npg

ORIGINAL ARTICLE

Detection of *APC* mosaicism by next-generation sequencing in an FAP patient

Kiyoshi Yamaguchi[1], Mitsuhiro Komura[2], Rui Yamaguchi[3], Seiya Imoto[2], Eigo Shimizu[2], Shinichi Kasuya[2], Tetsuo Shibuya[3], Seira Hatakeyama[1], Norihiko Takahashi[1], Tsuneo Ikenoue[1], Keisuke Hata[4], Giichiro Tsurita[4], Masaru Shinozaki[4], Yutaka Suzuki[5], Sumio Sugano[6], Satoru Miyano[2,3] and Yoichi Furukawa[1]

Familial adenomatous polyposis (FAP) of the colon is characterized by multiple polyps in the intestine and extra-colonic manifestations. Most FAP cases are caused by a germline mutation in the tumor-suppressor gene *APC*, but some cases of adenomatous polyposis result from germline mutations in *MUTYH*, *POLD1* or *POLE*. Although sequence analysis of *APC* by the Sanger method is routinely performed for genetic testing, there remain cases whose mutations are not detected by the analysis. Next-generation sequencing has enabled us to analyze the comprehensive human genome, improving the chance of identifying disease causative variants. In this study, we conducted whole-genome sequencing of a sporadic FAP patient in which we did not find any pathogenic *APC* mutations by the conventional Sanger sequencing. Whole-genome sequencing and subsequent deep sequencing identified a mosaic mutation of c.3175G > T, p.E1059X in ~ 12% of his peripheral leukocytes. Additional deep sequencing of his buccal mucosa, hair follicles, non-cancerous mucosa of the stomach and colon disclosed that these tissues harbored the *APC* mutation at different frequencies. Our data implied that genetic analysis by next-generation sequencing is an effective strategy to identify genetic mosaicism in hereditary diseases.
Journal of Human Genetics (2015) **60**, 227–231; doi:10.1038/jhg.2015.14; published online 26 February 2015

INTRODUCTION

Approximately 5% of colorectal cancer is caused by hereditary tumor syndromes including Lynch syndrome and familial polyposis of the colon. Their diagnosis is crucial not only for the treatment of patients, but also the healthcare of their relatives. Genetic testing is, therefore, of great importance for the diagnosis and the identification of the mutation carriers in the relatives.[1] Familial adenomatous polyposis (FAP) is an autosomal-dominant colorectal cancer predisposition syndrome that accounts for ~ 1% of newly diagnosed colorectal cancer cases. FAP is characterized by the development of multiple adenomatous polyps ranging from hundreds to thousands in the large intestine. A majority of the patients carry a germline mutation in the tumor-suppressor gene *APC* (adenomatous polyposis coli), but a small number of adenomatous polyposis cases are caused by germline mutations in *MUTYH*, *POLD1* or *POLE*. Genetic tests are performed for the screening of *APC*, but pathogenic mutations are not detected in some FAP cases by conventional direct sequencing owing to several reasons; (i) the region of testing is often restricted to 5′-half in the coding region and does not usually include the 3′-half, promoter region, or the 5′- or 3′-UTR. (ii) Structural alteration comprising large deletions/insertions and inversions are hard to identify by the

conventional sequencing method and other testing methods are required. (iii) Some cases of adenomatous polyposis are caused by mutations in other genes such as *MUTYH*, *POLD1* and *POLE*.[2,3] (iv) Some FAP cases are caused by somatic mosaicism of *APC*.[4–7]

Large-scale genome sequencing, also known as next-generation sequencing (NGS), is applicable to germline genomic sequencing, as well as other purposes including sequencing of tumors, sequencing of mRNA to analyze gene expression (RNA-seq), sequencing of DNA enriched by chromatin immunoprecipitation to characterize elements in protein-DNA interactions (ChIP-seq) and others. The entire genome of an individual can be sequenced in less than 1 week for 5000 to 10 000 dollars.[8] Cost reduction, together with advanced bioinformatics capabilities, have led to increased opportunities for NGS usage in various clinical applications including the detection of rare hereditary mutations, individualized therapy, pharmacogenomics, preconception/prenatal screening and population screening for disease risk.[9,10]

Here we show that genetic testing by NGS facilitates to identify mosaic *APC* mutations in patients with attenuated polyposis. NGS will improve genetic diagnosis of hereditary diseases whose mutations have been overlooked by conventional direct sequencing.

