



OPEN ACCESS

Original research

Germline mutations in *WNK2* could be associated with serrated polyposis syndrome

Yasmin Soares de Lima,¹ Coral Arnau-Collell,¹ Jenifer Muñoz,¹ Cristina Herrera-Pariente,¹ Leticia Moreira,¹ Teresa Ocaña,¹ Marcos Díaz-Gay,^{1,2} Sebastià Franch-Expósito,^{1,3} Miriam Cuatrecasas,⁴ Sabela Carballal,¹ Anael Lopez-Novo,⁵ Lorena Moreno,¹ Guerau Fernández,⁶ Aranzazu Díaz de Bustamante,⁷ Sophia Peters,⁸ Anna K Sommer,⁸ Isabel Spier ^{8,9} Iris B A W te Paske,¹⁰ Yasmijn J van Herwaarden,¹¹ Antoni Castells,¹ Luis Bujanda,¹² Gabriel Capellà,¹³ Verena Steinke-Lange,^{14,15} Khalid Mahmood,^{16,17,18} JiHoon Eric Joo,^{16,17} Julie Arnold,¹⁹ Susan Parry,¹⁹ Finlay A Macrae,^{20,21,22} Ingrid M Winship,^{21,22} Christophe Rosty,^{16,17,23,24} Joaquin Cubiella,²⁵ Daniel Rodríguez-Alcalde,²⁶ Elke Holinski-Feder,^{14,15} Richarda de Voer ¹⁰ Daniel D Buchanan ^{16,17,21} Stefan Aretz,^{8,9} Clara Ruiz-Ponte,⁵ Laura Valle,¹³ Francesc Balaguer,¹ Laia Bonjoch,¹ Sergi Castellvi-Bel ¹

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/jmg-2022-108684>).

For numbered affiliations see end of article.

Correspondence to

Dr Sergi Castellvi-Bel, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Rosselló 153, Barcelona, 08036, Spain; sbel@recerca.clinic.cat and Dr Laia Bonjoch, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Rosselló 153, Barcelona, Spain; bonjoch@recerca.clinic.cat

LB and SC-B are joint senior authors.

Received 6 May 2022
Accepted 27 September 2022
Published Online First 21 October 2022



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY. Published by BMJ.

To cite: Soares de Lima Y, Arnau-Collell C, Muñoz J, et al. *J Med Genet* 2023;**60**:557–567.

ABSTRACT

Background Patients with serrated polyposis syndrome (SPS) have multiple and/or large serrated colonic polyps and higher risk for colorectal cancer. SPS inherited genetic basis is mostly unknown. We aimed to identify new germline predisposition factors for SPS by functionally evaluating a candidate gene and replicating it in additional SPS cohorts.

Methods After a previous whole-exome sequencing in 39 SPS patients from 16 families (discovery cohort), we sequenced specific genes in an independent validation cohort of 211 unrelated SPS cases. Additional external replication was also available in 297 SPS cases. The *WNK2* gene was disrupted in HT-29 cells by gene editing, and *WNK2* variants were transfected using a lentiviral delivery system. Cells were analysed by immunoblots, real-time PCR and functional assays monitoring the mitogen-activated protein kinase (MAPK) pathway, cell cycle progression, survival and adhesion.

Results We identified 2 rare germline variants in the *WNK2* gene in the discovery cohort, 3 additional variants in the validation cohort and 10 other variants in the external cohorts. Variants c.2105C>T (p.Pro702Leu), c.4820C>T (p.Ala1607Val) and c.6157G>A (p.Val2053Ile) were functionally characterised, displaying higher levels of phospho-PAK1/2, phospho-ERK1/2, CCND1, clonogenic capacity and MMP2.

Conclusion After whole-exome sequencing in SPS cases with familial aggregation and replication of results in additional cohorts, we identified rare germline variants in the *WNK2* gene. Functional studies suggested germline *WNK2* variants affect protein function in the context of the MAPK pathway, a molecular hallmark in this disease.

INTRODUCTION

Colorectal cancer (CRC) is one of the most common cancers worldwide with a significant

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Serrated polyposis syndrome (SPS) predisposes to colorectal cancer, and its inherited genetic basis is mostly unknown.

WHAT THIS STUDY ADDS

⇒ Germline *WNK2* variants altering the MAPK pathway, a hallmark of this disease, could predispose to SPS.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ The identification of a novel hereditary SPS factor would permit a more accurate diagnosis of SPS patients and facilitate genetic counselling and prevention.

associated mortality. Aside from lung cancer, with an avoidable environmental cause, CRC is responsible for more deaths than any other malignancy in Western countries.¹ The vast majority of CRC cases develop through an adenoma-carcinoma sequence.² In recent years, another carcinogenesis pathway has been identified: the serrated pathway, starting from a different precancerous lesion, the serrated polyp. Although serrated polyps were previously considered indolent, current evidence estimates they are the precursor lesion for up to 30% of CRC cases.³

Serrated polyposis syndrome (SPS) is a clinical condition characterised by the presence of multiple and/or large serrated polyps in the colon, as well as an associated higher risk of CRC.^{4,5} The following criteria were established by the WHO in 2010 in order to help identifying SPS patients: (1) at least five serrated lesions/polyps proximal to the sigmoid colon with two or more of these being >10 mm, (2) any number of serrated polyps

proximal to the sigmoid colon in an individual who has a first-degree relative with serrated polyposis and (3) >20 serrated polyps of any size but distributed throughout the colon.⁶ This arbitrary definition is not based on any genetic alteration and has been considered somehow restrictive, leading to underdiagnosis of this syndrome. Recently, it was updated to not include the second criterion. Also, new criterion I includes polyps proximal to the rectum and polyps now have to be ≥ 5 mm. The updated criterion II now explicitly states that ≥ 5 of the serrated polyps should be located proximal to the rectum.⁷ Although its prevalence in the population is unknown, it could be higher than expected according to data from CRC screening.^{8–11} It is also probably underrecognised due insufficient knowledge in the medical community, the difficult endoscopic detection of serrated lesions/polyps (small size and flat morphology), the lack of understanding regarding germline predisposition and the absence of associated symptoms.

CRC, including those cases associated with SPS, as for other complex diseases, are caused by both genetic and environmental factors.¹² Smoking, body mass index and alcohol intake have been highlighted as important environmental risk factors for serrated polyps.¹³ Importantly, twin studies also showed that around 13%–30% of the variation in CRC susceptibility involves inherited genetic factors.¹⁴ *APC*, *MUTYH* and the mismatch repair (MMR) genes are among the most relevant genes involved in the main forms of hereditary CRC, familial adenomatous polyposis, *MUTYH*-associated polyposis (MAP) and Lynch syndrome.¹⁵ However, SPS is a disease with mostly unknown inherited genetic basis compared with other gastrointestinal polyposis syndromes. It has also been advocated that SPS may not be hereditary but mostly environmental. However, familial clustering and a high CRC risk for first-degree relatives of SPS patients have been described, which supports the involvement of germline predisposition in a subset of cases.¹⁶

First reported in 2014, germline loss-of-function variants in *RNF43* were associated with the development of multiple sessile serrated adenomas.¹⁷ *RNF43*, as regulator of the DNA damage response and negative regulator of Wnt signalling, was considered a likely SPS susceptibility gene. Subsequent studies have highlighted that *RNF43* accounts for a small proportion (<3%) of the germline predisposition in SPS.^{18,19} Among the list of other genes reported so far to be potentially involved in SPS germline predisposition, *MUTYH* was reported in very few cases although its role in SPS is probably not relevant.^{20,21} Other genes involved in polyposis predisposition such as *BMPRIA*, *SMAD4*, *PTEN* and *GREM1* have been also screened with negative results.²²

Additional efforts have been undertaken to identify candidate genes for germline predisposition to SPS. Our research group has postulated further germline candidates for SPS by using combined whole-exome sequencing (WES) and linkage studies in families with multiple members affected by SPS²³ and by performing germline and somatic WES in 39 patients from 16 SPS families showing familial aggregation mainly compatible with an autosomal dominant pattern of inheritance.²⁴ This last study highlighted *ANXA10*, *ASXL1*, *CFTR*, *DOT1L*, *HIC1*, *INO80*, *KLF3*, *MCM3AP*, *MCM8*, *PDLIM2*, *POLD1*, *TP53BP1*, *WNK2* and *WRN* as candidate genes for SPS germline predisposition.

