Biochemical and Biophysical Research Communications 447 (2014) 503–507

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

A *de novo* germline mutation of *APC* for inheritable colon cancer in a Chinese family using multigene next generation sequencing

Yan Zhang [a,1], Guanting Lu [b,*,1], Qingtao Hu [c,d,1], Xueyan Wang [b], Chaohua Li [b], Yuegan Mao [b], Manhua Cui [a,*]

[a] Department of Gynaecology and Obstetrics, Second Hospital of Jinlin University, Changchun 130041, PR China
[b] Beijing DnaLead Science and Technology Co., LTD, Beijing 102600, PR China
[c] Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100101, PR China
[d] University of Chinese Academy of Sciences, Beijing 100049, PR China

ARTICLE INFO

Article history:
Received 28 March 2014
Available online 13 April 2014

Keywords:
Inheritable colon cancer
de novo germline mutation
Premature stop codon

ABSTRACT

Inheritable colorectal cancers (CRC) accounted for about 20% of the CRC cases, such as hereditary nonpolyposis colorectal cancer (HNPCC), Gardner syndrome and familial adenomatous polyposis (FAP). A four-generation Han Chinese family was found affected with polyposis in colons. Inferred from the pedigree structure, the disease in this family showed an autosomal dominant inheritance model. To locate the causal mutations in this family, genomic DNAs were extracted and the next generation sequencing for 5 genes relating to colon cancer performed by Ion Torrent Personal Genome Machine with a 314 chip. The reads were aligned with human reference genome hg19 to call variants in the 5 genes. After analysis, 14 variants were detected in the sequenced sample and 13 been collected in dbSNP database and assigned with a rs identification number. In these variants, 9 were synonymous, 4 missense and 1 non-sense. In them, 2 rare variants (c.694C>T in *APC* and c.1690A>G in *MSH2*) might be the putative causal mutations for familial adenomatous polyposis (FAP) since the rarity of the mutated allele in normal controls. c.694C>T was detected in only affected members and generated a premature stop codon in *APC*. It should be a *de novo* germline mutation making *APC* containing this stop codon as targets for nonsense-mediated mRNA decay (NMD). c.1690A>G in *MSH2* was not only detected in affected members, but also in normal ones in the family. Functional prediction revealed that the amino acid affected by this variant had no effect on the function of *MSH2*. Here, we report a *de novo* germline mutation of *APC* as the causal variant in a Chinese family with inheritable colon cancer by the next generation sequencing.

1. Introduction

Colorectal cancer (CRC) is a cancer from uncontrolled cell growth in the colon or rectum or in the appendix. It is the third most commonly diagnosed cancer in the world and more common with around 60% of cases in developed countries [1]. It was estimated that in 2008, 1.23 million new cases of CRC were clinically diagnosed in the world, and killed 608,000 people [2]. Most CRCs occurred due to lifestyle and increasing age with only a minority of cases associated with underlying genetic disorders.

Those with a family history in two or more first-degree relatives had a two to threefold greater risk of disease and accounted for about 20% of all cases. A number of genetic syndromes were also associated with higher rates of colorectal cancer. The most common was hereditary nonpolyposis colorectal cancer (HNPCC) which was present in about 3% of CRC cases [3]. Other syndromes that were strongly associated with colon cancer including Gardner syndrome [4], and familial adenomatous polyposis (FAP) in which colon cancer nearly always occurred and accounted for approximately 1% of cases [5,6].

In familial inherited colon cancers, mutations in genes responsible for mismatch repair (MMR) and antagonizing for the WNT signalling pathway played dominant role in the development of the condition. *MLH1*, *MSH2*, *MSH6* and *PMS2* were 4 important MMR genes and had been related to HNPCC1 (OMIM: 120435) and mismatch repair cancer syndrome (OMIM: 276300) [7,8]. *APC* (Adenomatous polyposis coli) was an determinant suppressor for the WNT signalling pathway relating to brain tumor-polyposis syndrome [9], gastric cancer [10], hepatoblastoma [11], gardner syndrome [12] and familial adenomatous polyposis (FAP) [13].