[1]Division of Clinical Genome Research, Advanced Clinical Research Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan; [2]Laboratory of DNA Information Analysis, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan; [3]Laboratory of Sequence Analysis, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan; [4]Department of Surgery, Research Hospital, Institute of Medical Science, The University of Tokyo, Tokyo, Japan; [5]Department of Computational Biology, Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa, Japan and [6]Laboratory of Functional Genomics, Graduate School of Frontier Sciences, The University of Tokyo, Tokyo, Japan

Correspondence: Dr Y Furukawa, Division of Clinical Genome Research, Advanced Clinical Research Center, Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan.
E-mail: furukawa@ims.u-tokyo.ac.jp

MATERIALS AND METHODS

Patients and ethical approval

This project was approved by the ethical committee of Institute of Medical Science, the University of Tokyo (IMSUT-IRB, 23-18-0929 and 23-19-0929). Written informed consent was obtained from the patient in this study.

Genetic testing

Genomic DNA was extracted from peripheral blood of the patient according to the standard phenol extraction/purification procedure. *APC* coding exons were amplified with M13-tailed target specific primers, and the PCR products were sequenced on the Applied Biosystems 3730×l DNA Analyzer (Thermo Fisher Scientific, Waltham, MA, USA) using the BigDye Direct Cycle Sequencing Kit (Thermo Fisher Scientific). The primer sequences used for sequencing are available on request.

Whole-genome sequencing

We prepared insert libraries of 250–350 bp from 1.0 µg of genomic DNA from lymphocytes and sequenced them using the HiSeq 2000 platforms with paired-end reads of 101 bp according to the manufacturer's instructions (Illumina, San Diego, CA, USA). For the data processing, fastq files were aligned to human reference sequence (hg19) by BWA[11] (ver. 0.5.10) and a bam file was created. For the detection of variants, we compared the bam file with the reference sequence by a Bayesian approach. We used a beta non-informative prior for representing the probability of existence of variant and the posterior distribution was obtained using the information of observed reads at each candidate position. We supposed that a random variable 'X' follows the posterior distribution representing the probability of variant. We used $\Pr(X \geqslant 0.05)$ as the score, and regarded the candidate positions whose scores were greater than 0.9 as statistically significant.

Validation of variants detected by NGS

To confirm the mutation, DNA of peripheral blood was amplified independently by PCR (KOD-Plus kit, TOYOBO, Osaka, Japan) with a set of primers encompassing the possible pathogenic mutation, and direct sequence was performed by the Sanger method. The primer sequences used for amplification are shown in Supplementary Table 1. To confirm mosaicism, genomic DNA was extracted from hair follicles, two sites of buccal mucosa, three sites of non-tumorous stomach mucosa, five sites of non-tumorous colonic mucosa and five adenomatous polyps from the patient by the standard phenol extraction/purification procedure. Deep sequencing was carried out using IonPGM Sequencing 200 kit and Sequencing 400 kit (Thermo Fisher Scientific) with libraries of PCR products prepared using Ion Plus Fragment Library Kit (Thermo Fisher Scientific). Variants were identified using the Variant Caller deployed with Torrent Suite (Thermo Fisher Scientific).

RESULTS

In clinical genetic testing, we encountered a male, 41 years of age, who suffered from multiple polyps in his large intestine. He earlier visited a hospital for the secondary screening of his intestine because of occult blood in his fecal test. Colonoscopy detected <100 adenomatous polyps in his large intestine and subsequent histological examination of the polyps diagnosed adenomatosis. As he had no family history of polyposis or colorectal cancer, he was suspected to be a *de novo* case of FAP or a patient of MUTYH-associated polyposis. Direct sequencing of *APC* was performed using DNA extracted from his lymphocytes to examine the 5′-half of the coding region where most of the *APC* mutations occur; however, no pathogenic mutations were detected. Although structural analysis of *APC* by Multiplex Ligation-dependent Probe Amplification, or sequencing of *MUTYH*, *POLD* and *POLE* by the Sanger method were available for the second screening, we performed whole-genome sequencing to test the usefulness of NGS in clinical testing.