Accordingly, the main objectives of this study were the identification of novel germline causal genes for SPS predisposition by replicating a candidate gene in independent SPS cohorts and performing a functional evaluation of the detected rare variants. Confirmation of germline predisposition to SPS would permit a more accurate and adequate diagnosis of patients, as well as facilitating genetic counselling and prevention.

MATERIALS AND METHODS

Patients

The discovery cohort comprised 39 patients from 16 families (≥ 2 patients per family) diagnosed with SPS and fulfilling the 2010 WHO criteria.⁶ The updated 2019 WHO criteria were not available when this study was initiated and developed for the discovery cohort.⁷ A complete clinical characterisation of this discovery cohort was previously published.²⁴ No patients in the discovery cohort presented with pathogenic variants in *MUTYH*, *APC* or the DNA MMR genes, when analysed using gene panel sequencing and screening for point mutation, copy-number variants and potential splicing alterations.

Two hundred and eleven unrelated Spanish SPS patients were recruited in high-risk CRC clinics at Hospital Clínic de Barcelona, Institut Català d'Oncologia-IDIBELL and Fundació Pública Galega de Medicina Xenómica and were used as validation cohort. Additional external SPS cohorts ($n=297$) with available sequencing data for unrelated patients from the University of Bonn and the Medical Genetics Center Munich in Germany ($n=168$), the Radboud University Medical Centre in the Netherlands ($n=29$) and the Genetics of Colonic Polyposis Study from Australia and New Zealand ($n=100$) were also accessed. The 2010 WHO criteria were also used in this cohort for consistency.

Variant identification

For more details on variant identification and validation, see online supplemental material.

Variant prioritisation

Variant prioritisation was carried out considering several aspects. First, we only took into consideration those variants present in the canonical transcripts. Also, dominant and recessive analysis were pursued. Homozygous/compound heterozygous variants in relevant genes were not identified. Therefore, only heterozygous variants were further considered. In addition, a minimum allele frequency of 0.1% was required for variant filtering and only non-synonymous and/or truncating variants were prioritised. The missense variants had to fulfil at least three out of six pathogenic predictions used for analysis (PhyloP, SIFT, Polyphen, MutationTaster, CADD and LRT). The next crucial step of variant prioritisation considered data integration with the first cohort results. We prioritised genes that presented germline variants in both cohorts and conducted an extensive literature research over possible connections between candidate genes and SPS. Only candidate genes with rare, nonsynonymous/truncating or missense variants fulfilling at least three out of six pathogenic predictions detected in the discovery and the validation cohort were further considered. Among them, only those with a function compatible with SPS, CRC or cancer were selected.

Gene panel sequencing

For the validation cohort, a panel of 20 genes was designed using the DesignStudio online tool (Illumina, San Diego, USA). We included 14 genes suggested as plausible candidates to SPS in the discovery cohort comprising *ANXA10*, *ASXL1*, *CFTR*, *DOT1L*, *HIC1*, *INO80*, *KLF3*, *MCM3AP*, *MCM8*, *PDLIM2*, *POLD1*, *TP53BP1*, *WNK2* and *WRN*.²⁴ Additional genes were also included, such as *ANXA1* and *ANXA2* (histological markers for SPS, same family as *ANXA10*), *ASXL2* (same gene family as *ASXL1*), *POLE*, *FBLN2* (previously suggested by our group)²³ and *RNF43*, previously involved in germline SPS predisposition.

Gene panel analysis

The raw sequencing data were first analysed using the Miseq Reporter software (Illumina, San Diego, USA). First, the data were aligned to the hg19 human genome using the Burrows-Wheeler Aligner (BWA-MEM).²⁵ Then, variant calling was conducted using the Germline Variant Caller (Illumina, San Diego, USA). Variant annotation was performed as previously described using SnpEff and SnpSift software (<https://pcingola.github.io/SnpEff/>).^{24 26}

Functional characterisation of genetic variants

For details on the development of a cellular model for variant characterisation, ERK1/2 and PAK1 assays, see online supplemental material. All plasmids, antibodies, restriction enzymes and Taqman probes used in this study are listed in online supplemental table 1. Primer details are listed in online supplemental table 2. If not indicated otherwise, functional assays were developed with HT-29 cells cultured in McCoy 5A media supplemented with 5% FBS and 1 µg/mL of doxycycline.

MAPK pathway activity: ERK1/2 and PAK1 phosphorylation

To detect total phospho-ERK1/2, cells were stimulated with 1 ng/mL human epidermal growth factor (hEGF) for 10 min and assayed with the Phospho-ERK1 (T202/Y204)/ERK2 (T185/Y187) DuoSet IC ELISA kit according to the manufacturer's protocol (Bio-Techne, Minnesota, USA). Phospho-ERK1/2 levels between EGF-stimulated and non-stimulated conditions were quantified.

To detect phosphorylated PAK1, an In-Cell ELISA assay was done. The 96-well plates were coated with 50 µg/mL Poly-L-Lysine before cell seeding. Cells were stimulated with 10 ng/mL of hEGF for 5 min, immediately fixed and assayed with phospho-PAK1 (rabbit) and β-actin (mouse) antibodies overnight at 4°C. The multiplexed detection of both targets was performed with the antimouse Dylight 800 (ThermoFisher, Waltham, Massachusetts, USA) and antirabbit IRDye 680RD (LI-COR, Lincoln, Nebraska, USA) antibodies. Plates were scanned with Odyssey (LI-COR) and analysed with Image Studio 4.0 software.

CCND1 and MMP2 expression

Cells were seeded in P60 dishes at 600 000 cells per plate and left to grow for 2 days. Then, cells were stimulated with 1 ng/mL of hEGF for 16 hours in serum-free McCoy 5A media with 1 µg/mL of doxycycline. The next day, cells were detached using a cell-scrapper, and the RNA was extracted using the RNeasyMini Kit according to the manufacturer's protocol (Qiagen, Hilden, Germany). Retrotranscription and quantitative PCR were done as described in online supplemental material.

Clonogenic assay

Cells were seeded at low density, at 200 cells per well in a 6-well plate. Cells were maintained either in the presence of 1 ng/mL of hEGF or without hEGF. After 16 days, cells were fixed with methanol for 10 min and stained with a 0.5% crystal violet solution.

Adhesion assay

Before cell seeding, 96-well plates were coated with 5 µg/mL fibronectin in PBS 1X and incubated for 1 hour at room temperature. Next, the plate was dried out and blocked with BSA 1% for an additional hour. A total of 40 000 cells were seeded per well and left to attach for 60 min. Subsequently, unattached cells were removed by inversion; the plate was washed

carefully with serum-free McCoy 5A and fixed with methanol for 10 min. Finally, fixed cells were stained with a 0.5% crystal violet solution. Images were captured on an AID EliSpot reader system and analysed with the ReadPlate 3.0 plugin for Image J.

RESULTS

After variant prioritisation, only candidate genes with rare, non-synonymous/truncating or missense variants fulfilling at least three out of six pathogenic predictions detected in the discovery and the validation cohort were further considered. The WNK lysine deficient protein kinase 2 (WNK2) gene stood out among others for being a negative regulator of the mitogen-activated protein kinase (MAPK) pathway.²⁷ MAPK cascades are central signalling pathways that regulate basic processes, including cell proliferation, differentiation, stress responses and apoptosis. Mutations in these pathways lead to their constitutive activation and uncontrolled cell proliferation.²⁸ One of the cascades, MAPK/ERK, is of particular interest in SPS because one of its components, *BRAF*, shows activating mutations in approximately 75% of sessile serrated polyps.²⁹ Due to its role as a negative regulator of this pathway, *WNK2* was considered a promising candidate gene for germline SPS predisposition.

In the discovery cohort, two *WNK2* variants were detected including c.4820C>T (p.Ala1607Val) in family CAR_SPS.4, and c.6157G>A (p.Val2053Ile) in family CAR_SPS.6 (figure 1). These variants were classified as potentially damaging by five out of six missense pathogenicity prediction tools. Families showed CRC family history, and variants were detected in two family members affected with SPS, although no additional segregation analysis was possible. MMR system was preserved in the analysed serrated lesions from both families and the CAR_SPS.6 polyp was *BRAF* mutated. Loss of heterozygosity seeking a potential deletion of the wild-type (WT) *WNK2* allele was not detected in the analysed serrated lesions. We also performed WES on the most advanced serrated lesion available in one individual from each family, which allowed performing somatic mutational profiles. The single-base substitution (SBS) signatures SBS.1 and SBS.5, considered clock-like mutational signatures, were the most represented in both samples, and no other distinctive signature was apparent.²⁴

In a validation SPS cohort of 211 unrelated patients, gene panel sequencing revealed four additional rare, missense variants in *WNK2* including c.2105C>T (p.Pro702Leu) in PanSPS_044, c.2792C>T (p.Thr931Met) in PanSPS_095, c.3341C>T (p.Thr1114Met) in PanSPS_078 and c.5588T>C (p.Leu1863Pro) in PanSPS_055. The c.2792C>T variant was detected in a male SPS patient (onset at 67 years) in a family where an additional sample for variant segregation was available. It corresponded to a sister of the proband with CRC (72 y.o.). However, this variant was finally excluded for further studies since segregation was not confirmed. The tumour sample in PanSPS_044 (III-1) showed loss of expression for MLH1/PMS2 and was *BRAF* mutated and *MGMT* methylated. No similar information was available regarding MMR system or somatic alterations for any serrated lesion from PanSPS_055 and PanSPS_078. Pedigrees and the five *WNK2* variants that were selected are summarised in figure 1 and table 1.