* Corresponding authors.
E-mail addresses: guantlv@126.com (G. Lu), cuimanhua@126.com (M. Cui).
[1] These authors contributed equally to this project.

http://dx.doi.org/10.1016/j.bbrc.2014.04.014
0006-291X/

In a 4 generation Chinese family with inherited colon cancer, 6 affected members died of colon cancer and 1 of gastric cancer (also affected by systemic lupus erythematosus). According the pedigree structure, it was an autosomal dominant inheritance model. In order to locate the causal variants, all of the 5 genes related to HNPCC1 and FAP were selected for targeted multigene sequencing by Ion Torrent Personal Genome Machine (PGM, Life Technologies) loaded with a 314 chip. 2 rare variants (c.1690A>G in MSH2 and c.694C>T in APC) were detected in *MSH2* and *APC*, respectively. After systematic analysis, the c.694C>T in *APC* caused the codon 232 for Arginine (**C**GA) to a premature stop codon (**T**GA) making mRNAs with this stop codon targets of nonsense-mediated mRNA decay (NMD) [14,15]. Besides, the mutation also was co-segregated with the affected members in the family and not detected in the 351 normal Chinese peoples. After searching the ClinVar database [16], this mutation was shown to be a *de novo* germline variant.

2. Materials and methods

2.1. Patients information

A pregnant woman consulted with the genetic analyst in the Department of Gynaecology and Obstetrics, The 306th Hospital of People's Liberation Army for inheritable disease in her family. According to her narration, her father (II-3) died of colon cancer at 49. Two aunts also died of cancer, one of colon cancer at 52 (II-5) and one of gastric cancer at 56 (II-7) and also obsessed with systemic lupus erythematosus (SLE). The consultant really wanted to know whether she was prone to this inheritable disease. After taking clinical check, 3 members in the third generation were diagnosed to have polyposis in the colons and were recruited for this project (Fig 1 and Table 1). III-7 and III-11 were found to have numerous polyposis (more than 50) in colons at 16 and 19 years old, respectively. III-3 was with more than 150 polyposis in the premalignant condition at 32 years old.

As shown in Fig. 1, the disease in this pedigree was consistent with an autosomal dominant inheritance model. As the information provided, the disease had the high probability as familial adenomatous polyposis (FAP). In order to detect the putative causal mutations for this disease, 5 genes with inheritable colon cancer (*MLH1*, *MSH2*, *MSH6*, *PMS2* and *APC*) were selected for targeted sequencing. All participants were recruited after obtaining informed consent.

2.2. Sequencing and data analysis of the amplicons

The peripheral blood samples were stored in EDTA anticoagulant vacuum tubes and genomic DNA was extracted using QIAamp DNA Blood Midi Kit (QIAGEN, Germany). Multiplex primers for the coding exons of the *MLH1* (NM_000249), *MSH2* (NM_000251), *MSH6* (NM_000179), *PMS2* (NM_000535) and *APC* (NM_000038) were designed by ultrahigh-multiplex primer designing tool, Ion Ampliseq Designer (https://www.ampliseq.com/browse.action), exon padding setted as 5 with hg19 as the reference human genome. The sequence of primers were compiled in the Supplementary file. PCR amplifications for the targeted exons were performed manually using Phusion high-fidelity DNA polymerase (NEB, UK) according to the manufacturer's instructions. The amplicons were sequenced on the Ion Torrent Personal Genome Machine (PGM, Life Technologies) loaded with a 314 chip (314, 100 Mb output) following the recommended protocol. The variant calling was performed using the Torrent Suite Software v.3.2 by aligned to hu19 (Life Technologies) under the standard settings. Filtered variants were annotated using the Ion Reporter software v1.2 (Life Technologies). The frequencies of the detected variants were compared with SNPs collected in dbSNP database [17], HapMap project [18], 1000 genome project [19] and 351 whole-genome sequenced

