



# Genetic and epigenetic profiling of a solitary Peutz–Jeghers colon polyp

Heinz Linhart,<sup>1,2</sup> Felix Bormann,<sup>1</sup> Barbara Hutter,<sup>3</sup> Benedikt Brors,<sup>3</sup> and Frank Lyko<sup>1</sup>

<sup>1</sup>Division of Epigenetics, DKFZ-ZMBH Alliance, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany; <sup>2</sup>Asklepios Klinik Lindau, 88131 Lindau, Germany; <sup>3</sup>Division of Applied Bioinformatics, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany

**Abstract** Colon polyps represent precursor lesions of colon cancers and their malignant potential varies according to histological subtype. A rare subtype of colon polyps is the Peutz–Jeghers (PJ) polyp. PJ polyps mostly occur in the context of Peutz–Jeghers syndrome, which is characterized by the development of multiple polyps in the intestinal tract and hyperpigmentation of oral mucosa and lips. Peutz–Jeghers is an autosomal dominant disorder caused by pathogenic variants of the serine threonine kinase *STK11*. PJ polyps very rarely occur outside of the syndrome and are then referred to as solitary PJ polyps. Contrary to the situation in Peutz–Jeghers, the genetic basis and the malignant potential of solitary PJ polyps are currently unknown. Here we describe a detailed and comprehensive genetic profile of a solitary PJ polyp. Pathological examination revealed a high tissue homogeneity with >80% epithelial cells. Whole-genome sequencing failed to identify any clonal mutations but demonstrated a significant number of subclonal mutations. No somatic or germline mutations were found at the *STK11* locus, suggesting that solitary PJ polyps are genetically distinct from Peutz–Jeghers polyps. In addition, methylome analysis revealed global hypomethylation and CpG island hypermethylation, two features that have been described as hallmarks of the colorectal cancer epigenome. These results provide an example of a premalignant lesion that is defined by epigenetic, rather than genetic changes. Furthermore, our findings support the notion that solitary PJ polyps constitute neoplastic tissue with malignant potential that should be removed for cancer prevention.

Corresponding authors:  
h.linhart@asklepios.com;  
f.lyko@dkfz.de

© 2017 Linhart et al. This article is distributed under the terms of the Creative Commons Attribution-NonCommercial License, which permits reuse and redistribution, except for commercial purposes, provided that the original author and source are credited.

**Ontology terms:** intestinal polyp

Published by Cold Spring Harbor Laboratory Press

doi: 10.1101/mcs.a001610

## INTRODUCTION

Colon polyps are outgrowths of the colon mucosa and can progress to colon cancer. Polyps with malignant potential are categorized as neoplastic polyps, whereas polyps that are considered to have no cancer risk are categorized as nonneoplastic. Each of these categories is further subdivided into multiple histological subtypes. The Peutz–Jeghers (PJ) polyp constitutes a rare subtype with a characteristic histological appearance showing tree-like branching of smooth muscle cells and lobular organization of colonic crypts (Tse et al. 2013). PJ polyps almost exclusively occur in the context of the Peutz–Jeghers syndrome, which is characterized by mucocutaneous hyperpigmentation and the development of multiple intestinal polyps in the first two to three decades of life. Peutz–Jeghers is an autosomal dominant disorder caused by pathogenic variants of the *STK11* gene located on human Chromosome 19p13 (Yoo et al. 2002).

Even though PJ polyps are categorized as nonneoplastic, the disorder is associated with a 93% cancer risk (Mehenni et al. 2006), and the risk of colon cancer in Peutz–Jeghers patients has been estimated at 39% (van Lier et al. 2010). Also, neoplastic transformation of PJ polyps has been documented (Defago et al. 1996). Contrary to the polyps associated with Peutz–Jeghers syndrome, however, the molecular pathology and malignant potential of solitary PJ polyps is unknown. Only one study reported molecular analysis of a solitary PJ polyp finding no loss of heterozygosity at the 19p13.3 locus and no mutation of *STK11* (Kitaoka et al. 2004), suggesting that solitary PJ polyps constitute a separate entity. To further characterize the molecular pathology and malignant potential we here conducted the first comprehensive DNA sequencing and DNA methylation analysis of a solitary PJ polyp.

