Comprehensive analysis of the genetic mutation spectrum of Koreans being evaluated for familial adenomatous polyposis (FAP)

Seo-Jin Park

Department of Medicine

The Graduate School, Yonsei University

Comprehensive analysis of the genetic mutation spectrum of Koreans being evaluated for familial adenomatous polyposis (FAP)

Seo-Jin Park

Department of Medicine

The Graduate School, Yonsei University

Comprehensive analysis of the genetic mutation spectrum of Koreans being evaluated for familial adenomatous polyposis (FAP)

Directed by Professor Kyung-A Lee

The Doctoral Dissertation
submitted to the Department of Medicine,
the Graduate School of Yonsei University
in partial fulfillment of the requirements for the degree
of Doctor of Philosophy

Seo-Jin Park

December 2020

This certifies that the Doctoral Dissertation of Seo-Jin Park is approved.

Thesis Supervisor : Kyung-A Lee

Thesis Committee Member : Jong Rak Choi

Thesis Committee Member : Chang Seok Ki

Thesis Committee Member : Jae Jun Park

Thesis Committee Member : Sangwoo Kim

The Graduate School
Yonsei University

December 2020

ACKNOWLEDGEMENTS

First and foremost, I would like to express my immense gratitude to my thesis advisor, Dr. Kyung-A Lee, who patiently guided me throughout my long and extended dissertation process with unfaltering support. She is an invaluable teacher and a true mentor. I am also grateful to my thesis chairman, Dr. Jong Rak Choi for his encouragement and helpful feedback as he guided me through my research. I would also like thank my thesis committee members, Dr. Chang Seok Ki, Dr. Jae Jun Park, and Dr. Sangwoo Kim who offered insightful comments and sound advice on my research which enabled me to make important revisions to improve and refine my thesis.

In addition, I would like to give special thanks to Dr. Sung-Hee Han, Dr. Saeam Shin, Dr. Dongju Won, and Dr. Seung-Tae Lee for their help and support at various stages of my research. I am extremely lucky to have such great colleagues who extended their expertise and help.

Last but definitely not the least, I would like to express my deepest appreciation to my family for their unconditional love and support: my parents (who have been an unlimited source of love and encouragement throughout my whole life) and my older sister (who has been my go-to person in any crisis and her family who have cheered me on from afar) and my younger brother, my incredibly generous and graceful parents-in-law, and finally my two adorable children (Y and J) and my dearest husband. You have given me the strength, motivation, and inspiration to finally finish my thesis and strive to become a better version of myself everyday.

Thank you all for your support.

<TABLE OF CONTENTS>

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Comprehensive analysis of the genetic mutation spectrum of Koreans being evaluated for familial adenomatous polyposis (FAP)

Seo-Jin Park

Department of Medicine
The Graduate School, Yonsei University

(Directed by Professor Kyung-A Lee)

Familial adenomatous polyposis (FAP) is a cancer predisposition syndrome inherited in an autosomal dominant pattern. The presence of hundreds to thousands of colorectal adenomas in the colon is a characteristic manifestation of FAP, with various extracolonic manifestations. Mutations in the adenomatous polyposis coli (*APC*) gene are known to be associated with FAP. The majority of the *APC* gene mutations are frameshift or nonsense mutations leading to a truncation of the *APC* protein. As a known tumor suppressor gene, a deleterious gene produced by inactivating germline mutations in the *APC* gene is unable to suppress cellular growth and lead to formation of adenomas, which can subsequently show malignant potential. Characterization and detailed analysis of the accumulated genetic data has become a valuable asset with an increased interest in genetic screening along with the widespread recognition of precision medicine. Therefore, the aim of the study was to evaluate the mutation spectrum of the *APC* gene in the Korean population through analysis of the *APC* gene mutation tests in the most recent 15-year period. A total of 420 patients were referred for

APC gene mutation tests between May 2006 and June 2020. Clinical data including clinical diagnosis, family history, colonoscopic findings, and pathologic findings were reviewed. Genomic DNA was extracted from the peripheral blood of patients suspected of having FAP and were tested with either conventional Sanger sequencing or a NGS panel test. *APC* gene mutations were found in 167 out of 420 patients tested (39.8%). In addition to the well-known frameshift and nonsense mutations, there were 3 cases of whole gene deletions and 3 cases of single or multiple exon deletions. Novel *APC* gene mutations were detected in 17 patients diagnosed with FAP. VUS was detected in 19 patients and 1 case was reclassified as a likely pathogenic variant and 6 cases were reclassified as a likely benign variant, according to the ACMG criteria. NGS panel was able to detect more *APC* gene mutations than the conventional direct sequencing method, in which the discrepancy was mostly associated with large deletions and somatic mosaicism which are due to limitations of the conventional sequencing method. Many of the patients with negative *APC* gene mutation study results who were tested before the introduction of the NGS may consider getting tested again to confirm a clinical diagnosis of FAP. Likely pathogenic mutations in *BMPR1A*, *MUTYH*, *PMS1*, and *POLE* genes were detected in 8 patients without an *APC* gene mutation with NGS testing. VUS was detected in a heterogeneous group of genes in 35 patients without *APC* gene mutations and suspicious of FAP. A frameshift variant of *MSH6* gene (c.4068_4071dupGATT, p.Lys1358AspfsTer2) was detected in 7 patients with few to multiple adenomas, but showed conflicting evidence for classification. A VUS in *ALK* gene detected in one patient was reclassified as likely benign. Genetic testing has become an essential component in the diagnosis and management of FAP with continuous accumulation of genetic mutations in the database. Detection of previously undetectable mutations should

help in the diagnosis of FAP patients without a confirmed genetic cause with the introduction of NGS panels. In conclusion, the use of NGS panels requires a robust bioinformatics algorithm for the interpretation of genetic variants, as well as periodic review of new clinical evidence and revised recommendations for accurate ACMG classification, which will be an important process in diagnosis of hereditary cancer predisposition syndromes such as FAP.

Key words: *APC* gene, familial adenomatous polyposis (FAP), gene mutation test, next generation sequencing

Comprehensive analysis of the genetic mutation spectrum of Koreans being evaluated for familial adenomatous polyposis (FAP)

Seo-Jin Park

Department of Medicine
The Graduate School, Yonsei University

(Directed by Professor Kyung-A Lee)

I. INTRODUCTION

Familial adenomatous polyposis (FAP; MIM#175100) is a cancer predisposition syndrome inherited in an autosomal dominant pattern. The presence of hundreds to thousands of colorectal adenomas in the colon is a characteristic manifestation of FAP, with various extracolonic manifestations.[1-3] Mutations in the adenomatous polyposis coli (*APC*) gene are known to be associated with FAP. The *APC* gene is located on the long arm of chromosome 5 (5q22.2), and is composed of 15 exons with 8,532 base pairs. Exon 15 is the largest and comprises about 77% of the *APC* coding sequence, and hence a bulk of the mutations are localized within this exon. The majority of the *APC* gene mutations are frameshift or nonsense mutations leading to a truncation of the *APC* protein.[2-3] The Human Gene Mutation Database (HGMD, www.hgmd.cf.ac.uk/ac) has accumulated more than 2000 different mutations in the *APC* gene to date, which has doubled over the past decade. As a known tumor suppressor gene, a deleterious gene produced by inactivating germline mutations

in the *APC* gene is unable to suppress cellular growth and lead to formation of adenomas, which can subsequently show malignant potential.

Genetic testing is an essential component in the diagnosis and management of FAP. Previous studies relevant to *APC* gene mutations have shown a mutation positive rate ranging from 60 to 90% (Table 1).[4-9] Genetic studies of FAP patients are somewhat limited due to the low incidence rates and difficulties in gathering enough patient samples for a large-scale, comprehensive study. Characterization and detailed analysis of the accumulated genetic data has become a valuable asset with an increased interest in genetic screening along with the widespread recognition of precision medicine as well as the increasing use of next generation sequencing (NGS) in laboratories. In addition, the American College of Medical Genetics and Genomics (ACMG), Association for Molecular Pathology (AMP), and College of American Pathologists (CAP) worked together to meet the challenges of interpretation of sequence variations by publishing a guideline using various evidence-based criteria in 2015.[10] This recommendation has become an essential and important part of analysis of genetic testing due to its incorporation of population statistics, functional studies, and in silico analysis. Therefore, the aim of the study was to evaluate the mutation spectrum of the *APC* gene, reviewing of the classification of various *APC* sequence variations with the ACMG criteria, and further analysis of sequence variants identified in NGS studies in the Korean population with sequencing data from the most recent 15-year period.

Table 1. *APC* gene mutation positive rates in various countries detected by direct sequencing and MLPA when available

Country	Patients (n)	*APC* gene mutation (positive rate)	Reference
China	14	Sequencing: 9/14 (64.3%) MLPA: 2/14 (14.3%)	Sheng JQ et al. (2010)[4]
Korea	83	Sequencing: 59/83 (71.1%)	Kim DW et al. (2005)[5]
Singapore	53	Sequencing: 46/53 (86.8%) MLPA: 3/53 (5.7%)	Cao X et al. (2006)[6]
Sweden	96	Sequencing: 81/96 (84.4%)	Kanter-Smoler G et al. (2008)[7]
Taiwan	47	Sequencing: 31/47 (66.0%) MLPA: 5/47 (10.6%)	Chiang JM et al. (2010)[8]
United States	1591	Sequencing: 431/1591 (27.1%) Mayo clinic patients (n=31) with classic FAP: 27/31 (87.1%)	Kerr SE et al. (2013)[9]

* Abbreviations: *APC*, adenomatous polyposis coli; MLPA, multiplex ligation-dependent probe amplification

II. MATERIALS AND METHODS

1. Study population

A total of 420 patients were referred for *APC* gene mutation tests between May 2006 and June 2020. Non-Korean patients were excluded from the study population. Written informed consent for genetic testing was obtained from the patients according to the ethical guidance of the institutional guidelines. Among the tested samples, 63.5% were referred to a large commercial laboratory from diverse locations throughout Korea and 55% of the samples were from a single institution healthcare system (Figure 1). Retrospective chart

review was done on the 231 patients that visited two tertiary hospitals within the same institution health care system. Institutional review board approval was obtained for review of clinical data including clinical diagnosis, family history, colonoscopy findings, and pathologic findings.

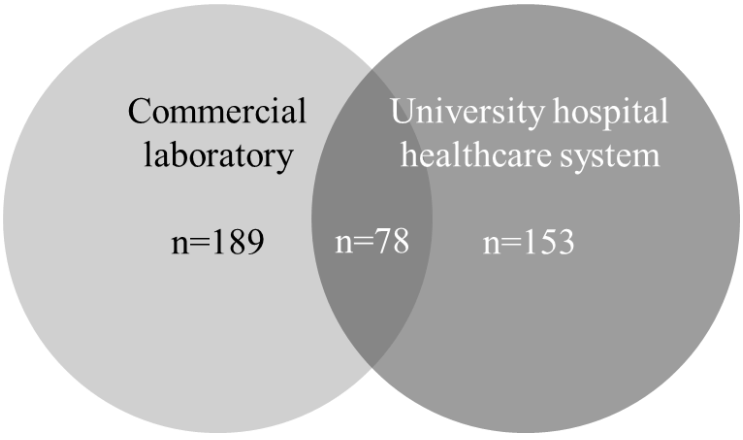

Figure 1. Distribution of patients referred for *APC* gene mutation tests between May 2006 and June 2020 (n=420).

2. Direct sequencing

Genomic DNA was extracted from the peripheral blood of patients suspected of having FAP. Prior to 2010, *APC* gene testing was done at a commercial laboratory with DHPLC screening method using a WAVE Maker System (Transgenomic Inc., San Jose, CA). WAVE MAKER Software v4.1 (Transgenomics Inc.) was used to predict the elution gradient and temperatures, with experimental results determined the precise gradients and optimal temperatures of each fragment. Confirmation with direct sequencing

was done when the PCR product was suggestive of heterozygosity.[11] Direct sequencing of the entire coding region of the *APC* gene including all the intron-exon boundaries were done starting from 2010. Sequencing was carried out on the ABI 3100 DNA sequencer (PE Applied Biosystems, Foster City, CA, USA) using the Applied Biosystems (ABI) Prism Big-Dye Terminator Cycle Sequencing kit with amplification conditions described for the HPLC method. Variants were annotated according to the Human Genome Variation Society nomenclature system (HGVS, available at http://www.hgvs.org/mutnomen).

