Received: 9 January 2021 | Revised: 10 March 2021 | Accepted: 11 March 2021

DOI: 10.1002/jcla.23768

RESEARCH ARTICLE

WILEY

Mutational screening through comprehensive bioinformatics analysis to detect novel germline mutations in the *APC* gene in patients with familial adenomatous polyposis (FAP)

Faranak Ghadamyari[1] | Mohammad Mehdi Heidari[1] | Sirous Zeinali[2] | Mehri Khatami[1] | Shahin Merat[3] | Hamideh Bagherian[4] | Leili Rejali[4] | Farzaneh Ghasemi[1]

[1]Department of Biology, Faculty of Science, Yazd University, Yazd, Iran

[2]Biotechnology Research Center, Pasteur Institute of Iran, Tehran, Iran

[3]Digestive Disease Research Center, Tehran University of Medical Sciences, Tehran, Iran

[4]Medical Genetics Laboratory, Kawsar Human Genetics Research Center, Tehran, Iran

Correspondence
Mohammad Mehdi Heidari, Department of Biology, Faculty of Science, Yazd University, Yazd, Iran.
Email: heidarimm@yazd.ac.ir

Abstract

Background: Familial adenomatous polyposis (FAP) as a colon cancer predisposition syndrome is an autosomal-dominant inherited condition and is diagnosed by the progress of hundreds or thousands of adenomatous colonic polyps in the colon. This study aims at the nature and effect of Adenomatous Polyposis Coli (APC) gene mutations in FAP tumorigenesis.

Methods: The genetic screening of 59 FAP Iranian patients in 10 families was performed by polymerase chain reactions and the direct sequencing of the entire coding exons of the APC gene. To do linkage haplotype analysis and multiplex PCR-based microsatellite examination, six short tandem repeat loci were selected in this gene. To evaluate and predict the potentially deleterious effects, comprehensive bioinformatics pathogenicity assays were used.

Results: A total of 12 germline heterozygous and homozygous nucleotide variations were identified. They included two missense mutations, four nonsense mutations, which would lead to the truncated and nonfunctional protein products, four synonymous or silent variations, and two nucleotide deletions of 1 to 5 bp or frameshift mutations. In addition, three novel heterozygous nonsense mutations were found in exons 10, 14, and 15 of the gene. There was also p.Arg653Met as a novel heterozygote mutation in exon 14 of the gene.

Conclusions: Bioinformatics analysis and three-dimensional structural modeling predicted that these missense and nonsense mutations generally are associated with the deleted or truncated domains of APC and have functional importance and mainly affected the APC protein. These findings may provide evidence for the progress of potential biomarkers and help to understand the role of the APC gene in FAP.

KEYWORDS
APC, cancer, familial adenomatous polyposis, germline mutation

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
© 2021 The Authors. *Journal of Clinical Laboratory Analysis* published by Wiley Periodicals LLC

1 | INTRODUCTION

Familial adenomatous polyposis (FAP), the most common gastrointestinal syndrome, is identified with multiple adenomatous polyps in the colon and rectum.[1-3] FAP is an inherited syndrome with an autosomal dominant manner, and its first manifestation occurs in teenage years.[4] As a colon cancer predisposition syndrome, FAP accounts for 1% of colorectal carcinoma (CRC) cases.[5] The worldwide incidence rate of FAP is 3–10 cases per 10 000–100 000 live births.[1,6,7] The risk of extra-colonic malignancies such as medulloblastoma, childhood hepatoblastoma, duodenal ampullary cancer, and follicular or papillary thyroid cancer increases in FAP patients.[8] Recently, other cases than FAP have also been designated as adenomatous polyposis syndromes. They include the autosomal recessive MUTYH-associated polyposis (MAP) syndrome, the autosomal recessive NTHL1-associated polyposis (NAP) syndrome, and the autosomal dominant polymerase proofreading-associated polyposis (PPAP) syndrome created through mutations in POLD1 and POLE genes.[5]

The majority of patients undergo germline mutations in the Adenomatous Polyposis Coli (APC) gene located at 5q21–q22.[9] Then, based on the two-hit theory, there may be next acquired mutations to inactivate the second native allele and lead to loss of heterozygosity (LOH) of *APC*. Finally, the LOH of *APC* causes polyps in the epithelial cells in the large intestine.[10] *APC* is a tumor suppressor gene and encodes a major protein with 2843 amino acids consisting of several motifs and domains that play very important roles in transcriptional activation, cell adhesion, cell migration, polarization, and apoptosis.[11] The main role of the APC protein in the cytoplasm is to regulate the level of the β-catenin protein. Mutations of the APC gene lead to the formation of a dysfunctional or truncated APC protein that increases the level of the β-catenin protein in the cytoplasm. Raised levels of the β-catenin protein in association with activated transcription factors such as, Tcf, cause the Wnt signaling pathway to be aberrant, leading to uncontrolled cell proliferation, and the development of cancer. More than 1000 germline mutations have been identified in the *APC* gene in FAP patients.[12] Point mutations in this gene were detected in about 60%–70% of FAP cases, while the large deletion (frameshift) mutations of the gene were identified in 10%–15% of those patients.[13] In some FAP cases, there is no APC mutation; instead, there are germline mutations in the MUTYH gene. Genetic analysis shows that only a small proportion of cases are explained by the recently described NAP and PPAP syndromes.[14] So far, only a few studies have been reported on Iranian FAP patients, and the main problems in these studies are small sample sizes and low diagnostic rates.[15,16] In this study, to better understand the APC gene mutations in Iranian FAP patients, a total of 10 FAP families (including proband cases and all the affected members) were assessed through direct sequencing associated with linkage haplotype analysis and comprehensive bioinformatics pathogenicity assays.

2 | MATERIALS AND METHODS

2.1 | Patients

In this research, 59 patients, 29 females and 30 males, with a confirmed diagnosis of classic FAP were enrolled from 10 Iranian families. These cases with a mean age of 28.8 years (ranging from 10 to 51 years, and the median age of 31.7 years at diagnosis) had referred to Imam Khomeini (Cancer Institute), Shariati (Gastroenterology Department), and Khatamolanbia (Pathology Department) hospitals in Tehran from May 2017 to April 2019. The accurate family history and the clinical features of all the affected cases were collected. The FAP patients were identified based on positive family history and clinical findings in colonoscopy and pathology. The typical diagnostic criteria of classical FAP depend on a dominant inheritance in family history and the identification of >100 adenomatous polyps in the colon and the rectum (Figure 1). As a control group, 55 unrelated healthy individuals with age- and sex-matched criteria (9–49 years: $p = 0.37$, and 28 male and 27 female: $p = 0.66$) were selected from the same geographic region and without any known familial history of cancer. In both the FAP and the healthy normal groups, gender distribution was somehow similar; there were 54 (51.92%) males and 60 (57.69%) females.

All of the protocols in this study were practiced in accordance with the guidelines of the Ethics Committee of Yazd University and the ethical standards of the Declaration of Helsinki (2000 and 2008) developed by the World Medical Association. After written informed consents were obtained from all the participants, during genetic counseling, genomic DNA was extracted from 5 ml of peripheral blood samples belonging to each patient, family member, and healthy case, and purified by using a DNA isolation kit (Qiagen) and according to the manufacturer's protocols.

2.2 | Mutation analysis

Polymerase chain reactions and direct sequencing were performed to detect the point nucleotide variations in the APC gene by the use of 26 pairs of specific primers, which were evaluated by the primer design software (Primer Premier 5.0). The possible secondary structures of the primers were surveyed with Gene Runner version 3.05 (http://www.generunner.com). The length of the PCR products ranged from 376 to 874 bp. Mutational analysis was accomplished with amplification of 14 exons, 12 fragments of exon 15, and the adjacent intronic regions of the APC gene to screen the entire coding region for mutations (Table S1).

All the polymerase chain reactions were performed in 25 μl of a reaction mixture that contained 5 pmol of each forward and reverse primer, 5 μl of a 1×polymerase chain reaction buffer (15 Mm magnesium chloride, 100 mM Tris HCl, Ph 8.8, 500 Mm potassium chloride, and 0.1% Triton X-100), 0.8 mM of each dNTP, 0.5 U Taq polymerase (Qiagen) and 100–200 ng of the total genomic DNA.