The average depth of sequence coverage was ~ 26x, and a total of 4.6 million variants were identified in the patient. Because three to four million variants are generally detected by whole-genome sequencing in an individual,[9] we considered that the number of variants was reasonable. Among the 4.6 million, 30 501 variants were located in the exonic regions, and 8 were detected in the exons or splicing regions of *APC* (Table 1). Importantly, one of the eight variants was a nonsense mutation (c.3175G>T p.E1059X) that was determined in 6 of 50 reads (12%). Although the frequency was much lower than 50%, we suspected that it may be the causative mutation of his polyposis because the nonsense mutation truncates the APC protein leading to the loss of domains involved in β-catenin degradation. We re-sequenced the region by the Sanger method, and found a low peak of mutant allele compared with the wild allele in his DNA from peripheral blood. Taken together with his family history and the number of polyps, we suspected the possibility of *APC* mosaicism. The Sanger method is apparently less sensitive for the detection of mosaicism, compared with NGS (Figure 1). Notably, NGS identified no deleterious mutation in *MUTYH*, *POLD1* or *POLE* (data not shown). The remaining seven *APC* variants had been detected in the initial screening.

We further performed a deep sequencing of his DNA isolated form peripheral blood, hair follicles and two sites of oral mucosa. The average depth of coverage achieved with amplicon sequencing was 34 699x ranging from 3726 to 87 133. As shown in Table 2, c.3175G>T mutation was observed in 453 of 3726 reads (12.2%) in peripheral blood, 3774 of 83 679 (4.5%) in hair follicles, and 2099 of 69 169 (3.0%) and 4860 of 66 557 (7.3%) in buccal mucosa. Because the patient underwent gastric endoscopy and total colectomy, we also examined the mutation in three spots of non-tumorous gastric mucosa, five spots of non-tumorous colonic mucosa and five adenomatous polyps in the colon. Interestingly, we found different frequencies of c.3175G>T mutation in non-tumorous gastric (18.9, 22.7 and 27.7%) and colonic (9.2, 3.4, 12.3, 5.8 and 9.0%) tissues (Table 2). In addition, the mutation was found at higher frequencies (32.3, 28.6, 29.8, 32.5 and 24.7%) in the colonic polyps

Table 1 Summary of variations in all *APC* exons detected by WGS

Type	Mutation	Protein alteration	dbSNP131	Ref (No of reads)	Obs (No of reads)	Mismatch ratio
Synonymous	c.T1458C	p.Y486Y	rs2229992	T (0)	C (75)	1
Synonymous	c.G1635A	p.A545A	rs351771	G (0)	A (35)	1
Nonsense	c.G3175T	p.E1059X	—	G (44)	T (6)	0.12
Synonymous	c.G4479A	p.T1493T	rs41115	G (0)	A (55)	1
Synonymous	c.G5034A	p.G1678G	rs42427	G (0)	A (67)	1
Synonymous	c.T5268G	p.S1756S	rs866006	T (0)	G (43)	1
Non synonymous	c.T5465A	p.V1822D	rs459552	T (0)	A (39)	1
Synonymous	c.G5880A	p.P1960P	rs465899	G (0)	A (23)	1