## RESULTS

---

### Clinical Presentation and Family History

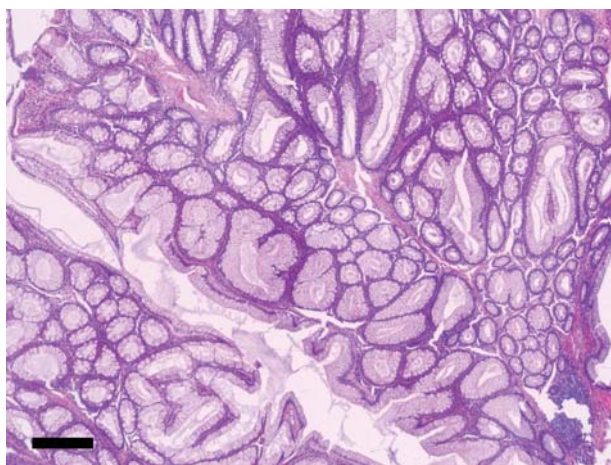
The patient complained of intermittent midabdominal pain that had started ~6 mo prior to admission. In addition, he reported a recent change of bowel habits with intermittent loose stools, no loss of weight, and no rectal bleeding. Past medical history included coronary heart disease and chronic obstructive pulmonary disease. The family history was negative for gastrointestinal tumors. Outpatient colonoscopy had shown a pedunculated polyp located in the distal section of the sigmoid colon and sigmoid diverticulosis. The patient was then referred to our clinic for polypectomy. The age of the patient at the time of polypectomy was 77 yr.

At admission clinical examination and laboratory studies were unremarkable and mucocutaneous hyperpigmentation was not detectable. Repeat colonoscopy in our hospital confirmed sigmoid diverticulosis and a pedunculated polyp of the distal sigmoid colon located at 20 cm from the anocutaneous line with a diameter of ~3 cm. No other polyps were detected by colonoscopy. The polyp was removed by an endoscopic snare and submitted for pathological assessment. Hematoxylin and eosin staining showed the characteristic lobular grouping of crypts surrounded by branching smooth muscle fibers (Fig. 1). The fraction of epithelial cells in the polyp was estimated to be >80%.

### Genomic Analyses

Whole-genome sequencing was performed at 70× genome coverage, respectively, for the polyp sample and for a blood sample from the same patient (Table 1). Data analysis identified 2079 high-confidence somatic single-nucleotide variants (SNVs), most of which correspond to C > T substitutions in an ACG trinucleotide context. Allele frequencies show a sharp peak at 0.1, indicating that most of the mutations are nonclonal. This pattern is reflected in the subset of 17 missense variants (Table 2). Notably, the *ERBB3* p.P306H mutation has been described in the COSMIC (Catalogue of Somatic Mutations in Cancer) database and was also identified in a non-small-cell lung adenocarcinoma in our in-house tumor cohorts. Located in the growth factor receptor/furin-like cysteine-rich domain, it is annotated as pathogenic by various functional impact prediction tools.

Analysis of mutational signatures (Alexandrov et al. 2013) identified AC1 as the dominant contributor. This signature corresponds to the spontaneous deamination of 5-methylcytosine, correlating with the age of the patient. We did not find any signatures related to DNA mismatch repair, which is in line with the stable genomic profile. Copy-number analysis based on read depth plots and structural variation detection tools failed to reveal any reliable focal or large-scale somatic aberrations. We observed a subclonal loss of the Y Chromosome in both the blood and, to a lesser extent, the polyp. This phenomenon can be explained by age-related loss of Y in a fraction of white blood cells (Pierre and Hoagland 1972; Forsberg et al. 2014).



**Figure 1.** A representative section of a solitary PJ polyp (hematoxylin and eosin stain). The image shows characteristic lobular grouping of crypts surrounded by branching smooth muscle fibers. Scale bar, 200  $\mu$ m.