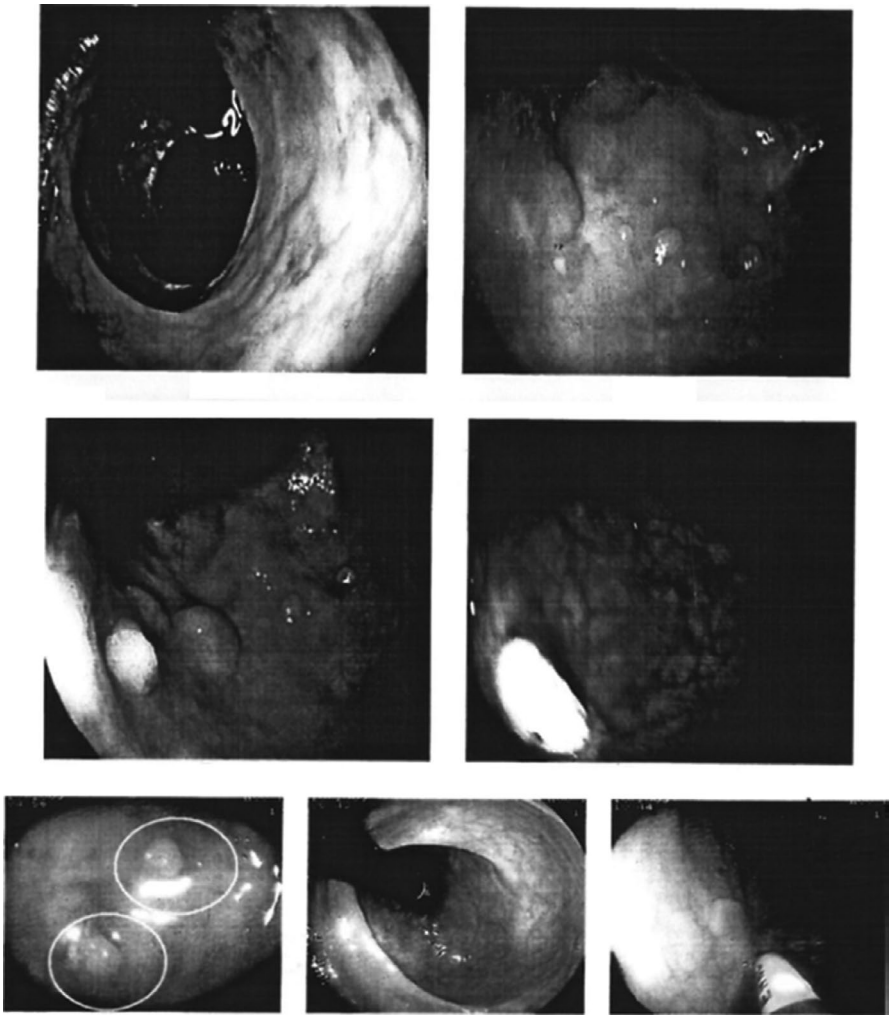

FIGURE 1 Colonoscopy images of several patients: There are nearly 100 polyps in the sigmoid colon and rectum (lower rectum and upper rectum), the polyps of the rectum are transformed to malignant, and the sigmoid colon consists of multiple polyps

The PCR was carried out in an automatic thermal cycler (Techne) and programmed under certain conditions including: preliminary denaturation at 94°C for 5 min, 30–35 cycles at 94°C for 1 min, recommended annealing temperatures of 58–63°C for 1 min, elongation at 72°C for 1 min, and an ultimate extension at 72°C for 10 min. The PCR products were electrophoresed on a 1–2% agarose gel, studied by direct sequencing, and then examined by the Geneious software. The nucleotide variations in the APC gene were named and referred to according to the nomenclature standard used by Human Genome Variation Society, NCBI, and Human Single Nucleotide Polymorphism database. Then, the variations were further verified with a search in other databases as well as the literature.

2.3 | Linkage analysis

The Tandem Repeat Finder software was used to find short tandem repeats (STRs). It suggested a list of STRs for the APC gene analysis. Based on several parameters such as the number of repetitions (at least 10), percentage of matches (100%), percentage of indels (0%), and entropy (1), six STR loci were selected for the Multiplex PCR-based microsatellite analysis of the APC gene (Table S2).

2.4 | *In Silico* analysis

The nucleotide and amino acid sequence alignments were carried out by the MEGA6 (multiple sequence alignment software), and the Blastp (Standard Protein Blast; http://www.ncbi.nlm.nih.gov/Blastp) program. To evaluate and predict the potentially deleterious effects and the pathogenicity scores of the detected missense and nonsense variants, those variants were analyzed by PolyPhen-2 (polymorphism phenotyping v2; http://genetics.bwh.harvard.edu/pph2/), SIFT (scale-invariant feature transform; http://sift.jcvi.org/), I-Mutant 2.0 (http://gpcr2.biocomp.unibo.it/cgi/predictors/I-Mutant2.0), P-Mut (http://mmb.irbbarcelona.org/PMut/), PANTHER (http://pantherdb.org/), CRAVAT (http://hg19.cravat.us/CRAVAT/) for missense mutations, ENTPRISE-X (http://cssb2.biology.gatech.edu), Mutpred-LOF (http://mutpred.mutdb.org/), and CRAVAT for nonsense mutations. To determine the hydrophobicity or hydrophilicity scales in the proteins, a plot was created by the ExPASy (Expert Protein Analysis System; http://web.expasy.org/protocol), and hydrophobicity indexes were calculated with the Peptide-2 prediction tool (www.peptide-2.com). Furthermore, to evaluate the potential alternative splicing effect of the synonym variations, HSF version 3.0 (Human Splicing Finder; http://umd.be/Redirect.html) and

NNSPLICE (https://www.fruitfly.org/seq_tools/splice.html) were used. A variation in at least two algorithms or databases was considered as having a damaging effect; it is referred to as a pathogenic mutation. Finally, the PyMol software (https://www.pymol.org/) was used for the imaging of the 3D shapes and the binding alterations in the structure of the protein.

2.5 | Statistical tests

Statistical analyses were carried out using Fisher's exact test version X2. *p*-value <0.05 was considered as statistically significant.

3 | RESULTS

3.1 | Mutation analysis

The entire coding sequence of the APC gene was sequenced in 59 related patients with FAP diagnosis in 10 families with a history of adenomatous polyposis. They were evaluated against 55 cancer-free control individuals. Finally, 12 typical nucleotide changes were identified with a higher substitution frequency in the FAP patients than in the healthy controls (Figures 2 and 3). The changes included two missense pathogenic variations (16.67%), four nonsense pathogenic variations, which lead to predicted truncated protein products (33.33%), four synonymous or silent variations (33.33%), and two nucleotide deletions of 1 to 5 bp or frameshift mutations (16.67%). Table 1 presents the types and distributions of the nucleotide variations identified in 10 FAP families. These variants occur alone in each patient, but not in combination with other variants and none of the patients had no several nucleotide variants, simultaneously. To assess the association between the rates of the pathogenic variations of the APC gene in the patients and the healthy controls, Fisher's exact test was employed. The result showed a statistically significant difference between patients and control cases (*p*-value = 0.000) (Table 2).

The nucleotide variations were detected between codons 463 and 1822 of the APC gene. Exon 15, as a hot spot region harbored 58.33% of all the mutations (one nonsense, two frameshifts, one missense, and three synonym variations), exon 14 presented 25% of the total pathogenic variations (two nonsense and one missense mutation), and exon 11 had only one synonymous variation (p. Tyr486 Tyr) (Figure 4). No other mutations were identified in the remaining exons. Of the nucleotide variations detected through direct sequencing, four were considered novel in that they have not been reported in the literature or in any database, yet. They included three nonsense mutations (p. Arg463Term, p. Glu601Term, and p. Lys1199Term) in six female patients and five affected males in five different families and a heterozygote missense mutation (p.