Functional characterisation of *WNK2* depletion

To unequivocally assess the functional effect of the previous candidate variants and assess their link to SPS, it was important to reduce the masking effect of the endogenous WT *WNK2* expression in the selected cellular model. For this reason, we

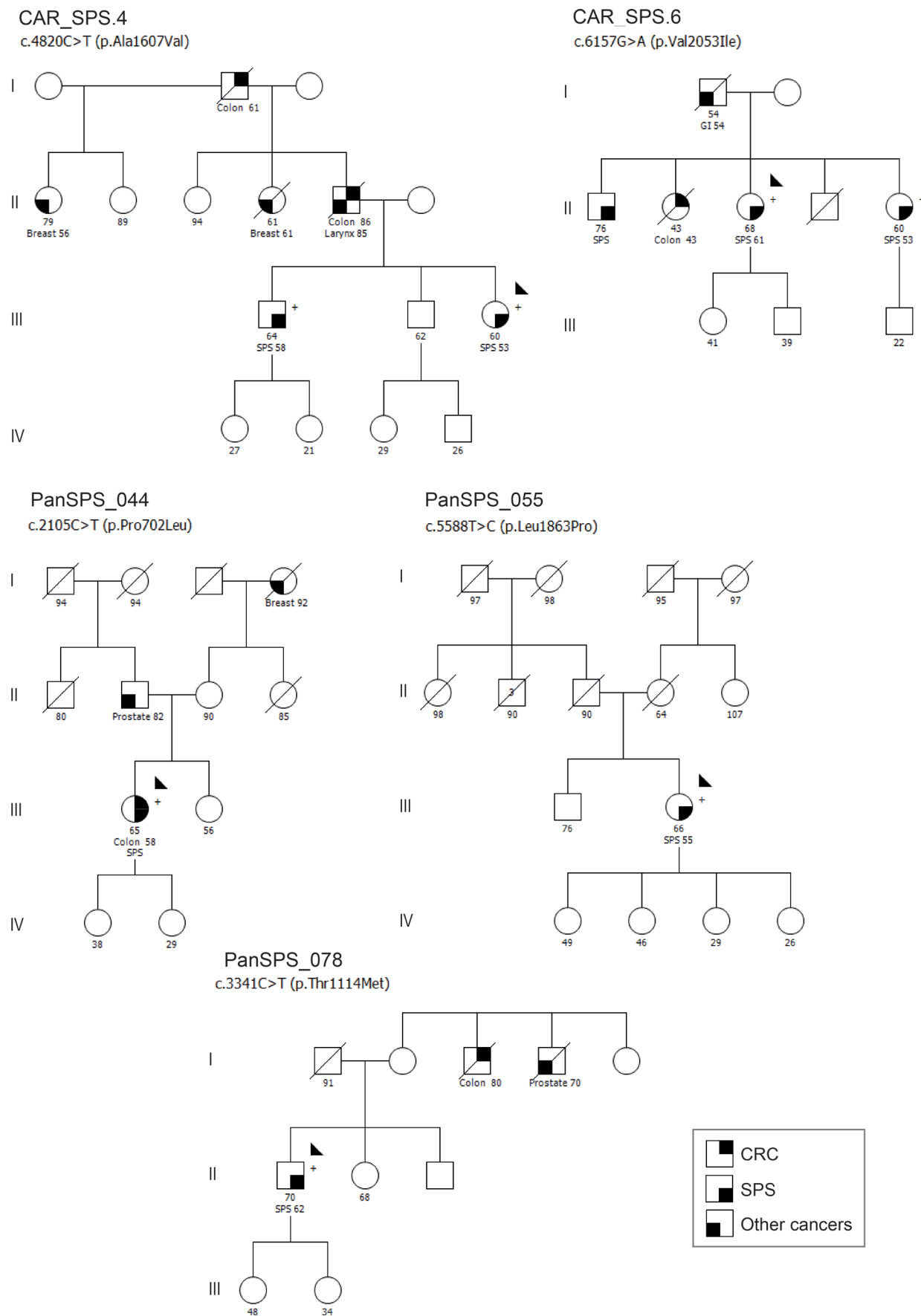


Figure 1 Pedigrees of five SPS families. Filled symbol indicates affected for CRC (upper right quarter), SPS (lower right quarter) or other types of cancer (lower left quarter). CRC, colon, breast, larynx, GI (gastrointestinal) and prostate refer to the type of cancer. Ages at diagnosis are depicted. The proband is indicated by an arrow. Variant carriers are indicated by (+). CRC, colorectal cancer; SPS, serrated polyposis syndrome.

Table 1 Rare germline variants identified in the *WNK2* gene in the discovery and validation SPS cohorts

Variant	Exon	Prediction tools	gnomAD	Allele count	Number of homzygotes	Family	SPS criteria	CRC index case	CRC family history	Cohort
c.2105C>T (p.Pro702Leu)	9	5	0.000125	19/152,210	0	PanSPS_044	3	Y	N	Validation
c.3341C>T (p.Thr1114Met)	12	3	0.0000526	8/152,182	0	PanSPS_078	1	N	Y	Validation
c.4820C>T (p.Ala1607Val)	20	5	0.000138	21/152,228	0	CAR_SPS.4	1,3	N	Y	Discovery
c.5588T>C (p.Leu1863Pro)	23	5	0	0	0	PanSPS_055	3	N	N	Validation
c.6157G>A (p.Val2053Ile)	25	5	0.0000657	10/152,200	0	CAR_SPS.6	2,3	N	Y	Discovery

Prediction tools: number of pathogenicity positive predictions according to missense bioinformatics prediction tools (PhyloP, SIFT, Polyphen, MutationTaster, CADD and LRT). SPS criteria: according to the 2010 WHO SPS clinical criteria. CRC family history: defined as presence of any CRC case in the family besides the index case.

CRC, colorectal cancer; gnomAD, Genome Aggregation Database; N, no; SPS, serrated polyposis syndrome; Y, yes.

performed a two-step genetic engineering strategy. First, we knocked-out *WNK2* in the human cell line HT-29 by CRISPR-Cas9. Then, we reintroduced each of the *WNK2* variants of interest using a lentiviral delivery system, and specific functional studies were carried out. We confirmed the CRISPR-mediated *WNK2* gene editing by Sanger sequencing and selected two clones (*WNK2*^{KO2} and *WNK2*^{KO7}) with no expression of *WNK2* at both mRNA and protein levels (online supplemental figure 1).

Afterwards, we evaluated the phenotype of the selected *WNK2*^{KO} clones. *WNK2* is a negative regulator of the ERK1/2 MAPK signalling cascade, where growth factors trigger signal transduction through a series of sequential protein phosphorylations. Specifically, *WNK2* modulates the Rac1/PAK1-mediated activation of ERK1/2 (figure 2A). Therefore, we assessed the phosphorylated status of both ERK1/2 and PAK1/2. Treatment with hEGF promoted a dose-dependent stimulation of these targets in HT-29 control cells. We observed that the lack of *WNK2* facilitated the activation of this pathway, as both clones displayed a higher increase in the phosphorylation of ERK1/2 and PAK1/2 (figure 2B,C), even though HT-29 cells are *BRAF* mutated. The effect was even more outstanding in clone *WNK2*^{KO2}, which showed the highest ERK1/2 and PAK1/2 phosphorylation levels at all tested doses.

We further characterised *WNK2*^{KO2} cells to determine the functional consequences of the alteration in the MAPK pathway. Since this pathway promotes cell cycle progression, we tested Cyclin D1 (*CCND1*) expression, one of the main PAK/ERK targets in mitogenic signalling. We detected increased *CCND1* expression levels (figure 2D) and a higher clonogenic capacity (figure 2E) of *WNK2*^{KO2} cells in comparison with control cells after hEGF induction, indicating that *WNK2* depletion altered cell cycle progression and cell proliferation.