APC gene direct sequencing at the single institution healthcare system was done genomic DNA was extraction from EDTA anti-coagulated whole blood samples using the QIAsymphony DSP DNA Mini Kit (QIAGEN, Hilden, Germany) automated protocol for QIAsymphony SP (QIAGEN). The concentration and quality of genomic DNA was evaluated by Nanodrop (ND-1000, Thermo Scientific, Wilmington, DE, USA). Primers were designed to amplify all coding exons and flanking introns of *APC* using the Primer3 software.[12] PCR was performed on 100 ng of genomic DNA using an AccuPower^{TM} HotStart PCR PreMix (Bioneer, Daejeon, Korea) under the following amplification conditions: 94°C for 5 min followed by 35 cycles of 94°C for 30 sec, 60°C for 30 sec and 72°C for 60 sec, and final extension at 72°C for 7 min. The PCR products were treated with Exo-AP PCR Clean-up Mix (MGmed, Inc., Seoul, Korea) and then sequenced in both directions on the Applied Biosystems 3730 Genetic Analyzer (Thermo Fisher Scientific, Foster City, CA, USA) and the BigDye Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific). The raw sequence data were analyzed using

Sequencher 5.3 software (Gene Codes, Ann Arbor, MI, USA). The results were aligned against reference sequence (NM_000038.5). Identified variants were annotated according to nomenclature recommendations of the Human Genome Variation Society (HGVS, available at http://www.hgvs.org/mutnomen).

3. Targeted next generation sequencing (NGS) gene panel and data analysis

For the customized NGS panel, we selected 60 genes related to hereditary cancer syndromes (*APC, ATM, AXIN1, AXIN2, BARD1, BLM, BMPR1A, BRCA1, BRCA2, BRIP1, CDH1, CDK4, CDKN2A, CHEK2, EPCAM, FANCM, FLCN, GALNT12, GREM1, MEN1, MLH1, MLH3, MRE11, MSH2, MSH6, MUTYH, NBN, NF1, NF2, NTHL1, PALB2, PMS1, PMS2, POLD1, POLE, PPM1D, PRSS1, PTEN, RAD50, RAD51, RAD51C, RAD51D, RB1, RET, RNF43, SCG5, SDHA, SDHAF2, SDHB, SDHC, SDHD, SLX4, SMAD4, STK11, TP53, TSC1, TSC2, VHL, WT1, XRCC2*). Custom probe capture panel was designed targeting all coding exons and flanking introns of target genes.

Intact dsDNA was quantified and adjusted to a concentration of 5 ng/μL using Qubit dsDNA BR Assay Kit on a Qubit 2.0 Fluorometer (Thermo Fisher Scientific). Genomic DNA was fragmented with the Bioruptor Pico Sonication System (Diagenode, Liège, Belgium) in 200 and 250 bp fragments. The size and concentration of sheared DNA were analyzed by TapeStation 4200 (Agilent Technologies, Santa Clara, CA, USA). Fragmented DNA were then purified using Agencourt Ampure XP beads (Beckman Coulter, Brea, CA, USA). The DNA fragments were end-repaired, phosphorylated, and adenylated on the 3′ ends. The index adaptors were ligated to the repaired ends,

DNA fragments were amplified, and fragments of 200−500 bp were isolated. Sequencing libraries were then hybridized with the custom target capture probe. Sequencing was done with the NextSeq 550Dx instrument (Illumina, San Diego, CA, USA) and the NextSeq 550 High-Output v2 Kit (300 cycles).

Data analysis was performed primarily through our custom pipeline. Raw sequence data were mapped to human genomic reference sequence (hg19) using the Burrows-Wheeler Aligner algorithm, followed by removal of duplicate reads, realignment of insertions and deletions, base quality recalibration, and variant calling using the Genome Analysis Toolkit (GATK). Single-nucleotide variants and small insertions or deletions were called and crosschecked using GATK Haplotypecaller and VarScan. Detected variants were further examined by visual verification using the Integrative Genomics Viewer (IGV) software (Broad Institute, Cambridge, MA, USA). ExomeDepth in the R package was used to detect exon-level copy number variants in target regions,[13] followed by visualization using a base-level read depth normalization algorithm implemented in the DxSeq Analyzer (Dxome, Seoul, Korea). Identified variants were classified into five categories - "pathogenic," "likely pathogenic," "uncertain significance," "likely benign," and "benign"- according to the guideline from American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG/AMP).[10] The following databases were used for variant annotation: dbSNP (http://www.ncbi.nlm.nih.gov/snp), Human Gene Mutation Database (HGMD, http://www.hgmd.org), ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/), 1000 Genome (http://browser.1000genomes.org), the Exome Aggregation

Consortium (ExAC, http://exac.broadinstitute.org), and the Korean Reference Genome Database (KRGDB, http://coda.nih.go.kr/coda/KRGDB/index.jsp). The pathogenicity of missense variants was predicted using five in silico prediction algorithms, including Sorting Tolerant from Intolerant (SIFT), Polymorphism Phenotyping v2 (PolyPhen-2), MutationTaster, MutationAssessor, and Functional Analysis through Hidden Markov Models (FATHMM) implemented in dbNSFP version 3.0a.

4. Multiplex ligation-dependent probe amplification (MLPA)

To confirm copy number variants (CNVs) of *APC* gene identified by next-generation sequencing, MLPA was done with P043 *APC* probemix (MRC Holland, Amsterdam, Netherlands). MLPA expermients were conducted by following the manufacturer's instructions. First, the 5 μL of DNA denatured at 98 °C for 5 min and cooled down to 25°C. 1.5 μL of probemix and 1.5 μL of MLPA buffer were added to each sample, heat-denatured for 1 min at 95 °C, followed by hybridization for 16 h at 60 °C. The ligation reaction with ligase-65 enzyme was performed at 54 °C for 15 min, followed by 5 min at 98 °C for heat inactivation of the enzyme. PCR amplification was carried out in a GeneAmp PCR System (Thermo Fisher Scientific). The amplicons were then analyzed by capillary electrophoresis by the Applied Biosystems 3730 Genetic Analyzer (Thermo Fisher Scientific) for fragment length determination. Raw data was analyzed using the GeneMarker software (SoftGenetics, State College, PA, USA). Each specific probe's peak height was normalized by dividing it with the combined heights of the control probes. The relative peak

height of each probe was compared with the same probe's relative peak height in the control samples. Peak ratios <0.75 were considered to be deletions and peak ratios >1.30 were considered to be duplications.

5. Mutant enrichment with 3'-modified oligonucleotide (MEMO)-PCR

For confirmation of low-level mutants identified from target panel sequencing, MEMO-PCR and sequencing analysis was performed using the previously described method.[14] PCR amplification was performed using two generic primers and one blocking primer designed to encompass the target mutation site and to overlap with one of the generic primers. The 3' ends of the blocking primers were modified by the addition of a C3 spacer. The PCR reaction was performed using the AccuPower HotStart PCR PreMix (Bioneer, Daejeon, Korea). The reaction mixture included 200 ng of DNA, 10 pmol of each generic primer, and 50 pmol of the blocking primer. The PCR was performed using a GeneAmp PCR System (Thermo Fisher Scientific), and the cycling conditions were as follows: 94°C for 5 minutes, 50 cycles of the main reaction (94°C for 30 sec, 59°C for 30 sec, and 72°C for 60 sec), and 72°C for 7 minutes. After the amplification reaction, the amplicons were purified with Exo-AP PCR Clean-up Mix (MGmed, Inc.). Cycle sequencing was performed on the Applied Biosystems 3730 Genetic Analyzer (Thermo Fisher Scientific) with the BigDye™ Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific). The obtained sequence results were analyzed with the Sequencher 5.3 software (Gene Codes, Ann Arbor, MI).

6. Data analysis

Further data analysis to determine the pathogenicity of the genetic variants were done by re-examining the American College of Medical Genetics and Genomics (ACMG) guidelines established in 2015, variant databases such as HGMD , population databases, as well as various in silico analysis software. The ExAC exclusively contains exome data and is known as the first release of the Genome Aggregation Database (gnomAD, https://gnomad.broadinstitute.org/). The subsequent gnomAD was released to integrate previous sequencing projects with exome and genome sequencing data. The secondary analysis was done with gnomAD v2.1.1 and updated the data that used previous versions of population data.

Various in silico analysis tools are available to estimate the pathogenicity of coding variants. Sorting Intolerant from Tolerant (SIFT, https://sift.bii.a-star.edu.sg/) predicts that variations in well-conserved protein families will tend to have deleterious effects on the basis that important amino acids are presumed to be conserved within the protein family.[15-16] SIFT generates a score that ranges between 0 and 1, which predicts the effects on proteins of an amino acid substitution. A SIFT score of less than 0.05 is predicted to be a “deleterious” mutation, whereas a value greater than or equal to 0.05 is expected to be a “tolerated” amino acid change. Prediction of functional effects of human nsSNPs (PolyPhen-2, http://genetics.bwh.harvard.edu/pph2/) is a tool that predicts the effects of a coding nonsynonymous SNP.[17] The PolyPhen-2 scoring system is divided into “probably damaging” (> 0.908), “possibly damaging” (0.446 < score ≤ 0.908), and “benign” (≤0.446). The HumDiv-

trained PolyPhen-2 model was built with all coding variants causing human Mendelian disease with functionally damaging effects. Evaluation of rare alleles at loci with potential involvement of complex phenotypes should use the HumDiv-trained model. The HumVar-trained PolyPhen-2 model was compiled from all human disease-causing mutations and common human nsSNPs (MAF>1%). The HumVar-trained model is recommended for the diagnosis of Mendelian disease requiring differentiation between mutations with drastic effects and variations including mildly deleterious alleles. MutationTaster (http://www.mutationtaster.org/) predicts the disease potential of coding DNA sequence variations using a Bayes classifier.[18] MutationAssessor (http://mutationassessor.org/r3/) is a predictive algorithm that estimates the functional impact based on amino acid conservation in the protein homologs.[19] The functional impact score is determined by a prediction algorithm that combines a conservation score and a specificity score, with higher scores more likely to be deleterious. Variants predicted to not impact protein function are classified as “neutral” or “low” and variants predicted to alter function are classified as “medium” or “high”. Combined annotation-dependent depletion (CADD, https://cadd.gs.washington.edu/) is an integrative model that can effectively discriminate casual variants as well as insertion/deletions identified in genetic analyses.[20-21] Raw CADD scores are generated straight from the computational model and higher scores are more indicative of deleterious effects. In general, Scaled CADD scores is typically used for reviewing individual or small sets of variants. The PHRED-like “scaled C-score” relatively ranks a variant to all possible substitutions. The

bottom 90% of reference SNVs are classified into scaled CADD C-scores below 10. Scaled C-score of 20 would indicate a variant in the top 1%, and a score of 30 would indicate the top 0.1%, which could potentially be considered to be a clinically relevant nucleotide variant. Rare exome variant ensemble learner (REVEL, https://sites.google.com/site/revelgenomics/) incorporates the scores of 13 different in silico analysis tools (MutPred, FATHMM v2.3, VEST 3.0, PolyPhen-2, SIFT, PROVEAN, MutationAssessor, MutationTaster, LRT, GERP++, SiPhy, phyloP, phastCons) to predict the effects of an amino acid substitution.[22] The REVEL score ranges between 0 and 1, with a disease-causing variant more likely to have a higher score. Generally, scores below 0.5 indicate a “likely benign” variation.

III. RESULTS

1. Spectrum of *APC* gene mutations

Pathogenic or likely pathogenic *APC* gene mutations were found in 167 out of 420 patients tested (39.8%, Figure 2). The majority of mutations were found in exon 15 (54.5% of cases), as expected, with a predilection for the well-known hotspots (codons 1062 and 1309). The most common mutation c.3927_3931delAAAGA, p.Glu1309AspfsTer4, which produces a truncated protein caused by a 5 base-pair deletion, was detected in 7.8% of mutation-positive cases. The distribution of the detected mutations is listed in Table 2. The proportion of mutations types are illustrated in Figure 3, which includes

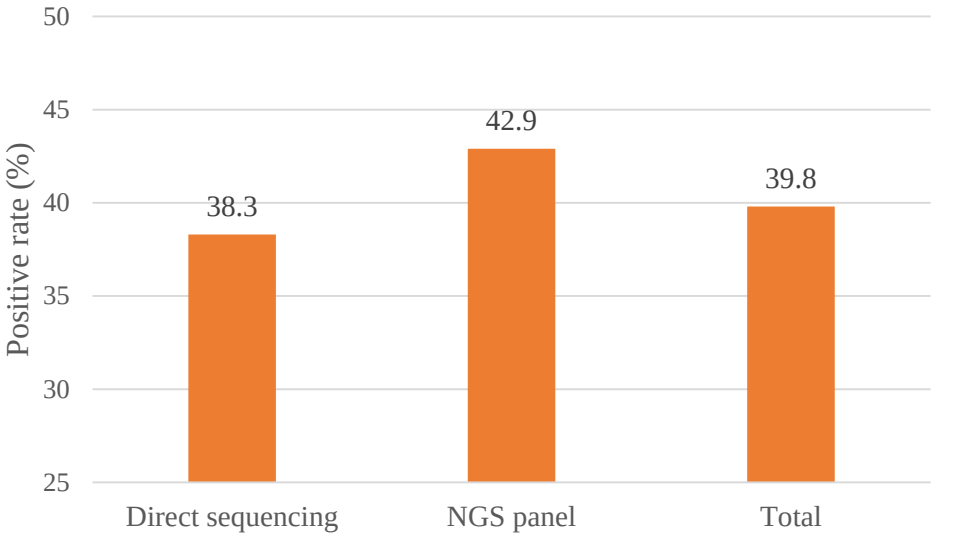

Figure 2. Positive rates of direct sequencing and NGS panel methods for detecting pathogenic or likely pathogenic APC gene mutations (n=420).

variants of unknown significance (VUS) found in the APC gene (n=186). In addition to the well-known nonsense (38.7%) and frameshift (36.0%) mutations, there were 3 cases of whole gene deletions and 3 cases of single or multiple exon deletions (3.6%, classified as large deletions). One of the cases of multiple exon deletions is illustrated in Figure 4. The patient showed a normalized depth near -0.5 in comparison to the normalized depth of controls by NGS, and was suspicious of exon deletions of the *APC* gene. Further analysis with MLPA revealed decreased peak ratios near 0.5 for exons 1-4, which confirmed multiple exon deletions in the *APC* gene.