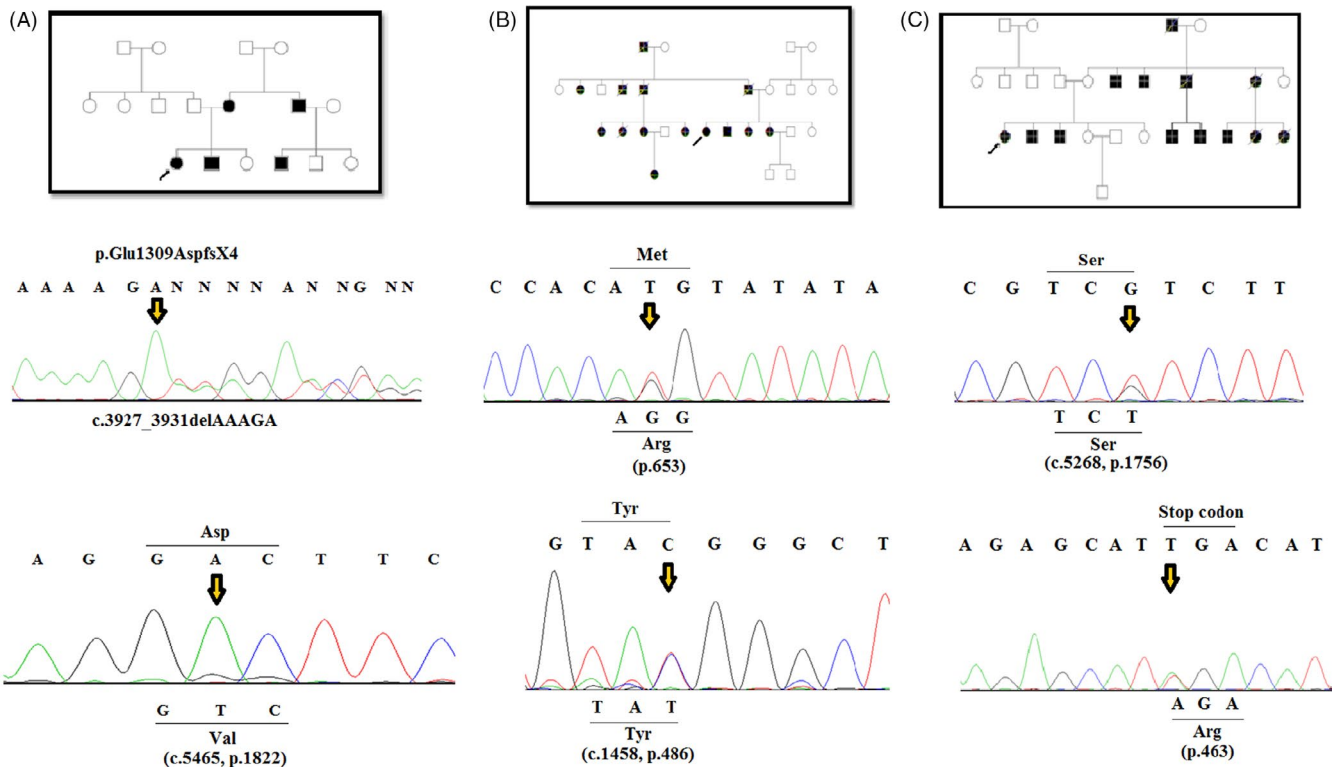

FIGURE 2 DNA sequencing of the mutations observed in the pedigrees of familial adenomatous polyposis cases (FAP). (A) Family 1 carrying the APC mutations c.3927_3931delAAAGA (p. Glu1309AspfsX4) and c.5465 GTC>GAC (p. Val1822Asp); (B) Family 2 carrying the APC mutations c.1958 AGG>ATG (p. Arg653Met) and c.1458 TAT>TAC (p. Tyr486 Tyr), and (C) Family 3 carrying the APC mutations c.5268 TCT>TCG (p. Ser1756Ser) and c.1387 AGA>TGA (p. Arg463Stop codon)

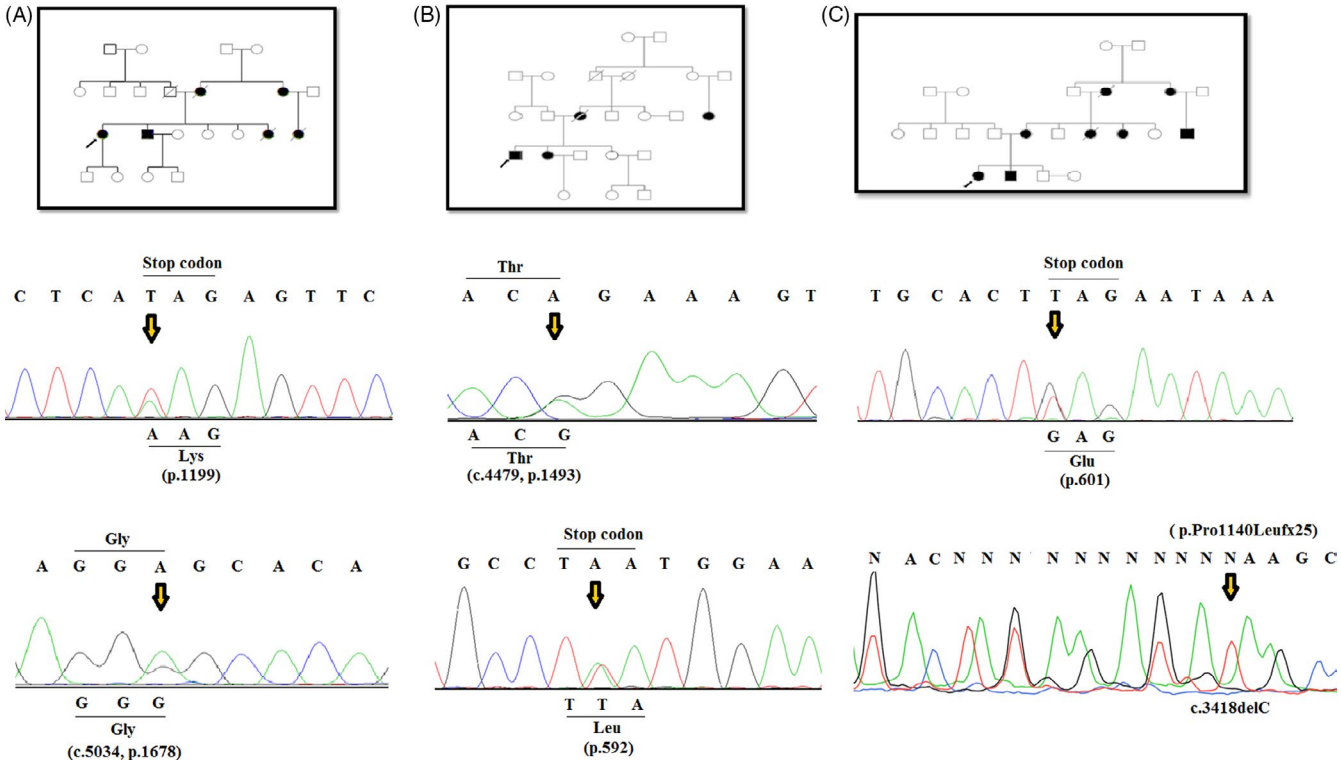

FIGURE 3 The heterozygous mutations in the FAP patients were validated using Sanger sequencing. (A) Family 4 carrying the APC mutations c.3595 AAG>TAG (p. Lys1199Stop codon) and c.5034 GGG>GGA (p. Gly1678Gly); (B) Family 6 carrying the APC mutations c.4479 ACG>ACA (p. Thr1493 Thr) and c.1775 TTA>TAA (p. Leu592Stop codon), and (C) Family 8 carrying the APC mutations c.1801 GAG>TAG (p. Glu601Stop codon) and c.3418delC (p. Pro1140Leufx25)

Arg653Met) in two sisters (15- to 25-year-old) and their brother (32-year-old) all of whom had colonoscopy-diagnosed FAP and belonged to family 2. These novel pathogenic variations were missing in the unaffected family members as well as all the 55 healthy controls.

3.2 | Genotype-phenotype correlation

In the patient cohort of this study, 11 nucleotide alterations were detected in the monoallelic (heterozygote) state in the APC gene, whereas one missense mutation (p. Val1822Asp) was observed in the biallelic (homozygote) state. Two heterozygous frameshift mutations (c.3418delC, p. Pro1140Leufx25, and c.3927-3931delAAAGA, p. Glu1309AspfsX4) localized in exon 15 involved the deletion of cytosine at the nucleotide position of 3418, which leads to a subsequent reading frameshift in the codons of this exon, and the deletion of AAAGA between the nucleotides from 3927 to 3931. These two frameshift mutations were observed in 10 related patients. There was c.3418delC, p. Pro1140Leufx25 mutation observed in a 48-year-old mother affected with intermediate FAP and in two of her sick children (one girl and one boy with FAP, diagnosed at the age of 20 and 18, respectively). They belonged to family 8. Also c.3927-3931delAAAGA, p. Glu1309AspfsX4 was detected in two families with the FAP phenotype. In family 1, there were a 30-year-old female patient who had surgical treatment with colectomy at the age of 20 and two male patients at 16 and 21. In family 9, there were two females at 19 and 31 years of age affected with FAP and two affected males at 29 and 41. The mutational analysis was extended to the unaffected and healthy family members in these families; no one had any deletion mutation.