The activation of the MAPK pathway promotes the expression of metalloproteinases, which degrade the extracellular matrix and are associated with cellular adhesion and a more aggressive phenotype. Specifically, *WNK2* has been described to negatively regulate two metalloproteinases, MMP2 and MMP9.³⁰ Therefore, we also analysed the effects of *WNK2* depletion on cell adhesion and matrix metalloproteinase-2 (*MMP2*) expression, for which *WNK2* has already been described to work as a negative regulator. *WNK2*^{KO2} exhibited increased fibronectin-mediated cell attachment (figure 2F) and higher *MMP2* expression levels (figure 2G). Altogether, these results show that deletion of *WNK2* in HT29 cells alters the regulation of downstream mediators of the MAPK pathway, one of the main pathways that drive serrated tumorigenesis, suggesting that alterations on this gene may be associated with the development of the serrated polyposis phenotype.

Functional characterisation of *WNK2* germline variants

Three *WNK2* candidate variants were selected to investigate their functional effect including p.Ala1607Val and p.Val2053Ile from the discovery cohort and p.Pro702Leu from the validation cohort (table 1). Families carrying these variants had the most severe clinical presentation including SPS aggregation or SPS and CRC in the index case. These variants were also classified as potentially damaging by five out of six pathogenicity prediction tools. These variants were designed by site-directed mutagenesis and individually reintroduced in both *WNK2*^{KO2} and *WNK2*^{KO7} clones. As a control, the WT *WNK2* sequence was also introduced in both clones to rescue the original phenotype. We selected the optimal doxycycline dose (1 µg/mL) to obtain successful gene expression levels for each of the introduced *WNK2* sequences. All *WNK2* variants were equally expressed at both RNA and protein levels (online supplemental figure 2).

We then proceeded to do the functional characterisation of the selected variants. We first examined the phosphorylation status of ERK1/2 and PAK1/2, two key components of the MAPK pathway (figure 2A). *WNK2*^{KO2} cells re-expressing either *WNK2*-WT or each of the selected variants were treated with hEGF, and both ERK1/2 and PAK1/2 phosphorylation levels were assessed. The activation of the pathway was determined before and after hEGF treatment (EGF +/– ratio). hEGF induced the phosphorylation of both ERK1/2 (figure 3A) and PAK1/2 (figure 3B) kinases. When comparing with the rescued WT *WNK2* phenotype, the MAPK activation was present to some extent for the three variants, being more noticeable with cells expressing the p.Pro702Leu variant.

Next, as the MAPK pathway promotes cell cycle progression, we tested whether *WNK2* variants would be implicated in *CCND1* expression. All missense *WNK2* variants promoted a moderate increase in *CCND1* expression in comparison with cells expressing the WT counterpart (figure 3C). A similar trend was observed when performing the clonogenic cell survival assay, in which cells expressing *WNK2* variants showed a higher survival rate (figure 3D).

Finally, we focused on cell adhesion and metalloproteinase *MMP2* expression. We first analysed the expression of *MMP2* after hEGF treatment. We observed a tendency for a higher expression of this marker for the three variants in comparison with cells expressing with *WNK2*-WT (figure 3E). Then, cells were subjected to an adhesion test to the extracellular matrix protein fibronectin, which is cleaved by *MMP2*. By doing so, we detected a greater adhesion capacity of cells expressing the

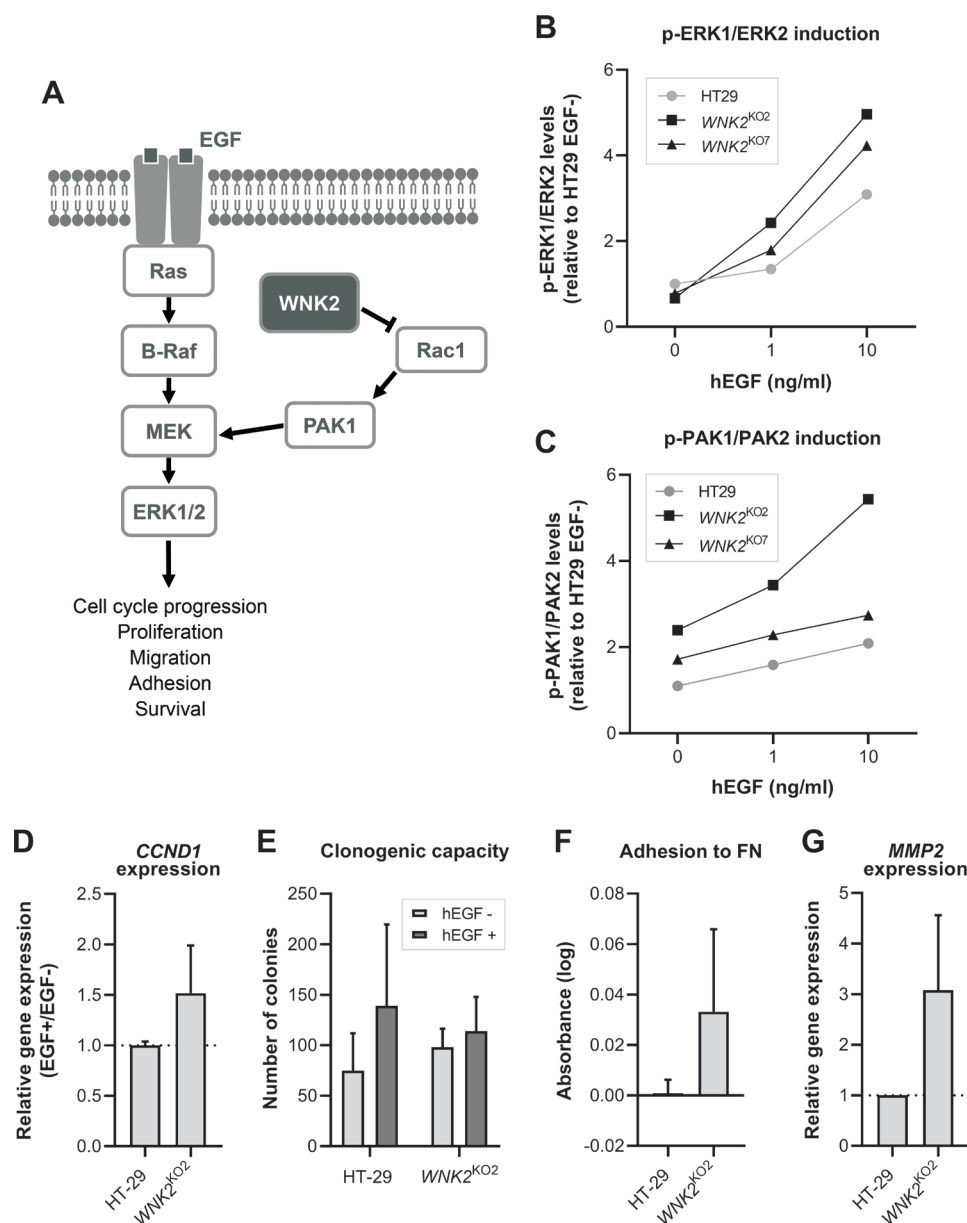


Figure 2 Functional characterisation of $WNK2^{KO}$ clones. (A) Diagram depicting the role of WNK2 in the ERK pathway. WNK2 affects GTP loading of Rac1, interfering in the cascade of the ERK signal transduction pathway. Adapted from Moniz and Jordan.³² (B) Dose-response of hEGF-induced ERK1/2 and (C) PAK1/2 phosphorylation levels in both $WNK2^{KO2}$ and $WNK2^{KO7}$ clones. Samples were assayed in triplicate. (D) *CCND1* mRNA relative expression after treatment with 1 ng/mL of hEGF. Data are expressed as EGF+/EGF- ratio, and mean \pm SD is represented (n=3). (E) Clonogenic capacity of cells cultured during 16 days in the presence or absence of 1 ng/mL of hEGF. Data represent mean \pm SD (n=4). (F) Fibronectin (FN)-mediated cell adhesion assay using crystal violet staining. Data represent mean \pm SD (n=3). (G) *MMP2* mRNA relative expression after treatment with 1 ng/mL of hEGF. Data represent mean \pm SD (n=3). The experiments were performed in triplicate and repeated three or four times, as indicated in each case. hEGF, human epidermal growth factor.

p.Pro702Leu and p.Ala1607Val variants (figure 3F), compared with the rescued WT phenotype.