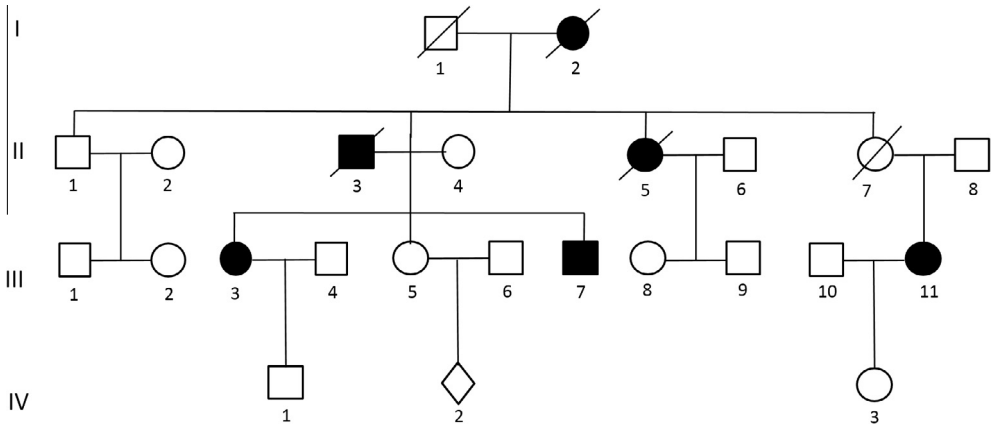

Fig. 1. Pedigree structure of the Han family with familial inheritable colon cancer. Squares and circles denoted males and females respectively with filled ones as affected members. Diamond represented the individual with unknown sex. Individuals labeled with a solidus were deceased.

Table 1
Informaiton of patients in this pedigree.

Patients	Gender	Age at diagnosis	Live/Dead	Disease	Status of polyposis
I-2	Female	N.A.	Dead	N.A.	N.A.
II-3	Male	N.A.	Dead at 49	Colon cancer	N.A.
II-5	Female	N.A.	Dead at 52	Colon cancer	N.A.
II-7	Female	N.A.	Dead at 56	Gastric cancer	N.A.
III-3	Female	32	Live	Polyposis	>150 polyposis/premalignant
III-7	Male	16	Live	Polyposis	>50 polyposis/noncancerous
III-9	Female	19	Live	Polyposis	>50 polyposis/noncancerous