Abbreviations: APC, adenomatous polyposis coli; No, number; Obs, observed nucleotide; Ref, reference nucleotide; WGS, whole-genome sequencing.
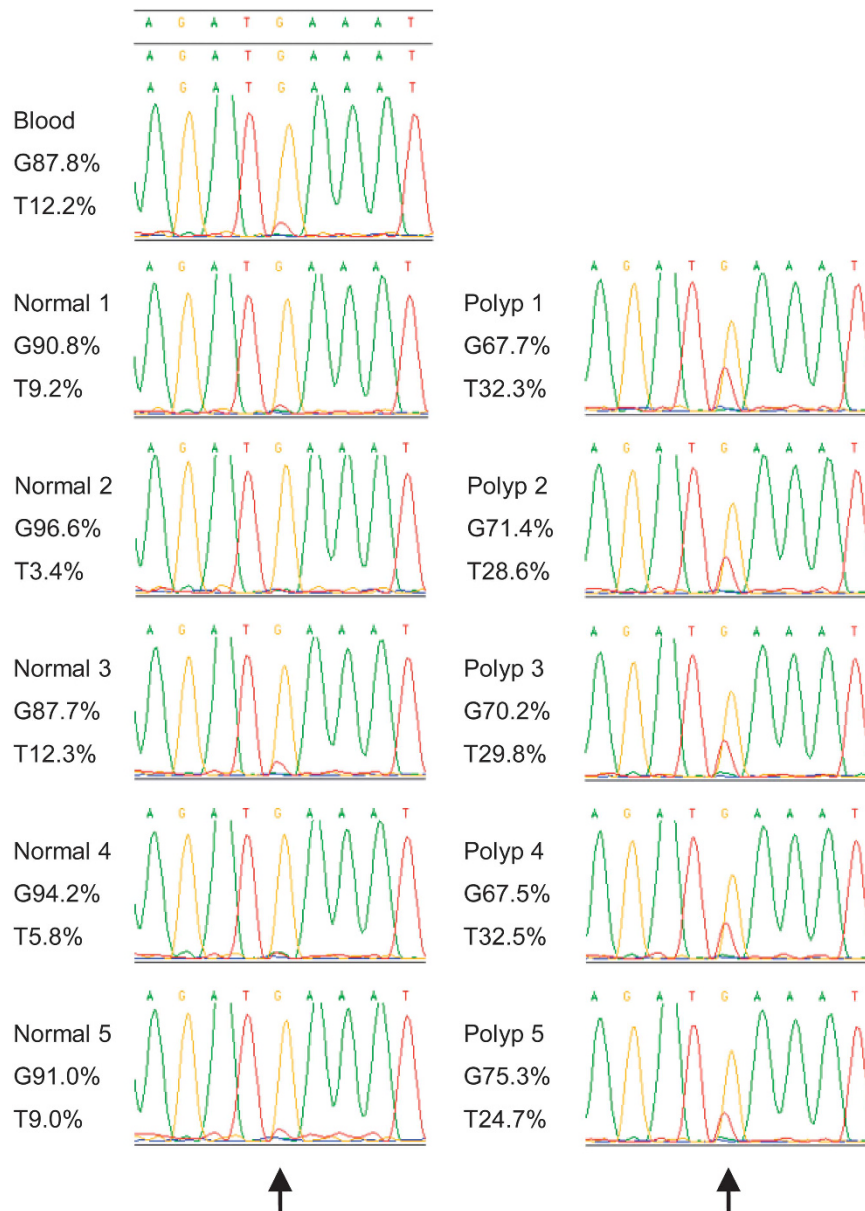
Figure 1 Direct sequencing of the mutant allele (c.3175G>T) in peripheral blood, non-cancerous colonic mucosae and adenomas of the colon. The arrows indicate the position of mutation. The corresponding results obtained from deep sequencing are shown on the left.

than the non-tumorous mucosa (Table 2). These data implied that the c.3175G>T mutation has a crucial role in the development of adenoma, and are consistent with the view of a clonal expansion of cells carrying *APC* mutation. As the average frequency of *APC* mutation in the polyps is ∼30%, we can estimate that between 50 and 60% of the polyps are composed of tumor cells if loss of heterozygosity is not involved in the tumorigenesis.