Table 2. Distribution of patients with pathogenic or likely pathogenic mutations according to location in the *APC* gene (n=167)

Location	Size (base pair)	Positive cases (n)	Types of mutations (n)
Exon 1	135	-	
Exon 2	85	1	1
Exon 3	202	5	3
Exon 4	109	3	3
Intron 4		1	1
Exon 5	114	9	5
Intron 5		1	1
Exon 6	84	12	3
Exon 7	105	1	1
Exon 8	99	1	1
Intron 8		1	1
Exon 9	379	10	6
Intron 9		1	1
Exon 10	96	1	1
Exon 11	140	2	2
Exon 12	78	2	2
Exon 13	117	4	3
Intron 13		6	2
Exon 14	215	8	7
Intron 14		1	1
Exon 15	6574	91	5
Large deletions		6	4

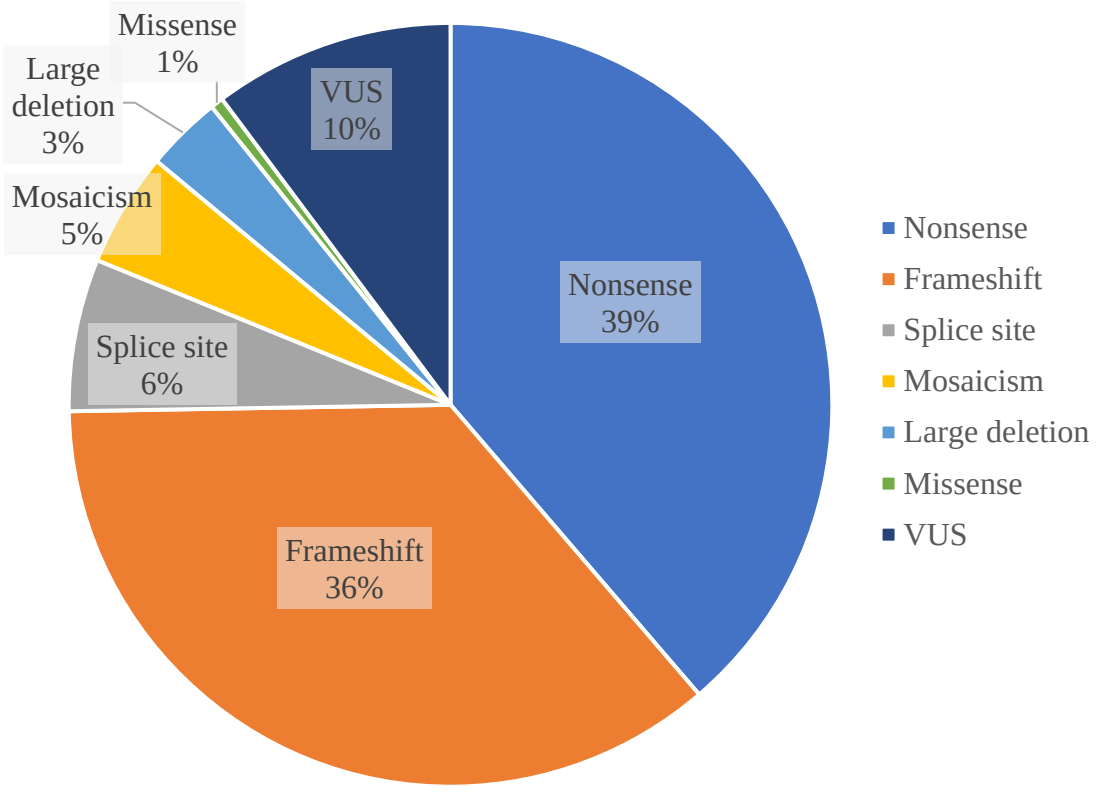

Figure 3. Proportion of mutation types detected in the *APC* gene, including variants of unknown significance (VUS) (n=186). Nonsense and frameshift mutations were the most commonly found mutation types in the *APC* gene.

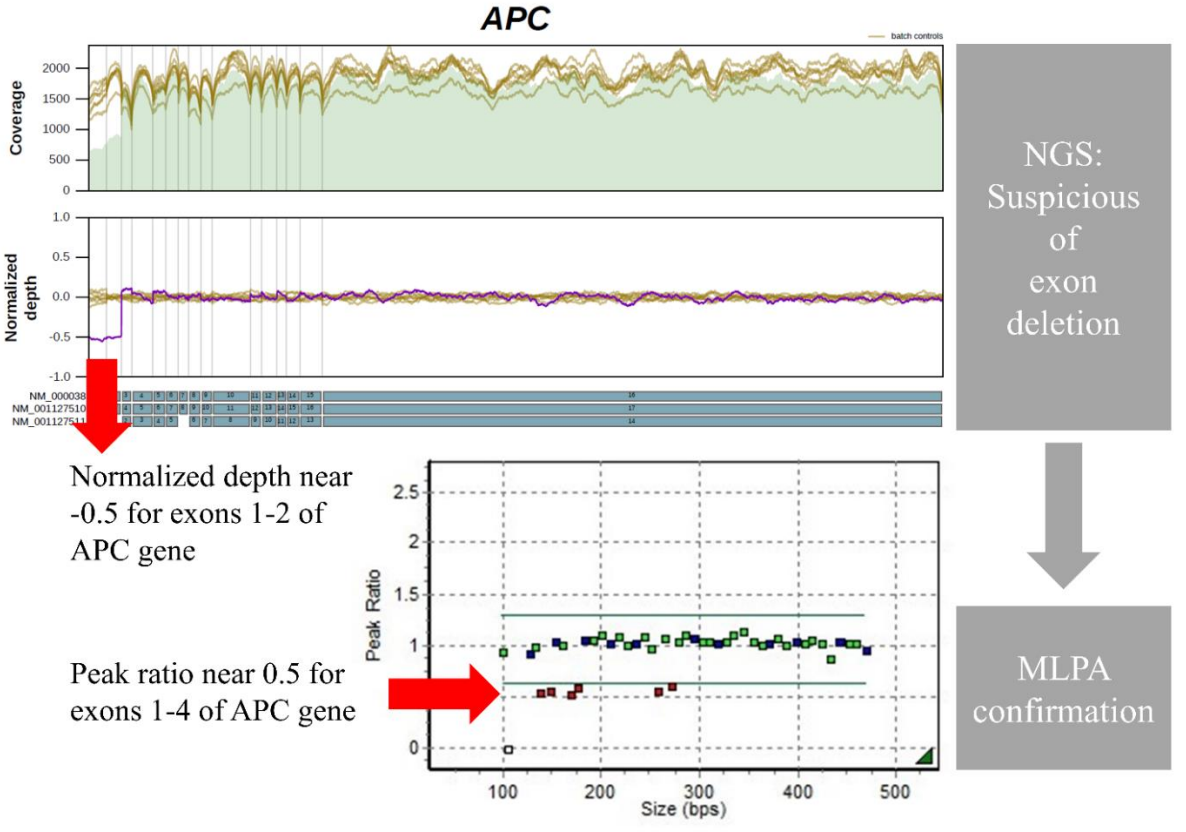

Figure 4. An example of a patient with suspected exon deletion on NGS, with normalized depth decreased by half compared to the normalized depth of controls. The patient showed peak ratios near 0.5 for exons 1-4 of *APC* gene on MLPA, which confirmed multiple exon deletion of the *APC* gene.

Twelve of the patients were tested in duplicate due to the emergence of NGS panels as an alternative method for genetic testing (Table 3). Four of the patients had surgery due to development of colon cancer and the pathologic findings confirmed a diagnosis of adenocarcinoma with multiple tubular adenomas, suggestive of FAP. Eight of the patients showed multiple to numerous adenomas in the colon on colonoscopy, which were clinically suggestive of FAP. A patient was suspected of having whole gene deletion of the *APC* gene with normalized depth of near -0.5 for the entire *APC* gene (Table 3, patient 6). Decreased peak ratios near 0.5 for all exons confirmed whole gene deletion of the *APC* gene by testing with MLPA (Figure 5). Two patients showed *APC* gene mosaicism (Table 3, patients 7 and 10) and likely pathogenic mutations in *BMPR1A* and *MUTYH* gene were detected in two patients (Table 3, patients 4 and 5). The remaining 5 patients had 1 to 4 genetic variations of unknown significance (VUS) in one of the 60 genes tested in the NGS panel.

Novel *APC* gene mutations were detected in 17 patients (Table 4). None of the mutations were detected in the known databases such as HGMD, ClinVar, or gnomAD. Approximately half of the mutations were located in exon 15 of the *APC* gene (52.9%) and a majority of the novel mutations were frameshift mutations (70.6%). Due to the formation of truncating mutations of the *APC* gene in FAP patients, mutation classification assumed pathogenicity for frameshift or nonsense mutations prior to the incorporation of the ACMG criteria. All of the novel mutations were re-classified as likely pathogenic variants according to the ACMG guidelines. All 17 mutations fulfilled the null

Table 3. Clinical and laboratory findings of patients tested by both direct sequencing and NGS panel tests (n=12).

No.	Sex/ Age	Direct Sequencing	NGS panel test	Clinical diagnosis	Colonoscopy or surgical findings
1	M/33	Negative	*APC* c.1262G>A, p.Trp421Ter, hetero, P	FAP	Total colectomy: Multiple synchronous adenocarcinomas with numerous adenomatous polyps, consistent with FAP
2	F/55	Negative	*BRIP* c.2830C>G, p.Gln944Glu, hetero, VUS	Attenuated FAP, Sigmoid colon cancer	Anterior resection: Adenocarcinoma, moderately differentiated; multiple polypoid lesions are noted
3	M/35	Negative	*MLH1* c.677+7C>T, hetero, VUS *CDH1* c.1223C>T, p.Ala408Val, hetero, VUS	FAP	Colonoscopy: numerous adenomas in entire colon (>100)
4	F/28	Negative	*BMPR1A* c.335del, p.Asp112ValfsTer11, hetero, LP	FAP	Colonoscopy: numerous adenomas in entire colon
5	F/57	Negative	*MUTYH* c.857G>A, p.Gly286Glu, hetero, LP *MUTYH* c.842C>T, p.Ala28Val, hetero, VUS	FAP	Colonoscopy: numerous adenomas in entire colon
6	M/32	Negative	*APC* whole gene deletion	FAP	Colonoscopy: numerous adenomas in entire colon
7	F/44	Negative	*APC* c.4348C>T, p.Arg1450Ter, hetero, P (mosaicism)	FAP, Colon cancer	Lower anterior resection: Adenocarcinoma, moderately differentiated; multiple tubular adenomas

8	M/36	Negative	*SLX4* c.3583_3585del, p.Ile1195del, hetero, VUS *POLD1* c.216A>G, p.Pro72=, hetero, VUS	FAP	Colonoscopy: numerous adenomas in entire colon
9	M/56	Negative	Negative	FAP, Rectal cancer	Total proctocolectomy: Adenocarcinoma, moderately differentiated; multiple tubular adenomas (>100)
10	F/44	Negative	*APC* c.3295_3296delGT, Val1099PhefsTer19,LP (mosaicism)	FAP, Thyroid cancer, Retinal disorder	Colonoscopy: numerous adenomas in entire colon
11	F/54	Negative	*BMPR1A* c.1243G>A, p.Glu415Lys, hetero, VUS *BLM* c.2839A>G, p.Ile947Val, hetero, VUS *NF1* c.5160G>T, p.Glu1720Asp, hetero, VUS *NF2* c.240+15C>T, hetero, VUS	r/o FAP	Colonoscopy: multiple adenomas
12	F/40	Negative	*NF1* c.1740_1742delTTT, p.Phe580del, hetero, VUS	r/o FAP	Colonoscopy: multiple adenomas

Abbreviations: NGS, next generation sequencing; FAP, familial adenomatous polyposis; P, pathogenic; VUS, variant of unknown significance; LP, likely pathogenic.