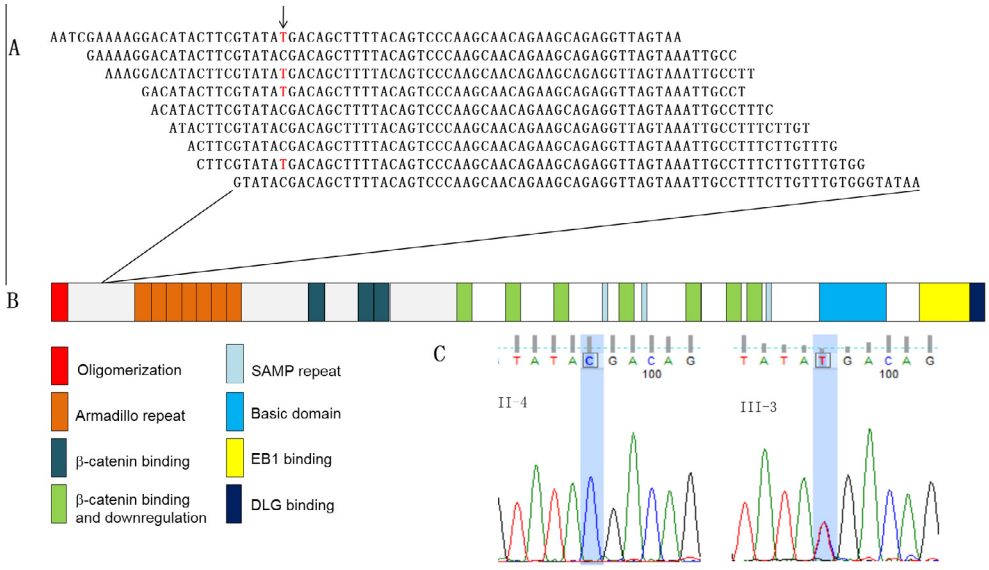

Fig. 2. Detection and verification of the c.694C>T in APC. (A) Reads aligned with the APC around c.694C>T, red represented the mutated variant. (B) Diagram of the protein structure of APC. (C) Sanger sequencing of the sequence containing the mutation in II-4 and III-3 members. The light green indicated of the mutated nucleotide.

normal Han Chinese. The effects of the variants on protein function were predicted by PolyPhen-2 [20].

2.3. Validation of the putative causal variants

To validate the mutations detected by deep sequencing, the genomic DNAs containing the putative variants were amplified and sequenced by traditional Sanger sequencing and analyzed by the Lasergene package, SeqMan Pro v7.1.0. The sequence of primers for Sanger sequencing were provided in the Supplementary File.

3. Results

3.1. Variants detected by Ion Torrent

After alignment with hg19 by Torrent Suite Software and annotated by Ion Reporter software, 14 variants were detected in the sequenced sample and 13 been collected in dbSNP database and assigned with a rs identification number. (Table 2). In these variants, 9 were synonymous, 4 missense and 1 non-sense. Except two variants, the frequencies of variant alleles of the other 12 were greater than at least 10%, indicative of being polymorphisms. The remaining two variants (c.694C>T in *APC* and c.1690A>G in *MSH2*) were with either no information of frequency or being a rare SNP in Chinese (MAF = 0.4%).

3.2. Verification of the detected rare variants

The genomic region containing the two rare variants were amplified and sequenced by Sanger sequencing in all of the members in the family. c.1690A>G was detected in only 2 affected members (III-7 and III-11) and also in normal members (II-1, III-5 and III-9). The c.694C>T was only detected in all of 3 affected members (Fig. 2).

4. Discussion

Colorectal cancer was the third most commonly diagnosed cancer in the world and occurred due to lifestyle and increasing age, with only a minority of cases associated with underlying genetic disorders. Due to the existence of polyposis in colons and the autosomal dominant inheritance mode in this Han Chinese family, the disease might be classified as familial adenomatous polyposis (FAP). It had been reported that FAP accounted for a small portion of CRCs (about 1%) and mainly caused by the mutation in APC gene.

Table 2
Deteced variants and frequencies in different databases.

Genes	Acc. No.	Nucleotiede variants	Amino acid variants	rs ID	Frequencies of detected variants			
					dbSNP	HapMap	1000 Genomes	Sequenced normal Chinese
PMS2	NM_000535	c.1621G>A	p.Lys541Glu	rs2228006	0.879	0.869	0.783	0.087
		c.1408C>T	p.Pro470Ser	rs1805321	0.107	0	0.168	0.178
		c.780C>A	p.Ser260Ser	rs1805319	0.594	N.A.	0.723	0.875
MSH6	NM_000179	c.116G>A	p.Gly39Glu	rs1042821	0.307	N.A.	0.315	0.359
		c.3306T>A	p.Thr1102Thr	rs2020910	0.053	0.255	0.089	0.303
MSH2	NM_000251	c.1690A>G	p.Thr564Ala	rs55778204	N.A.	N.A.	0	0.004
APC	NM_000038	c.694C>T	p.Arg232Ter	Novel	N.A.	N.A.	0	0
		c.1458T>C	p.Tyr486Tyr	rs2229992	0.467	0.438	0.429	0.425
		c.1635G>A	p.ALa545Ala	rs351771	0.48	0.326	0.416	0.301
		c.4479G>A	p.Thr1493Thr	rs41115	0.339	0.328	0.414	0.303
		c.5034G>A	p.Gly1678Gly	rs42427	0.474	0.313	0.415	0.302
		c.5268T>G	p.Ser1756Ser	rs866006	0.478	0.338	0.414	0.3
		c.5465T>A	p.Val1822Asp	rs459552	0.743	0.825	0.758	0.801
		c.5880G>A	p.Pro1960Pro	rs465899	0.437	0.328	0.416	0.302