DISCUSSION

In this report, we have shown that NGS is a powerful tool to identify mosaicism of *APC* mutation in a patient with polyposis. Although reports of somatic *APC* mosaicism are limited, recent studies have demonstrated that the mosaicism (10–20%) is present in a significant number of FAP patients harboring *de novo* germline mutation of *APC*.[4,6] Therefore, *APC* mosaicism may have been underestimated

compared with other tumor-suppressor genes such as *RB1*, *TSC1* and *TSC2*.[12–14]

According to the number of polyps and age of onset, FAP is divided into three clinical subtypes: attenuated (mild), classical (typical) and severe.[15,16] A genotype–phenotype correlation has been well known in FAP. A recent study confirmed that patients with *APC* mutations between codon 1249–1549 certainly developed polyposis at an early age and exhibited a worse survival. On the other hand, patients with *APC* mutations in codon <178 or 312–412 have a later onset of polyposis and exhibited an improved survival.[17] Aretz *et al.*[4] previously reported eight cases with *APC* mosaic mutations. Although the eight mutations were located between codons 216–1464 associated with classical or severe phenotype, most of the patients exhibited mild phenotype. As the cells carrying *APC* mutation are scattered at a lower frequency in the epithelia of colonic mucosa compared with the classical FAP, the patients with mosaic mutations likely have a milder

Table 2 Frequency of the mutation (c.3175G>T) in various tissues

Tissue	Subject	Frequency (%)	No of reads		
			Total	Ref allele (G)	Variant allele (T)
Blood(control)	1	0.0	64 489	31 023+ 33 386−	6+ 0−
Blood	1	12.2	3726	1705+ 1566−	302+ 151−
Hair follicles	1	4.5	83 679	39 527+ 40 249−	1918+ 1856−
Buccal mucosa	1	3.0	69 169	32 605+ 34 368−	1047+ 1052−
	2	7.3	66 557	30 510+ 31 082−	2477+ 2383−
Gastric mucosa	1	18.9	81 041	29 167+ 36 446−	7025+ 8228−
	2	22.7	74 101	27 245+ 29 871−	8089+ 8654−
	3	27.7	84 151	27 281+ 33 392−	11 009+ 12 284−
Normal coloic mucosa	1	9.2	5149	2419+ 2252−	290+ 182−
	2	3.4	4698	2248+ 2291−	83+ 76−
	3	12.3	4321	1978+ 1807−	327+ 206−
	4	5.8	4514	2226+ 2026−	165+ 96−
	5	9.0	4097	2123+ 1602−	267+ 103−
Polyp	1	32.3	4797	1687+ 1557−	1013+ 536−
	2	28.6	4543	1883+ 1356−	937+ 362−
	3	29.8	4402	1537+ 1551−	841+ 469−
	4	32.5	4554	1628+ 1442−	981+ 494−
	5	24.7	4166	1581+ 1549−	685+ 344−

+,− indicate strand reads.

polyposis phenotype than the expected phenotype based on the site of the mutation.[4,6] In agreement with this notion, our case showed an attenuated form of polyposis demonstrating less than 100 polyps in the intestine, although the mutation of *APC* c.3175G>T, p.E1059X was reported to exhibit a florid form of adenomatous polyps in the large intestine at a young age and additional extra-colonic manifestations (duodenal adenoma and fundic gland polyps) in a FAP patient.[18] Application of NGS in genetic testing for patients with polyposis may increase the frequency of *APC* mosaicism in cases without familial history and those with mild phenotype.

After the surveillance of large intestine for eight years, the patient underwent subtotal colectomy in combination with ileorectal anastomosis because two of four biopsies of the colonic polyps histologically showed severe atypia. As we were unable to examine the DNA of his parents, we could not confirm *de novo APC* mutation in the patient. In addition, although his children have not been investigated, we should consider the possibility that germline mosaicism may lead to more severe phenotypes in the next generation.

On the basis of our result from deep sequencing, the frequency of the mutant allele was not constant among different types of tissues, and even in different sections isolated from the same tissue. The mutant allele was also found to a relatively higher extent in normal gastric mucosa (18.9–27.7%) compared with that in peripheral blood (12.2%), hair follicles (4.5%), buccal mucosa (3.0–7.3%) and normal colonic mucosa (3.4–12.3%). During embryogenesis, a zygote starts development and forms three germ layers; ectoderm, endoderm and mesoderm. Peripheral blood originates from mesoderm, hair follicles and buccal mucosa from ectoderm, and gastric and colonic mucosa from endoderm. Therefore the mutation should occur before the separation of these three layers. Although cells carrying the mutation were delivered into different tissues at different frequencies, we may assume from the frequencies that the mutation occurred at the four- or eight-cell stage. As the patient carried the mutation in non-tumorous gastric mucosa at a relatively high frequency, future surveillance paying special attention to his stomach is essential because the relative risk for gastric cancer in FAP patients is higher than the

normal population in Asia.[19,20] Interestingly we also found that the mutation frequency is different at the location even in the same tissue, suggesting a possibility of intra-organic mosaicism.

