



ORIGINAL ARTICLE

A truncated protein product of the germline variant of the *DUOX2* gene leads to adenomatous polyposis

Mengyuan Yang^{1,2}, Yingxin Zhao^{1,2}, Yuwei Ding¹, Juan Wang^{1,2}, Yinuo Tan^{1,2}, Dong Xu³, Ying Yuan^{1,2}

¹Department of Medical Oncology, Zhejiang University School of Medicine, Hangzhou 310009, China; ²Cancer Institute, Zhejiang University School of Medicine, Hangzhou 310009, China; ³Department of Colorectal Surgery and Oncology, Key Laboratory of Cancer Prevention and Intervention, Ministry of Education, The Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310009, China

ABSTRACT

Objective: In some patients with adenomatous polyposis, an identifiable pathogenic variant of known associated genes cannot be found. Researchers have studied this for decades; however, few new genes have been identified.

Methods: Adenomatous polyposis coli (*APC*) negative polyposis patients were identified through next-generation sequencing and multiplex ligation-dependent probe amplification. Then, whole-exome sequencing (WES) was used to determine candidate genes harboring pathogenic variants. Functional experiments were performed to explore their effects. Subsequently, using Sanger sequencing, we found other polyposis patients carrying variants of the *DUOX2* gene, encoding dual oxidase 2, and analyzed them.

Results: From 88 patients with suspected familial adenomatous polyposis, 25 unrelated *APC* negative polyposis patients were identified. Based on the WES results of 3 patients and 2 healthy relatives from a family, the germline nonsense variant (c.1588A>T; p.K530X) of the *DUOX2* gene was speculated to play a decisive role in the pedigree in relation to adenomatous polyposis. During functional experiments, we observed that the truncated protein, hDuox2 K530, was overexpressed in the adenoma in a carrier of the *DUOX2* nonsense variant, causing abnormal cell proliferation through endoplasmic reticulum (ER) retention. In addition, we found two unrelated *APC* negative patients carrying *DUOX2* missense variants (c.3329G>A, p.R1110Q; c.4027C>T, p.L1343F). Given the results of the *in silico* analysis, these two missense variants might exert a negative influence on the function of hDuox2.

Conclusions: To our knowledge, this is the first study that reports the possible association of *DUOX2* germline variants with adenomatous polyposis. With an autosomal dominant inheritance, it causes ER retention, inducing an unfolded protein response.

KEYWORDS

Adenomatous polyposis; *DUOX2*; whole-exome sequencing; endoplasmic reticulum retention; unfolded protein response

Introduction

Familial adenomatous polyposis (FAP, OMIM 175100) is an autosomal dominant disorder in which hundreds to thousands of adenomatous polyps develop¹. Depending on the severity of polyposis, patients are classified as either classic FAP (CFAP) or attenuated FAP (AFAP). Compared to CFAP, AFAP is characterized by a smaller number of polyps (10–99 polyps), a later age of onset, and a lower probability of malignant transformation of the adenomas².

The germline variants in the tumor suppressor gene, *APC* (adenomatous polyposis coli), are widely recognized as the leading genetic cause of this disease. The mutation detection rate of the *APC* gene in patients with FAP is closely related to the phenotype, as it is up to 82% in CFAP patients, and only 24% in patients with AFAP³. This highlights that the genetic background of some of the FAP patients remains unclear, especially in people with attenuated phenotypes. We named these patients with polyposis without an identifiable *APC* pathogenic variant as *APC* negative polyposis. In the recent decades, several genes have been reported to play important roles in *APC* negative polyposis. Germline biallelic pathogenic variants in *MUTYH* (mutY DNA glycosylase) were detected in 17% of the individuals with *APC* negative polyposis, but no biallelic variant was noted in typical cases⁴.

For suspected patients of adenomatous polyposis, early genetic screening is necessary to identify disease-causing gene variants in order to guide the treatment, follow-up, and family

Correspondence to: Ying Yuan

E-mail: yuanying1999@zju.edu.cn

ORCID iD: <http://orcid.org/0000-0002-3922-9553>

Received June 11, 2020; accepted September 08, 2020.

Available at www.cancerbiomed.org

©2021 Cancer Biology & Medicine. Creative Commons Attribution-NonCommercial 4.0 International License

management strategies⁵. For patients with *APC* negative polyposis, as the cause of the disease is unknown, individualized therapeutic strategy cannot be formulated. Over the decades, numerous studies have been conducted worldwide in an attempt to identify genes responsible for adenomatous polyposis. A vast majority of those studies still focused on the *APC* gene, and very few new genes have been found. Therefore, we conducted a multi-center clinical study of the Chinese population to collect pedigrees of *APC* negative polyposis, aiming to find a novel disease-causing gene for adenomatous polyposis.

Materials and methods

Patients and DNA samples

In this study, patients suspected of adenomatous polyposis were included. The criteria used were more than 10 polyps observed under colonoscopy, and pathological confirmation of adenoma. Clinical data and pedigree information were collected. The protocol of this study is shown in **Figure 1**. The study was approved by the Ethics Committee of the Second Affiliated Hospital of Zhejiang University School of Medicine (Approval No. 2017.066). Written informed consent was obtained from all patients.

Genomic DNA was extracted from peripheral EDTA-anticoagulated blood samples using the standard salting-out procedure (QIAamp DNA blood midi kit; Qiagen, Hilden, Germany). For pedigree members, if blood samples were

not available, DNA was extracted from oral swabs using a TIANamp swab DNA kit (Tiangen, Beijing, China) or the negative margins of surgical specimens using a QIAamp DNA FFPE tissue kit (Qiagen). High quality DNA (OD260/280 = 1.8–2.0) samples were used for further genetic testing.

Next-generation sequencing (NGS) and analysis

Variants of 139 genes associated with different hereditary cancers and polyposis were screened by NGS, which was performed by Genetron Health (Beijing, China) on the HiSeqX-ten sequencing platform (Illumina, San Diego, USA). These 139 genes are shown in **Supplementary Table S1**. We classified germline variants according to the American College of Medical Genetics and Genomics standards and guidelines for sequence variant interpretation⁶. *APC* mutant polyposis (FAP) was diagnosed if the patient was accompanied by germline pathogenic or likely pathogenic variants in the *APC* gene.