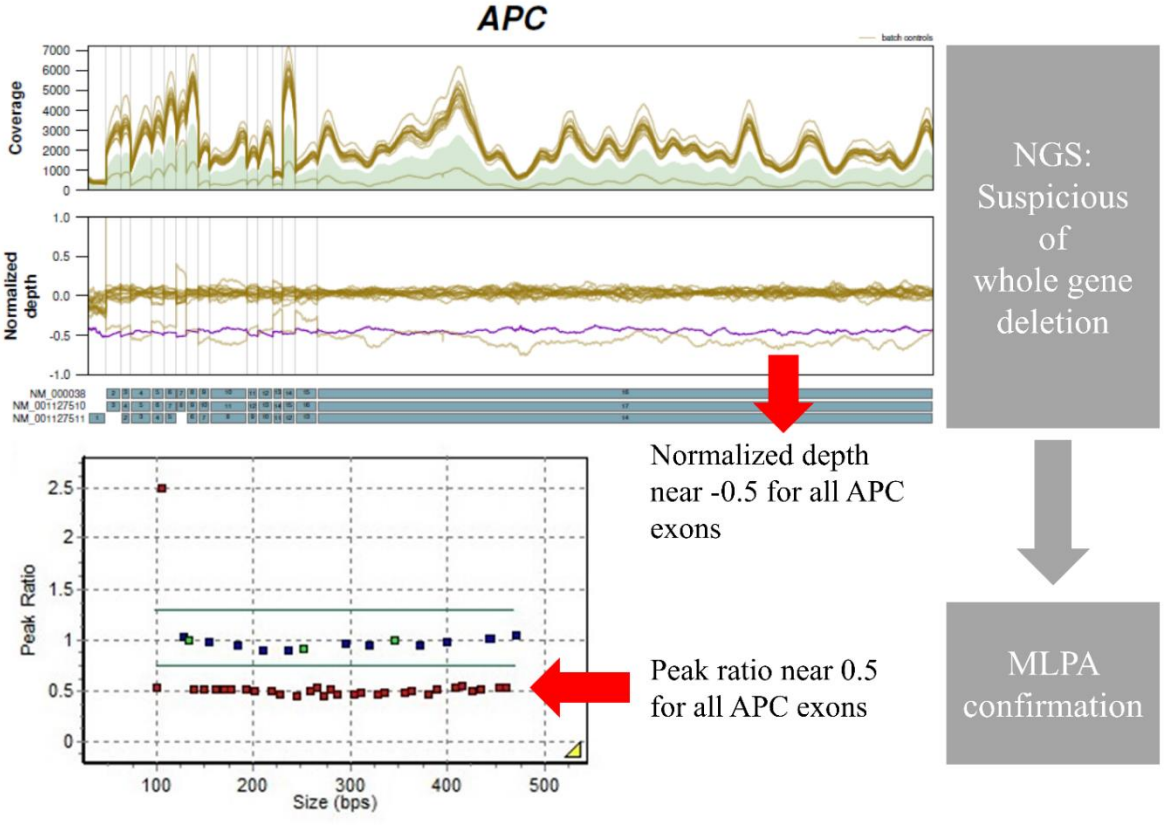

Figure 5. An example of a patient with suspected whole gene deletion on NGS, with normalized depth decreased by half compared to the normalized depth of controls. The patient showed peak ratios near 0.5 for all *APC* exons on MLPA, which was confirmatory of *APC* gene deletion.

Table 4. Clinical and laboratory findings of FAP patients with novel mutations detected in the *APC* gene (n=17).

No.	Sex/ Age	Exon	*APC* gene mutation	Type of mutation	Effect	Diagnosis	Family history	ACMG	Effect
13	F/28	5	c.611delT, p.Leu204GlnfsTer15, hetero	FS	P	FAP	None	PVS1 PM2	LP
14	M/63	9	c.1257delC, p.Cys420ValfsTer34, hetero	FS	LP (PVS1 +PM2)	FAP	None	=	LP
15	F/32	11	c.1541_1542delCC, p.Ala514Glufs22, hetero	FS	P	FAP	None	PVS1 PM2	LP
16	M/30	12	c.1587delG, p.Val530TrpfsTer4, hetero	FS	LP (PVS1 +PM2)	FAP	Mother (colon cancer), Aunt (colon cancer)	=	LP
17	M/46	13	c.1708A>T, p.Lys570*, hetero	NS	P	r/o FAP	None	PVS1 PM2	LP
18	M/28	14	c.1771_1777delGCCTTAT, p.Ala591GlyfsTer17, hetero	FS	P	FAP	None	PVS1 PM2	LP
19	F/43	14	c.1844_1857delinsCCATCTT, p.Phe615SerfsTer13, hetero	FS	P	FAP	None	PVS1 PM2	LP
20	M/14	14	c.1928_1958+14del, p.Ser643AlafsTer20, hetero	FS	P	Unknown	-	PVS1 PM2	LP
21	M/51	15	c.2149dupA, p.Met717AsnfsTer17, hetero	FS	P	Unknown	-	PVS1 PM2	LP
22	M/37	15	c.2327_2328dupTA, p.Asp777*, hetero	NS	P	Unknown	-	PVS1 PM2	LP

23	M/32	15	c.2492T>A, p.Leu831*, hetero	NS	P	Unknown	-	PVS1 PM2	LP
24	M/28	15	c.2887dupA, p.Ser963LysfsTer4, hetero	FS	P	FAP, colon cancer	None	PVS1 PM2	LP
25	M/33	15	c.2923A>T, p.Lys975*, hetero	NS	P	Unknown	-	PVS1 PM2	LP
26	M/27	15	c.3169delG, p.Glu1057LysfsTer4, hetero	FS	P	Unknown	-	PVS1 PM2	LP
27	F/24	15	c.3610C>T, p.Gln1204*, hetero	NS	LP (PVS1 +PM2)	FAP, Papillary microcarcinoma (cribriform-morular variant)	Grandfather (colon cancer)	=	LP
28	F/55	15	c.4148delT, p.Met1383SerfsTer32, hetero	FS	P	Unknown	-	PVS1 PM2	LP
29	F/54	15	c.4429dupC, p.Gln1477ProfsTer10, hetero	FS	P	Unknown	-	PVS1 PM2	LP

*Abbreviations: ACMG, American College of Medical Genetics and Genomics; FS, frameshift mutation; NS, nonsense mutation; P, pathogenic; LP, likely pathogenic; FAP, familial adenomatous polyposis; =, same evidence for ACMG classification.

criteria. All of the novel mutations were re-classified as likely pathogenic variants according to the ACMG guidelines. All 17 mutations fulfilled the null variant evidence of PVS1 criteria since they were either a frameshift or nonsense mutation. In addition, all mutations were novel findings that were absent from previously known databases which fulfilled the PM2 criteria. The combination of PVS1 and PM2 are classified as "likely pathogenic". Three of the patients had been reported with the ACMG criteria as likely pathogenic, and their status did not change (expressed as "=" in Table 4).

2. Variants of unknown significance (VUS) in the *APC* gene

VUS of the *APC* gene was identified in 19 patients. The population data was re-examined due to the updates in the database, including ExAC database, gnomAD v2.1.1, and KRGDB. The variants were evaluated with various in silico algorithms for additional evidence of the ACMG criteria. Eight patients did not show any new evidence and remained a VUS according to the ACMG criteria. One patient with c.423-8A>G was originally classified as VUS (PM2), but secondary review suggested a shifting of the splice acceptor site which may disrupt the normal splicing event and create a splice variant (Table 4, patient 14), which may be supportive evidence of PP3. A recent publication with RNA sequencing results confirmed the shifting of the splice acceptor site causing a frameshift mutation, which was evidence of PS3 and upgraded the previous VUS to a likely pathogenic variant.[23] Six patients with VUS near the latter half of exon 15 showed likely benign results on multiple lines of in silico algorithms, which could be supportive evidence of BP4 (Table 5, patients 39-

41,18,45,47). Four of the cases were patients diagnosed with FAP who had likely pathogenic variants in the *APC* gene with an additional VUS in the *APC* gene (Table 5, patients 40-41,18,47). A synonymous variant classified as VUS was detected in a patient with cribriform morular variant of papillary thyroid cancer (Table 5, patient 38). Two patients with a diagnosis of FAP or aFAP did not show any additional evidence for reclassification (Table 5, patients 36, 43). Additional variants were identified in the NGS panel and is described in the appropriate section (Table 5, patient 39).

3. Somatic mosaicism in *APC* gene mutations

A total of 9 patients with somatic mosaicism was detected in the NGS panel tests with variant allele frequency (VAF) below 20% (Table 6). Two thirds of the mutations were detected in exon 15 without clustering near the mutation cluster region (MCR). All 9 patients showed nucleotide substitutions, small deletions or duplications, causing a frameshift or nonsense mutation. All patients showed findings of multiple adenomas on colonoscopy or colon cancer, which was consistent with a diagnosis of FAP. Confirmation test with direct sequencing or MEMO-PCR was done on available samples. The results of a patient suspected mosaicism on NGS due to low variant allele frequency is illustrated in Figure 6 (Table 6, patient 10). MEMO-PCR of the peripheral blood detected the c.3295_3296del, p.Val1099PhefsTer19 mutation. Further analysis with tissue (neoplastic polyp) revealed identical c.3295_3296del, p.Val1099PhefsTer19 mutation with MEMO-PCR. The tissue NGS results

Table 5. Clinical and laboratory findings of VUS in the *APC* gene with additional data for reclassification of ACMG criteria (n=19).

No	Sex/ Age	*APC* gene variant	Type of mutation	Effect (initial)	ExAC total /E.Asian (%)[1]	gnomAD genome/ E.Asian (%)[1]	KRGDB (%)[1]	Additional evidence	ACMG	Diagnosis
30	F/51	c.92C>G, p.Ser31Cys	MS (VUS)	PM2 BP1	-	-	-	-	PM2 BP1	Colon polyps, colon cancer
31	F/56	c.423-8A>G	Sp (VUS)	PM2	-	-	-	Recent report (PMID 32067438)	PM2 PP3 PS3 (LP)	FAP, colon cancer (2 brothers: colon cancer)
32	M/23	c.1276G>T, p.Ala426Ser	MS (VUS)	BP1	0.0032 / 0.0981	0.0096 / 0.0641	0.1608	-	BP1	Colon polyps (Family Hx(-))
33	M/57	c.1276G>T, p.Ala426Ser	MS (VUS)	BP1	0.0032 / 0.0981	0.0096 / 0.0641	0.1608	-	BP1	N/A
34	M/48	c.1276G>A, p.Ala426Thr	MS (VUS)	BP1	-	-	-	PM2	PM2 BP1	Colon polyps, rectal cancer
35	F/52	c.3378C>G, p.Ser1126Arg	MS (VUS)	BP1	0.0060 / 0.0762	-	0.1608	-	BP1	Colon cancer
36	F/37	c.3964G>A, p.Glu1322Lys	MS (VUS)	BP1	0.0024 / 0	0.0032 / 0	-	PM2	PM2 BP1	FAP, colon cancer
37	N/A	c.4142C>T, p.Pro1381Leu	MS (VUS)	PM2 BP1	-	-	-	-	PM2 BP1	N/A

38	F/46	c.4782A>G, p.Pro1594=	Syn (VUS)	PM2 BP1	-	-	-	-	PM2 BP1	Papillary thyroid cancer (cribriform morular variant)
39	F/70	c.5257G>C, p.Ala1753Pro	MS (VUS)	BP1	0.0020 / 0.0272	0.0032 / 0.0641	-	BP4	BP1 BP4 (LB)	Multiple adenomas, r/o aFAP
40	M/43	c.5257G>C, p.Ala1753Pro	MS (VUS)	BP1	0.0020 / 0.0272	0.0032 / 0.0641	-	BP4	BP1 BP4 (LB)	FAP (+*APC*, LP)
41	M/36	c.5378C>G, p.Ala1793Gly	MS (VUS)	BP1	0.0004 / 0.0055	-	-	PM2 BP4	PM2 BP1 BP4 (LB)	FAP (+*APC*, LP)
18	M/28	c.5708A>G, p.Asn1903Ser	MS (VUS)	BP1	0.0016 / 0.0016	-	-	PM2 BP4	PM2 BP1 BP4 (LB)	FAP (+*APC*, LP)
42	M/19	c.6380A>G, p.Gln2127Arg	MS (VUS)	PM2 BP1	-	-	-	-	PM2 BP1	N/A
43	M/35	c.6896C>T, p.Pro2299Leu	MS (VUS)	PM2 BP1	0.0008 / 0	-	-	-	PM2 BP1	FAP, rectal cancer (FHx: -)
44	M/44	c.7112G>C, p.Arg2371Thr	MS (VUS)	PM2 BP1	0.0004 / 0.0054	-	-	PM2	PM2 BP1	N/A

45	F/58	c.7150T>A, p.Leu2384Ile	MS (VUS)	BP1	0.0028 / 0.0381	-	0.0804	BP4	BP1 BP4 (LB)	Colon polyps
46	F/28	c.7433A>C, p.Gln2478Pro	MS (VUS)	BP1	0.0004 / 0.0054	-	-	PM2 PP3	PM2 PP3 BP1	Family history of neoplasm
47	F/47	c.7969G>A, p.Val2657Ile	MS (VUS)	BP1	0.0004 / 0.0054	-	0.03	BP4	BP1 BP4 (LB)	FAP (Mother: colon cancer) (+*APC*, LP)

*Abbreviations: ACMG, American College of Medical Genetics and Genomics; E.Asian, East Asian; ExAC, Exome Aggregation Consortium; F, female; FAP, familial adenomatous polyposis; gnomAD, genome aggregation database; KRGDB, Korean Reference Genome Database; LB, likely benign; LP, likely pathogenic; M, male; MS, missense variant; N/A, not available; Sp, splice site variant; Syn, synonymous variant; VUS, variant of unknown significance.
[1]Population frequencies are converted to % and rounded off to the 4th digit below the decimal point.