Sequencing by the Sanger method is a gold standard of genetic testing for FAP. However, it is reported that the sensitivity of mutation by the direct sequence method is ~ 15%.[21] Consistent with this data, we failed to identify the mutation in our initial screening. Although PTT, DGGE, DHPLC and HRM are also applicable for the initial screening and may have higher sensitivities compared with Sanger's method, they are indirect detection systems and additional confirmatory sequencing is essential. In the near future, sequencing by NGS will replace the screening strategies for polyposis because the cost for NGS is dramatically decreasing. Previous reports also revealed the effectiveness of NGS for the detection of mosaic mutations.[21,22] Consistently, our data show that amplicon sequence with NGS is useful for the quantification of mosaic mutation (Figure 1 and Table 2). We additionally confirmed the mutation using a different set of primers, and the degree of mosaicism was comparable to that using the initial set of primers (data not shown), suggesting that the ratio is not affected by the set of primers used. Therefore, the amplicon-based NGS approach is reliable for confirming low-level mutation and quantifying mosaic mutations. It is of note that the sensitivity of detection of mosaic mutation by NGS totally depends on the number of reads. Although we identified the mosaic mutation in 6 out of 50 reads, we might miss the mutation if the depth of reads at the region was less than 10 or so. Although the sensitivity to identify low levels of mosaic mutation will be increased by the use a high depth sequence method such as amplicon sequence or whole-exome sequence, these methods may overlook structural changes such as translocation and deletion/amplification of large regions. Therefore, we should take advantage of the most appropriate method for the detection of different types of alterations. In addition, it may be possible to analyze multiple polyps, if available, in the patients, because the tumors are largely composed of relatively homogeneous cell population, and the mosaic mutation should be shared in the DNA from the polyps.

In conclusion, we successfully identified a mosaic mutation in a patient with a fraction of 12% mutated allele in peripheral blood by whole-genome sequencing covered at 26x average depth. As mutational mosaicism of the *APC* gene has relevance to cancer risk, genetic diagnosis is useful for the decision of surveillance and personalized treatment of the patients, and may be applied for the pre-symptomatic diagnosis of their children. These results will accelerate the application of NGS in clinical settings.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

1 Stoffel, E. M. & Kastrinos, F. Familial colorectal cancer, beyond Lynch syndrome. *Clin. Gastroenterol. Hepatol.* **12**, 1059–1068 (2014).

2 Al-Tassan, N., Chmiel, N. H., Maynard, J., Fleming, N., Livingston, A. L. & Williams, G. T. *et al.* Inherited variants of MYH associated with somatic G:C−>T:A mutations in colorectal tumors. *Nat. Genet.* **30**, 227–232 (2002).

3 Palles, C., Cazier, J. B., Howarth, K. M., Domingo, E., Jones, A. M. & Broderick, P. *et al.* Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas. *Nat. Genet.* **45**, 136–144 (2013).

4 Aretz, S., Stienen, D., Friedrichs, N., Stemmler, S., Uhlhaas, S. & Rahner, N. *et al.* Somatic APC mosaicism: a frequent cause of familial adenomatous polyposis (FAP). *Hum. Mutat.* **28**, 985–992 (2007).

5 Farrington, S. M. & Dunlop, M. G. Mosaicism and sporadic familial adenomatous polyposis. *Am. J. Hum. Genet.* **64**, 653–658 (1999).

6 Hes, F. J., Nielsen, M., Bik, E. C., Konvalinka, D., Wijnen, J. T. & Bakker, E. *et al.* Somatic APC mosaicism: an underestimated cause of polyposis coli. *Gut.* **57**, 71–76 (2008).