The conservation indexes (CIs) were compared for missense and nonsense mutations and amino acid changes in the APC genes of different species (including Mus musculus: house mouse, Rattus norvegicus, Drosophila melanogaster, zebrafish, Canis lupus familiaris: dog, Bos taurus: cattle, Gallus: chicken, Pan troglodytes: chimpanzee, Callorhinchus milii: elephant shark, and Camelus ferus). The results showed that the missense and nonsense mutations (especially, p. Arg653Met and p. Val1822Asp) were highly conserved with a CI of at least 95%–100% (results not shown).

3.3 | Pathogenicity of the nucleotide variations

To evaluate the potentially damaging effects of mutations, pathogenicity-based studies were performed for nonsense, missense, and silent variations by several bioinformatics tools (Tables 3, 4, and 5). According to the results, p. Arg653Met (rs106050331) missense pathogenic variation was deleterious, and its pathogenicity scores were higher than 0.05. It was also found that this

TABLE 1 Point mutations and large fragment deletions of the APC gene germline mutation in 10 family members with familial adenomatous polyposis detected in this study

Family	Number of patients	Proband/age	Mutation/Exon	Homo/ Hetero	Type	Novel/ Reported
1	4: A sibling and their children (one girl and two boys)	Female/30	c.5465 GTC>GAC (p. Val1822Asp) /15	Homo	Missense	rs459552
			c.3927_3931delAAAGA (p. Glu1309AspfsX4) /15	Hetero	Frameshift	Reported
2	14: Five died and eight women and one sick man	Female/35	c.1458 TAT>TAC (p. Tyr486 Tyr) /11	Hetero	Synonym	rs2229992
			c.5034 GGG>GGA (p. Gly1678Gly) /15	Homo	Synonym	rs42427
			c.5268 TCT>TCG (p. Ser1756Ser) /15	Hetero	Synonym	rs866006
			c.1958 AGG>ATG (p. Arg653Met) /14	Hetero	Missense	Novel
3	13: Five died and 1one woman and seven sick men	Female/23	c.5268 TCT>TCG (p. Ser1756Ser) /15	Hetero	Synonym	rs866006
			c.4479 ACG>ACA (p. Thr1493 Thr) /15	Hetero	Synonym	rs41115
			c.1387 AGA>TGA (p. Arg463Stop codon) /10	Hetero	Nonsense	Novel
4	6: Three died and two women and one sick man	Female/25	c.3595 AAG>TAG (p. Lys1199Stop codon) /15	Hetero	Nonsense	Novel
5	3: Two brothers and one sister	Female/15	-	-	-	-
6	4: One died and two women and one sick man	Male/18	c.5465 GTC>GAC (p. Val1822Asp) /15	Homo	Missense	rs459552
			c.4479 ACG>ACA (p. Thr1493 Thr) /15	Hetero	Synonym	rs41115
			c.5034 GGG>GGA (p. Gly1678Gly) /15	Homo	Synonym	rs42427
			c.1775 TTA>TAA (p. Leu592Stop codon) /14	Hetero	Nonsense	Reported
7	11: One died and six women and four sick men	Female/18	-	-	-	-
8	8: Two died and four women and two sick men	Female/20	c.1801 GAG>TAG (p. Glu601Stop codon) /14	Hetero	Nonsense	Novel
			c.3418delC (p. Pro1140Leufx25) /15	Hetero	Frameshift	Reported
9	14: Two died and three women and nine sick men	Male/28	c.3927_3931delAAAGA (p. Glu1309AspfsX4) /15	Hetero	Frameshift	Reported
10	4: Three died and one sick woman	Female/17	c.5465 GTC>GAC (p. Val1822Asp) /15	Homo	Missense	rs459552
			c.1458 TAT>TAC (p. Tyr486 Tyr) /11	Hetero	Synonym	rs2229992
			c.1801 GAG>TAG (p. Glu601Stop codon) /14	Hetero	Nonsense	Novel

mutation leads to a substitution of arginine (a hydrophilic amino acid) with methionine (a very hydrophobic amino acid) at codon 653. Furthermore, p. Val1822Asp (re459552) missense mutation affects the polarity of the protein due to the replacement of a very hydrophobic amino acid (valine) with a hydrophilic amino acid (aspartic acid). The plots created by the ExPASy tool revealed obvious changes in protein hydrophobicity for both missense mutations, as compared with the wild type protein (Figure 5). Using prediction software programs such as Mutpred-LOF and ENTPRISE-X, all the four nonsense mutations in this study (p. Arg463Term, p. Leu592Term, p. Glu601Term, and p. Lys1199Term) were found to be pathogenic and disease-associated. To assess the potentially splicing effects of synonymous variations, *in silico* analyses were performed by different splice site prediction programs. Of the

TABLE 2 Association of *APC* gene mutations in the control cases and the FAP patients

	FAP patients (*n* = 56)	Control cases (*n* = 55)	*p*-value
Missense mutations (Frequency %)	15 (26.78)	0/ 0%	0.000
Nonsense mutations (Frequency %)	17 (30.35)	0/ 0%	0.000
Synonymous mutations (Frequency %)	18 (32.14)	5/ 9%	0.610
Frameshift mutations (Frequency %)	10 (17.85)	0/ 0%	0.000
p-value for the total mutations	0.000		

Note: p-value <0.05 was considered statistically significant.

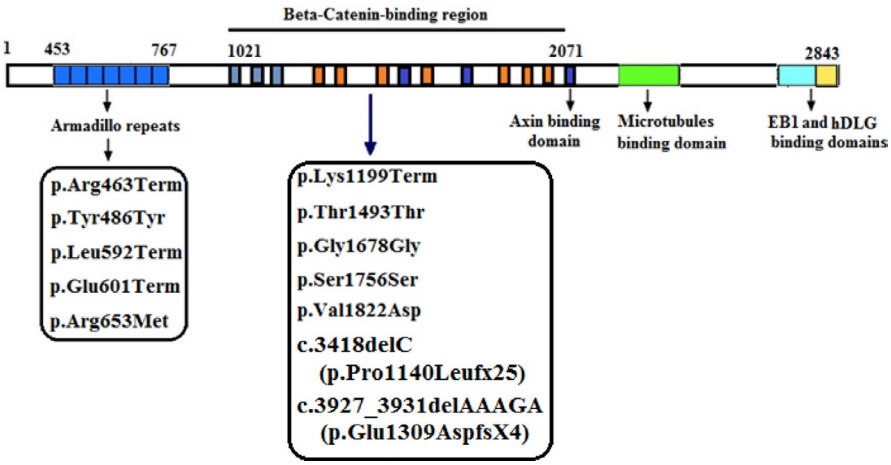

FIGURE 4 Position of the mutations identified in the functional domains of the APC protein in the familial patients with FAP

four nucleotide changes as synonyms, splicing alterations were predicted for only one (p. Thr1493 Thr: rs41115).

3.4 | Linkage analysis

To do haplotype analysis for the identified APC germline mutations, six STR markers (D5SUAPC16.79, D5SUAPC15.17, D5SUAPC9.1, D5SDAPC10.23, D5SDAPC10.66, and D5SDAPC16.57) were typed for the APC loci (Figure 6). The paternal and maternal haplotypes were defined based on Mendelian inheritance using informative markers. They were then compared among the affected offspring within the families. This analysis included individuals from two or three generations that had shared haplotype overlapping regions in the pedigrees. The haplotypes associated with nucleotide variations were analyzed for four unrelated families. Using the mentioned markers, strong co-segregation was detected in the APC gene of the parents and their affected offspring in families 8 and 9 (Figure 6C,D). In the other two families (families 5 and 7), however, no common haplotype was identified for the members without nucleotide variations; haplotypes were found in parents, but not in their affected children, and vice versa (Figure 6A,B). Pedigree 5 (Figure 6A) was a two-generation family with a co-segregating common haplotype. This haplotype is present in the unaffected father and segregates all three affected children in the family. Pedigree 7 (Figure 6B) is a three-generation family without a co-segregating common haplotype in four affected individuals. These results prove the existence of a segregation pattern of variants in relevant pedigrees. Through sequencing, potential sequence variants were found in all the pedigrees except families 5 and 7, and there were no nucleotide changes in the APC gene.

