknown | nil | N/A | nil | 191 | male |

del, deletion; FHx, family history; dn, de novo; inh, inherited; * oncogene; [#] presence of a complex chromosomal rearrangement; F, female; M, male

**Table S4. CNVs detected involving a CPG in both the Stage 1 and Stage 2 Cohort Analyses.**

| Patient | CPG | Del / gain of CPG | Co-ordinates of CNV affecting CPG | CNV: benign / pathological / unrelated to clinical features | CNV dn or inh | Other array / cytogenetic findings | Other array / cytogenetic findings: benign / pathological / unrelated | Additional notes | Age at array (m) | Sex |
|---|---|---|---|---|---|---|---|---|---|---|
| P101 | APC | del | arr[GRCh37] 5q15-q23.1(95,864,493-115,563,383)x1 | Pathogenic | dn | t(5;8)(q15;q13) dn | N/A | Nil | 205 | F |
| P103 | BMPR1A | del | arr[GRCh37] 10q23.2(88,197,182-88,680,492)x1 | Unrelated | inh | t(1;10)(p32;q23.2) inh | unrelated | No FHx relevant to cancer gene. Deletion exons 1-10 BMPR1A on minimum co-ordinates. Maximum co-ordinates, arr[GRCh37] 10q23.2(88,150,393-88,828,085)x1 encompass BMPR1A. Pathogenic deletions of exons 1-3 and 2-11 have been described (34-35) | 26 | M |
| P104 | BMPR1A | del | arr[GRCh36] 10q22.3q23.2(81,631,898-88,930,398)x1 | Pathogenic | dn | nil | N/A | Nil | 175 | M |
| P003 | BMPR1A | del | arr[GRCh37] 10q22.3q23.2(81,489,155-88,847,906)x1 | Pathogenic | dn | Nil | N/A | Nil | 13 | M |
| P004 | BMPR1A | del | arr[GRCh37] 10q22.3q23.2(81,489,155-89,189,038)x1 | Pathogenic | dn | Nil | N/A | Nil | 1 | M |
| P005 | BMPR1A | del | arr[GRCh37] 10q22.3q23.2(81,628,948-89,189,038)x1 | Pathogenic | inh | Nil | N/A | No relevant FHx | 0 | M |
| P006 | BMPR1A & PTEN | del | arr[GRCh37] 10q23.1q23.31(87,785,872-90,301,307)x1 | Pathogenic | dn | Nil | N/A | Nil | 0 | F |
| P123 | KIT* and PDGFRA* | del | arr[GRCh37] 4q12q13.1(53,850,058 | Pathogenic | dn | nil | N/A | nil | 292 | F |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | -65,512,919)x1 | | | | | | | |
| P124 | MAX | del | arr[GRCh37] 14q23.3(65,538,709-67,351,711)x1 | uncertain significance | dn | nil | N/A | Not known if del on pat or mat chr. | 100 | M |
| P125 | MET* | del | arr[GRCh37] 7q22.3q31.2(106,871,009-122,986,270)x1 | Pathogenic | dn | nil | N/A | nil | 162 | F |
| P126 | MET* | del | arr[GRCh37] 7q31.1q31.31(109,958,035-117,946,370)x1 | Pathogenic | dn | nil | N/A | nil | 272 | F |
| P127 | MET* | gain# | arr[GRCh37] 7q21.3q36.3(94,257,592-159,124,141)x3 | Pathogenic | dn | 46,X,der(X)t(X;7)(p21.3;q21.3). arr[GRCh37] Xp22.33p21.3(310,953-28,515,804)x1 | pathogenic | nil | 29 | F |
| P007 | MET* | gain | arr[GRCh37] 7q31.2(114,687,978-116,355,918)x3 | Unrelated | inh | Nil | N/A | No relevant FHx | 0 | F |
| P129 | PDGFRα * | gain | min: arr[GRCh37] 4q12(54,873,967-55,081,970)x3 max: arr[GRCh37] 4q12(54,830,498-55,133,838)x3 | uncertain significance | unknown | nil | N/A | Maximum region of CNV contains all of exons 1-6 and part of exon 7 of transcript NM_006206.4 of PDGFRα | 33 | M |
| P131 | PMS2 | del | arr[GRCh37] 7p22.1(5,151,574-6,745,570)x1 | pathogenic | dn | nil | N/A | nil | 6 | M |
| P132 | PMS2 | del | arr[GRCh36] 7p22.1(5,337,139-6,263,352)x1 | pathogenic | dn | nil | N/A | nil | 30 | M |
| P133 | PMS2 | del | arr[GRCh37] 7p22.1(5,617,810-6,350,273)x1 | uncertain significance | unknown | nil | N/A | nil | 80 | M |
| P136 | RB1 | del | arr[GRCh37] 13q13.1q31.1(33,786,863-83,585,442)x1 | pathogenic | dn | nil | N/A | subsequently developed RB | 0 | F |