Some FAP was caused by mutations in *MUTYH* gene at an autosomal recessive model [21]. Inferred from the pedigree struture of this family, it might be caused by mutations in APC gene. But genes in the MMR family (*MLH1*, *MSH2*, *MSH6* and *PMS2*) also could produce similar symptoms with this family.

In order to detect the causal variants efficiently, a ultrahigh multiplex amplification of the genes with colon cancers coupled with second-generation sequencing was performed for one of the affected members (III-7). After analysis, 14 variants were detected in the sequenced sample. 13 of the them had been collected in the dbSNP database and assigned with a rs identification number (Table 2). As for the 14 variants, 9 were synonymous, 4 missense and 1 non-sense. 12 were polymorphic with the frequencies of variant alleles greater than 10%. As for the 2 remaining variants, c.694C>T in APC and c.1690A>G in MSH2 were either lack of information of frequency in any databases mentioned above or being a rare SNP in Chinese (MAF = 0.4%). c.694C>T in APC caused the codon at position 232 for Arginine (**C**GA) to a premature translational stop codon (**T**GA). Transcripts containing this premature stop codon would be degraded through nonsense-mediated mRNA decay (NMD) [14,15]. The c.1690A>G in MSH2 changed the codon at 564 for Threonine (**A**CC) to Alanine (**G**CC).

According to the frequencies in public databases, these 2 variants might be putative causal mutations in this family. They were sequenced in the whole family members to check the cosegregation in affected members. c.694C>T in APC was only detected in affected individuals. The c.1690A>G in MSH2 was not only detected in 2 patients, but also in a few normal individuals in this family. But it could not ruled out the possibility of the two variants as causal mutations for this family. The 2 rare variants were also analyzed in the 351 whole-genome sequenced normal Han Chinese. The c.694C>T did not detected in any samples and c.1690A>G in only 2 individuals heterozygously (0.4%). Since the frequency of c.1690A>G was too low, the possibility as the causal factor could still not be completely rejected. So the impact of this variant on the function of MSH2 was predicted by PolyPhen-2. Unexpectedly, the variant was predicted to be BENIGN with a score of 0.000. This represented that c.1690A>G had no functional affect at all, since amino acid at this position varied greatly among different species including rat, mouse, rabbit, turkey, African malaria mosquito, Xenopus tropicalis and Florida lancelet. This indicated that the c.1690A>G possessed the slightest possibility as the causal variants.

As for the *de novo* germline mutation, c.694C>T in APC caused a premature stop codon and made the transcripts produced by this allele as targets for NMD. It only found in the 3 affected members in the family. Besides, the variant was not detected in any of the public databases and in the 351 normal Han Chinese. This showed that c.694C>T in APC was the causal factor for this FAP family.

5. Conclusion

Using the next generation sequencing for targeted regions of 4 gens related to colon cancer, a *de novo* germline mutation in the tumor suppressor gene, APC was discovered to be the causal mutation for the FAP family. So, next generation high-throughput sequencing might be an effective method to discover mutations for disease-affected pedigrees.

6. Competing interests

The authors declare no competing interests with respect to the present article.

7. Authors' contributions

G. Lu analyzed the data and wrote the manuscript. Y. Zhang and M. Cui provided the samples. Q. Hu, X. Wang and Y. Mao performed the genomics DNA extraction and sequencing.