The functional characterisation of $WNK2$ variants was replicated in clone $WNK2^{KO7}$ (online supplemental figure 3), focusing on the assays more directed to the MAPK pathway itself (ERK1/2 and PAK1/2), and similar results were obtained. Consistently, the activation of the MAPK pathway and the adhesion capacity of cells were higher when $WNK2$ variants were expressed, again with a prominent effect in the case of variant p.Pro702Leu.

All in all, these results suggested that $WNK2$ variants partially failed to repress the activation of this molecular pathway and were concordant in some measure with the alteration of the

MAPK pathway, supporting the malfunctioning of $WNK2$ variants.

Screening of the candidate gene variants in additional SPS cohorts

International SPS cohorts from Germany, Australia and the Netherlands with available WES or whole-genome sequencing data were further consulted seeking for additional rare, non-synonymous/truncating or missense $WNK2$ alterations in their patients (n=297). Genetic variants were selected based on low population frequency (<0.25%), deleterious effect on the protein and pathogenic bioinformatics prediction (Combined Annotation

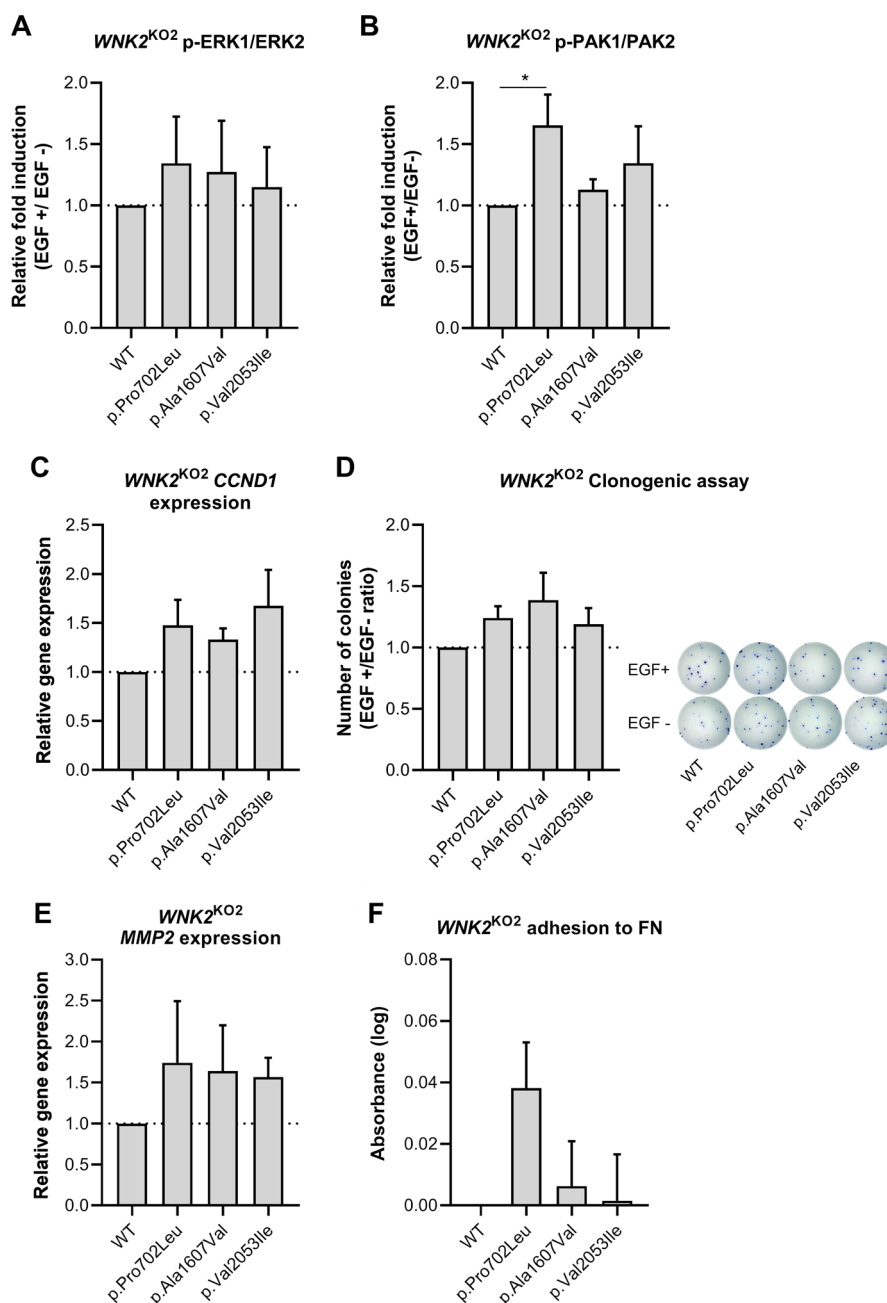


Figure 3 Variants in *WNK2* activate the MAPK pathway and contribute to cell cycle deregulation, cell proliferation and altered cellular adhesion. *WNK2*^{KO2} cells re-expressing either *WNK2* WT or each of the selected variants were functionally characterised. The variants were expressed successfully under the inducible promoter. The experiments were performed in triplicate and repeated three or four times, as indicated in each case. (A) ERK1/2 phosphorylation levels measured by ELISA in cell samples treated or untreated with 1 ng/mL hEGF. Data are displayed as EGF+/- ratio (n=3; mean±SD). (B) PAK1/2 phosphorylation levels measured by an In-Cell ELISA (ICE) assay in cell samples treated or untreated with 10 ng/mL hEGF. Data are displayed as EGF+/- ratio (n=3; mean±SD). (C) *CCND1* mRNA relative expression after a 16-hour treatment with 1 ng/mL of hEGF (n=3; mean±SD). (D) Clonogenic capacity of cells cultured during 16 days in the presence or absence of 1 ng/mL of hEGF. Data are displayed as EGF+/- ratio (n=4; mean±SD). On the right, representative images of methanol-fixed, crystal violet stained colonies. *P<0.05, analysis of variance with Fisher's LSD post hoc test. (E) *MMP2* mRNA relative expression after a 16-hour treatment with 1 ng/mL of hEGF (n=3; mean±SD). The variants were expressed successfully under the inducible promoter. (F) *WNK2*^{KO2} cells expressing either the WT sequence or each of the candidate variants were cultured on fibronectin (FN)-coated plates for 1 hour, and the adherent cells were detected by crystal violet staining (n=3; mean±SD). hEGF, human epidermal growth factor; WT, wild-type.

Dependent Depletion Phred (CADD) score >15). In summary, 10 additional rare, protein-altering *WNK2* genetic variants were identified in 12 SPS patients (table 2). Considering the number of variants found in all analysed SPS cohorts, the frequency of germline *WNK2* alterations in SPS could be considered ~3% (17/524, 3.24%).

To further test the association of *WNK2* with SPS predisposition, we also performed a gene-based burden test, where the aggregate burden of rare, protein-altering variants in *WNK2* was compared between our cases and control subjects. To do so, we accessed the data available from 262 healthy controls at the Collaborative Spanish Variant Server CSVS data (<http://csvs>).

Table 2 Rare, non-synonymous/truncating or missense variants in *WNK2* identified in additional international SPS cohorts

Variant	Exon	gnomAD	Allele number	Number of homozygotes	CADD	Patient	SPS criteria	Onset age	CRC index case	Cohort
c.106_107insG (p.Pro36Argfs*121)	1	0	0	0	N/A	BN-174	1,3	22	N	DE
c.1853G>A (p.Ser618Asn)	8	0	0	0	26.4	RB-1	1	57	N	NL
c.2758G>A (p.Ala920Thr)	11	0.0001315	20/152 056	0	15	GCP208001	1,3	27	N	AU
c.3418G>A (p.Gly1140Ser)	14	0	0	0	23.9	MUC-6	1,3	42	NA	DE
c.3623C>T (p.Thr1208Met)	15	0.000006582	1/151 930	0	26.6	MUC-4	*	30	N	DE
c.5476C>T (p.Arg1826Trp)	23	0	0	0	25.1	BN-210	3	55	N	DE
c.5656C>T (p.Arg1886Trp)	23	0.002233	219/152 194	1	21.2	RB-2	†	51	N	NL
						BN-145	3	29	NA	DE
						MUC-1116-01	1	20	N	DE
c.5906C>G (p.Pro1969Arg)	24	0	0	0	25.3	GCP038001	3	59	N	AU
c.6080C>G (p.Ala2027Gly)	25	0	0	0	25.1	BN-83	3	18	NA	DE
c.6512G>A (p.Ser2171Asn)	28	0	0	0	23	BN-104	1,3	21	NA	DE

A cut-off of 15 was used to select possible pathogenic variants. SPS criteria: according to the 2010 WHO SPS clinical criteria.