Multiplex ligation-dependent probe amplification

Genomic DNA samples from patients without an identified pathogenic variant were examined for large deletions or duplications of *APC* and *MUTYH* genes using multiplex ligation-dependent probe amplification (MLPA) (SALSA[®] MLPA[®] probemix P043-E1; MRC, Amsterdam, The Netherlands) according to the manufacturer's instructions. The results were analyzed

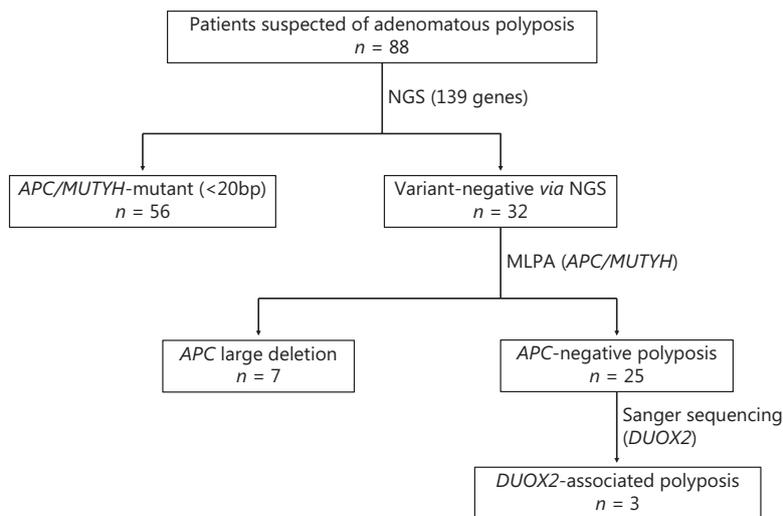


Figure 1 Flow chart of the multi-center clinical study. Patients enrolled in this study underwent multiple genetic tests to identify germline pathogenic variants. NGS, next-generation sequencing; MLPA, multiplex ligation-dependent probe amplification.

against controls using Coffalyser software (MRC). A dosage quotient of 0.8–1.2 was interpreted as normal; 0, 0.4–0.65 and 1.3–1.65 were interpreted as homozygous deletion, heterozygous deletion or heterozygous duplication, respectively. For each sample, we performed 3 independent experiments.

Whole-exome sequencing (WES)

WES was performed by XY Biotechnology (Shanghai, China) for genome analysis by capture using the Agilent SureSelect Human All Exon V6 (Agilent, Santa Clara, CA, USA). After capturing and enriching all exon regions, sequencing was performed on a HiSeqX-ten sequencing platform (Illumina). An average of 95% of all bases were covered 20 or more times. The quality control of raw exome data was conducted using fastp (version 0.14.1). Only high quality raw data (> 80% of the sequenced reads having a Qphred of 30) were subjected to further analysis. Sequenced reads were aligned to the GRCh37/hg19 human reference genome.

Sanger sequencing

The variant NM_014080.4: c.1588A>T (p.K530X) of *DUOX2* was validated by Sanger sequencing according to standard protocols. Genomic DNA was used to amplify the corresponding region (primers: 5'-CCCCTCTGGGTCTCTTTTCT-3' and 5'-TTGATCCTTCTCTGGCCACT-3'). PCR products were purified and sequenced on an ABI 3730 DNA Analyzer (Applied Biosystems, Foster City, CA, USA). We designed 21 pairs of primers (Supplementary Table S2) to sequence the whole exon region of the *DUOX2* gene for patients with *APC* negative polyposis.

Immunohistochemistry (IHC) staining

IHC staining was performed according to standard protocols. After deparaffinization, antigen retrieval, and blocking, the paraffin-embedded slides were incubated with anti-DUOX2 (1-100 amino acids) antibody (1:500 dilution, ab97266; Abcam, Cambridge, UK), anti-non-phospho (active) β -catenin antibody (1:1,000 dilution; Cell Signaling Technology, Danvers, MA, USA) or anti-GRP78 BiP antibody (1:100 dilution, ab108615; Abcam,). The immunoreactivities were visualized using 3,3'-diaminobenzidine (Dako, Carpinteria, CA, USA).

Immunofluorescence (IF)

For IF, pretreatment of paraffin-embedded slides was similar to IHC analysis. The slides were incubated simultaneously with rabbit anti-DUOX2 (400–500 amino acids; 1:500 dilution; NB110-61576; Novus Biological, Centennial, CO, USA) and mouse anti-DUOX2 (628–680 amino acids; 1:500 dilution, sc-398681; Santa Cruz Biotechnology, Santa Cruz, CA, USA) antibodies. Cy3-conjugated goat anti-rabbit IgG and 488-conjugated goat anti-mouse IgG (GB21303, and GB25301; Servicebio, Wuhan, China) were used as secondary antibodies. Nuclear counterstaining was performed using 4',6-diamidino-2-phenylindole. The slide was scanned using the Panoramic 250/MIDI (3DHISTECH, Budapest, Hungary).

Cell culture and establishment of cell lines

RKO cells, provided by Stem Cell Bank, Chinese Academy of Sciences (Beijing, China) were grown in Minimum Essential Medium (Gibco, Gaithersburg, MD, USA) supplemented with 10% fetal bovine serum (Gibco), NaHCO_3 (1.5 g/L), and 1% sodium pyruvate (Invitrogen, Carlsbad, CA, USA).

A 1,588 bp fragment coding hDuox2 K530 with a 6 × His tag was amplified from human *DUOX2* cDNA and was connected to the lentiviral expression vector pLVX-IRES-puro. The plasmid pLVX-IRES-puro-hDuox2-K530 with pMD2.G and pSPAX2 were packaged into viral particles in HEK293T cells. Lentiviral infection was performed according to the manufacturer's protocols. To obtain a RKO cell line with stable expression of hDuox2 K530, puromycin was used to select the cells. Overexpression was verified by Western blot. Empty vector was transfected in the same way as the negative controls. Plasmids used in the above experiments were all purchased from Youbio (Hunan, China).

Cell proliferation

Cell Counting Kit-8 (CCK-8; Dojindo, Tokyo, Japan) was used to evaluate cell proliferation. Experiments were performed according to the manufacturer's instructions. An optical density (OD) value of 450 nm was used to measure cell proliferation. The ratio of the OD value at each time point to the baseline OD value gave the relative proliferation activity. Five replicate wells were set for each group, and 3 independent experiments were performed to calculate the mean and 95% confidence interval (CI).

Immunocytochemistry

Cells were fixed with 4% paraformaldehyde, permeabilized with 0.1% Triton-X 100 (Sigma-Aldrich, St. Louis, MO, USA), and incubated with anti-6 × His tag (1:500 dilution, MA1-21315; Invitrogen) and rabbit anti-calreticulin (1:400 dilution, 12238; Cell Signaling Technology) antibodies. The cells were then incubated with Alexa Fluor Plus 555-labeled goat anti-mouse IgG (1:200 dilution, A32727; Invitrogen) and Alexa Fluor® 488-labeled anti-rabbit IgG (1:200 dilution, A32731; Invitrogen). Nuclear counterstaining was performed using DAPI.