7 Mongin, C., Coulet, F., Lefevre, J. H., Colas, C., Svrcek, M. & Eyries, M. *et al.* Unexplained polyposis: a challenge for geneticists, pathologists and gastroenterologists. *Clin. Genet.* **81**, 38–46 (2012).

8 Johansen Taber, K. A., Dickinson, B. D. & Wilson, M. The promise and challenges of next-generation genome sequencing for clinical care. *JAMA Intern. Med.* **174**, 275–280 (2014).

9 Biesecker, L. G. Opportunities and challenges for the integration of massively parallel genomic sequencing into clinical practice: lessons from the ClinSeq project. *Genet. Med.* **14**, 393–398 (2012).

10 Green, R. C., Berg, J. S., Grody, W. W., Kalia, S. S., Korf, B. R. & Martin, C. L. *et al.* ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genet. Med.* **15**, 565–574 (2013).

11 Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

12 Lohmann, D. R. RB1 gene mutations in retinoblastoma. *Hum. Mutat.* **14**, 283–288 (1999).

13 Sippel, K. C., Fraioli, R. E., Smith, G. D., Schalkoff, M. E., Sutherland, J. & Gallie, B. L. *et al.* Frequency of somatic and germ-line mosaicism in retinoblastoma: implications for genetic counseling. *Am. J. Hum. Genet.* **62**, 610–619 (1998).

14 Verhoef, S., Bakker, L., Tempelaars, A. M., Hesseling-Janssen, A. L., Mazurczak, T. & Jozwiak, S. *et al.* High rate of mosaicism in tuberous sclerosis complex. *Am. J. Hum. Genet.* **64**, 1632–1637 (1999).

15 Knudsen, A. L., Bisgaard, M. L. & Bulow, S. Attenuated familial adenomatous polyposis (AFAP). A review of the literature. *Fam. Cancer.* **2**, 43–55 (2003).

16 Nagase, H., Miyoshi, Y., Horii, A., Aoki, T., Ogawa, M. & Utsunomiya, J. *et al.* Correlation between the location of germ-line mutations in the APC gene and the number of colorectal polyps in familial adenomatous polyposis patients. *Cancer Res.* **52**, 4055–4057 (1992).

17 Newton, K. F., Mallinson, E. K., Bowen, J., Lalloo, F., Clancy, T. & Hill, J. *et al.* Genotype-phenotype correlation in colorectal polyposis. *Clin. Genet.* **81**, 521–531 (2012).

18 Kanter-Smoler, G., Fritzell, K., Rohlin, A., Engwall, Y., Hallberg, B. & Bergman, A. *et al.* Clinical characterization and the mutation spectrum in Swedish adenomatous polyposis families. *BMC Med.* **6**, 10 (2008).

19 Iwama, T., Mishima, Y. & Utsunomiya, J. The impact of familial adenomatous polyposis on the tumorigenesis and mortality at the several organs. Its rational treatment. *Ann. Surg.* **217**, 101–108 (1993).

20 Park, J. G., Park, K. J., Ahn, Y. O., Song, I. S., Choi, K. W. & Moon, H. Y. *et al.* Risk of gastric cancer among Korean familial adenomatous polyposis patients. Report of three cases. *Dis. Colon Rectum* **35**, 996–998 (1992).

21 Rohlin, A., Wernersson, J., Engwall, Y., Wiklund, L., Bjork, J. & Nordling, M. Parallel sequencing used in detection of mosaic mutations: comparison with four diagnostic DNA screening techniques. *Hum. Mutat.* **30**, 1012–1020 (2009).

22 Pagnamenta, A. T., Lise, S., Harrison, V., Stewart, H., Jayawant, S. & Quaghebeur, G. *et al.* Exome sequencing can detect pathogenic mosaic mutations present at low allele frequencies. *J. Hum. Genet.* **57**, 70–72 (2012).

Supplementary Information accompanies the paper on Journal of Human Genetics website (http://www.nature.com/jhg)