From the 42 high-confidence somatic insertions and deletions (indels) that were identified genome-wide, none fall into a coding gene or splice site, probably because of allele frequencies below the default reliability threshold. Lowering the threshold identified 243 indels, three of which were located in coding regions (Table 2). The *FLT3* frameshift insertion is located in the protein kinase domain, just 30 amino acids from the hotspot position at D835.

Notably, neither somatic nor germline mutations or copy-number alterations were found to be present at the *STK11* locus, even though this gene is often found mutated in Peutz–Jeghers patients (Yoo et al. 2002). The only deviation from the reference sequence is the common intronic variant rs4807008, for which the patient is homozygous. We could also not identify any rare or private coding germline SNV or indel in known cancer-predisposing genes. These results are consistent with the findings from the only other published study investigating the *STK11* mutation status of a solitary PJ polyp (Kitaoka et al. 2004).

### DNA Methylation Analyses

Because our genomic analysis did not detect any clonal mutations, we also analyzed the epigenome of the polyp. To this end, we used array-based methylation profiling of 450,000 methylation sites in the human genome (Infinium HumanMethylation450 arrays). The same platform was also used by The Cancer Genome Atlas (TCGA) for the analysis of 373 colorectal tumors and 45 control colon mucosa samples. Principal component analysis of the combined data sets showed that the solitary PJ polyp clustered separately from the majority of the normal colonic mucosa samples (Fig. 2A), thus indicating an aberrant methylome in the adenoma sample.

To further analyze these methylation changes, methylation values of individual probes were compared between the polyp and the TCGA normal data set. This identified 306

**Table 1.** Sequencing coverage of whole-genome sequencing

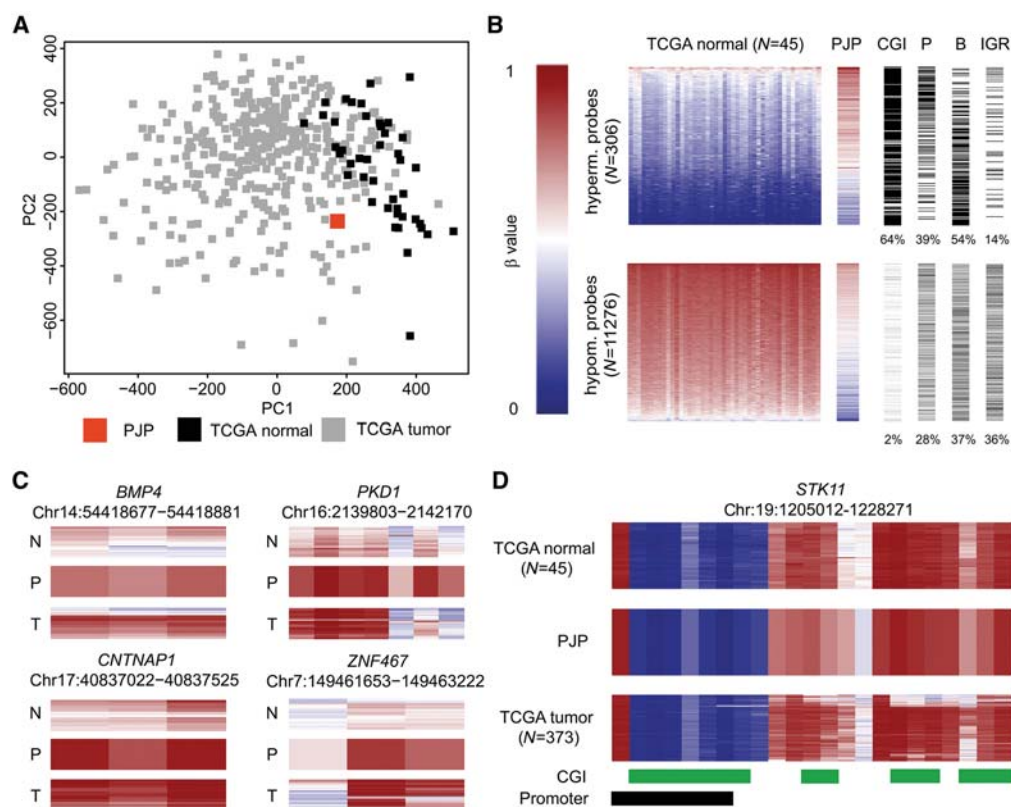
Sample	Read pairs sequenced	Reads aligned	PCR duplicates	Average read coverage
Polyp	1,680,950,965	99.73%	8.70%	69.97 $\times$
Blood	1,764,374,344	99.56%	8.50%	73.81 $\times$