In silico analysis of missense variants

The functional effects of two missense mutations were predicted using SIFT (version 1.03)⁷, Mutation Taster²⁸, and Mutation assessor⁹ algorithms. The stability change of 2 missense variants were predicted by using the following online available *in silico* algorithms: I-Mutant 2.0¹⁰, MUpro¹¹, and iStable¹². Three-dimensional (3D) models of the ferric oxidoreductase domain and flavin adenine dinucleotide (FAD)-binding ferredoxin reductase (FR)-type domain of wild-type and mutated hDuox2 were generated using the homology modeling program, SWISS-MODEL¹³. The quality of the 3D models obtained by SWISS-MODEL was verified using the following online available *in silico* algorithms: Verify3.0¹⁴, PROCHECK¹⁵, and ProQ¹⁶. Analysis and visualization of the wild-type and mutated models were performed with the PyMOL visualizer. Default parameters were adopted for all of the above programs. Detailed information about the *in silico* software is summarized in **Supplementary Table S3**.

Statistical analysis

All statistical analyses were performed using R version 3.6.2 (<https://www.r-project.org/>). The difference was considered statistically significant when the two-sided *P* value was less than 0.05. Illustrator of biological sequences (IBS) was used to create a schematic of the protein domain¹⁷. A schematic diagram of the pathogenic mechanism was created using BioRender.com.

Results

Identification of *APC* negative polyposis patients

A total of 88 patients suspected of having adenomatous polyposis were enrolled in this study. Using the 139 gene NGS

panel and MLPA, we identified 25 unrelated individuals without any specific *APC* pathogenic variants, named *APC* negative polyposis. The clinical characteristics, including the age at first visit, sex, FAP classification, extra-intestinal manifestation, and the family history of patients with *APC* negative and *APC* mutant polyposis are summarized in **Supplementary Table S4**. Consistent with our expectations, patients with *APC* negative polyposis were diagnosed at a later age, and had a lighter polyp load, with fewer extra-intestinal manifestations and a family history related to polyposis.

Identification of 3 unrelated individuals with *DUOX2* variants

We performed WES on one of the families with *APC* negative polyposis (the pedigree is shown in **Figure 2A**). The proband (III-1 in **Figure 2A**) was a female with 20–30 tubulovillous adenomas in the proximal colon at the age of 26 years. Meanwhile, her younger sister (III-2) also underwent colonoscopy, which revealed many polyps at the age of 24 years. She subsequently underwent total colectomy with ileorectal anastomosis. Pathology confirmed hundreds of adenomas scattered over the colon, mainly in the right half. Their mother (II-2) died from colon cancer with distant metastases at the age of 47 years, but the history of polyps in the intestine was uncertain. During our follow-up, one of their mother's brothers (II-6) underwent an endoscopy revealing an ascending colon adenocarcinoma and dozens of adenomas in the sigmoid colon at the age of 51 years.

The reasons for choosing this family were as follows: 1. four patients (II-2, II-6, III-1, and III-2) with colorectal adenomatous polyposis or colorectal cancer diagnosed in the family; 2. III-2 had a classic polyposis phenotype (> 100 polyps); and 3. we were able to obtain a complete genetic sample to help analyze the co-segregation of the family. WES showed that a total of 17,654 variants were shared by 3 patients (II-2, III-1, and III-2), and 2 unaffected individuals (II-1 and II-4) did not carry them. Synonymous or intronic variants other than those affecting consensus splice sites were excluded from further analysis. According to the American College of Medical Genetics and Genomics standards and guidelines, we selected the pathogenic or the likely pathogenic variants, and excluded the high occurrence variants (> 0.1% in the 1000 Genomes database). We identified a heterozygous nonsense variant c.1588A>T; p.K530X in the *DUOX2* gene, which might have been critical to pathogenesis in this family with *APC* negative polyposis.

By sequencing the whole exon region of the *DUOX2* gene from other patients with *APC* negative polyposis using Sanger