*This patient did not fulfil SPS criteria but presented a large serrated polyp at 30 and strong CRC family history.

†This patient presented a serrated polyp count between 1 and 10.

AU, Australia; CADD, Combined Annotation Dependent Depletion Phred score; CRC, colorectal cancer; DE, Germany; gnomAD, Genome Aggregation Database variant frequency; N, no; NA, not available; NL, Netherlands; SPS, serrated polyposis syndrome; Y, yes.

babelomics.org/).³¹ By applying the previous filters of frequency and pathogenicity, we identified two rare, protein-altering variants in this control dataset (2/262, 0.76%) and confirmed an enrichment for rare, nonsynonymous/truncating or missense variants in the *WNK2* gene in our SPS cohort ($\chi^2=4.55$, p value=0.03).

DISCUSSION

In this study, we initially analysed two independent SPS cohorts resulting in the identification of five rare, protein-altering germline variants in the *WNK2* gene. Additionally, examination of international SPS cohorts yielded 10 additional *WNK2* genetic variants in 12 SPS patients. To assess the impact on *WNK2* of three of the genetic variants identified in the original cohort, we developed a cellular model using CRISPR-Cas9 technology and further performed functional assays.

WNK2 is a member of the WNK 'With-No-Lysine(K)' kinase subfamily, in which four different kinases have been identified (*WNK1-4*). This class of kinases was described to play a role in organism development and osmoregulation, but also in cancer, such as gliomas, hepatocellular carcinoma and CRC.^{27 32 33} This protein is predominantly expressed in heart, brain and colon and, unlike the other three WNKs, is not expressed in kidney.

Noteworthy, the *WNK2* gene is located at chromosomal region 9q22.31, a region previously linked to familial CRC.³⁴ *WNK2* somatic mutations are found in several cancer types, including CRC, ovarian, hepatic and gastric cancer.^{35 36} *WNK2* has also been reported to be epigenetically silenced in pancreatic

adenocarcinoma and gliomas.^{37 38} Interestingly, *WNK2* downregulation has also been detected in serrated polyps.³⁹

WNK2 regulates the phosphorylation and activation of ERK1/2, one of the best studied MAPK cascades.^{40 41} The MAPK/ERK pathway controls many cellular processes, including cellular proliferation, cell survival, migration, invasion and adhesion. This pathway is deregulated in around one-third of all human cancers, being remarkable in CRC.⁴² Most of the alterations constitutively activate it and occur in the upstream elements of the signalling pathway, such as mutations in *BRAF* or *KRAS*.²⁸ Due to its central role in many basic cellular processes, this pathway is tightly regulated at different levels of the cascade. In this sense, *WNK2* negatively regulates it by controlling the GTP-loading of Rac1.³⁷ Rac1, when activated, triggers a signalling cascade resulting in the activation of PAK1, phosphorylation of the S298 residue of MEK1 and consequent activation of ERK1/2 (figure 2A).^{40 41}

We functionally characterised three rare, missense germline *WNK2* variants detected in SPS cohorts including c.2105C>T (p.Pro702Leu), c.4820C>T (p.Ala1607Val) and c.6157G>A (p.Val2053Ile) by focusing on whether they altered the MAPK signalling cascade. The *WNK2* knockout cellular models displayed higher phospho-PAK1/2 and phospho-ERK1/2 levels, implying that *WNK2* depletion promoted the activation of the MAPK pathway and in agreement with previous results in HeLa and HT-29 cells,^{40 41} pancreatic adenocarcinoma tissue³⁸ and in hepatocarcinoma cell lines.⁴³ All tested *WNK2* variants showed the same tendency, with a prominent *WNK2* malfunction

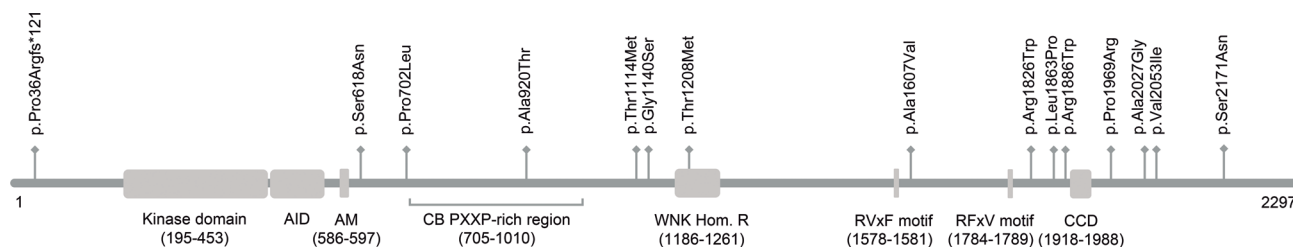


Figure 4 Schematic representation of *WNK2* indicating the location of the identified variants. The *WNK2* protein has a main kinase domain, an autoinhibitory domain (AD) and short homology regions shared with the other WNK kinases: an acidic motif (AM), the WNK homology region (WNK Hom R) and a coiled-coil domain (CCD). Some motifs and protein binding sites are also indicated, such as the compositionally biased (CB) PXXP-rich region, the RVxF motif and the RFxV motif. Additional predicted eukaryotic linear motifs (ELMs) can be found in online supplemental table 3.

detected for the p.Pro702Leu variant, which displayed the most significant phosphorylation levels of both PAK1/2 and ERK1/2 in *WNK2*^{KO2} and *WNK2*^{KO7} clones.

We next evaluated some of the main downstream cellular processes affected by MAPK/ERK deregulation, such as cell cycle progression and cell survival. *CCND1* is an important downstream effector of the MAPK pathway and has its expression levels regulated in response to mitogenic signals.⁴⁴ We observed that *CCND1* expression levels were increased, although not statistically significant for all *WNK2* variants, and that the activation profile was similar to the observed PAK1/2 phosphorylation levels. As multiple signalling pathways can converge on *CCND1* transcriptional activation, we hypothesise that *WNK2* malfunction can influence *CCND1* expression by ERK dependent and independent pathways driven by PAK1.⁴⁵

Cell survival is also influenced by regulation of the MAPK pathway. Cells harbouring the *WNK2* variants showed an increase in their clonogenic capacity compared with *WNK2*-WT cells. Our results agree with previous results in hepatocellular carcinoma *WNK2*-silenced cells, in which re-expression of *WNK2*-WT suppressed colony formation, whereas introducing mutated *WNK2* increased colony formation capacity.⁴³ In addition, upregulation of *WNK2* expression has also been linked to apoptosis, senescence and autophagy in colon cancer and glioma cells, which are processes focused on cell cycle control.^{46 47}

Cell adhesion is an important feature of the cell malignancy process and is closely regulated by PAK1 activation.⁴⁸ Matrix metalloproteinases, which have been widely described as MAPK transcriptional targets, are responsible for extracellular matrix degradation and have an important role in cellular invasion processes. For this reason, we assessed both *MMP2* expression and fibronectin-mediated adhesion of our cellular models. Cells expressing tested variants showed higher *MMP2* levels than *WNK2*-WT cells. Moreover, the adhesion capacity of p.Pro702Leu *WNK2* variant stood out among others, suggesting a possible impact of *WNK2* impairment in extracellular matrix remodelling. Previous work in glioma cell lines had already described the negative correlation between *WNK2* and *MMP2* expression and activity, highlighting the *WNK2* importance in cell invasion and migration.^{30 49}

WNK protein kinases have a conserved kinase domain, an autoinhibitory domain, one or two coiled-coil domains, and numerous protein interaction motifs, including PXXP proline-rich motifs and RFX/V/I motifs.³² Further protein motif prediction in *WNK2* was performed with the eukaryotic linear motif resource to search for motifs affected by the identified genetic variants.⁵⁰ According to this resource, a summary of the predicted effect of the 15 *WNK2* genetic variants is available in online supplemental table 3, and their location in the protein structure is depicted in figure 4. Most variants are located to motifs predicted to have a functional meaning, including protein binding and phosphorylation recognition sites.⁵¹ Overall, the multiple protein–protein interaction motifs in *WNK2* seem to reveal that, apart from its kinase activity, it could be also considered a scaffolding protein that facilitates protein–protein interactions in the MAPK cascade. In this sense, mutations in the kinase domain and those located along the *WNK2* sequence could impair its role as a MAPK regulator.