PCR, polymerase chain reaction.

**Table 2.** Somatic mutations in a solitary PJ polyp detected by whole-genome sequencing

Chr	Position GRCh37	dbSNP ID	Ref. base	Alt. base	Gene	Variant type	Predicted effect	Coding DNA exchange	Amino acid exchange	Genotype	Allele freq.	COSMIC ID
1	38082222	rs200218812	C	T	RSPO1	Substitution	Missense	c.G220A	p.V74I	het.	0.07	
1	158622275		C	A	SPTA1	Substitution	Missense	c.G3357T	p.K1119N	het.	0.06	
3	73432532		G	A	PDZRN3	Substitution	Missense	C3185T	p.S1062	het.	0.07	
4	156860666	rs199672156	C	T	CTSO	Substitution	Missense	c.G409A	p.V137M	het.	0.10	
5	52097492		CGGA	C	PELO	Deletion	No frameshift	c.977_979del	p.326_327del	het.	0.05	
5	52097497		C	CT	PELO	Insertion	Frameshift	c.982_983insT	p.R328fs	het.	0.05	
5	148753978		C	T	IL17B	Substitution	Missense	c.G497A	p.R166H	het.	0.08	
6	102124509		T	A	GRIK2	Substitution	Missense	c.T553A	p.L185M	het.	0.06	
7	95906629		C	T	SLC25A13	Substitution	Missense	c.G91A	p.G31S	het.	0.06	COSM3883588
8	25745417		C	T	EBF2	Substitution	Missense	c.G823A	p.G275R	het.	0.07	
8	67355035		C	T	ADHFE1	Substitution	Missense	c.C100T	p.P34S	het.	0.07	
11	73179461		A	G	FAM168A	Substitution	Missense	c.T59C	p.M20T	het.	0.09	
12	56482369		C	A	ERBB3	Substitution	Missense	c.C917A	p.P306H	het.	0.09	COSM4043445
13	27257046	rs373084396	G	A	WASF3	Substitution	Missense	c.G1286A	p.R429Q	het.	0.11	
13	28589790		T		FLT3	Insertion	Frameshift	c.2597_2598ins GACCACCGAG	p.G866fs	het.	0.07	
13	50141395		C	A	RCBTB1	Substitution	Missense	c.G21T	p.W7C	het.	0.10	
14	24045075		C	A	JPH4	Substitution	Missense	c.G970T	p.D324Y	het.	0.10	
19	48533818		C	T	CABP5	Substitution	Missense	c.G518A	p.R173H	het.	0.09	COSM299924
19	50393133		C	T	IL41	Substitution	Missense	c.G1498A	p.A500T	het.	0.09	
19	52147182		G	A	SIGLEC14	Substitution	Missense	c.C862T	p.R288W	het.	0.10	

dbSNP, Database for Short Genetic Variations; COSMIC, Catalogue of Somatic Mutations in Cancer; het., heterozygous.