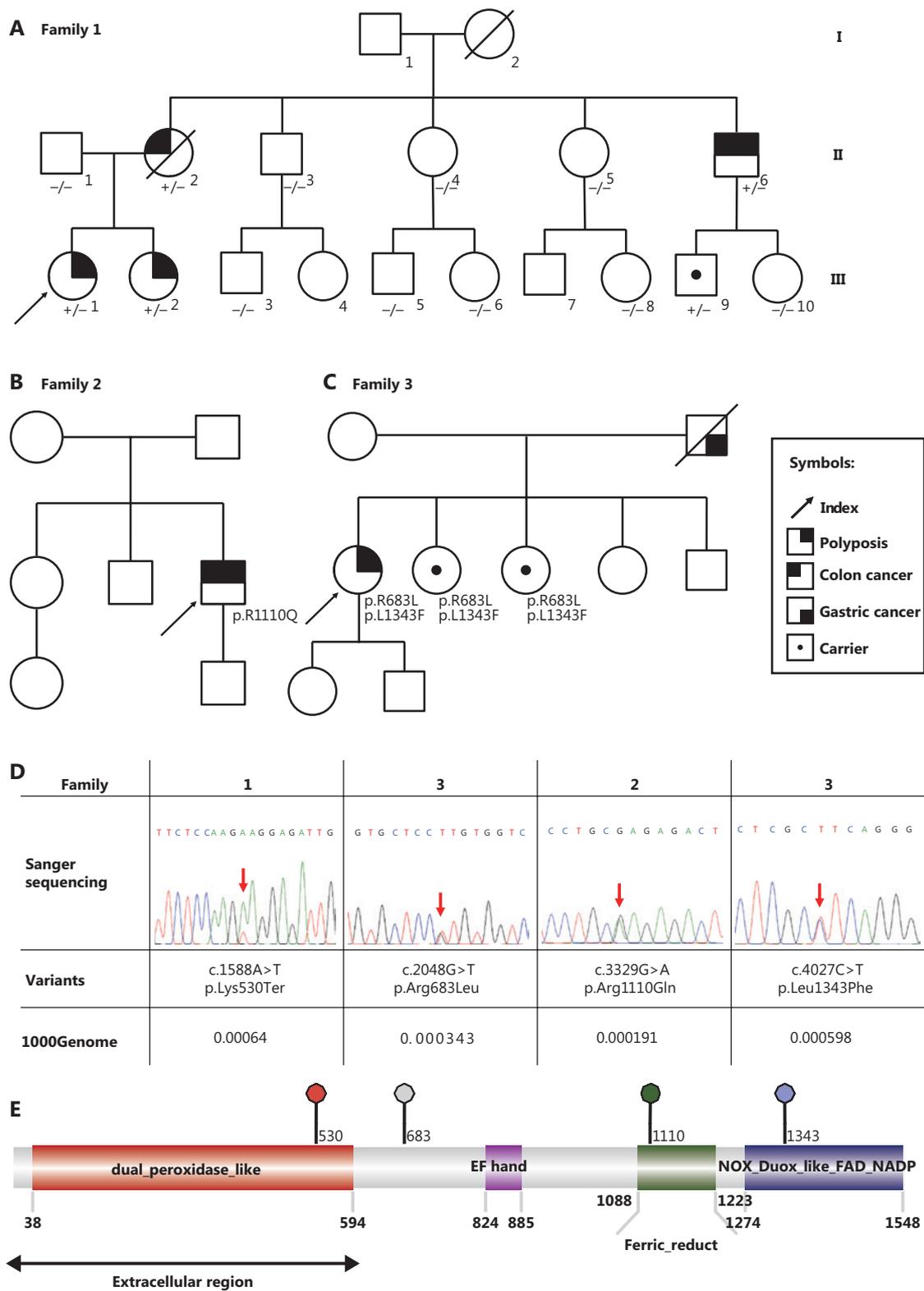


Figure 2 Pedigrees and variants of 3 unrelated families carrying *DUOX2* variants. Pedigrees of family 1(A), 2(B), and 3(C). The genotype and phenotype information are displayed, +/- represents a heterozygous nonsense (c.1588A>T; p.K530X) variant in the *DUOX2* gene, and -/- represents the wild type. (D) The four *DUOX2* variants in these 3 families were validated by Sanger sequencing. Their locations are depicted in a corresponding lollipop plot of *DUOX2* (created with IBS) and the conserved domains of the hDuo2 protein are also shown (E).

sequencing, we additionally found 2 unrelated patients (pedigrees are shown in **Figure 2B and 2C**) carrying missense variants in the *DUOX2* gene. The Sanger validation is shown in **Figure 1D**, and the 2 missense variants (p.R1110Q and p.L1343F) were both located in highly conserved protein domains (**Figure 2E**).

The effect of truncated proteins from the *DUOX2* nonsense variant on polyposis

The hDuox2 is overexpressed in adenomas independent of the Wnt pathway

We obtained paraffin-embedded adenoma and adjacent normal tissues from the affected case (III-2 in Family 1). IHC staining with an antibody targeting 1–100 amino acids of Duox2 revealed that its expression level in adenoma tissue was significantly higher than that in normal colon mucosa (**Figure 3A**). In addition, we also found that there was no obvious difference in the expression of non-phospho (active) β -catenin between the 2 samples, which indicated that the Wnt pathway was not activated abnormally in this adenoma tissue.

Truncated protein hDuox2 K530 was overexpressed in the adenoma, instead of the wild-type protein

In the patient that was a heterozygous carrier, the adenoma was able to express both wild-type (hDuox2-WT) and truncated

proteins (hDuox2 K530) of *DUOX2*. However, the antibodies used in IHC analysis could bind to both these proteins, making it impossible to determine which protein was overexpressed in the adenoma. To overcome this issue, IF analysis was conducted using two antibodies with different immunogenicities (**Figure 3B**) to determine whether the wild-type protein or the truncated protein was overexpressed. The antibody targeting 400–500 amino acids of hDuox2 enabled the detection of both wild-type and truncated proteins, and another antibody targeting 628–680 amino acids could only bind to the wild-type hDuox2. In the FFPE samples from the *DUOX2* p.K530X carrier (III-2 in Family 1), normal colon mucosa did not react with both antibodies, which was consistent with the results observed by IHC analysis. Moreover, we found that adenoma tissue only showed positive results with the antibody targeting 400–500 amino acids, but not with another antibody, indicating that the truncated protein, hDuox2 K530, was overexpressed in the adenoma. In addition, the tumors from patients with *DUOX2* wild-type colon cancer served as a positive control (**Figure 3C**).

Overexpression of the truncated protein hDuox2 K530 activates cell proliferation through the unfolded protein response

To determine the effect of the truncated protein on polyposis, we generated a human colorectal cancer cell line, RKO, stably

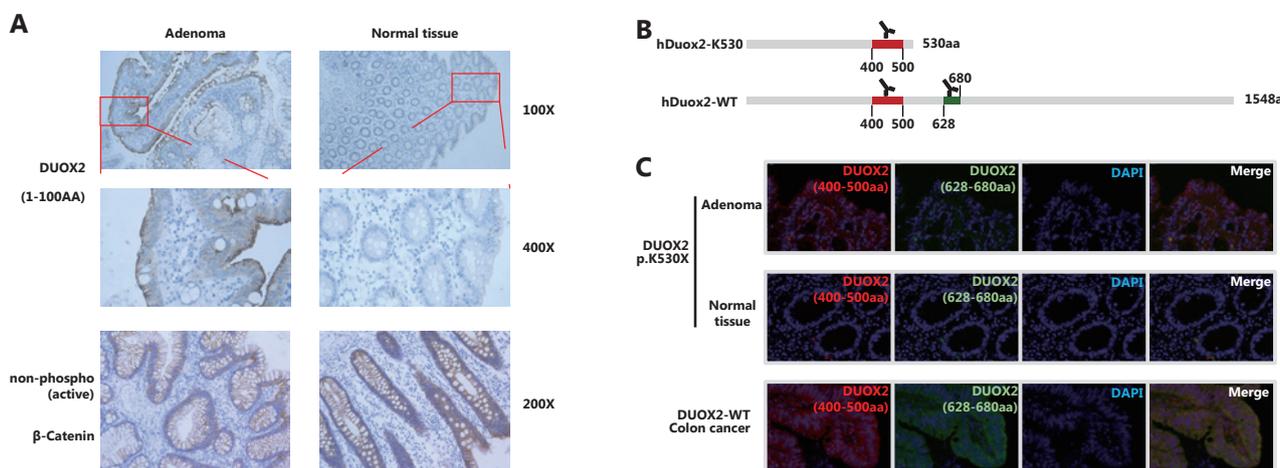


Figure 3 The truncated protein, hDuox2 K530, is overexpressed in adenomas. (A) Immunohistochemistry (IHC) staining of hDuox2 with an antibody targeting the N-terminal 1–100 amino acids of human Duox2 protein in the adenoma and normal tissues from a *DUOX2* p.K530X carrier. The expression of non-phospho (active) β -Catenin was also detected in these 2 samples by IHC analysis. (B) The positions of the amino acid sequence recognized by 2 different anti-Duox2 antibodies are marked. (C) Immunofluorescence staining of wild-type and truncated proteins of hDuox2 in adenomas and normal tissues from a *DUOX2* p.K530X carrier as well as the tumor from a *DUOX2* wild-type patient. AA, amino acid.

overexpressing hDuox2 K530 with a 6 × His tag (**Figure 4A**). CCK-8 analysis showed that from day 3 to day 5, the relative absorbance of RKO-hDuox2 K530 was significantly higher than that of the control group ($P < 0.0001$; **Figure 4B**). This indicated that overexpression of truncated protein could cause abnormal cell proliferation, which explained the occurrence of adenomas. Regarding the underlying mechanism, we hypothesized that the truncated protein could not be folded into a higher level protein structure due to the incomplete amino acid sequence, and was therefore retained in the endoplasmic reticulum (ER)¹⁸. The continuous expression of truncated proteins activated the unfolded protein response (UPR) of the ER, resulting in abnormal cell proliferation¹⁹. Through IHC, we observed that the truncated protein was co-localized with the ER marker, calreticulin (**Figure 4C**). Moreover, the expression level of GRP78 (BiP, one of the UPR markers²⁰) increased in adenomas from the *DUOX2*

p.K530X carrier, which preliminarily confirmed our hypothesis (**Figure 4D**).