Acknowledgments

We thank members of the family for their participation in this project. This work was supported by the National Natural Science Foundation of China (81101490 to Guanting Lu).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bbrc.2014.04.014.

References

[1] N. Cancer Genome Atlas, Comprehensive molecular characterization of human colon and rectal cancer, Nature 487 (2012) 330–337.
[2] I. Soerjomataram, J. Lortet-Tieulent, D.M. Parkin, J. Ferlay, C. Mathers, D. Forman, F. Bray, Global burden of cancer in 2008: a systematic analysis of disability-adjusted life-years in 12 world regions, Lancet 380 (2012) 1840–1850.
[3] D. Cunningham, W. Atkin, H.J. Lenz, H.T. Lynch, B. Minsky, B. Nordlinger, N. Starling, Colorectal cancer, Lancet 375 (2010) 1030–1047.
[4] E. Juhn, A. Khachemoune, Gardner syndrome: skin manifestations, differential diagnosis and management, Am. J. Clin. Dermatol. 11 (2010) 117–122.
[5] E. Half, D. Bercovich, P. Rozen, Familial adenomatous polyposis, Orphanet J. Rare Dis. 4 (2009) 22.
[6] G.T. Torrezan, F.C. da Silva, E.M. Santos, A.C. Krepischi, M.I. Achatz, S. Aguiar Jr., B.M. Rossi, D.M. Carraro, Mutational spectrum of the APC and MUTYH genes and genotype-phenotype correlations in Brazilian FAP, AFAP, and MAP patients, Orphanet J. Rare Dis. 8 (2013) 54.
[7] J. Auclair, D. Leroux, F. Desseigne, C. Lasset, J.C. Saurin, M.O. Joly, S. Pinson, X.L. Xu, G. Montmain, E. Ruano, C. Navarro, A. Puisieux, Q. Wang, Novel biallelic mutations in MSH6 and PMS2 genes: gene conversion as a likely cause of PMS2 gene inactivation, Hum. Mutat. 28 (2007) 1084–1090.
[8] J.W. Poley, A. Wagner, M.M. Hoogmans, F.H. Menko, C. Tops, J.M. Kros, R.E. Reddingius, H. Meijers-Heijboer, E.J. Kuipers, W.N. Dinjens, Rotterdam initiative on gastrointestinal hereditary, biallelic germline mutations of mismatch-repair genes: a possible cause for multiple pediatric malignancies, Cancer 109 (2007) 2349–2356.
[9] S.R. Hamilton, B. Liu, R.E. Parsons, N. Papadopoulos, J. Jen, S.M. Powell, A.J. Krush, T. Berk, Z. Cohen, B. Tetu, et al., The molecular basis of Turcot's syndrome, N. Engl. J. Med. 332 (1995) 839–847.
[10] Y.L. Wallis, D.G. Morton, C.M. McKeown, F. Macdonald, Molecular analysis of the APC gene in 205 families: extended genotype-phenotype correlations in FAP and evidence for the role of APC amino acid changes in colorectal cancer predisposition, J. Med. Genet. 36 (1999) 14–20.
[11] H. Oda, Y. Imai, Y. Nakatsuru, J. Hata, T. Ishikawa, Somatic mutations of the APC gene in sporadic hepatoblastomas, Cancer Res. 56 (1996) 3320–3323.
[12] L.K. Su, E.K. Abdalla, C.H. Law, W. Kohlmann, A. Rashid, J.N. Vauthey, Biallelic inactivation of the APC gene is associated with hepatocellular carcinoma in familial adenomatous polyposis coli, Cancer 92 (2001) 332–339.
[13] S. Aretz, D. Stienen, S. Uhlhaas, C. Pagenstecher, E. Mangold, R. Caspari, P. Propping, W. Friedl, Large submicroscopic genomic APC deletions are a common cause of typical familial adenomatous polyposis, J. Med. Genet. 42 (2005) 185–192.
[14] A.J. Ward, M. Norrbom, S. Chun, C.F. Bennett, F. Rigo, Nonsense-mediated decay as a terminating mechanism for antisense oligonucleotides, Nucleic Acids Res. (2014) [Epub ahead of print].
[15] Y.F. Chang, J.S. Imam, M.F. Wilkinson, The nonsense-mediated decay RNA surveillance pathway, Annu. Rev. Biochem. 76 (2007) 51–74.
[16] M.J. Landrum, J.M. Lee, G.R. Riley, W. Jang, W.S. Rubinstein, D.M. Church, D.R. Maglott, ClinVar: public archive of relationships among sequence variation and human phenotype, Nucleic Acids Res. 42 (2014) D980–D985.
[17] E.M. Smigielski, K. Sirotkin, M. Ward, S.T. Sherry, DbSNP: a database of single nucleotide polymorphisms, Nucleic Acids Res. 28 (2000) 352–355.
[18] C. International HapMap, K.A. Frazer, D.G. Ballinger, D.R. Cox, D.A. Hinds, L.L. Stuve, R.A. Gibbs, J.W. Belmont, A. Boudreau, P. Hardenbol, S.M. Leal, S. Pasternak, D.A. Wheeler, T.D. Willis, F. Yu, H. Yang, C. Zeng, Y. Gao, H. Hu, W. Hu, C. Li, W. Lin, S. Liu, H. Pan, X. Tang, J. Wang, W. Wang, J. Yu, B. Zhang, Q. Zhang, H. Zhao, H. Zhao, J. Zhou, S.B. Gabriel, R. Barry, B. Blumenstiel, A. Camargo, M. Defelice, M. Faggart, M. Goyette, S. Gupta, J. Moore, H. Nguyen, R.C. Onofrio, M. Parkin, J. Roy, E. Stahl, E. Winchester, L. Ziaugra, D. Altshuler, Y. Shen, Z. Yao, W. Huang, X. Chu, Y. He, L. Jin, Y. Liu, Y. Shen, W. Sun, H. Wang, Y. Wang, Y. Wang, X. Xiong, L. Xu, M.M. Waye, S.K. Tsui, H. Xue, J.T. Wong, L.M.