**Figure 2.** Aberrant DNA methylation in a solitary PJ polyp. (A) Principal component analysis of all probes in the merged TCGA (The Cancer Genome Atlas) and polyp data set. (B) Identification of hyper- and hypomethylated probes in the PJ polyp.  $\beta$  values are colored from blue ( $\beta = 0$ ) over white ( $\beta = 0.5$ ) to red ( $\beta = 1$ ). Probe annotation features are shown on the right. CGI, CpG island; P, promoter; B, gene body; IGR, intergenic region. (C) Methylation patterns of selected “tumor prone CpG islands” in TCGA normal colonic mucosa (N,  $N = 45$ ), the PJ polyp (PJP,  $N = 1$ ), and TCGA colorectal tumors (T,  $N = 373$ ). Each column represents one probe of the CpG island,  $\beta$  values are indicated as in B. (D) Methylation pattern of the *STK11* locus. CpG island–associated probes are indicated by green bars; promoter-associated probes are indicated by the black bar.

hypermethylated probes and 11,276 hypomethylated probes (Fig. 2B). Interestingly, the hypermethylated probes were strongly associated with CpG islands, whereas the hypomethylated probes were often associated with intergenic regions (Fig. 2B). These features have been described as the defining hallmarks of the colon cancer methylome (Berman et al. 2012). We also identified “tumor prone CpG islands” (Berman et al. 2012) using the TCGA colorectal cancer data set and subsequently compared their methylation with the polyp data (see Methods). This identified several CpG islands that showed tumor-associated hypermethylation in the solitary PJ polyp (Fig. 2C). Of note, no methylation changes were detected at the *STK11* locus (Fig. 2D), indicating that the locus is not affected by a cancer-specific epigenetic mutation.

## DISCUSSION

Solitary PJ polyps are a rare type of hamartomatous polyps of unknown etiology. Our report provides the first genome sequence of a solitary PJ polyp and reveals a number of genetic variants with relatively low allele frequencies (~10%) and in genes that are not commonly

associated with colorectal carcinogenesis. Consistent with a previous report (Kitaoka et al. 2004), no genetic alterations were detected at the *STK11* locus, which is commonly associated with Peutz–Jeghers. These findings indicate that the molecular profile of solitary PJ polyps differs from Peutz–Jeghers syndrome. Our report also provides the first methylome analysis of a solitary PJ polyp and reveals moderate epigenetic alterations that are similar to the methylation changes of colorectal tumors. Together, our findings establish the first comprehensive genetic profile of a solitary PJ polyp.

Colorectal polyps are precursor lesions of colorectal cancer and are generally considered to evolve clonally from the colorectal mucosa (Fearon et al. 1987). The absence of clonal mutations in the solitary PJ polyp therefore suggests that the initial signal for clonal expansion is not provided by genetic changes in epithelial cells but by other mechanisms. Indeed, our data can be interpreted to reflect a substantial intrapolyp heterogeneity for subclonal mutations, which is consistent with the “Big Bang” model of colorectal carcinogenesis (Sottoriva et al. 2015). The acquisition of a tumor-driving mutation in a specific subclone could then represent the initiating event for tumor growth.

Interestingly, our methylome analysis revealed additional molecular alterations of the polyp genome. A comparative methylation pattern analysis with data sets from normal colorectal mucosa and colorectal tumor samples revealed moderate but clearly detectable epigenetic changes in the polyp that resemble the methylation changes commonly associated with colorectal cancer (Berman et al. 2012). More specifically, we identified “tumor-prone CpG islands” that showed a similar degree of hypermethylation in the polyp and in colorectal cancer. These findings are consistent with the notion that epigenetic changes represent early events during colorectal carcinogenesis (Lao and Grady 2011).

It is well established that Peutz–Jeghers patients carry a significant cancer risk, and it has been shown that cancer can develop in PJ polyps (McGarrity et al. 2000). Accordingly, surveillance colonoscopy is recommended in affected individuals starting at age 8, and removal of polyps of >5 mm is recommended (Syngal et al. 2015). Contrary to the polyps from Peutz–Jeghers patients, solitary PJ polyps are considered to have no malignant potential (Oncel et al. 2003). However, because solitary PJ polyps constitute a rare entity it is not possible to reliably determine malignant potential by clinical studies. Rather, cancer risk has to be extrapolated from molecular data. Our finding of tumor-like CpG island hypermethylation and genomic hypomethylation in a solitary PJ polyp supports the notion that such polyps do carry a cancer risk and should be removed.