In silico analysis of two *DUOX2* missense variants

We identified 3 missense variants in the *DUOX2* gene (p.R683L, p.R1110Q, and p.L1343F) from 2 unrelated families with unexplained polyposis. Because the variant p.R683L did not occur in the conserved domain, we conducted an *in silico* analysis on the other 2 variants. First, 3 different online servers (SIFT, Mutation Taster2, and Mutation assessor) based on multiple algorithms, were used for structural and functional annotation of these 2 missense variants (results are shown in **Table 1**). Next, stability analysis was conducted using I-Mutant2.0, MUpro, and iStable. Both variants were predicted to decrease protein stability. Finally, we performed

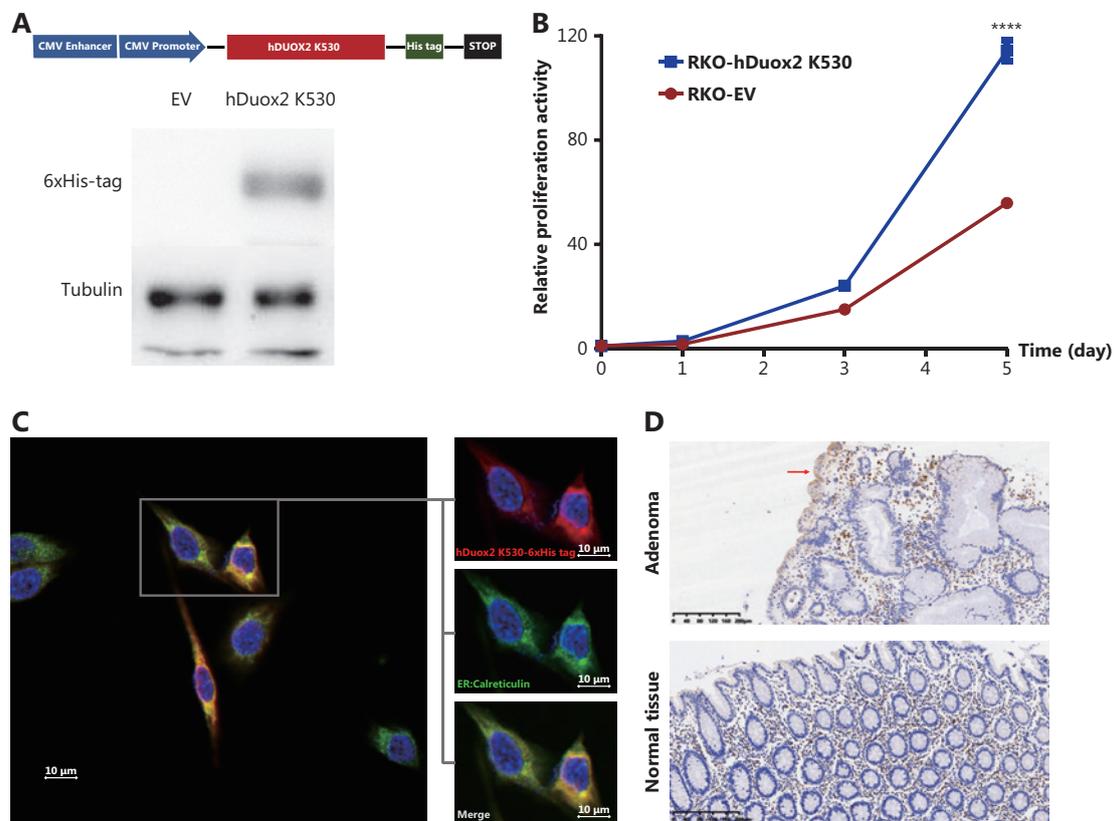


Figure 4 Overexpression of truncated protein, hDuox2 K530, promotes cell proliferation through the unfolded protein response. (A) Construction of the RKO cell line with stable expression of hDuox2 K530. (B) Growth curve of RKO-hDuox2 K530 and the control group. **** $P < 0.0001$ (C) Confocal microscopy with anti-6 × His-tag (red) and anti-calreticulin (green) in the RKO-hDuox2 K530 cell line. (D) Immunohistochemistry staining (100×) of GRP78 in the adenoma and normal tissues from a *DUOX2* p.K530X carrier. EV, empty vector.

Table 1 Prediction of effects of missense variants on protein function and stability

Variant	Function prediction			Stability prediction					
	SIFT [†]	Mutation Taster2 [‡]	Mutation assessor [§]	i-Mutant2.0 [¶]		MUpro ^{**}		iStable ^{**}	
				Prediction	DDG value	Prediction	Conf. score	Prediction	Conf. score
c.3329G > A (p.Arg1110Gln)	Deleterious	Disease causing		Decrease	-1.02	Decrease	-0.68	Decrease	0.72
c.4027C > T (p.Leu1343Phe)	Deleterious	Disease causing	Medium	Decrease	-1.23	Decrease	-1	Decrease	0.87

[†]SIFT classifies an amino acid substitution as “deleterious” or “tolerated” according to the normalized probability. A substitution is predicted as deleterious if the score is < 0.05 and tolerated if the score is ≥ 0.05.

[‡]Mutation Taster2 predicts an alteration as one of 4 possible types: “disease causing,” “disease causing automatic,” “polymorphism,” and “polymorphism automatic.”

[§]Mutation assessor outputs the annotations of functional impact of a variant as functional (high, medium), predicted non-functional (low, neutral).

[¶]i-Mutant2.0 outputs free energy change value (DDG) between mutant and wild-type protein. Stability decreases if DDG < 0 and increases if DDG > 0.

^{**}MUpro uses a confidence score between -1 and 1 to predict the effects of mutation on protein stability and measure the confidence of the prediction. A score < 0 means the mutation decreases the protein stability, and the smaller the score, the more confident the prediction. Conversely, a score > 0 means the mutation increases the protein stability, and the bigger the score, the more confident the prediction.

^{**}iStable outputs the prediction as “stability decrease” or “stability increase.”