Moreover, it should be highlighted that the *WNK2* gene seems to be intolerant to loss-of-function genetic variation as evident by gene constraint scores (pLI=1, LOEUF ratio=0.12 (0.07–0.21)).⁵² Together with our gene-burden test results, it would be supporting its potential role in germline predisposition to SPS.

Taking into account the number of variants found in our cohorts, the frequency of germline *WNK2* alterations in SPS could be considered ~3%. Undoubtedly, the present study is preliminary, and analysis of additional larger familial SPS cohorts and further functional studies are needed to provide more information about the prevalence and implication of germline *WNK2* mutations in SPS. As a limitation, our study used WES in the discovery cohort and alterations outside the coding sequence, in non-canonical transcripts or epimutations cannot be ruled out. Additionally, it is important to perform continued segregation analyses of the reported *WNK2* variants in the affected families to confirm (or rule out) their pathogenicity. Finally, further studies in somatic tissue of serrated lesions or CRC in *WNK2* carriers could shed light regarding mutational signatures associated with this genetic defect, as well as organoid modelling could also help to confirm the involvement of this gene in the sequence of events moving towards a serrated phenotype.

In summary, our findings indicate that germline *WNK2* variants in SPS patients may be implicated in inherited predisposition to SPS and postulate that the disruption of the role of *WNK2* as a MAPK regulator could be the plausible underlying mechanism. However, a thorough assessment of the evidence for and against pathogenicity is still needed, as well as replication in additional SPS cohorts, in order to clarify a causative role for germline *WNK2* variants in SPS.

Author affiliations

¹Department of Gastroenterology, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBERehd), Hospital Clínic, Barcelona, Spain

²Department of Cellular and Molecular Medicine, University of California San Diego (UCSD), San Diego, CA, USA

³Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY, USA

⁴Department of Pathology, Hospital Clínic de Barcelona, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBERehd) and Tumor Bank-Biobank, Barcelona, Spain

⁵Fundación Pública Galega de Medicina Xenómica (FPGMX), Grupo de Medicina Xenómica-USC, Instituto de Investigación Sanitaria de Santiago (IDIS), Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Santiago de Compostela, Spain

⁶Department of Genetic and Molecular Medicine-IPER, Hospital Sant Joan de Déu and Institut de Recerca Sant Joan de Déu, Center for Biomedical Research Network on Rare Diseases (CIBERER), Barcelona, Spain

⁷Department of Genetics, Hospital Universitario de Mostoles, Mostoles, Spain

⁸Institute of Human Genetics, Medical Faculty, University of Bonn, Bonn, Germany

⁹National Center for Hereditary Tumor Syndromes, University Hospital Bonn, Bonn, Germany

¹⁰Department of Human Genetics, Radboud Institute for Molecular Life Sciences, Radboud University Medical Center, Nijmegen, The Netherlands

¹¹Department of Gastroenterology and Hepatology, Radboud University Medical Centre, Nijmegen, The Netherlands

¹²Gastroenterology Department, Hospital Donostia-Instituto Biodonostia, Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBERehd), Basque Country University (UPV/EHU), San Sebastian, Spain

¹³Hereditary Cancer Program, Institute of Oncology, Oncobell, Institut d'Investigació Biomèdica de Bellvitge (IDIBELL), Centro de Investigación Biomédica en Red de Cáncer (CIBERONC), L'Hospitalet de Llobregat, Barcelona, Spain

¹⁴Medizinische Klinik und Poliklinik IV, Campus Innenstadt, Klinikum der Universität München, Munich, Germany

¹⁵MGZ – Center of Medical Genetics Center, Munich, Germany

¹⁶Colorectal Oncogenomics Group, Department of Clinical Pathology, Melbourne Medical School, The University of Melbourne, Parkville, Victoria, Australia

¹⁷University of Melbourne Centre for Cancer Research, The University of Melbourne, Parkville, Victoria, Australia

¹⁸Melbourne Bioinformatics, The University of Melbourne, Carlton, Victoria, Australia

¹⁹New Zealand Familial Gastrointestinal Cancer Service, Auckland, New Zealand

²⁰Colorectal Medicine and Genetics, Royal Melbourne Hospital, Parkville, Victoria, Australia

²¹Genomic Medicine and Family Cancer Clinic, Royal Melbourne Hospital, Parkville, Victoria, Australia

²²Department of Medicine, The University of Melbourne, Parkville, Victoria, Australia

²³Envoi Specialist Pathologists, Brisbane, Queensland, Australia

²⁴University of Queensland, Brisbane, Queensland, Australia

²⁵Gastroenterology Department, Complexo Hospitalario Universitario de Ourense, Instituto de Investigación Sanitaria Galicia Sur, Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBERehd), Ourense, Spain

²⁶Digestive Disease Section, Hospital Universitario de Móstoles, Móstoles, Spain

Correction notice This article has been corrected since it was first published. The open access licence has been updated to CC BY.

Twitter Laura Valle @LValleResearch

Acknowledgements We are sincerely grateful to the patients and Biobanks of Hospital Clínic–IDIBAPS, IDIBELL and Biobanco Vasco. We also thank members of the Colorectal Oncogenomics Group for their support of this manuscript and the Clinical Geneticists and Genetic Counsellors from the Family Cancer Clinics across Australia for their contributions to patient recruitment. The work was carried out (in part) at the Esther Koplowitz Centre.

Contributors Conceptualisation: Ysdl, CA-C, MD-G, SF-E, LB and SC-B; funding acquisition: AC, RdV, DDB, SA, CR-P, LV and SC-B; investigation: Ysdl, CA-C, JM, CH-P, LeM, TO, MD-G, SF-E, MC, SC, AL-N, LoM, GF, AddB, SP, AKS, IS, IBAWtP, YJvH, VS-L, EH-F, RdV, KM, JEI, JA, SP, FAM, IMW, CR, DDB, SA, CR-P, FB, LV, LB and SC-B; resources: LeM, TO, MC, SC, LoM, AC, LBU, GC, JC, DR-A, CR-P, LV, FB, LB and SC-B; supervision: LB and SC-B; visualisation: Ysdl, CA-C, JM, LoM, SP, DDB, LB and SC-B; guarantor: SCB; writing – original draft: Ysdl, LB and SC-B; writing – review and editing: all authors.

Funding YSL was awarded an INPhINIT fellowship (LCF/BQ/DI18/11660058) from 'la Caixa' Foundation (ID 100010434 funded by the EU Horizon 2020 Programme Marie Skłodowska-Curie grant agreement no. 713673). CA-C, JM, CH-P and SF-E were supported by a contract from CIBERehd. MD-G was supported by a contract from Agència de Gestió d'Ajuts Universitaris i de Recerca -AGAUR- (Generalitat de Catalunya, 2019FI_B2_00203). LB was supported by a Juan de la Cierva postdoctoral contract (FICI-2017-32593). IBAWtP and AKS are supported by European Union's Horizon 2020 research and innovation programme under agreement no. 779257 (Solve-RD). CIBERehd, CIBERer and CIBERONC are funded by the Instituto de Salud Carlos III. This research was supported by grants from Fundació La Marató de TV3 (2019-202008-10), Fondo de Investigación Sanitaria/FEDER (17/00878, 17/00509, 20/00113, 20/00226), Fundación Científica de la Asociación Española contra el Cáncer (PRYGN211085CAST), Spanish Ministry of Science, Innovation and Universities, co-funded by FEDER funds (SAF2016-80888-R, PID2020-112595RB-I00), European Reference Network on Genetic Tumor Risk Syndromes (ERN GENTURIS, project ID No 739547; partly cofunded by the European Union within the framework of the Third Health Programme 'ERN-2016-Framework Partnership Agreement 2017–2021'), NHMRC Investigator grant (GNT1194896), University of Melbourne Dame Kate Campbell Fellowship, PERIS (SLT002/16/00398, Generalitat de Catalunya), CERCA Program (Generalitat de Catalunya), Xarxa de Bancs de Tumors de Catalunya (XBTC, Pla Director d'Oncologia de Catalunya) and Agència de Gestió d'Ajuts Universitaris i de Recerca (Generalitat de Catalunya, GRPRE 2017SGR21, GRC 2017SGR653, 2017SGR1282). This article is based on work from COST Action CA17118, supported by COST (European Cooperation in Science and Technology). www.cost.eu.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by Hospital Clínic de Barcelona Clinical Research Ethics committee (registration number 2013/8286). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given,

and indication of whether changes were made. See: <https://creativecommons.org/licenses/by/4.0/>.