Table 6. Characteristics of patients with suspected somatic mutations in the *APC* gene (n=9).

No	Sex/ Age	Exon	Mutation	Type	VAF	Clinical finding	Confirmation test
47	F/47	6	c.694C>T, p.Arg232Ter	NS	3.4%	Colonoscopy: polyps Total colectomy: Numerous tubular adenomas, consistent with FAP (>70)	Sequencing (adenoma)
48	F/44	8	c.902del, p.Pro301LeufsTer4	FS	5.1%	Colonoscopy: numerous adenomas in entire colon	MEMO-PCR (PB)
49	F/52	14	c.1754delT, p.Leu585ProfsTer5	FS	2.0%	Colonoscopy: adenomas Anterior resection of sigmoid colon: Adenocarcinoma, MD	- (sample unavailable)
50	M/36	15	c.2626C>T, p.Arg876Ter	NS	11.8%	Colonoscopy: numerous adenomas in entire colon	-
51	M/36	15	c.3211_3238dup, p.Glu1080AlafsTer 10	FS	19.5%	Colonoscopy: rectal cancer	Sequencing (PB)
10	F/45	15	c.3295_3296delGT, p.Val1099PhefsTer 19	FS	6.8% (adenoma 20.6%)	Colonoscopy: numerous adenomas in entire colon	MEMO-PCR (PB and adenoma)
52	F/55	15	c.3566C>G, p.Ser1189Ter	NS	11.4%	Colonoscopy: numerous adenomas in entire colon	Sequencing (PB)
53	F/44	15	c.3860_3861dup, p.Gly1288Ter	NS	9.4%	Total proctocolectomy: Adenocarcinoma, MD + multiple tubular and tubulovillous adenomas (>100)	MEMO-PCR (PB)
7	F/44	15	c.4348C>T, p.Arg1450Ter	NS	6.4%	Colonoscopy: numerous adenomas in entire colon	-

*Abbreviations: F, female; FAP, familial adenomatous polyposis; FS, frameshift mutation; M, male; MD, moderately

differentiated; MEMO-PCR, mutant enrichment with 3’-modified oligonucleotide polymerase chain reaction; NS, nonsense mutation; PB, peripheral blood; VAF, variant allele frequency.

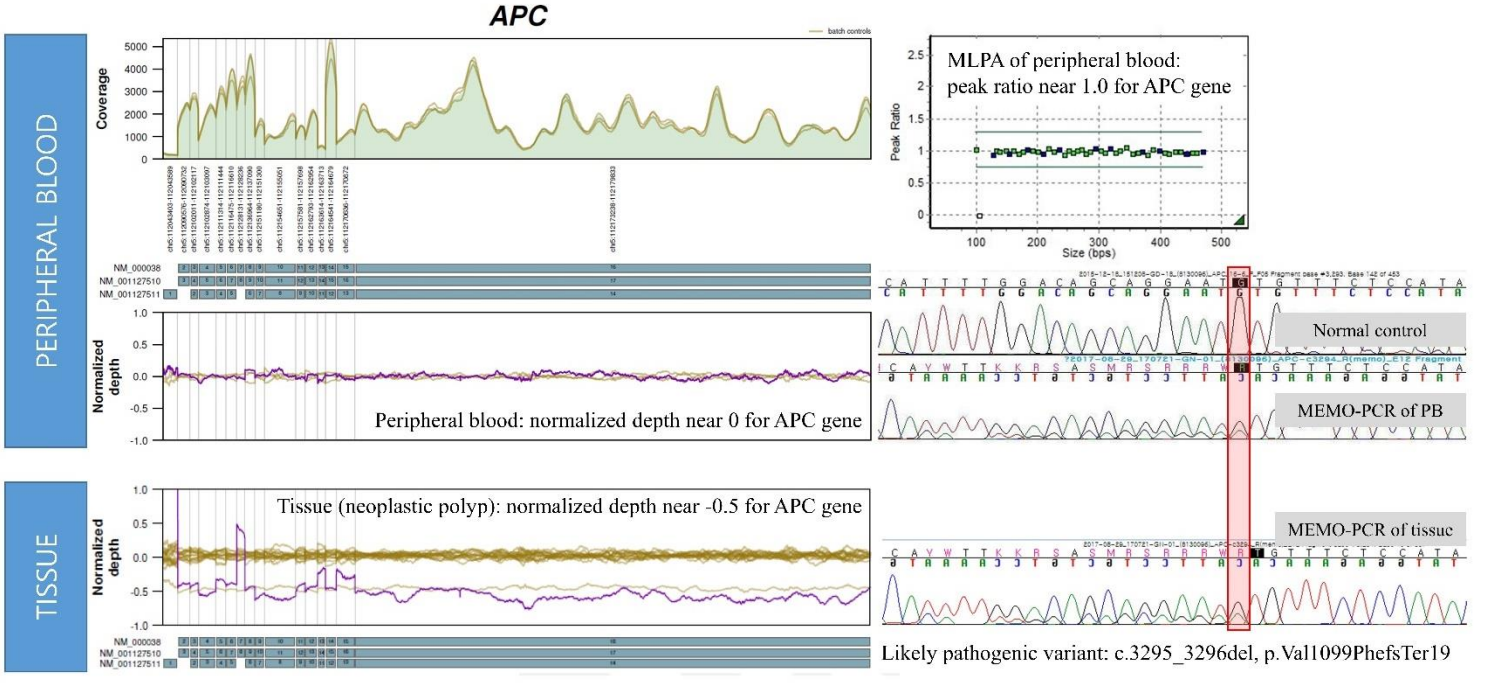

Figure 6. An example of a patient with suspected mosaicism on NGS due to low variant allele frequency. MEMO-PCR of the peripheral blood detected the c.3295_3296del, p.Val1099PhefsTer19 mutation. NGS and MLPA results confirmed no copy number variations of the *APC* gene in the peripheral blood. Further analysis with tissue (neoplastic polyp) revealed identical mutation with MEMO-PCR. MLPA analysis of the neoplastic polyp revealed a normalized depth near 0.5, which was suggestive of whole gene deletion.

showed normalized depth near -0.5, which was suggestive of whole gene deletion, whereas the peripheral blood NGS results did not show any copy number variations. Subsequent MLPA tests confirmed the copy number variation detected in this patient.

4. FAP-diagnosed or FAP-suspicious patients without *APC* gene mutations

Eight patients with multiple to numerous adenomas in the colon were negative for the *APC* gene mutation but showed other pathologic variants (Table 7). A *BMPR1A* gene mutation was the most frequently detected, consisting of frameshift and nonsense mutations and an exon deletion. *MUTYH* gene mutation was detected in two patients, and *PMS1* and *POLE* gene mutations were also found. All 8 patients had additional 1-3 variants which were classified as VUS. A patient with *BMPR1A* exon 3 deletion was initially classified as likely pathogenic, but was able to be reclassified into a pathogenic variant with the additional evidence of PP5 (Table 7, patient 59).

The NGS results of 35 patients with negative *APC* gene mutations who had multiple to numerous adenomas on colonoscopy and were suspected of having FAP were examined (Table 8) for other genetic variations. One to six different variants were detected in 12 patients with numerous adenomas in entire colon on colonoscopy (Table 8, patients 3, 8, 60-69). Only one variant (*SLX4* c.3583_3585del, p.Ile1195del) overlapped among these patients (Table 8, patient 8, 64). Even with the addition of other ACMG criteria, there was no difference in ACMG classification in all 12 patients. Similar to the previous group of patients, one to six different variants were identified in 23 patients

Table 7. Clinicopathologic characteristics of *APC* gene mutation negative, FAP suspected patients with likely pathogenic mutations in other genes (n=8).

No	Sex/ Age	Gene	Mutation	Type of mutation	ACMG classification	Other VUS	Diagnosis
4	F/28	*BMPR1A*	c.335del, p.Asp112ValfsTer11, hetero	FS	PVS1 PM2 (LP)	RAD50 (PM2+BP1)	Numerous adenomas in entire colon
5	F/57	*MUTYH*	c.857G>A, p.Gly286Glu, hetero (rs730881833)	MS	PS3 PP3 PP5 (LP)	*MUTYH* (PP3) *CDH1* (BS1) PTCH (-)	Numerous adenomas in entire colon
54	F/58	*PMS1*	c.1258delC, p.His420IlefsTer22, hetero	FS	PVS1 PM2 (LP)	BRCA (PM2) AXIN2 (PM2+BP7) FANCM (BS1)	Multiple adenomas
55	M/18	*BMPR1A*	c.236delA, p.Asn79MetfsTer8, hetero	FS	PVS1 PM2 (LP)	*MLH1* (PP3+BS1) *BLM* (-) KRAS (BP6)	Multiple adenomas
56	M/57	*POLE*	c.5063delC, p.Pro1688LeufsTer73, hetero	FS	PVS1 PM2 (LP)	PTCH1 (PM2)	Multiple adenomas
57	F/31	*MUTYH*	c.799C>T, p.Gln267*, hetero (rs786203115) in trans-compound heterozygote with *MUTYH* c.842C>T, p.Ala281Val	NS	PVS1 PP5 (LP)	*MUTYH* (PP3)	Multiple adenomas

58	M/72	*BMPR1A*	c.682C>T, p.Arg228*, hetero (rs587782682)	NS	PVS1 PM2 PP5 (P)	*MLH1* (-)	Multiple adenomas
59	M/31	*BMPR1A*	Exon 3 deletion	DEL	- (LP) → PVS1, PM2, PP5 (P)	BRCA (BP6) PTEN (BP7) *BRIP*1 (PM2+BP7)	Numerous adenomas in entire colon

*Abbreviations: ACMG, American College of Medical Genetics and Genomics; DEL, large deletion; F, female; FAP, familial adenomatous polyposis; FS, frameshift mutation; LP, likely pathogenic; M, male; MS, missense mutation; NS, nonsense mutation; P, pathogenic; VUS, variant of unknown significance.

Table 8. Clinicopathologic characteristics of *APC* gene mutation negative patients with numerous adenomas in entire colon (n=12) and multiple adenomas (n=23) who had other gene variants detected in the NGS panel.