3D modeling analysis of the wild-type and mutant domains (**Figure 5**). As for the variant p.R1110Q, we found that in the wild-type hDuox2 protein, a network of 3 intrahelical hydrogen bonds existed between Arg1110 and the other 2 residues, Ile1106 and Asp1121. When the Arg1110 was mutated to Gln due to a mutation, the 2 hydrogen bonds between Arg1110 and Asp1121 were lost and replaced by a novel hydrogen bond with Asn1115, which could influence the stability of the α -helix (**Figure 5A**). Additionally, comparing tertiary structures of the wild-type and mutant caused by p.L1343F, we observed that the directions of the amino acid molecules around the mutation site changed slightly (arrows in **Figure 5B**), and distant random coils showed a significant difference between the 2 domains (red circle in **Figure 5B**).

Discussion

APC negative polyposis refers to patients with clinical features of FAP, but conventional genetic testing, such as NGS and MLPA, failed to define their pathogenic variants. The causes may include the following 3 aspects. 1. Some variants in the non-coding region of the *APC* gene, such as the promoter region²¹, could be excluded by NGS, which only covers the exon regions as well as the first and last 20 bps of the

intron regions. 2. No genetic factors, but some environmental factors or personal habits cause polyps in the intestine²². 3. There are other genes with pathogenic variants, even those that have never been reported to be associated with polyposis. Pathogenic genes related to adenomatous polyposis are mainly divided into these 2 fields. First, the loss-of-function of key components in the Wnt pathway causes abnormal activation, such as *APC*³ and *AXIN2*²³. Second, several repair-related genes during DNA replication are also thought to be associated with polyposis, such as *MUTYH*⁴, *NTHL1*²⁴, *MSH3*²⁵, *POLE*, and *POLD1*^{26–28}.

This is the first study to report that germline variants of the *DUOX2* gene could also be implicated in a fraction of adenomatous polyposis cases. Dual oxidase 2 (Duox2) is one of the important members of the nicotinamide adenine dinucleotide phosphate (NADPH) oxidase family. It forms a complex with its maturation factor, Duoxa2, to catalyze the synthesis of hydrogen peroxide (H₂O₂). At first, Duox2 was thought to be active only in the thyroid, being involved in thyroid hormone synthesis²⁹. Later, Duox2 was found to be expressed in various tissues, including the entire intestinal mucosa³⁰. Most importantly, Duox2 acts as the first barrier of the intestinal epithelium and participates in the innate immune response of intestinal mucosa to maintain immune homeostasis by producing H₂O₂^{31,32}. Previous studies have reported that several missense



Figure 5 Three-dimensional (3D) models of the wild-type (blue) and mutant (green) domains. (A) Superimposition of the 3D model between wild-type (blue) and mutant (green) Ferric oxidoreductase domains are displayed. A magnified view of the region containing the wild-type Arg1110 residue (yellow) and its hydrogen bonds (yellow) formed with other residues is shown in the right upper panel. The magnified view of the mutant Gln1110 residue (pink) is shown in the right lower panel. (B) The 3D models and the superimposition between wild-type (blue) and mutant (green) FAD-binding FR-type domains are displayed. Magnified views of the region containing the wild-type Leu1343 residue (yellow) and the mutant Phe1343 residue (pink) are shown in the right upper and lower panels, respectively. Note that 3 residues around the mutation site have changed directions (red arrow).

variants (p.R1211C, p.R1492C, and p.P303R) of *DUOX2* are thought to be crucial in some patients with inflammatory bowel disease (IBD), especially those with very early onset IBD^{33–35}. In recent years, there have been several studies on the role of *DUOX2* in cancer, especially colon cancer, suggesting that *DUOX2* is an oncogene that facilitates metastatic ability³⁶ and chemoresistance³⁷ of CRC cells. However, there is no published report indicating the association of *DUOX2* germline variants with adenomatous polyposis. After an extensive literature review, we found that in 2006, a Japanese research

team had already observed that the expression of Duox2 in flat polyps was more than 10 times higher than in the adjacent normal epithelial tissues³⁸. This suggested that overexpression of Duox2 might be an important factor in the occurrence of polyps.

Duox2 protein is a 7-transmembrane protein composed of an N-terminal peroxidase-like domain (outside the cell), 2 EF-hand domains, and a C-terminal region binding to NADPH and FAD (inside the cell)³⁹. Under physiological conditions, immature and partially glycosylated Duox2 is in the ER and is

transported from the ER to the Golgi apparatus during certain types of stimulations⁴⁰. After processing and glycosylation, it is finally localized on the cell membrane, acting as a transmembrane protein. Due to the nonsense variant, p.K530X, the translation of the Duox2 protein is terminated at the Tyr530. IHC and IF assays also confirmed that the truncated protein, hDuox2 K530, was overexpressed in an adenoma from a heterozygous variant carrier. As it contains only the N-terminal peroxidase-like domain, the truncated protein cannot be processed and modified as expected; thus, it is retained in the ER. We found that the truncated protein was located in the ER of RKO-hDuox2 K530 cells. The high expression level of hDuox2 K530 caused ER stress and adversely affected cells, leading to the activation of the UPR to remove these adverse effects, maintaining cell stability by stopping protein translation, improving folding ability, degrading abnormally folded protein, or eventually inducing an apoptosis state⁴¹. During this time, the expression of GRP78, a UPR marker, would be

significantly upregulated²⁰, and we were able to detect this change by IHC staining of the adenoma from the patient.

However, some studies have suggested that continued excessive activation of the UPR is closely related to the occurrence and development of tumors, including promoting cell proliferation and migration, inducing DNA damage and mediating immune escape¹⁹. Activating transcription factor 6 (ATF6) and its downstream pathways play a key role after UPR activation⁴². Coleman et al.⁴³ observed that 12-week-old transgenic nATF6^{IEC} mice spontaneously developed adenomatous polyps in the intestine without being stimulated. We also found that the RKO-hDuox2 K530 cell line had significantly faster cell proliferation than the control group. Whether this was due to the abnormal activation of the UPR and its downstream pathways remains to be clarified.