4 | DISCUSSION

APC, as a tumor suppressor gene, is a member of the destruction complex in the Wnt signaling pathway and the negative β-catenin regulator. Wnt signaling has a highly influential role in important cell pathways, such as cell proliferation and differentiation of intestinal stem cells, cell migration, cell adhesion, cytoskeletal reorganization, and apoptosis.[17,18] In the absence of Wnt ligand, β-catenin in the cytoplasm is phosphorylated. So, β-catenin is targeted for degradation via a destruction complex including Axin, APC, and kinases GSK3, CKIα. Due to the binding of the Wnt ligand with frizzled receptor (Wnt receptor) and its co-receptor, β-catenin is protected from phosphorylation and degradation. Then, β-catenin and DNA-binding factors lead to the transcription of several Wnt target genes.[19] Loss of function of APC, as in mutated APC, leads to β-catenin promotion and the overexpression of Wnt target genes and oncogenes.[11,20] An enormous number of pathogenic APC gene mutations have been reported and described in different populations of many countries.[21-23] The rate of point mutations (substitutions) in the FAP disease is

TABLE 3 Bioinformatics analysis of missense mutations and their correlation with the computational functional prediction tools

AA. change	SNP	SIFT	PROVEAN	PolyPhen−2 /Score	PANTHER	P-Mut	I-Mutant	CRAVAT	
								VEST	CHASM
p.R653 M	rs1060503318	APF/0.01[a]	Deleterious/−5.75	Probably Damaging/1.000	Probably Damaging	Deleterious/ 0.69	Decrease/−0.59	0.9	0.598
p.V1822D	rs459552	T/0.48	Neutral/ 0.59	Benign/0.000	Probably Benign	Neutral/0.00	Decrease/−0.38	0.031	0.572

[a]There is low confidence in this prediction. Notice that AA: Amino acid, T: Tolerated, APF: Affected Protein Function, CHASM/Score: From 0 (likely passenger) to 1 (likely driver). SIFT/Score: score <0.05 is predicted to be deleterious, PROVEAN: Score < −2.5: deleterious, P-Mut/Score: score >0.5 is predicted to be deleterious, I-Mutant/DDG: DDG <0: Decrease Stability and DDG >0: Increase Stability.

about 87% while the rate of frameshift, deletion, or insertion mutations in pathogenicity is just about 10%–15%.[6,24]

In the present study, among 59 familial DNA specimens, we reported 12 heterozygous and homozygous nucleotide variations in the APC gene of 10 Iranian families with FAP. We detected deleterious germline mutations in 86.6% of the patients; this mutation rate is higher than the range of 50%–70% designated for classical FAP in different European populations. [25-27] Most of the mutations in the APC gene are protein-truncating mutations including nonsense and deletion/insertion (frameshift) mutations, but whole-gene deletions or whole-exon deletions have not been observed,[25] which is consistent with the results of this study. Based on the variants indicated in the Human Gene Mutation Database (HGMD), four out of twelve nucleotide variations identified in the patients were described here for the first time with a frequency of about 45.76% (27/59). They included p. Arg463term, Glu601term, p.lys1199term, and p. Arg653Met.

Among 12 discovered hereditary variations, two frameshift mutations (p. Pro1140Leufx25 and p. Glu1309AspfsX4) were localized in exon 15 and at amino acid positions 1140 and 1309 of the APC gene. These deletion mutations had a frequency of 37.28% (22/59) and resulted in premature stop-codon formation. Also, a 5 nucleotide deletion at the hot spot codon 1309 (c.3927_3931delAAAGA) in 15% of the FAP cases is reported to be the most frequent germline mutations.[2,28,29] Many studies have also reported positive correlations between experimental manifestations and APC mutations.[30-32] Friedl et al,[33] Plawski et al,[34] and Kashfi et al[35] Investigated APC mutations in unrelated FAP patients and found different nucleotide deletions in hot spot codons. According to their results, nucleotide deletions at codons 1250–1464, mainly at codon 1309 (p. Glu1309AspfsX4), are responsible for the progression of colorectal cancer in a carrier at the age of 10, which was earlier than in patients carrying other mutations. The remarkable point about the genotype-phenotype correlation in our patients was that two deletion (frameshift) mutations in codons 1140 and 1309 in exon 15 were associated with the earliest age of onset in colorectal adenomas (polyps) without any extra-colonic manifestations. In our patients with these deletion mutations, the median age of onset was 20 years (range 8–31), while the average age at polyp diagnosis was 31.7 years (ranging from 15 to 58) and our results obtained for both groups were statistically significant (p = 0.005). Interestingly, these deletion mutations in our patients were frequently associated with the presence of >1000 polyps, but in the remaining patients with other point mutations, we did not observe such a phenotypic correlation.

Comparative studies of the genotype–phenotype correlations at the onset age of the adenomatous colorectal polyp development in other patients with other point mutations did not show a significant difference (p = 0.76). There is, however, a very important and alarming fact about our patients; those who had these nucleotide variations were diagnosed with a severe clinical phenotype of FAP at a high frequency (40/59, 67.79%). Moreover, in the patient group without mutations of the APC gene (families 5

TABLE 4 Potential pathogenic features of nonsense mutations in the APC gene identified in the protein-coding region

Nonsense mutations	Genomic location (Chr5)	SNP	Mutpred-LOF/Score	ENTPRISE-X/Score	CRAVAT/ VEST-Score
p. Arg463Term	g.112821970A>T	Novel	Pathogenic/0.61694	Disease-associated/0.90703	0.985
p. Leu592Term	g.112834982 T > A	Novel	Pathogenic/0.62191	Disease-associated /0.95401	0.975
p. Glu601Term	g.112835008G>T	Novel	Pathogenic/0.62144	Disease-associated /0.98371	0.954
p. Lys1199Term	g.112839189A>T	Novel	Pathogenic/0.61693	Disease-associated /0.94824	0.807

Note: Mutpred-LOF: Score>0.5 is the predicted pathogenic, ENTPRISE-X: score >0.5 is disease-associated.
Abbreviations: CRAVAT, Cancer-Related Analysis of Variants Toolkit; VEST, Variant Effect Scoring Tool.

TABLE 5 *In silico* study of synonymous APC variants: Scores represent the values of the splice donor or acceptor sites

Polymorphism	SNP	codon	Human splicing finder (HSF)	NNSPLICE/Score
p. Tyr486 Tyr	rs2229992	c.1458TAT>TAC	No difference was found (+ESE) Creation SRp55 site	No difference was found
p. Thr1493 Thr	rs41115	c.4479ACG>ACA	Creation acceptor site Disruption donor site Creation branch point site	Creation acceptor site/0.89
p. Gly1678Gly	rs42427	c.5034GGG>GGA	Creation acceptor site	No difference was found
p. Ser1756Ser	rs866006	c.5268TCT>TCG	No difference was found	No difference was found

Note: NNSPLICE: Splice site prediction for 1 sequence with the donor score cutoff of 0.40 and the acceptor score cutoff of 0.40.

and 7), 13.55% (8/59) of the cases had developed colorectal adenomas (polyps). These unexpected results on the relationship of mutations in the APC gene and the clinical manifestations can be due to the environmental influences and/or other modifying endogenous factors such as methylation of the promoter region of the gene.[36,37]

The other types of pathogenic and deleterious mutations observed in this study were four nonsense mutations (including p. Arg463term, p. Leu592term, Glu601term, and p.lys1199term) in 21 familial FAP cases (35.59%). Such mutations result in the formation of premature stop codons, and subsequently, most of the normal protein is lost in the truncating protein, it can be concluded that these nonsense mutations are deleterious. Also, through the *in silico* analysis conducted in this study, pathogenic effects were predicted for all these mutations in our patients. We did not observe these deleterious pathogenic variations in the healthy controls of the same ethnic origin, gender, and age range. Therefore, the patients and the control group were significantly different in this case ($p < 0.05$).