No	Sex/ Age	Gene	Other gene variants	Chromo-some location	ACMG (initial)	ExAC total /E.Asian (%)[1]	gnomAD genome/ E.Asian (%)[1]	KRGDB (%)[1]	Additional evidence	ACMG (revised)	Diagnosis
3	M/35	*MLH1*	c.677+7C>T	chr3: 37053597	PM2, BP6	0.0009 / 0	-	-	-		Numerous adenomas in entire colon
		CDH1	c.1223C>T, p.Ala408Val	chr16: 68847301	BP6	0.02 / 0	0.0065 / 0	-	-		
8	M/36	*SLX4*	c.3583_3585 del, p.Ile1195del	chr16: 3640054-3640056	PM4, BS1	0.07 / 0.98	0.0382 / 0.7702	-	-		
		POLD 1	c.216A>G, p.Pro72=	chr19: 50902641	BP7	0.001 / 0.01	-	-	PM2	PM2 +BP7 (VUS)	
60	F/55	*GALN T12*	c.829G>A, p.Gly277Ser	chr9: 101594151	N/A	0.0038 / 0.01	0.0032 / 0	0.3215	PP3	PP3 (VUS)	
		ALK	c.2210C>T, p.Ser737Leu	chr2: 29462691	N/A	0.01 / 0.14	0.0032 / 0.0642	-	-		
		STK11	c.1190C>T, p.Ala397Val	chr19: 1226534	PP2, BP4	0.02 / 0.06	-	-	-		
		MSH2	c.1886A>G, p.Gln629Arg	chr2: 47702290	BS1	0.14 / 1.81	0.0893 / 1.667	1.5273	-		

		RET	c.833C>A, p.Thr278Asn	chr10: 43600607	BS1	0.22 / 2.83	0.1341 / 2.632	3.1351	-	
		NBN	c.1657A>G, p.Met553Val	chr8: 90965660	PM2, BP4	0.0009 / 0.01	-	-	-	
61	M/47	*FANCM*	c.4931G>A, p.Arg1644Gln	chr14: 45658156	BS1	0.15 / 1.77	0.0987 / 1.861	0.8039	-	
		BLM	c.1785A>T, p.=	chr15: 91304388	BP7	0.0042 / 0.05	-	-	-	
62	M/34	*POLE*	c.4290+5C>T	chr12: 133220418	BS1	2.35 / 3.3	2.724 / 2.5	3.3762	-	
		SLX4	c.4057C>T, p.His1353Tyr	chr16: 3639582	PM2, BP4	0.0019 / 0.03	0.0032 / 0.0641	-	-	
		NF1	c.1740_1742 delTTT, p.Phe580del	chr17: 29548963-29548965	PM4	0.001 / 0.12	-	-	-	
		PHOX2B	c.765_779del GGCAGCGGCGGCAGC, p.Ala256_Ala260del	chr4: 41747990-41748004	PM4, BS1	0.15 / 1.91	0.2843 / 4.376	-	-	
63	M/89	*MSH2*	c.1168C>T, p.Leu390Phe	chr2: 47656972	PP3, BS1	0.18 / 2.3	0.0766 / 1.54	2.8939	-	
64	F/33	*CHEK2*	c.1160C>T, p.Thr387Ile	chr22: 29091797	N/A	0.0009 / 0	-	-	PM2 PP3	PM2 +PP3 (VUS)

		BRCA2	c.6029T>G, p.Val2010Gly	chr13: 32914521	BP4	0.0028 / 0.04	-	0.0804	-	
		SLX4	c.3583_3585 del, p.Ile1195del	chr16: 3640054-3640056	PM4, BS1	0.07 / 0.98	0.0382 / 0.7702	-	-	
65	M/46	*SDHB*	c.488C>T, p.Ser163Phe	chr1: 17354296	PP3, PP2	0.0009 / 0.01	-	-	PM2	PM2 +PP2 +PP3 (VUS)
		AXIN2	c.1908-11T>A	chr17: 63532682	N/A	0.001 / 0.01	-	-	PM2	PM2 (VUS)
		GALNT12	c.850G>A, p.Val284Met	chr9: 101594172	N/A	0.0009 / 0	0.0064 / 0	-	PM2	PM2 (VUS)
66	M/58	*CDK2NA*	c.501G>A, p.Ala167=	chr9: 21968727	BP7	0.0091 / 0	0.0064 / 0.1282	0.2412	-	
		BARD1	c.1972C>T, p.Arg658Cys	chr2: 215595164	BS1	0.8 / 1.12	0.6213 / 0.8333	0.9646	-	
67	M/39	*CHEK2*	c.246_260del, p.Asp82_Glu86del	chr22: 29130450-29130464	PM4	0.02 / 0.05	0.02 / 0	-	-	
		POLD1	c.1846C>T, p.Pro616Ser	chr19: 50912034	BP6	0.0009 / 0	-	0.0804	-	
		NBN	c.323T>C, p.Ile108Thr	chr8: 90993119	PM2	0.0048 / 0	0.0032 / 0	-	-	

		MLH3	c.4032C>T, p.Gly1344=	chr14: 75489575	BP7	0.02 / 0.2	0.0097 / 0.18	0.2412	-		
68	F/31	*MSH2*	c.1894A>G, p.Ile632Val	chr2: 47702298	N/A	-	0.0032 / 0.06	-	-		
		MLH1	c.649C>T, p.Arg217Cys (Likely benign)	chr3: 37053562	PP3, BP6	0.03 / 0.38	0.0097 / 0.12	0.4823	-		
		MLH1	c.704A>T, p.Asp235Val (VUS)	chr3: 37055949	N/A	0.0039 / 0.05	-	0.08039	-		
69	M/50	*CHEK 2*	c.1111C>T, p.His371Tyr	chr22: 29091846	N/A	0.06 / 0.42	0.03 / 0.49	0.2412	-		
2	F/56	*BRIP1*	c.2830C>G, p.Gln944Glu	chr17: 59763272	BP4	0.03 / 0.25	0.0032 / 0.0641	0.2412	-		Multiple adenomas
70	F/28	*MLH1*	c.649C>T, p.Arg217Cys	chr3: 37053562	PP3, BP6	0.03 / 0.38	0.0096 / 0.1284	0.4823	-		
11	F/55	*BMPR 1A*	c.1243G>A, p.Glu415Lys	chr10: 88681353	BP6	0.07 / 0	0.0223 / 0	-	PP3	PP3 +BP6 (VUS)	
		BLM	c.2839A>G, p.Ile947Val	chr15: 91333894	N/A	0.02 / 0.29	0.0159 / 0.3209	-	BP4	BP4 (VUS)	
		NF1	c.5160G>T, p.Glu1720Asp	chr17: 29653162	PP2	0.0009 / 0.01	-	0.08039	-		
		NF2	c.240+15C>T	chr22: 30032880	BS1	0.03 / 0.32	0.0255 / 0.3213	0.9646	-		

12	F/40	*NF1*	c.1740_1742 delTTT, p.Phe580del	chr17: 29548963-29548965	PM4	0.0095/ 0.12	-	-	-	
71	M/49	*BRIP1*	c.2554A>G, p.Asn852Asp	chr17: 59770812	N/A	0.0038 / 0.05	-	0.1608	-	
		MLH3	c.3488G>A, p.Gly1163Asp	chr14: 75506696	BS1	0.2 / 2.64	0.1688 / 2.807	2.17	-	
72	M/49	*AXIN1*	c.853C>T, p.Arg285Trp	chr16: 396173	PP3	0.0038 / 0	0.0032 / 0	-	PM2	PM2 +PP3 (VUS)
		KRAS	c.556G>A, p.Val186Ile	chr12: 25362740	PM2, PP2	-	-	-	-	
		NTRK1	c.97G>T, p.Ala33Ser	chr1: 156830823	N/A	0.008 / 0.26	-	-	-	
73	M/53	*MSH6*	c.4068_4071 dupGATT, p.Lys1358AspfsTer2	chr2: 48033981	PVS1, BS1	0.24 / 3.19	0.1504 / 3.013	-	-	
		POLE	c.6135C>T, p.=	chr12: 133209251	BP7	0.02 / 0.06	0.0032 / 0	-	-	
74	M/69	*FLCN*	c.205G>A, p.Val69Ile	chr17: 17131247	N/A	0.05 / 0	0.0032 / 0	-	-	
		ALK	c.487G>T, p.Val163Leu	chr2: 30143039	BS1	0.03 / 0.41	030414 / 0.5769	0.0804	BP4	BS1 +BP4

										(LB)
75	M/65	*POLE*	c.1101T>G, p.Phe367Leu	chr12: 133252326	PM2	-	-	-	-	
		ATM	c.8265T>C, p.Tyr2755=	chr11: 108206685	PM2, PP5, BP6	0.0019 / 0.03	0.0032 / 0.0643	-	-	
76	M/59	*BRCA2*	c.4320A>C, p.Lys1440Asn	chr13: 32912812	PM2, BP6	0.0019 / 0.03			-	
77	M/52	*MSH6*	c.4068_4071dupGATT, p.Lys1358AspfsTer2	chr2: 48033981	PVS1, BS1	0.24 / 3.19	0.1504 / 3.013	-	-	
		SDHB	c.541-3C>T	chr1: 17350572	N/A	0.008 / 0.11	-	0.4019	-	
		PTCH1	c.3964G>A, p.Ala1322Thr	chr9: 98209574	PM2, PP3	-	-	-	-	
		FANCM	c.925G>A, p.Glu309Lys	chr14: 45620606	N/A	0.0009 / 0.01	-	-	PM2	PM2 (VUS)
78	M/71	*BMPR1A*	c.452T>C, p.Ile151Thr	chr10: 88659805	PM2	-	-	-	-	
		POLE	c.4411C>T, p.Arg1471Cys	chr12: 133220026	N/A	0.01 / 0.06	-	0.0804	-	

		NTRK1	c.631G>A, p.Val211Met	chr1: 156838353	N/A	0.02 / 0.21	0.0223 / 0.1926	0.1608	-	
		*BRIP*1	c.2051G>A, p.Cys684Tyr	chr17: 59853808	PM2, PP3	-	-	-	-	
		SLX4	c.5248G>T, p.Ala1750Ser	chr16: 3632599	BP4	0.02 / 0.24	0.0382 / 0.7051	0.1608	-	
		SLX4	c.5249C>T, p.Ala1750Val	chr16: 3632600	BP4	0.02 / 0.24	0.035 / 0.7051	0.1608	-	
79	F/58	*PTCH1*	c.86G>T, p.Gly29Val	chr9: 98279017	N/A	0.04 / 0.48	0.0032 / 0	0.1608	-	
		PHOX 2B	c.765_779del GGCAGCGG CGGCAGC, p.Ala256_Ala 260del	chr4: 41747990-41748004	PM4, BS1	0.15 / 1.91	0.2843 / 4.376	-	-	
80	F/66	*AXIN2*	c.128_133du p, p.Gly43_Gln 44dup	chr17: 63554605	PM4	0.0029 / 0.04	-	-	-	
		SLX4	c.635G>A, p.Arg212Gln	chr16: 3656600	N/A	0.0019 / 0.01	0.0032 / 0	-	PM2	PM2 (VUS)
81	M/63	*MLH3*	c.277C>G, p.Arg93Gly	chr14: 75516082	PP3	0.0066 / 0.09	-	0.1608	-	
		BRCA1	c.3448C>T, p.Pro1150Ser	chr17: 41244100	PP3, BP6	0.0085 / 0.11	0.0032 / 0.0642	0.3215	-	

		MLH1	c.649C>T, p.Arg217Cys	chr3: 37053562	PP3, BP6	0.03 / 0.38	0.0096 / 0.1284	0.4823	-	
82	M/58	*MSH6*	c.4068_4071 dup, p.Lys1358AspfsTer2	chr2: 48033981	PVS1, BS1	0.24 / 3.19	0.1504 / 3.013	-	-	
83	M/47	*ATM*	c.1823T>A, p.Leu608Gln	chr11: 108123564	PM2	-	-	-	-	
84	M/71	*RET*	c.-6C>G	chr10: 43572701	PM2	-	-	-	-	
85	M/67	*BRCA2*	c.7052C>G, p.Ala2351Gly	chr13: 32929042	BP6	0.01 / 0.13	0.0064 / 0.1282	0.4823	-	
		BRIP1	c.2258A>G, p.Asp753Gly	chr17: 59820495	N/A	0.0028 / 0.04	-	0.2412	-	
86	M/62	*TSC2*	c.1939G>A, p.Asp647Asn	chr16: 2121610	PP3, BP6	0.04 / 0.03	0.0446 / 0	0.1608	-	
		FLCN	c.1177-9C>T	chr17: 17119826	N/A	-	-	0.0804	gnomAD Korean 0	PM2 (VUS)
		FANCM	c.682-6del	chr14: 45609829	N/A	0.0009 / 0.01	-	-	PM2	PM2 (VUS)
87	M/63	*GALNT12*	c.1015A>T, p.Asn339Tyr	chr9: 101597628	PM2	-	-	-	PP3	PM2 +PP3 (VUS)

		MSH2	c.14C>A, p.Pro5Gln	chr2: 47630344	PP3	0.0081 / 0.1	0.0032 / 0.0643	-	-	
		MSH6	c.4068_4071 dup, p.Lys1358AspfsTer2	chr2: 48033981	PVS1, BS1	0.24 / 3.19	0.1504 / 3.013	-	-	
88	M/38	*CDH1*	c.2494G>A, p.Val832Met	chr16: 68867247	BS1	0.03 / 0.27	0.0032 / 0	0.8842	PP3	PP3 +BS1 (VUS)
89	M/67	*MSH6*	c.4068_4071 dup, p.Lys1358AspfsTer2	chr2: 48033981	PVS1, BS1	0.24 / 3.19	0.1504 / 3.013	-	-	

*Abbreviations: ACMG, American College of Medical Genetics and Genomics; chr: chromosome; E.Asian, East Asian; ExAC, Exome Aggregation Consortium; F, female; FAP, familial adenomatous polyposis; gnomAD, genome aggregation database; KRGDB, Korean Reference Genome Database; LB, likely benign; LP, likely pathogenic; M, male; N/A, not available; VUS, variant of unknown significance.