There were 6 affected individuals in the 3 pedigrees with *DUOX2*-associated polyposis. Considering the clinical features, we summarized the following points. First, if it is not diagnosed

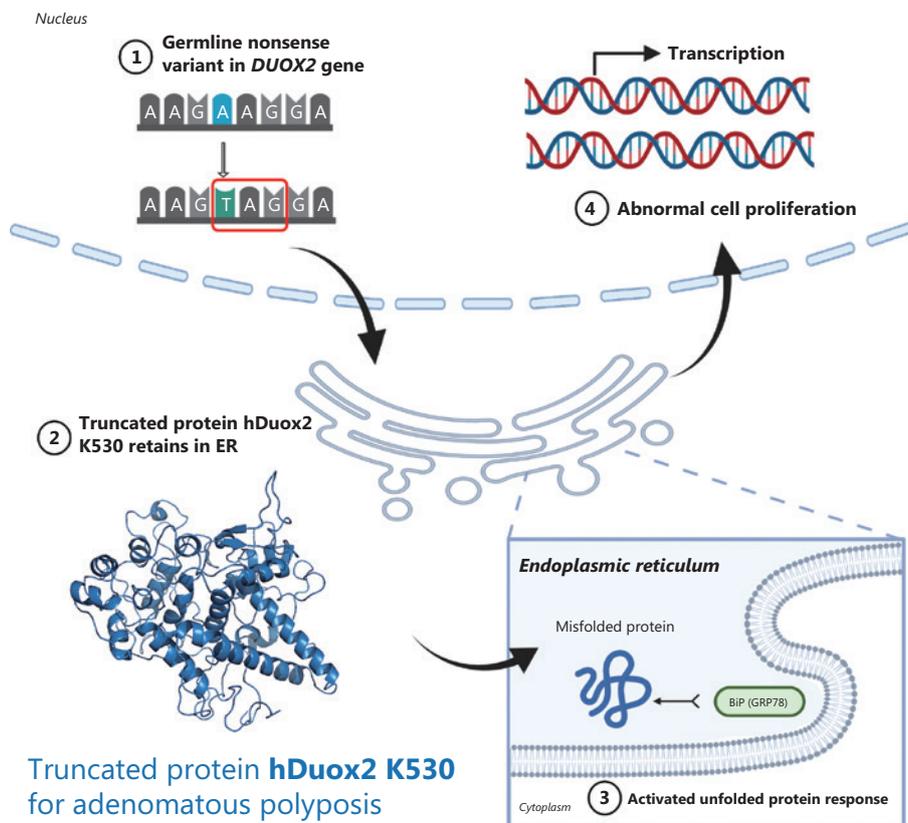


Figure 6 Schematic diagram of truncated protein, hDuox2 K530, causing adenomatous polyposis. The diagram shows that the truncated protein, hDuox2 K530, is translated due to a nonsense variant (c.1588A>T) occurring in the *DUOX2* gene. Therefore, this misfolded protein is retained in the endoplasmic reticulum and activates an unfolded protein response, which could induce abnormal cell proliferation.

early, the number of polyps can exceed 100 and may develop into malignant tumors. However, if diagnosed and treated in the early stage of the disease, the prognosis of the patient is significantly better than that of the FAP. For example, the first index (III-1 in Family 1) has been followed-up since the initial diagnosis in December 2017. After the 20–30 polyps were removed under the colonoscope, only a new 0.2 cm inflammatory polyp was noted during regular review. Second, none of these cases had extracolonic manifestations associated with FAP, such as congenital hypertrophy of the retinal pigment epithelium and desmoid tumor. Finally, through genetic testing of the members in families 1 and 3, we found that the penetrances of the *DUOX2* were 80% and 33.3%, respectively. We believe that the penetrance may depend on the type and location of the variants.

This study had several limitations. Due to the small population size, we found only 3 *DUOX2*-related polyposis families. Although bioinformatics tools were used to analyze the potential pathogenicity of the missense variants, we failed to explain the pathogenic mechanisms through functional experiments, and we could not adequately summarize the clinical characteristics of those patients. Thus, it is necessary to include the *DUOX2* gene in the genetic screening for adenomatous polyposis and to collect more patients with *DUOX2*-related polyposis in a larger population.

Conclusions

In conclusion, by evaluating patients with *APC*-negative polyposis, we conducted WES and Sanger sequencing to identify three unrelated patients carrying *DUOX2* germline variants, suggesting that the *DUOX2* gene might be a novel disease-associated gene for adenomatous polyposis with an autosomal dominant inheritance. Based on a literature search and functional experiments, we speculate that the mutant protein causes adenoma by activating the UPR in the ER (Figure 6). Further studies are required to validate this possibility.

Acknowledgements

We are grateful to Professor Jinrong Peng (Zhejiang University) for guiding the research design.

Grant support

This work was supported by the National Key R&D Program of China (Grant No. 2017YFC0908200); and the National

Natural Science Foundation of China (Grant Nos. 81872481 and 81902956).

Conflict of interest statement

No potential conflicts of interest are disclosed.

References

- Hagen CE, Setia N, Lauwers GY. Familial adenomatous polyposis: a review of gastrointestinal manifestations. *Diagn Histopathol.* 2015; 21: 152-60.
- Galiatsatos P, Foulkes WD. Familial adenomatous polyposis. *Am J Gastroenterol.* 2006; 101: 385-98.
- Friedl W, Aretz S. Familial adenomatous polyposis: experience from a study of 1164 unrelated German polyposis patients. *Hered Cancer Clin Pract.* 2005; 3: 95-114.
- Aretz S, Uhlhaas S, Goergens H, Siberg K, Vogel M, Pagenstecher C, et al. *MUTYH*-associated polyposis: 70 of 71 patients with biallelic mutations present with an attenuated or atypical phenotype. *Int J Cancer.* 2006; 119: 807-14.
- Gupta S, Provenzale D, Llor X, Halverson AL, Grady W, Chung DC, et al. NCCN Guidelines Insights: Genetic/Familial High-Risk Assessment: Colorectal, Version 2.2019. *J Natl Compr Canc Netw.* 2019; 17: 1032-41.
- 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.
- Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the functional effect of amino acid substitutions and indels. *PLoS One.* 2012; 7: e46688.
- Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods.* 2014; 11: 361-2.
- Reva B, Antipin Y, Sander C. Determinants of protein function revealed by combinatorial entropy optimization. *Genome Biol.* 2007; 8: R232.
- Capriotti E, Fariselli P, Casadio R. I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Res.* 2005; 33: 306-10.
- Bon L, Lucchetti C, Portolan F, Pagan M. MUPRO: a multipurpose robot. *Int J Neurosci.* 2002; 112: 855-68.
- Chen C-W, Lin J, Chu Y-W. iStable: off-the-shelf predictor integration for predicting protein stability changes. *BMC Bioinformatics.* 2013; 14: S5.
- Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 2018; 46: 296-303.