In addition, two missense mutations (p. Arg653Met and p. Val1822Asp) were identified in our patients. Indeed, p. Arg653Met mutation was observed in FAP patients for the first time. Previously, Campo et al. (2017) in an Italian population described that SNP rs459552 or p. Val1822Asp mutation in the APC gene. The mutation had occurred in a conserved amino acid position but it was predicted to be tolerated by SIFT and suggested to be benign by PolyPhen. In this respect, Picelli et al,[38] Cerasuolo et al,[39] and Fernández-Rozadilla et al[40] concluded that it was not a disease causing variant because it was found equally distributed among normal and FAP-affected individuals. Our analysis showed that only p. Arg653Met mutation is likely to have damaging and deleterious effects on the APC protein. Also, it was shown that p. Arg653Met mutations can alter the polarity of the protein. Through a three-dimensional structure model and by the use of PyMOL, it was proved that the 3D structure of the mutant protein and the number of bonds between amino acids had been altered. It is, thus, implied and proposed that the conformation of the protein may be altered by this missense mutation (Figure 7).

Furthermore, four silent nucleotide variations (p. Tyr486 Tyr, p. Thr1493 Thr, p. Gly1678Gly, and p. Ser1756Ser) were identified in 50 patients (84.74%) all of whom had previously been reported in other populations or Gene databanks. By a linkage analysis, it was demonstrated that haplotype linkages in all the families would affect the individuals who shared the same haplotype for SNP markers but no common haplotypes were identified in families 5 and 7. These results suggested that a common origin of FAP was unlikely for these families. It was also proposed that other mutations may have occurred in the promoter region of the gene or that other genes such as MUTYH could be involved in the FAP patients [41] and further mutation analyses, such as whole-exome sequencing, should be performed in the future to elucidate the genetic origin of FAP in such families.

The mutations identified in the present study were almost located in the coding sequences of the APC gene between codons 463

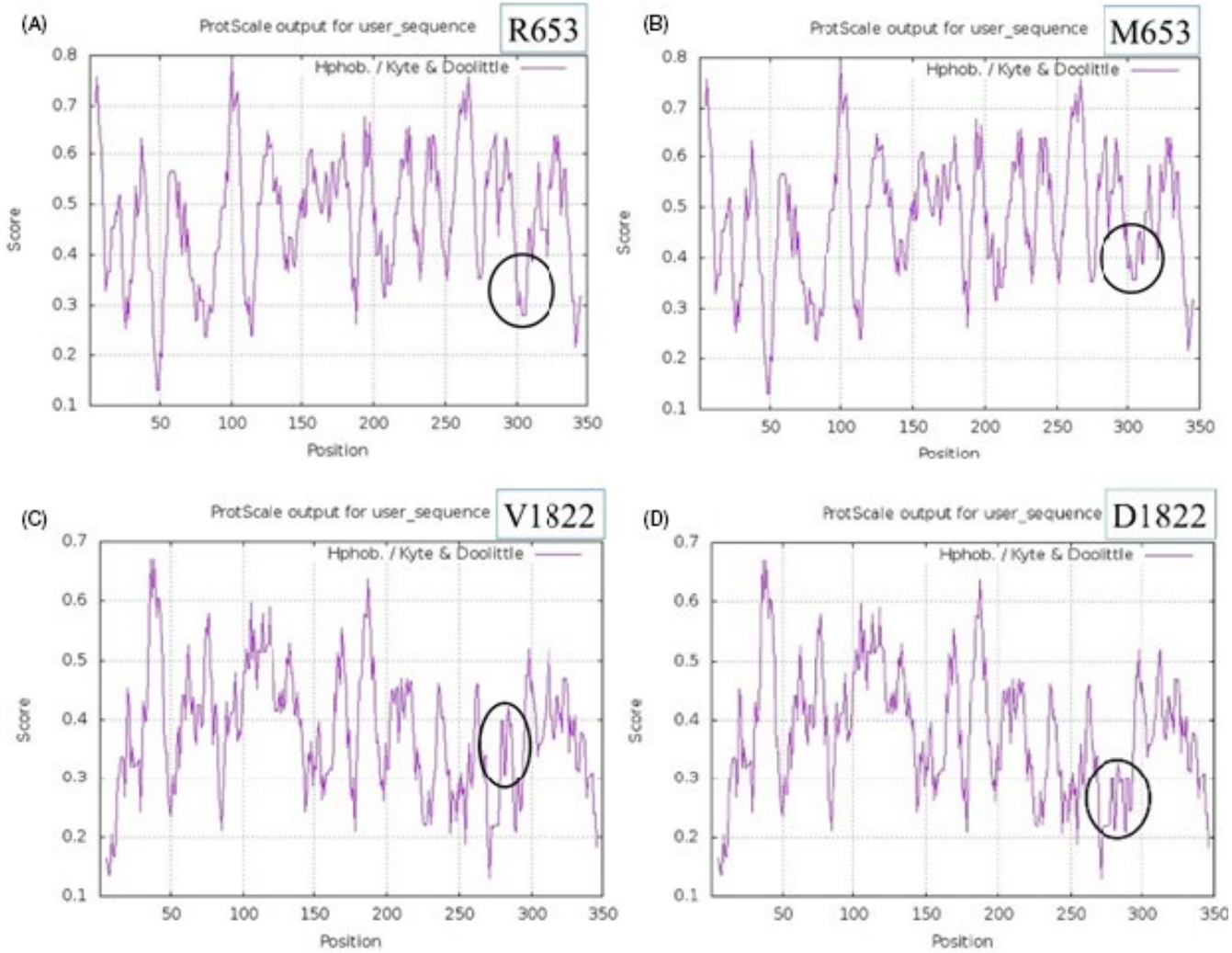

FIGURE 5 The plot of hydrophobicity of (A) a wild type in position 653 (Arg), (B) a mutant form in position 653 (Met), (C) a wild type in position 1822 (Val), and (D) a mutant form in position 1822 (Asp)

and 1822, and most of these mutations occurred remarkably in exon 15, which, according to other report in the literature, is a hotspot exon. Shahnazi Gerdehsang et al. investigated a part of exon 15 of the APC gene in 23 FAP patients in several provinces of Iran and identified new mutations in the three patients, including c.2910delT, c.3577–3578 delCA, and a substitution (N862 K) in patients with severe polyposis in colorectal.[15] Mostly, in classical FAP, 60% of germline mutations in the APC gene took place in the central region of the protein or the mutation cluster region (MCR) of APC (amino acids 1250–1580).[42]

There are various domains suggested for the APC protein, including the dimerization domain, the armadillo repeats (40-amino acid repeats between amino acids 453 and 764), β-catenin binding domain (three 15- amino acid repeats between amino acids 1020 and 1169 and seven 20-amino acid repeats between amino acids 1324 and 2075), the Axin-binding domain (SAMP repeats), the basic domain, and the C-terminal domains.[43] Zhang et al. showed that mutations in these essential regions of the APC gene play critical roles in intercellular adhesion, cell proliferation, and apoptosis and lead to tumorigenesis induction.yyy[13] Our results indicate that 5 mutations (p. Arg463term, p. Tyr486 Tyr, p. Leu592term, Glu601term, and p. Arg653Met) were located at armadillo repeats, and the remaining seven nucleotide alterations were observed in the β-catenin binding domain of the APC protein (p. Pro1140Leufx25, p.lys1199term, p. Glu1309AspfsX4, p. Thr1493 Thr, p. Gly1678Gly, p. Ser1756Ser, and p. Val1822Asp).

The important role of the armadillo domain in cell adhesion and the stimulation of cell migration lies in the binding to Asef1, Asef2, and IQGAP1. Mutations in this domain interrupt cell adhesion or increase the ability of cell growth and migration; as a result, the APC gene, which inhibits the growth and proliferation of tumor, is disrupted. The beta-catenin domain is also involved in cell adhesion and the control of cell proliferation and differentiation. Mutations in the beta-catenin domain lead to the constitutive activation of beta-catenin, uncontrolled cell proliferation and, finally, severe polyposis development in FAP patients.