Galver, J.B. Fan, K. Gunderson, S.S. Murray, A.R. Oliphant, M.S. Chee, A. Montpetit, F. Chagnon, V. Ferretti, M. Leboeuf, J.F. Olivier, M.S. Phillips, S. Roumy, C. Sallee, A. Verner, T.J. Hudson, P.Y. Kwok, D. Cai, D.C. Koboldt, R.D. Miller, L. Pawlikowska, P. Taillon-Miller, M. Xiao, L.C. Tsui, W. Mak, Y.Q. Song, P.K. Tam, Y. Nakamura, T. Kawaguchi, T. Kitamoto, T. Morizono, A. Nagashima, et al., A second generation human haplotype map of over 3.1 million SNPs, Nature 449 (2007) 851–861.

[19] C. Genomes Project, G.R. Abecasis, A. Auton, L.D. Brooks, M.A. DePristo, R.M. Durbin, R.E. Handsaker, H.M. Kang, G.T. Marth, G.A. McVean, An integrated map of genetic variation from 1,092 human genomes, Nature 491 (2012) 56–65.

[20] I.A. Adzhubei, S. Schmidt, L. Peshkin, V.E. Ramensky, A. Gerasimova, P. Bork, A.S. Kondrashov, S.R. Sunyaev, A method and server for predicting damaging missense mutations, Nat. Methods 7 (2010) 248–249.

[21] S. Jones, P. Emmerson, J. Maynard, J.M. Best, S. Jordan, G.T. Williams, J.R. Sampson, J.P. Cheadle, Biallelic germline mutations in MYH predispose to multiple colorectal adenoma and somatic G: C T: a mutations, Hum. Mol. Genet. 11 (2002) 2961–2967.