| P137 | RB1 | del | arr[GRCh37] 13q14.11q21.32(41,717,885-66,044,521)x1 | pathogenic | dn | nil | N/A | subsequently developed RB as did MZ twin | 3 | M |
| P139 | SDHA | del | arr[GRCh37] 5p15.33p14.1(22,149- | pathogenic | unknown | nil | N/A | nil | 0 | M |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | 24,835,505)x1 | | | | | | | |
| P140 | SDHA | del | arr[GRCh37] 5p15.33p15.1(22,149-18,133,322)x1 | pathogenic | unknown | arr[GRCh37] 5p15.1p14.1(18,215,614-26,972,861)x3 | uncertain significance | nil | 3 | F |
| P141 | SDHA | del | arr[GRCh37] 5p15.33p14.3(22,149-21,217,291)x1 | pathogenic | dn | nil | N/A | nil | 6 | F |
| P142 | SDHA | del | arr[GRCh36] 5p15.33(75,149-3,292,621)x1 | uncertain significance | inh | nil | N/A | nil | 16 | M |
| P143 | SDHA | del | arr[GRCh37] 5p15.33p15.1(22,149-15,151,743)x1 | pathogenic | dn | nil | N/A | nil | 26 | M |
| P144 | SDHA | del | arr[GRCh37] 5p15.33p15.2(22,149-11,213,706)x1 | pathogenic | unknown | nil | N/A | nil | 460 | M |
| P121 | SDHA | del# | arr[GRCh37] 5p15.33p15.31(22,149-9,700,223)x1 | pathogenic | inh | 46,XX,del(5)(p14.3p15.1)inh, der(5)t(5;11)(p15.31;p15.2)inh. arr[GRCh37] 5p15.1p14.3(16,168,384-22,925,861)x1,11p15.5p15.2(113,082,14,430,533)x3 | pathogenic | Child has inherited from each parent (i) unbalanced form of familial t(5;11)(p15.31;p15.2) (ii) familial del(5)(p14.3p15.1) associated with learning difficulties. | 0 | F |
| P015 | SDHA | del | arr[GRCh37] 5p15.33(22,179-4,163,877)x1 | pathogenic | dn | Nil | N/A | Nil | 8 | F |
| P016 | SDHA | del | arr[GRCh37] 5p15.33(86,236-401,759)x1 | uncertain significance | inh | arr[GRCh37] 7q35(146,175,061-146,376,898)x1 46,XY,t(4;18)(q23;q23) | uncertain significance | FHx unknown. Child has inherited two dels inherited from one parent and a balanced translocation from the other parent. | 2 | M |
| P145 | SDHB | del | arr[GRCh36] 1p36.21-p36.12(15,422,198-21,235,203)x1 | pathogenic | dn | nil | N/A | nil | 74 | M |
| P146 | SDHD | del | arr[GRCh37] 11q22.3q23.3(105,289,128-116,644,381)x1 | pathogenic | dn | nil | N/A | del on pat chr | 13 | F |
| P148 | SMARCB1 | del | arr[GRCh36] arr | pathogenic | unknown | nil | N/A | nil | 148 | M |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | 22q11.21q11.23(20,052,254-22,923,820)x1 | | | | | | | |
| P149 | TMEM127 | del | arr[GRCh36] 2q11.1q11.2(94,892,764-100,915,841)x1 | pathogenic | dn | nil | N/A | nil | 97 | M |
| P150 | TMEM127 | del | arr[GRCh37] 2q11.1q11.2(96,766,564-97,039,740)x1 | uncertain significance | unknown | nil | N/A | nil | 191 | M |
| P022 | TMEM127 | del | arr[GRCh37] 2q11.2(95,910,348-97,570,502)x1 | uncertain significance | unknown | Nil | N/A | Nil | 35 | M |
| P023 | TMEM127 | del | arr[GRCh37] 2q11.1q11.2(96,766,590-98,206,184)x1 | unrelated | inh | Nil | N/A | No relevant FHx | 6 | M |
| P029 | TSC2 | del | arr[GRCh37] 16p13.3(2,097,019-2,099,039)x1dn | pathogenic | dn | Nil | N/A | TS clinically diagnosed subsequently. Deletion exons 1-2 TSC2 on minimum co-ordinates. Deletion exons 1-2 and part of exon 3 on maximum co-ordinates, arr[GRCh37] 16p13.3(2,047,554-2,100,464)x1. Pathogenic deletions of exon 3 and 5' UTR to exon 3 been described in TS (32-33). | 7 | M |

del, deletion; FHx, family history; dn, de novo; inh, inherited; * oncogene; [#] presence of a complex chromosomal rearrangement; F, female; M, male