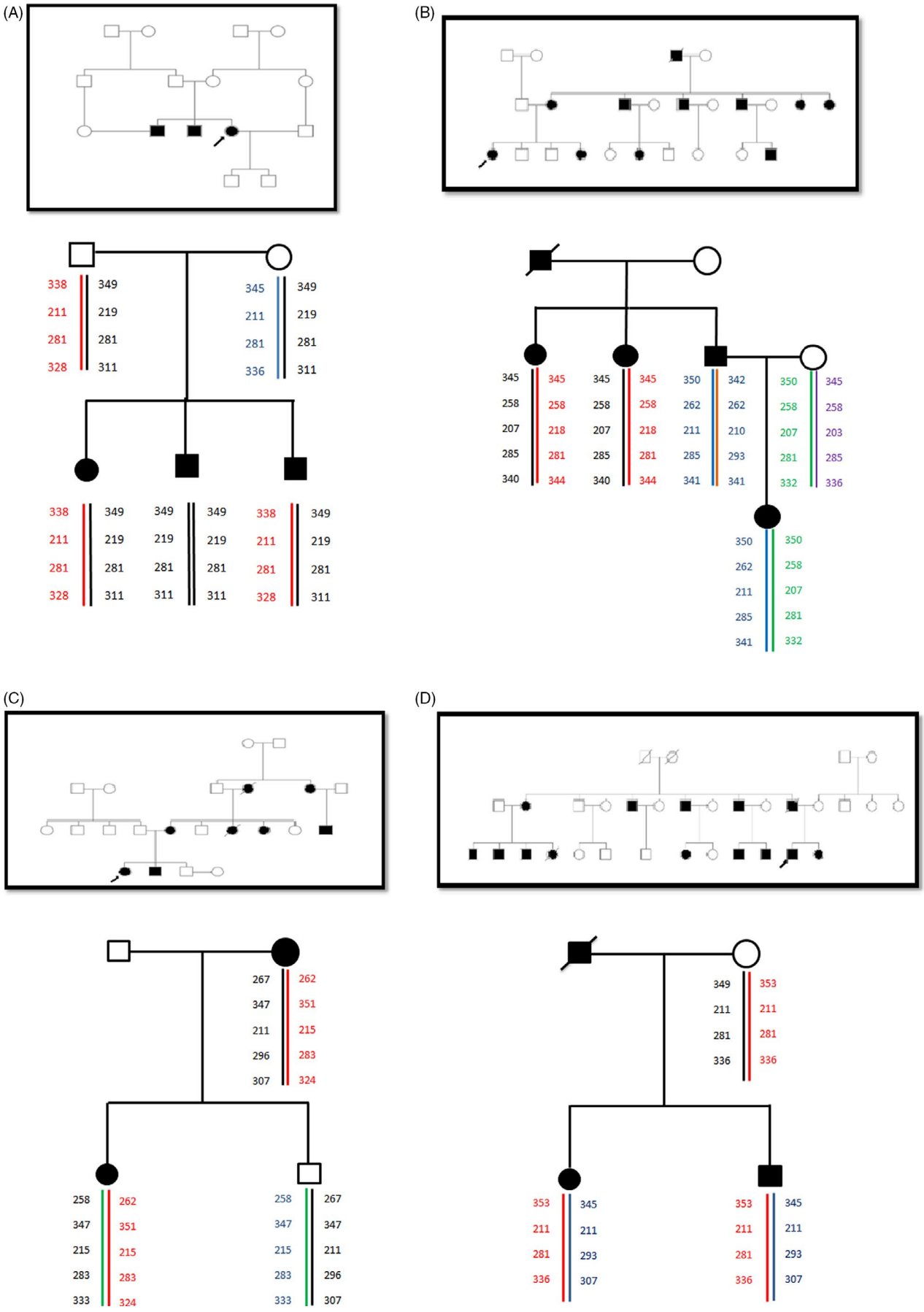

FIGURE 6 Pedigree and haplotype analysis with STR markers represented in 4 FAP family members: A, B) In families 5 and 7, there were no common haplotypes to identify, but C, D) families 8 and 9 showed a shared haplotype around the APC gene (red haplotype)

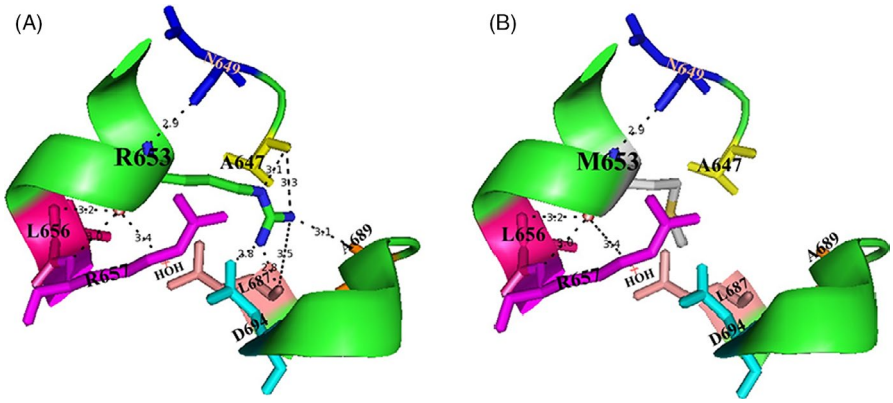

FIGURE 7 Structure modeling of the wild type and p. Arg653Met mutation of APC with the PyMOL software: Compared with the wild type (A), the mutant (B) leads to the loss of a number of hydrogen bonds. Arginine at position 653 has hydrogen bonds with the bond length of 2.9 angstroms with asparagine 649, 3.2 angstroms with leucine 656, 0.3 angstroms with arginine 657, and a bond with a length of 3.4 angstroms with a water molecule. In the mutant form, however, the methionine side chain at position 653 has two bonds with the lengths of 3.3 and 3.1 angstroms with alanine 647, one bond with the length of 3.1 angstroms with alanine 689, two bonds with the lengths of 2.8 and 3.5 angstroms with leucine 687 angstroms, and a 2.8-angstrom bond with aspartic acid 694

5 | CONCLUSION

With a focus on 59 patients from ten Iranian families with FAP, 12 nucleotide variations were identified in the APC gene. Bioinformatics evidence indicated that some of these variations affect the structure or function of the protein and half of them including four nonsense and two frameshift pathogenic variations produce a truncated APC protein. These findings suggest that deleterious modifications in the APC gene, participate in tumor development of FAP. The assessment of APC gene mutations in FAP patients in different populations can be a helpful marker for distinguishing cases at risk and diagnosis carriers and non-carriers. However, it should be noted that the importance and main role of these mutations in FAP pathogenesis of cannot be fully understood by sequencing analysis alone; rather further functional analyses should be performed in the laboratory.

ACKNOWLEDGMENTS

We thank all the patients for providing blood samples for our scientific research. We also appreciate by Yazd University Human Research Committee for approving this study.

CONFLICT OF INTERESTS

The authors declare that there are no conflicts of interest.

AUTHOR CONTRIBUTIONS

Mohammad Mehdi Heidari and Sirous Zeinali directed the study. Mohammad Mehdi Heidari, Mehri Khatami, and Sirous Zeinali contributed to the project design. Faranak Ghadamyari and Leili Rejali performed the experiments. Mohammad Mehdi Heidari, Faranak Ghadamyari, and Farzaneh Ghasemi analyzed the bioinformatics data. Shahin Merat, Hamideh Bagherian, and Leili Rejali contributed samples, data, and comments on the manuscript. Mohammad Mehdi Heidari and Sirous Zeinali analyzed and interpreted the data. Faranak Ghadamyari, Mohammad Mehdi Heidari1, and Sirous Zeinali contributed reagents, materials, and/or analysis tools. Mehri Khatami wrote the manuscript.

ETHICAL APPROVAL

All of the protocols in this study were performed in accordance with the guidelines of the Ethics Committee of Yazd University and with the World Medical Association Helsinki Declaration of 2000 and 2008 ethical standards.

DATA AVAILABILITY STATEMENT

The data used in this study are available from the corresponding author upon request.