[1]Population frequencies are converted to % and rounded off to the 4th digit below the decimal point.

with multiple adenomas on colonoscopy (Table 8, patients 2, 11, 12, 70-89). A missense variant of *ALK* c.487G>T, p.Val163Leu was initially classified as VUS but likely benign in multiple lines of in silico algorithms added evidence of BP4 and could be reclassified as a likely benign variant, according to the ACMG criteria (Table 8, patient 74). One frameshift variant was detected in 5 different patients (*MSH6* c.4068_4071dupGATT, p.Lys1358AspfsTer2) but the allele frequency was greater than 1% in both the ExAC and gnomAD population data (Table 8, patients 73, 77, 82, 87, 89). The *MSH6* variant was also detected in two additional patients with adenomas on colonoscopy and the clinicopathologic findings are described in Table 9.

Overall, 44 patients with negative *APC* gene mutations and adenomas, multiple adenomas, or numerous adenomas on colonoscopy were evaluated for VUS on the NGS panel. Variants were found in 40 different genes with *MSH6* c.4068_4071dupGATT, p.Lys1358AspfsTer2 frameshift variant as the most frequently detected gene (Figure 7). *MLH1* and *SLX4* gene variants were detected 6 times each, and *POLE* and *POLD1* gene variants were found 5 times each (Figure 7).

Table 9. Clinicopathologic features of *APC* gene mutation negative FAP suspected patients with the most commonly noted VUS, *MSH6* c.4068_4071dupGATT, p.Lys1358AspfsTer2 (rs267608142, rs55740729, PVS1+BS1) variant.

No	Sex/ Age	Diagnosis	Family history	Gene	Other gene variants	ACMG classification
73	M/53	Multiple adenomas	Mother: stomach cancer	*POLE*	c.6135C>T, p.2045= (rs368662693)	VUS (BP7)
77	M/52	Multiple adenomas	Unknown	*SDHB*	c.541-3C>T (rs751920183)	-
				PTCH1	c.3964G>A, p.Ala1322Thr	VUS (PM2+PP3)
				FANCM	c.925G>A, p.Glu309Lys (rs778377621)	-
82	M/58	Multiple adenomas	Unknown	-	-	-
87	M/63	Multiple adenomas	Mother: colon cancer	*GALNT12*	c.1015A>T, p.Asn339Tyr	VUS (PM2)
				MSH2	c.14C>A, p.Pro5Gln (rs56170584)	VUS (PP3)
89	M/67	Multiple adenomas	Mother: colon cancer	-		-
90	M/60	Adenomas	Mother: colon cancer	*ATM*	c.5063T>C, p.Ile1688Thr (rs199836342)	-
				SDHB	c.-6G>A (rs2295056)	-
91	F/73	Adenomas, Rectal cancer	Father: stomach cancer	*BMPR1A*	c.713G>A, p.Arg238Gln (rs191742018)	VUS (BP6)
				TSC2	c.4046C>T, p.Ala1349Val (rs201979616)	VUS (BP6)
			Mother: Liver cancer	*CDH1*	c.2494G>A, p.Val832Met (rs35572355)	VUS (BS1)
				RNF43	c.2268A>G, p.Pro756= (rs754064531)	VUS (BP7)

*Abbreviations: ACMG, American College of Medical Genetics and Genomics; F, female; FAP, familial adenomatous polyposis; M, male; VUS, variant of unknown significance.

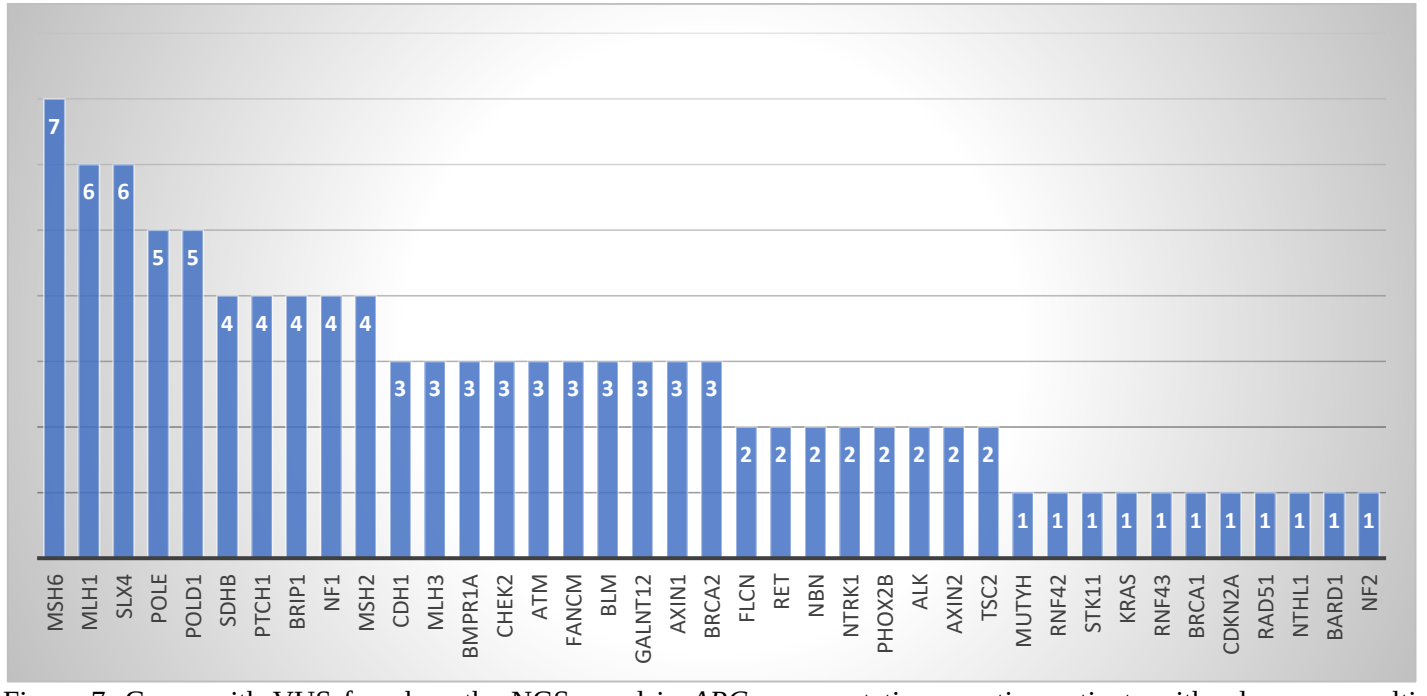

Figure 7. Genes with VUS found on the NGS panel in *APC* gene mutation negative patients with adenomas, multiple adenomas, or numerous adenomas (n=44). The most frequently detected variant was in the *MSH6* with *MLH1*, *SLX4*, *POLE*, and *POLD1* also commonly detected among patients suspicious for FAP.

IV. DISCUSSION

In our study, we were able to directly compare the conventional Sanger sequencing method with the NGS test in 12 patients that had duplicate tests. Many of the patients had been tested with the conventional method before the introduction of NGS into the clinical laboratory and a sequential test with the NGS panel was done in patients who were clinically highly suspected of having FAP. Whole gene deletion or *APC* gene mosaicism seen in three of the duplicate-tested patients could not have been found by the conventional method due to the limitation of the direct sequencing method (Table 3). In addition, five more patients were found to have whole gene deletions or exon deletions with the NGS panel. These patients would have also been found to have negative *APC* gene mutation studies with the conventional method, which may suggest the need for a sequential study with multiplex ligation-dependent probe amplification (MLPA) or methods to detect copy number variations to check for large deletions or duplications of the *APC* gene in patients with negative sequencing results but clinically suggestive of FAP. A patient was initially tested with the DHPLC screening method which showed negative results and no further testing with sequencing as was protocol (Table 3, patient 1). However, the patient was tested again with the NGS panel 10 years later, which revealed a pathogenic nonsense mutation. This discrepancy could be explained by the absence of a heterozygous peak which would result in a false negative DHPLC result. A patient who received total proctocolectomy due to adenocarcinoma and multiple tubular adenomas (>100) showed negative *APC* gene results for direct sequencing and NGS panel (Table 3, patient 9). No other variants were detected on the NGS panel.

Further evaluation may be warranted in such patients who are clinically consistent with FAP but do not show any associated genetic mutations.

APC gene mutations are known to have a relatively high incidence of de novo mutations and somatic mosaicism have been considered to have a considerable component of the sporadic FAP patients. Somatic mosaicism is difficult to identify with the conventional direct sequencing method due to its limitations in detecting allele frequencies below 15-20%. NGS testing could overcome this limitation and be able to detect these variations relatively easily. There were nine somatic mosaicism cases with the NGS panel within a span of 3.5 years. Many of the patients with negative *APC* gene mutation study results who were tested before the introduction of the NGS may consider getting tested again to confirm a clinical diagnosis of FAP.

The introduction of the ACMG guidelines for interpretation of variants was revolutionary, especially with the increasing amount of accumulated sequencing data with extensive use of NGS in the clinical and research setting. However, it is not always possible to easily apply the classification system to all variants, which may have been one of the reasons to develop further recommendations. ClinGen is a genome-based organization that is funded by the National Institutes of Health (NIH) and works to develop and support genomic resources to ultimately improve patient care (https://clinicalgenome.org/). As one of many working groups of the ClinGen, the sequence variant interpretation working group (SVI WG) provides general recommendations to help with the application of the ACMG criteria by consulting with expert panel groups and various professionals involved in biocuration. The first version of the SVI

recommendation for the PM2 criterion was recently approved. The recommendation proposes the PM2 criterion be downgraded to a supporting level, due to concerns that absent or rare variants in the general population is given too much weight. With the use of this criterion, a new rule for classification of "likely pathogenic" should be developed for cases of PVS1+1 supporting criterion. The novel *APC* gene mutations were all classified as likely pathogenic with a PVS1+PM2 combination (Table 4). With the proposed update, the novel mutations will still be able to be classified as a likely pathogenic variant. In addition, it is important that the population-based evidence is used in datasets comprised of unrelated individuals with at least 2,000 observed alleles.[24] In addition, the BA1 or BS1 criteria may need to be re-examined in some cases, since population datasets in ExAC may not necessarily be composed mainly of healthy controls. Using the ACMG classification system is the foundation for a more standardized approach to genetic test reports, but laboratories should always consider the updates and revisions in their reporting process.

FAP is known to be associated with a cribriform-morular variant of thyroid carcinoma, a rare form of thyroid cancer.[25-26] Two patients were diagnosed with papillary microcarcinoma, cribriform morular variant in this study. One patient was diagnosed as FAP with a novel likely pathogenic variant (c.3610C>T, p.Gln1204Ter, hetero) in the *APC* gene (Table 4, patient 27). A cribriform morular variant was identified on the pathologic diagnosis for total thyroidectomy. The second patient was referred for *APC* gene testing after undergoing total thyroidectomy which revealed a cribriform morular variant type but was not further evaluated for presence of colon polyps. A synonymous variant

(c.4782A>G, p.Pro1594=, hetero) was detected, and was classified as a VUS (Table 5, patient 38). As thyroid nodules are relatively commonly detected by ultrasound screening, it is important to notify clinicians of the subsequent tests that might be required if the patient is diagnosed with the cribriform-morular variant of thyroid cancer.[25-26]

The presence of *APC* mosaicism has been reported often in the literature but the limitations in detection of the low level variants has been somewhat of a problem until now. With the use of NGS technology, detection of low level variants has become more feasible. The NGS algorithm in our study consisted of detecting copy number variants with ExomeDepth in the R package and in addition to the DxSeq Analyzer algorithm.[13] Low level variants identified by this protocol was confirmed by MEMO-PCR or direct sequencing. In the present study, 9 patients with suspected somatic mutations of the *APC* gene were detected. All 9 patients were either diagnosed with FAP or highly suspicious of FAP. Further studies confirmed the presence of the *APC* mutations at a low VAF, with identical *APC* mutations detected in tissue samples (when available). Since the peripheral blood is of mesoderm origin and the colonic mucosa is developed from the endoderm, the mutational event is postulated to have occurred during early embryogenesis, before the separation of the two layers.[27-30] Depending on the timing of the mutational event, the offspring may have increased risk of inheriting the mutation if a mosaic variant is also detected in the tissue, which may entail genetic counseling in families of probands. The incidence of *APC* mosaicism was reported to be near 11%, which was similar to cases of somatic mosaicism in tuberous sclerosis, hemophilia A, and retinoblastoma (6-13%).[31-33]

Likely pathogenic mutations in the *BMPR1A* gene were detected in 4 patients with multiple to numerous adenomas and negative *APC* gene mutations (Table 7, patients 4, 55, 58, 59). The *BMPR1A* gene is associated with Hereditary mixed polyposis syndrome (HMPS) and juvenile polyposis syndrome (JPS).[34-35] The clinical features of HMPS include mixed hyperplastic, adenomatous, and atypical juvenile polyps in the colon. The presence of adenomatous polyps admixed with hyperplastic and juvenile polyps may lead to misdiagnosis of this disease as FAP.[36] In clinical practice, it may be difficult to discriminate adenomatous polyps from juvenile polyps with dysplasia. Although multiple to numerous adenomas were suggestive of FAP, these patients with *BMPR1A* gene mutations should be diagnosed as JPS. Biallelic mutations in the *MUTYH* gene is associated with *MUTYH*-associated polyposis (MAP).[37] MAP shows an autosomal recessive inheritance pattern, with a relative mild disease phenotype. 10-30% of *APC* gene mutation negative patients show a *MUTYH* mutation.[37-38] Although 2 patients harbored a likely pathogenic *MUTYH* gene mutation, both patients did not have biallelic mutations to make a diagnosis of MAP. In both cases, the 2nd *MUTYH* gene variation was a VUS. The proofreading-associated polyposis (PPAP) is associated with germline pathogenic mutations in the polymerase epsilon (*POLE* and polymerase delta (*POLD1*) exonuclease domain.[39-41] The patient in our study had a frameshift mutation of the *POLE* gene, which was easily diagnosed as likely pathogenic (Table 7, patient 56). However, *POLE* or *POLD1* missense variants can be challenging to interpret in terms of pathogenicity.