## METHODS

---

### Pathology and DNA Isolation

The polyp was removed by colonoscopy using an endoscopic snare. Subsequently, a representative part of the sample was removed and placed into RNAlater solution for stabilization of nucleic acids. The remaining portion of the polyp was formalin fixed and paraffin embedded for pathological analysis. DNA was isolated using the QIAGEN Blood and Tissue Kit. Following hematoxylin and eosin staining of paraffin embedded sections pathological assessment of polyp morphology was conducted according to current guidelines (Quirke et al. 2011).

### Genomic Analyses

Each sample was sequenced on two lanes of an Illumina X-Ten platform (see Table 1 for sequencing coverages). Reads were mapped to the 1000 Genomes phase 2 assembly of the human reference genome GRCh37 (NCBI build 37.1, downloaded from [ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/technical/reference/phase2\\_reference\\_assembly\\_sequence](ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/technical/reference/phase2_reference_assembly_sequence)) using

a hardware-accelerated version of BWA (Burrows–Wheeler alignment) (Li and Durbin 2009) version 0.7.8 mem with option `-T 0`. The sambamba package (Tarasov et al. 2015) version 0.4.6 and 0.5.9, respectively, was used for sorting and to mark polymerase chain reaction (PCR) duplicates during merging the per-lane BAM (binary alignment) files.

For detection of SNVs we used our in-house pipeline. The pipeline is based on SAMtools mpileup and BCFtools (Li et al. 2009) version 0.1.19 with parameter adjustments to allow calling of somatic variants with low allele frequency using heuristic filtering as described previously (Jones et al. 2013). We applied Platypus (Rimmer et al. 2014) version 0.7.4 to identify short indels with a similar reliability scoring as for SNVs. After not detecting any coding indels with the default high-confidence threshold of 8, the value was lowered to 7 to allow also medium-confidence events with variant frequency of  $<10\%$  and allele bias at the same time. All mutations were annotated with ANNOVAR (Wang et al. 2010) version November 2014, using the GENCODE version 19 gene model. Additional information was retrieved by overlapping the genomic positions with dbSNP (Database for Short Genetic Variations) version 141, 1000 Genomes phase1 integrated calls 20101123, COSMIC version 71, and dbNSFP (Database for Nonsynonymous SNPs' Functional Predictions) version 2.9 (Liu et al. 2011, 2013).

A supervised analysis of the genome-wide somatic SNVs for mutational signatures (Alexandrov et al. 2013) was conducted with the R package YAPSA. A detailed description of the algorithm is provided elsewhere (Dieter et al. 2016). We searched for structural variants with CREST (Wang et al. 2011) and applied in-house tools for copy-number identification.

### DNA Methylation Analysis

Available Infinium 450K raw data idat files from colon adenocarcinoma, rectal adenocarcinoma, and corresponding normal samples were downloaded from The Cancer Genome Atlas repository and preprocessed using Minfi (Aryee et al. 2014). Probes located on the sex chromosomes, or probes that failed detection *P*-value testing in at least one sample, were removed from the analysis. Probes showing cross-reactivity to at least one distant site, and thus allowing a mismatch of one base in the probe sequence, were also eliminated. Finally, probes containing a SNP with an allele frequency of at least 0.01 and located a maximum of 5 bases away from the single-base extension (SBE) site were also removed.  $\beta$  values were extracted from the remaining probes.

Principal component analysis using the R package FactoMiner (Lê et al. 2008) was performed with the extracted and merged  $\beta$  values of the TCGA data set and the PJ polyp data. Differentially methylated probes between the TCGA normal data set and the PJ polyp were addressed by identifying the probes in the PJ polyp data with an absolute  $\beta$ -value difference of  $\geq 0.1$  from the mean  $\beta$  value of the TCGA normal data set and with an absolute  $\beta$  value greater than the mean  $\beta$  value plus three times the standard deviation of the TCGA normal data set.

Hypermethylated tumor-prone CpG islands were identified in the TCGA data set by averaging the  $\beta$  values of all probes per CpG island and comparing the result of the tumor and normal data using a two-sided *t*-test. CpG islands showing a  $\beta$ -value difference of  $\geq 0.1$  and a *P*-value of  $<0.01$ , were considered hypermethylated in colorectal tumors.