14. Lüthy R, Bowie JU, Eisenberg D. Assessment of protein models with three-dimensional profiles. *Nature*. 1992; 356: 83-5.
 15. Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR*. 1996; 8: 477-86.
 16. Wallner B, Elofsson A. Can correct protein models be identified? *Protein Sci*. 2003; 12: 1073-86.
 17. Liu W, Xie Y, Ma J, Luo X, Nie P, Zuo Z, et al. IBS: an illustrator for the presentation and visualization of biological sequences. *Bioinformatics*. 2015; 31: 3359-61.
 18. Lin JH, Walter P, Yen TSB. Endoplasmic reticulum stress in disease pathogenesis. *Annu Rev Pathol*. 2008; 3: 399-425.
 19. Urra H, Dufey E, Avril T, Chevet E, Hetz C. Endoplasmic reticulum stress and the hallmarks of cancer. *Trends Cancer*. 2016; 2: 252-62.
 20. Wang M, Wey S, Zhang Y, Ye R, Lee AS. Role of the unfolded protein response regulator GRP78/BiP in development, cancer, and neurological disorders. *Antioxid Redox Signal*. 2009; 11: 2307-16.
 21. Li J, Woods SL, Healey S, Beesley J, Chen X, Lee JS, et al. Point mutations in Exon 1B of APC reveal gastric adenocarcinoma and proximal polyposis of the stomach as a familial adenomatous polyposis variant. *Am J Hum Genet*. 2016; 98: 830-42.
 22. Fu Z, Shrubsole MJ, Smalley WE, Wu H, Chen Z, Shyr Y, et al. Lifestyle factors and their combined impact on the risk of colorectal polyps. *Am J Epidemiol*. 2012; 176: 766-76.
 23. Rivera B, Perea J, Sánchez E, Villapún M, Sánchez-Tomé E, Mercadillo F, et al. A novel AXIN2 germline variant associated with attenuated FAP without signs of oligodontia or ectodermal dysplasia. *Eur J Hum Genet*. 2014; 22: 423-6.
 24. Weren RDA, Ligtenberg MJL, Kets CM, De Voer RM, Verwiel ETP, Spruijt L, et al. A germline homozygous mutation in the base-excision repair gene NTHL1 causes adenomatous polyposis and colorectal cancer. *Nat Genet*. 2015; 47: 668-71.
 25. Adam R, Spier I, Zhao B, Kloth M, Marquez J, Hinrichsen I, et al. Exome sequencing identifies biallelic MSH3 germline mutations as a recessive subtype of colorectal adenomatous polyposis. *Am J Hum Genet*. 2016; 99: 337-51.
 26. Spier I, Holzapfel S, Altmüller J, Zhao B, Horpaopan S, Vogt S, et al. Frequency and phenotypic spectrum of germline mutations in POLE and seven other polymerase genes in 266 patients with colorectal adenomas and carcinomas. *Int J Cancer*. 2015; 137: 320-31.
 27. Palles C, Cazier J-B, 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.
 28. Bellido F, Pineda M, Aiza G, Valdes-Mas R, Navarro M, Puente DA, 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: 325-32.
 29. Moreno JC, Bikker H, Kempers MJ, van Trotsenburg AS, Baas F, de Vijlder JJ, et al. Inactivating mutations in the gene for thyroid oxidase 2. *N Engl J Med*. 2002; 347: 95-102.
 30. El Hassani RA, Benfares N, Caillou B, Talbot M, Sabourin JC, Belotte V, et al. Dual oxidase2 is expressed all along the digestive tract. *Am J Physiol Gastrointest Liver Physiol*. 2005; 288: 51-5.
 31. Grasberger H, Gao J, Nagao-Kitamoto H, Kitamoto S, Zhang M, Kamada N, et al. Increased expression of DUOX2 is an epithelial response to mucosal dysbiosis required for immune homeostasis in mouse intestine. *Gastroenterology*. 2015; 149: 1849-59.
 32. Burgueño JF, Fritsch J, Santander AM, Brito N, Fernández I, Pignac-Kobinger J, et al. Intestinal epithelial cells respond to chronic inflammation and dysbiosis by synthesizing H2O2. *Front Physiol*. 2019; 10: 1-13.
 33. Hayes P, Dhillon S, O'Neill K, Thoeni C, Hui KY, Elkadri A, et al. Defects in nicotinamide-adenine dinucleotide phosphate oxidase genes NOX1 and DUOX2 in very early onset inflammatory bowel disease. *Cell Mol Gastroenterol Hepatol*. 2015; 1: 489-502.
 34. Parlato M, Charbit-Henrion F, Hayes P, Tiberti A, Aloï M, Cucchiara S, et al. First identification of biallelic inherited DUOX2 inactivating mutations as a cause of very early onset inflammatory bowel disease. *Gastroenterology*. 2017; 153: 609-11.
 35. Levine AP, Pontikos N, Schiff ER, Jostins L, Speed D, Lovat LB, et al. Genetic complexity of Crohn's Disease in two large ashkenazi jewish families. *Gastroenterology*. 2016; 151: 698-709.
 36. Zhang X, Han J, Feng L, Zhi L, Jiang D, Yu B, et al. DUOX2 promotes the progression of colorectal cancer cells by regulating the AKT pathway and interacting with RPL3. *Carcinogenesis*. 2020; bgaa056: 1-13.
 37. Kang KA, Ryu YS, Piao MJ, Shilnikova K, Kang HK, Yi JM, et al. DUOX2-mediated production of reactive oxygen species induces epithelial mesenchymal transition in 5-fluorouracil resistant human colon cancer cells. *Redox Biol*. 2018; 17: 224-35.
 38. Kita H, Hikichi Y, Hikami K, Tsuneyama K, Cui ZG, Osawa H, et al. Differential gene expression between flat adenoma and normal mucosa in the colon in a microarray analysis. *J Gastroenterol*. 2006; 41: 1053-63.
 39. Wu Y, Antony S, Hewitt SM, Jiang G, Yang SX, Meitzler JL, et al. Functional activity and tumor-specific expression of dual oxidase 2 in pancreatic cancer cells and human malignancies characterized with a novel monoclonal antibody. *Int J Oncol*. 2013; 42: 1229-38.
 40. Donkó A, Péterfi Z, Sum A, Leto T, Geiszt M. Dual oxidases. *Philos Trans R Soc Lond B Biol Sci*. 2005; 360: 2301-8.
 41. Ron D, Walter P. Signal integration in the endoplasmic reticulum unfolded protein response. *Nat Rev Mol Cell Biol*. 2007; 5: 1-11.
 42. Haze K, Yoshida H, Yanagi H, Yura T, Mori K. Mammalian transcription factor ATF6 is synthesized as a transmembrane protein and activated by proteolysis in response to endoplasmic reticulum stress. *Mol Biol Cell*. 1999; 10: 3787-99.
 43. Coleman OI, Lobner EM, Bierwirth S, Sorbie A, Waldschmitt N, Rath E, et al. Activated ATF6 induces intestinal dysbiosis and innate immune response to promote colorectal tumorigenesis. *Gastroenterology*. 2018; 155: 1539-52.
- Cite this article as:** Yang M, Zhao Y, Ding Y, Wang J, Tan Y, Xu D, et al. A truncated protein product of the germline variant of the *DUOX2* gene leads to adenomatous polyposis. *Cancer Biol Med*. 2021; 18: 215-226. doi: 10.20892/j.issn.2095-3941.2020.0305