V. CONCLUSION

In conclusion, we evaluated the mutation spectrum of 420 Korean patients with multiple colon adenomas referred for *APC* gene mutations. On retrospective analysis of the *APC* variations, there were discrepancies in the results due to limitations in various methods. With genetic testing becoming more popular, there has been a steep increase in the accumulation of genetic mutations in the database, which allows for continuous attention to VUS, which should be re-evaluated periodically for reclassification. The use of NGS panels in detection of genetic mutations in FAP patients is a good alternative, especially with a robust algorithm that could detect a wide array of genetic variations. With the advent of NGS during recent years, the screening of genetic aberrations has become more prevalent as well as become a significant method in the diagnostic process. The continued testing and expansion of mutational studies that are accompanied by clinical significance will help the diagnosis of hereditary neoplasms.

REFERENCES

1. Bisgaard ML, Fenger K, Bulow S, Niebuhr E, Mohr J. Familial adenomatous polyposis (FAP): frequency, penetrance, and mutation rate. Hum Mutat 1994;3:121-5.

2. Half E, Bercovich D, Rozen P. Familial adenomatous polyposis. Orphanet J of Rare Dis 2009;4:22.

3. Nieuwenhuis MH and Vasen HF. Correlations between mutation site in *APC* and phenotype of familial adenomatous polyposis (FAP): a review of the literature. Crit Rev Oncol Hematol 2007;61:153-61.

4. Sheng JQ, Cui WJ, Fu L, Jin P, Han Y, Li SJ, et al. *APC* gene mutations in Chinese familial adenomatous polyposis patients. World J Gastroenterol 2010;16:1522-6.

5. Kim DW, Kim IJ, Kang HC, Park HW, Shin Y, Park JH, et al. Mutation spectrum of the *APC* gene in 83 Korean FAP families. Hum Mutat 2005;26(3):281.

6. Cao X, Hong Y, Eu KW, Loi C, Cheah PY. Singapore Familial Adenomatous Polyposis (FAP) patients with classical adenomatous polyposis but undetectable *APC* mutations have accelerated cancer progression. Am J Gastroenterol 2006;101:2810-7.

7. 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 2008;6:10.

8. Chiang JM, Chen HW, Tang RP, Chen JS, Changchien CR, Hsieh PS, et al. Mutation analysis of the *APC* gene in Taiwanese FAP families: low incidence of

APC germline mutation in a distinct subgroup of FAP families. Fam Cancer 2010;9:117-24.

9. Kerr SE, Thomas CB, Thibodeau SN, Ferber MJ, Halling KC. *APC* germline mutations in individuals being evaluated for familial adenomatous polyposis. J Mol Diagn 2013;15(1):31-43.

10. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015;17:405-24.

11. Myoshi Y, Ando H, Nagase H, Nishisho I, Horii A, Miki Y, et al. Germ-line mutations of the *APC* gene in 53 familial adenomatous polyposis patients. Proc Natl Acad Sci USA 1992;89:4452-6.

12. Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 2000;132:365-86.

13. Plagnol V, Curtis J, Epstein M, Mok KY, Stebbings E, Grigoriadou S, et al. A robust model for read count data in exome sequencing experiments and implications for copy number variant calling. Bioinformatics 2012;28:2747-54.

14. Lee ST, Kim JY, Kown MJ, Kim SW, Chung JH, Ahn MJ, et al. Mutant enrichment with 3'-modified oligonucleotides a practical PCR method for detecting trace mutant DNAs. J Mol Diagn 2011;13:657-68.

15. Sim NL, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC. SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic Acids Research 2012:40:W452–7.

16. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-

synonymous variants on protein function using the SIFT algorithm. Nat Protoc 2009; 4(7):1073-81.

17. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nat Methods 2010 Apr;7(4):248-9.

18. Schwarz JM, Rödelsperger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. Nat Methods 2010 Aug;7(8):575-6.

19. Reva B, Antipin Y, Sander C. Predicting the functional impact of protein mutations: application to cancer genomics. Nucleic Acids Res 2011 Sep 1; 39(17):e118.

20. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic Acids Res 2018;47(D1):D886-94.

21. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet 2014 Mar;46(3):310-5.

22. Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, et al. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. Am J Hum Genet. 2016 Oct 6;99(4):877-85.

23. Kim A, Kim HK, Ahn S, Hong YS, Lim SB, Byeon JS, et al. Identification of a novel splice variant (c.423-8A>G) of *APC* by RNA sequencing. Ann Lab Med 2020;40:345-7.

24. Ghosh R, Harrison SM, Rehm HL, Plon SE, Biesecker LG, ClinGen Sequence Variant Interpretation Working Group. Updated recommendation for the benign

stand-alone ACMG/AMP criterion. Hum Mutat 2018;39:1525-30.

25. Cameselle-Teijeiro JM, Peteiro-Gonzalez D, Caneiro-Gomez J, Sanchez-Ares M, Abdulkader I, Eloy C, et al. Cribriform-morular variant of thyroid carcinoma: a neoplasm with distinctive phenotype associated with the activation of the WNT/β-catenin pathway. Mod Pathol 2018;31:1168-79.

26. Akaishi J, Kondo T, Sugino K, Ogimi Y, Masaki C, Hames KY, et al. Cribriform-morular variant of papillary thyroid carcinoma: clinical and pathological features of 30 cases. World J Surg 2018;42:3616-23.

27. 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 2007;28:985-92.

28. Hes FJ, Nielsen M, Bik EC, Konvalinka D, Wijnen JT, Bakker E, et al. Somatic *APC* mosaicism: an underestimated cause of polyposis coli. Gut 2008;57:71-6.

29. Iwaizumi M, Tao H, Yamaguchi K, Yamada H, Shinmura K, Kahyo T, et al. A novel *APC* mosaicism in a patient with familial adenomatous polyposis. Hum Genome Var 2015;2:15057.

30. Out AA, van Minderhout IJHM, van der Stoep N, van Bommel LSR, Kluijt I, Aalfs C, et al. High-resolution melting (HRM) re-analysis of a polyposis patients cohort reveals previously undetected heterozygous and mosaic *APC* gene mutations. Fam Cancer 2015;14:247-57.

31. Verhoef S, Bakker L, Tempelaars AM, Hesseling-Janssen AL, Mazurczak T, Jozwiak S, et al. High rate of mosaicism in tuberous sclerosis complex. Am J Hum Genet 1999;64:1632-7.

32. Leuer M, Oldenburg J, Lavergne JM, Ludwig M, Gregin A, Eigel A, et al. Somatic mosaicism in hemophilia A: a fairly common event. Am J Hum Genet 2001;69:75-87.

33. Lohmann DR, Gerick M, Brandt B, Oelschlager U, Lorenz B, Passarge E, et al. Constitutional RB1-gene mutations in patients with isolated unilateral retinoblastoma. Am J Hum Genet 1997;61:282-94.

34. Larsen Haidle J, Howe JR. Juvenile Polyposis Syndrome. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, editors. GeneReviews®. Seattle: University of Washington;1993-2020 (Available from: https://www.ncbi.nlm.nih.gov/books/NBK1469/, Updated 2017 Mar 9).

35. Cichy W, Klincewicz B, Plawski A. Juvenile polyposis syndrome. Arch Med Sci 2014;10:570-7.

36. Gao XH, Li J, Zhao ZY, Xu XD, Du YQ, Yan HL, et al. Juvenile polyposis syndrome might be misdiagnosed as familial adenomatous polyposis: a case report and literature review. BMG Gastroenterol 2020;20:167.

37. D'Elia G, Caliendo G, Casamassimi A, Cioffi M, Molinari AM, Vietri MT. *APC* and *MUTYH* analysis in FAP patients: a novel mutation in *APC* gene and genotype-phenotype correlation. Genes 2018;9:322.

38. Mazzei F, Viel A, Bignami M. Role of *MUTYH* in human cancer. Mutat Res 2013;743-744:33-43.

39. Palles C, Cazier JB, Howarth KM, Domingo E, Jones AM, Broderick P, et al. Germline mutations affecting the proofreading domains of *POLE* and *POLD1* predispose to colorectal adenomas and carcinomas. Nat Genet 2013;45:136-44.

40. Esteban-Jurado C, Gimenez-Zaragoza D, Munoz J, Franch-Exposito S,

Alvarez-Barona M, Ocana T, et al. *POLE* and *POLD1* screening in 155 patients with multiple polyps and early-onset colorectal cacner. Oncotarget 2017;8:26732-43.

41. Mur P, Garcia-Mulero S, del Valle J, Magraner-Pardo L, Vidal A, Pineda M, et al. Role of *POLE* and *POLD1* in familial cancer. Genet Med 2020 (Epub ahead of print. PMID: 32792570).

ABSTRACT (IN KOREAN)

가족성 선종성 용종증 의심 환자의 유전자 검사를 통한 돌연변이 양상 분석

<지도교수 이경아>

연세대학교 대학원 의학과

박서진

가족성 선종성 용종증은 대장암의 약 1%를 차지하는 상염색체 우성 유전질환이다. 10대 이후부터 100개 이상의 선종성 용종이 대장에 나타나며, 다발성으로 발생하여 대장암으로 진행의 위험이 있는 질환이다. 질환과 연관된 *APC* 유전자는 종양억제 유전자로, 염기서열의 변이 또는 결실 등으로 단백질이 조기종료되는 현상이 흔히 나타나며, 이는 다발성 용종의 발생을 억제하는 능력을 상실하게 된다. 본 연구에서는 한국인에서 가족성 선종성 용종증 의심환자 420명을 대상으로 15년간 시행한 염기서열분석법과 차세대염기서열분석법의 결과에서 돌연변이 양상을 분석하였다. 약 39.8%의 환자에서 *APC* 유전자 돌연변이가 발견되었고, 기존 보고된 바와 같이 틀이동돌연변이와 무의미돌연변이가 가장 흔하게 나타났다. 기보고 없는 돌연변이가 가족성 선종성 용종증 환자 17명에서 발견되었고, *APC* 유전자의 불확실성변이(VUS)가 19명에서 보였다. 미국의학유전학회 (ACMG)의 지침을 재분석하여 1명은 “likely pathogenic” 돌연변이로 재분류하였고, 6명은 양성변이로 재분류하였다. 차세대염기서열검사를 시행하였을 때 기존 염기서열분석법보다 더 많은 변이를 발견하였는데, 넓은 범위의 염기결손 또는 낮은 비율의 돌연변이를 찾아낼 수 없는 것은 기존 검사법에서 보일 수 있는 한계이다. 차세대염기서열검사를 시행한

환자 8명에서 *APC* 유전자는 정상이었지만 *BMPR1A, MUTYH, PMS1, POLE* 유전자의 돌연변이를 발견하였고, 추가로 35명의 환자에서 다양한 유전자에서 불확실성변이가 나타났다. 이 중 7명의 환자에서 *MSH6* 유전자의 틀이동변이(c.4068_4071dupGATT, p.Lys1358AspfsTer2)가 발견되었는데, 상반된 결과의 문헌들로 인해 재분류를 할 수 없었고, *ALK* 유전자에서 보인 불확실성변이 1개는 양성변이로 재분류되었다. 가족성 선종성 용종증 환자의 진단에는 유전자 검사가 필수적인 요소가 되었고, 차세대염기서열검사법의 도입으로 기존 검사법의 한계로 인해 검출될 수 없었던 다양한 종류의 변이를 발견할 수 있어 진단에 도움이 될 수 있다. 앞으로도 이와 같은 유전성 종양질환의 진단에 중요한 역할을 유지하기 위해서는 방대한 양의 정보를 처리하는 차세대염기서열검사법은 분석력이 입증된 생물정보학적 접근이 필요하며, 임상 정보와 ACMG 지침의 개정사항들에 대한 주기적인 재분석이 필요하다.

핵심되는 말 : *APC* 유전자, 가족성 선종성 용종증, 유전자 돌연변이검사, 차세대 염기서열검사