## ADDITIONAL INFORMATION

---

### Data Deposition and Access

Genome-sequencing data are available from the NCBI Sequence Read Archive (SRA; <https://www.ncbi.nlm.nih.gov/sra>) under the accession numbers SRR5028731 (polyp) and

SRR5028722 (blood). Methylation data are available from the EBI-EMBL ArrayExpress database (<https://www.ebi.ac.uk/arrayexpress/>) under the accession number E-MTAB-5255.

### Ethics Statement

Written consent for tissue sampling and analysis was obtained from the patient and was conducted as approved by the ethics committee of the University of Heidelberg (reference number S-101/2010). Also, permission in writing was obtained from the patient to allow for publication of genetic data.

### Acknowledgments

We thank Tanja Musch for technical support, Günter Raddatz for computational support, the DKFZ Genomics and Proteomics Core Facility for sequencing services, and Thomas Geisler (Pathology, Amedes, Göttingen) for support with histological assessment and imaging. We also thank Kortine Kleinheinz, Daniel Huebschmann, and Ivo Buchhalter for providing bioinformatics tools and help with data interpretation.

### Competing Interest Statement

The authors have declared no competing interest.

### Referees

Yehudit Bergman  
Anonymous

Received November 24, 2016;  
accepted in revised form January  
24, 2017.

### Author Contributions

H.L. provided patient samples. H.L. and F.L. designed the research. F.B. and B.H. performed the research and analyzed the data. B.B. contributed to the data analysis. H.L., F.B., B.H., and F.L. wrote the manuscript. All authors made significant contributions to, reviewed, and approved the final version of the manuscript.

### Funding

This work was supported by a grant from the Helmholtz-Israel-Cooperation in Personalized Medicine to H.L. and F.L.

## REFERENCES

- Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Børresen-Dale AL, et al. 2013. Signatures of mutational processes in human cancer. *Nature* **500**: 415–421.
- Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, Irizarry RA. 2014. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* **30**: 1363–1369.
- Berman BP, Weisenberger DJ, Aman JF, Hinoue T, Ramjan Z, Liu Y, Noushmehr H, Lange CP, van Dijk CM, Tollenaar RA, et al. 2012. Regions of focal DNA hypermethylation and long-range hypomethylation in colorectal cancer coincide with nuclear lamina-associated domains. *Nat Genet* **44**: 40–46.
- Defago MR, Higa AL, Campra JL, Paradelo M, Uehara A, Torres Mazzucchi MH, Videla R. 1996. Carcinoma in situ arising in a gastric hamartomatous polyp in a patient with Peutz–Jeghers syndrome. *Endoscopy* **28**: 267.
- Dieter SM, Heining C, Agaimy A, Huebschmann D, Bonekamp D, Hutter B, Ehrenberg KR, Fröhlich M, Schlesner M, Scholl C, et al. 2016. Mutant KIT as imatinib-sensitive target in metastatic sinonasal carcinoma. *Ann Oncol* doi:10.1093/annonc/mdw446.
- Fearon ER, Hamilton SR, Vogelstein B. 1987. Clonal analysis of human colorectal tumors. *Science* **238**: 193–197.
- Forsberg LA, Rasi C, Malmqvist N, Davies H, Pasupulati S, Pakalapati G, Sandgren J, Diaz de Stahl T, Zaghlool A, Giedraitis V, et al. 2014. Mosaic loss of chromosome Y in peripheral blood is associated with shorter survival and higher risk of cancer. *Nat Genet* **46**: 624–628.
- Jones DT, Hutter B, Jager N, Korshunov A, Kool M, Warnatz HJ, Zichner T, Lambert SR, Ryzhova M, Quang DA, et al. 2013. Recurrent somatic alterations of FGFR1 and NTRK2 in pilocytic astrocytoma. *Nat Genet* **45**: 927–932.