ORCID

Mohammad Mehdi Heidari https://orcid.org/0000-0002-3328-4746

REFERENCES

1. Leoz ML, Carballal S, Moreira L, et al. The genetic basis of familial adenomatous polyposis and its implications for clinical practice and risk management. *Appl Clin Genet*. 2015;8:95.
2. Half E, Bercovich D, Rozen P. Familial adenomatous polyposis. *Orphanet J Rare Dis*. 2009;4:22.
3. Rivera F, Mendoza Y, Medina H. Hereditary colorectal cancer syndromes. *Ann Colorect Res*. 2020;8:157-169.
4. Vandrovcová J, Štekrová J, Kebrdlová V, et al. Molecular analysis of the APC and MYH genes in Czech families affected by FAP or multiple adenomas: 13 novel mutations. *Hum Mutat*. 2004;23:397.
5. Khan N, Lipsa A, Arunachal G, et al. Novel mutations and phenotypic associations identified through APC, MUTYH, NTHL1, POLD1, POLE gene analysis in Indian Familial Adenomatous Polyposis cohort. *Sci Rep*. 2017;7:1-6.
6. Nallamilli BRR, Hegde M. Detecting APC gene mutations in familial adenomatous polyposis (FAP). *Current Protocols in Human Genetics*. 2017;92:10.18.11-10.18.16.
7. Afkhami E, Heidari MM, Khatami M, et al. Detection of novel mitochondrial mutations in cytochrome C oxidase subunit 1 (COX1)

in patients with familial adenomatous polyposis (FAP). *Clin Transl Oncol*. 2020;22:908-918.
8. Dinarvand P, Davaro E, Doan J, et al. Familial adenomatous polyposis syndrome: An update and review of extraintestinal manifestations. *Arch Pathol Lab Med*. 2019;143(11):1382–1398.
9. Giang H, Nguyen VT, Nguyen SD, et al. Detection of a heterozygous germline APC mutation in a three-generation family with familial adenomatous polyposis using targeted massive parallel sequencing in Vietnam. *BMC Med Genet*. 2018;19:188.
10. Pattison AM, Barton JR, Entezari AA, et al. Silencing the intestinal GUCY2C tumor suppressor axis requires APC loss of heterozygosity. *Cancer Biol Ther*. 2020;9:1-7.
11. Ghatak S, Chakraborty P, Sarkar SR, et al. Novel APC gene mutations associated with protein alteration in diffuse type gastric cancer. *BMC Med Genet*. 2017;18:61.
12. Tsukanov A, Pospekhova N, Shubin V, et al. Mutations in the APC gene in Russian patients with classic form of familial adenomatous polyposis. *Russian J Genet*. 2017;53:369-375.
13. Zhang L, Shay JW. Multiple roles of APC and its therapeutic implications in colorectal cancer. *J Natl Cancer Inst*. 2017;109: djw332.
14. Bellido F, Pineda M, Aiza G, et al. POLE and POLD1 mutations in 529 kindred with familial colorectal cancer and/or polyposis: review of reported cases and recommendations for genetic testing and surveillance. *Genet Med*. 2016;18:325-332.
15. Gerdehsang PS, Ranji N, Gorji M, et al. New mutations in APC gene among familial adenomatous polyposis (FAP) patients in Iran. *Int J Human Genet*. 2017;17:145-150.
16. Kashfi SMH, Golmohammadi M, Behboudi Farahbakhsh F, et al. Novel missense mutation at codon 2774 (c.8321 g>a) p.s2774n of apc gene in a denovo case of familial adenomatous polyposis. *Arch Iran Med*. 2015;18:446-449.
17. Hankey W, Frankel WL, Groden J. Functions of the APC tumor suppressor protein dependent and independent of canonical WNT signaling: implications for therapeutic targeting. *Cancer Metast Rev*. 2018;37:159-172.
18. Zhan T, Rindtorff N, Boutros M. Wnt signaling in cancer. *Oncogene*. 2017;36:1461-1473.
19. Regimbald-Dumas Y, He X. Wnt signalling: what the X@# is WTX? *The EMBO J*. 2011;30:1415-1417.
20. Heinen CD, Goss KH, Cornelius JR, et al. The APC tumor suppressor controls entry into S-phase through its ability to regulate the cyclin D/RB pathway. *Gastroenterology*. 2002;123:751-763.
21. Sameer AS, Shah ZA, Abdullah S, et al. Analysis of molecular aberrations of Wnt pathway gladiators in colorectal cancer in the Kashmiri population. *Hum Genom*. 2011;5:441.
22. Christie M, Jorissen RN, Mouradov D, et al. Different APC genotypes in proximal and distal sporadic colorectal cancers suggest distinct WNT/β-catenin signalling thresholds for tumourigenesis. *Oncogene*. 2013;32:4675-4682.
23. Chen QW, Zhang XM, Zhou JN, et al. Analysis of small fragment deletions of the APC gene in Chinese patients with familial adenomatous polyposis, a precancerous condition. *Asian Pac J Cancer Prev*. 2015;16:4915-4920.
24. Jass JR. Colorectal polyposes: from phenotype to diagnosis. *Pathology-Res Pract*. 2008;204:431-447.
25. Michils G, Tejpar S, Thoelen R, et al. Large deletions of the APC gene in 15% of mutation-negative patients with classical polyposis (FAP): a Belgian study. *Hum Mutat*. 2005;25:125-134.
26. Cao X, Hong Y, Eu KW, et al. Singapore familial adenomatous polyposis (FAP) patients with classical adenomatous polyposis but undetectable APC mutations have accelerated cancer progression. *Am J Gastroenterol*. 2006;101:2810-2817.
27. Rivera B, González S, Sánchez-Tomé E, et al. Clinical and genetic characterization of classical forms of familial adenomatous polyposis: a Spanish population study. *Ann Oncol*. 2011;22:903-909.
28. Pang M, Liu Y, Hou X, et al. A novel APC mutation identified in a large Chinese family with familial adenomatous polyposis and a brief literature review. *Mol Med Rep*. 2018;18:1423-1432.
29. Liao D-X, Li B, Du X-M, et al. Two Chinese pedigrees for adenomatous polyposis coli: new mutations at codon 1309 and predisposition to phenotypic variations. *Fam Cancer*. 2014;13:361-368.
30. Aceto GM, Fantini F, De Iure S, et al. Correlation between mutations and mRNA expression of APC and MUTYH genes: new insight into hereditary colorectal polyposis predisposition. *J Exp Clin Cancer Res*. 2015;34:131.
31. Li H, Zhang L, Jiang Q, et al. Identification a nonsense mutation of APC gene in Chinese patients with familial adenomatous polyposis. *Exp Ther Med*. 2017;13:1495-1499.
32. Nieuwenhuis M, Vasen H. Correlations between mutation site in APC and phenotype of familial adenomatous polyposis (FAP): a review of the literature. *Crit Rev Oncol/Hematol*. 2007;61:153-161.
33. Friedl W, Aretz S. Familial adenomatous polyposis: experience from a study of 1164 unrelated german polyposis patients. *Heredit Cancer Clin Pract*. 2005;3:95-114.
34. Plawski A, Slomski R. APC gene mutations causing familial adenomatous polyposis in Polish patients. *J Appl Genet*. 2008;49:407-414.
35. Kashfi H, Farahbakhsh F, Golmohammadi M, et al. Frameshift mutations (Deletion at codon 1309 and Codon 849) in the APC gene in Iranian FAP patients: a case series and review of the literature. *Int J Mol Cell Med*. 2014;3:196-202.
36. Armaghany T, Wilson JD, Chu Q, et al. Genetic alterations in colorectal cancer. *Gastrointest Cancer Res*. 2012;5:19-27.
37. Kwong LN, Dove WF. APC and its modifiers in colon cancer. *Adv Exp Med Biol*. 2009;656:85-106.
38. Picelli S, Zajac P, Zhou X-L, et al. Common variants in human CRC genes as low-risk alleles. *Eur J Cancer*. 1990;2010(46):1041-1048.
39. Cerasuolo A, Miele E, Russo M, et al. Sporadic pediatric severe familial adenomatous polyposis: A case report. *Mol Clin Oncol*. 2020;13:20.
40. Fernández-Rozadilla C, de Castro L, Clofent J, et al. Single nucleotide polymorphisms in the Wnt and BMP pathways and colorectal cancer risk in a Spanish cohort. *PLoS One*. 2010;5(9):e12673.
41. Jasperson K, Tuohy T, Neklason D, et al. Hereditary and familial colon cancer. *Gastroenterology*. 2010;138:2044-2058.
42. Groves C, Lamlum H, Crabtree M, et al. Mutation cluster region, association between germline and somatic mutations and genotype-phenotype correlation in upper gastrointestinal familial adenomatous polyposis. *Am J Pathol*. 2002;160:2055-2061.
43. Plawski A, Banasiewicz T, Borun P, et al. Familial adenomatous polyposis of the colon. *Hered Cancer Clin Pract*. 2013;11:15.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Ghadamyari F, Heidari MM, Zeinali S, et al. Mutational screening through comprehensive bioinformatics analysis to detect novel germline mutations in the *APC* gene in patients with familial adenomatous polyposis (FAP). *J Clin Lab Anal*. 2021;35:e23768. https://doi.org/10.1002/jcla.23768