- Kitaoka F, Shiogama T, Mizutani A, Tsurunaga Y, Fukui H, Higami Y, Shimokawa I, Taguchi T, Kanematsu T. 2004. A solitary Peutz–Jeghers-type hamartomatous polyp in the duodenum. A case report including results of mutation analysis. *Digestion* **69**: 79–82.
- Lao VV, Grady WM. 2011. Epigenetics and colorectal cancer. *Nat Rev Gastroenterol Hepatol* **8**: 686–700.
- Lê S, Josse J, Husson F. 2008. FactoMineR: an R package for multivariate analysis. *J Stat Softw* **25**: 1–18.
- Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**: 1754–1760.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078–2079.
- Liu X, Jian X, Boerwinkle E. 2011. dbNSFP: a lightweight database of human nonsynonymous SNPs and their functional predictions. *Hum Mutat* **32**: 894–899.
- Liu X, Jian X, Boerwinkle E. 2013. dbNSFP v2.0: a database of human non-synonymous SNVs and their functional predictions and annotations. *Hum Mutat* **34**: E2393–E2402.
- McGarrity TJ, Kulin HE, Zaino RJ. 2000. Peutz–Jeghers syndrome. *Am J Gastroenterol* **95**: 596–604.
- Mehenni H, Resta N, Park JG, Miyaki M, Guanti G, Costanza MC. 2006. Cancer risks in *LKB1* germline mutation carriers. *Gut* **55**: 984–990.
- Oncel M, Remzi FH, Church JM, Goldblum JR, Zutshi M, Fazio VW. 2003. Course and follow-up of solitary Peutz–Jeghers polyps: a case series. *Int J Colorectal Dis* **18**: 33–35.
- Pierre RV, Hoagland HC. 1972. Age-associated aneuploidy: loss of Y chromosome from human bone marrow cells with aging. *Cancer* **30**: 889–894.
- Quirke P, Risio M, Lambert R, von Karsa L, Vieth M. 2011. Quality assurance in pathology in colorectal cancer screening and diagnosis—European recommendations. *Virchows Arch* **458**: 1–19.
- Rimmer A, Phan H, Mathieson I, Iqbal Z, Twigg SR, Consortium WGS, Wilkie AO, McVean G, Lunter G. 2014. Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications. *Nat Genet* **46**: 912–918.
- Sottoriva A, Kang H, Ma Z, Graham TA, Salomon MP, Zhao J, Marjoram P, Siegmund K, Press MF, Shibata D, et al. 2015. A Big Bang model of human colorectal tumor growth. *Nat Genet* **47**: 209–216.
- Syngal S, Brand RE, Church JM, Giardiello FM, Hampel HL, Burt RW, American College of Gastroenterology. 2015. ACG clinical guideline: genetic testing and management of hereditary gastrointestinal cancer syndromes. *Am J Gastroenterol* **110**: 223–262; quiz 263.
- Tarasov A, Vilella AJ, Cuppen E, Nijman IJ, Prins P. 2015. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* **31**: 2032–2034.
- Tse JY, Wu S, Shinagare SA, Lauwers GY, Yilmaz O, Wu CL, Deshpande V. 2013. Peutz–Jeghers syndrome: a critical look at colonic Peutz–Jeghers polyps. *Mod Pathol* **26**: 1235–1240.
- van Lier MG, Wagner A, Mathus-Vliegen EM, Kuipers EJ, Steyerberg EW, van Leerdam ME. 2010. High cancer risk in Peutz–Jeghers syndrome: a systematic review and surveillance recommendations. *Am J Gastroenterol* **105**: 1258–1264; author reply 1265.
- Wang K, Li M, Hakonarson H. 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* **38**: e164.
- Wang J, Mullighan CG, Easton J, Roberts S, Heatley SL, Ma J, Rusch MC, Chen K, Harris CC, Ding L, et al. 2011. CREST maps somatic structural variation in cancer genomes with base-pair resolution. *Nat Methods* **8**: 652–654.
- Yoo LI, Chung DC, Yuan J. 2002. *LKB1*—a master tumour suppressor of the small intestine and beyond. *Nat Rev Cancer* **2**: 529–535.