ORCID iDs

Isabel Spier <http://orcid.org/0000-0003-2407-7427>

Richarda de Voer <http://orcid.org/0000-0002-8222-0343>

Daniel D Buchanan <http://orcid.org/0000-0003-2225-6675>

Sergi Castellvi-Bel <http://orcid.org/0000-0003-1217-5097>

REFERENCES

- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2021;71:209–49.
- Morson BC. The evolution of colorectal carcinoma. *Clin Radiol* 1984;35:425–31.
- Mankaney G, Roupheal C, Burke CA. Serrated polyposis syndrome. *Clin Gastroenterol Hepatol* 2020;18:777–9.
- Ijspeert JEG, Vermeulen L, Meijer GA, Dekker E. Serrated neoplasia-role in colorectal carcinogenesis and clinical implications. *Nat Rev Gastroenterol Hepatol* 2015;12:401–9.
- Carballal S, Rodríguez-Alcalde D, Moreira L, Hernández L, Rodríguez L, Rodríguez-Moranta F, Gonzalo V, Bujanda L, Bessa X, Poves C, Cubiella J, Castro I, González M, Moya E, Oquién S, Clófent J, Quintero E, Esteban P, Piñol V, Fernández FJ, Jover R, Cid L, López-Cerón M, Cuatrecasas M, López-Vicente J, Leoz ML, Rivero-Sánchez L, Castells A, Pellisé M, Balaguer F. Gastrointestinal Oncology Group of the Spanish Gastroenterological Association. Colorectal cancer risk factors in patients with serrated polyposis syndrome: a large multicentre study. *Gut* 2016;65:1829–37.
- Snover DC, Ahnen DJ, Burt RW, Odze RD. Serrated Polyps of the Colon and Rectum and Serrated Polyposis. In: *Who classification of tumours of the digestive system*. 4th ed. Lyon, France: IARC, 2010.
- Rosty C, Brosens L, Dekker E, Nagtegaal ID. Serrated Polyposis. In: *WHO classification of tumours of the digestive system*. 5th ed. Lyon, France: IARC, 2019.
- Moreira L, Pellisé M, Carballal S, Bessa X, Ocaña T, Serradesanferm A, Grau J, Macià F, Andreu M, Castells A, Balaguer F, PROCOLON research group. High prevalence of serrated polyposis syndrome in FIT-based colorectal cancer screening programmes. *Gut* 2013;62:476–7.
- Rivero-Sánchez L, Lopez-Ceron M, Carballal S, Moreira L, Bessa X, Serradesanferm A, Pozo A, Augé JM, Ocaña T, Sánchez A, Leoz ML, Cuatrecasas M, Grau J, Llach J, Castells A, Balaguer F, Pellisé M, Group P, Procolon Group. Reassessment colonoscopy to diagnose serrated polyposis syndrome in a colorectal cancer screening population. *Endoscopy* 2017;49:44–53.
- Ijspeert JEG, Bevan R, Senore C, Kaminski MF, Kuipers EJ, Mroz A, Bessa X, Cassoni P, Hassan C, Repici A, Balaguer F, Rees CJ, Dekker E. Detection rate of serrated polyps and serrated polyposis syndrome in colorectal cancer screening cohorts: a European overview. *Gut* 2017;66:1225–32.
- Rodríguez-Alcalde D, Carballal S, Moreira L, Hernández L, Rodríguez-Alonso L, Rodríguez-Moranta F, Gonzalo V, Bujanda L, Bessa X, Poves C, Cubiella J, Castro I, González M, Moya E, Oquién S, Clófent J, Quintero E, Esteban P, Piñol V, Fernández FJ, Jover R, Cid L, Saperas E, López-Cerón M, Cuatrecasas M, López-Vicente J, Rivero-Sánchez L, Jung G, Vila-Casadesús M, Sánchez A, Castells A, Pellisé M, Balaguer F. Gastrointestinal Oncology Group of the Spanish Gastroenterological Association. High incidence of advanced colorectal neoplasia during endoscopic surveillance in serrated polyposis syndrome. *Endoscopy* 2019;51:142–51.
- Murphy N, Moreno V, Hughes DJ, Vodicka L, Vodicka P, Aglago EK, Gunter MJ, Jenab M. Lifestyle and dietary environmental factors in colorectal cancer susceptibility. *Mol Aspects Med* 2019;69:2–9.
- Crockett SD, Nagtegaal ID. Terminology, molecular features, epidemiology, and management of serrated colorectal neoplasia. *Gastroenterology* 2019;157:949–66.
- Frank C, Sundquist J, Yu H, Hemminki A, Hemminki K. Concordant and discordant familial cancer: familial risks, proportions and population impact. *Int J Cancer* 2017;140:1510–6.
- Valle L, de Voer RM, Goldberg Y, Sijns W, Försti A, Ruiz-Ponte C, Caldés T, Garré P, Olsen MF, Nordling M, Castellvi-Bel S, Hemminki K. Update on genetic predisposition to colorectal cancer and polyposis. *Mol Aspects Med* 2019;69:10–26.
- Guarinos C, Sánchez-Fortún C, Rodríguez-Soler M, Alenda C, Payá A, Jover R. Serrated polyposis syndrome: molecular, pathological and clinical aspects. *World J Gastroenterol* 2012;18:2452–61.
- Gala MK, Mizukami Y, Le LP, Moriichi K, Austin T, Yamamoto M, Lauwers GY, Bardeesy N, Chung DC. Germline mutations in oncogene-induced senescence pathways are associated with multiple sessile serrated adenomas. *Gastroenterology* 2014;146:520–9.
- Buchanan DD, Clendenning M, Zhuoer L, Stewart JR, Joseland S, Woodall S, Arnold J, Semotiuk K, Aronson M, Holter S, Gallinger S, Jenkins MA, Sweet K, Macrae FA, Winship IM, Parry S, Rosty C. Genetics of colonic polyposis study. Lack of evidence for germline RNF43 mutations in patients with serrated polyposis syndrome from a large multinational study. *Gut* 2017;66:1170–2.
- Quintana I, Mejías-Luque R, Terradas M, Navarro M, Piñol V, Mur P, Belhadj S, Grau E, Darder E, Solanes A, Brunet J, Capellá G, Gerhard M, Valle L. Evidence

- suggests that germline *RNF43* mutations are a rare cause of serrated polyposis. *Gut* 2018;67:2230–2.
- 20 Chow E, Lipton L, Lynch E, D'Souza R, Aragona C, Hodgkin L, Brown G, Winship I, Barker M, Buchanan D, Cowie S, Nasioulas S, du Sart D, Young J, Leggett B, Jass J, Macrae F. Hyperplastic polyposis syndrome: phenotypic presentations and the role of MBD4 and MYH. *Gastroenterology* 2006;131:30–9.
 - 21 Boparai KS, Dekker E, Van Eeden S, Polak MM, Bartelsman JFWM, Mathus-Vliegen EMH, Keller JJ, van Noesel CJM. Hyperplastic polyps and sessile serrated adenomas as a phenotypic expression of MYH-associated polyposis. *Gastroenterology* 2008;135:2014–8.
 - 22 Clendenning M, Young JP, Walsh MD, Woodall S, Arnold J, Jenkins M, Win AK, Hopper JL, Sweet K, Gallinger S, Rosty C, Parry S, Buchanan DD. Germline mutations in the polyposis-associated genes BMPR1A, SMAD4, PTEN, MUTYH and GREM1 are not common in individuals with serrated polyposis syndrome. *PLoS One* 2013;8:e66705.
 - 23 Toma C, Díaz-Gay M, Soares de Lima Y, Arnau-Collé C, Franch-Expósito S, Muñoz J, Overs B, Bonjoch L, Carballal S, Ocaña T, Cuatrecasas M, Díaz de Bustamante A, Castells A, Bujanda L, Cubiella J, Balaguer F, Rodríguez-Alcalde D, Fullerton JM, Castellvi-Bel S. Identification of a novel candidate gene for serrated polyposis syndrome germline predisposition by performing linkage analysis combined with whole-exome sequencing. *Clin Transl Gastroenterol* 2019;10:e00100.
 - 24 Soares de Lima Y, Arnau-Collé C, Díaz-Gay M, Bonjoch L, Franch-Expósito S, Muñoz J, Moreira L, Ocaña T, Cuatrecasas M, Herrera-Pariente C, Carballal S, Moreno L, Díaz de Bustamante A, Castells A, Bujanda L, Cubiella J, Rodríguez-Alcalde D, Balaguer F, Castellvi-Bel S, Germline C-BS. Germline and somatic whole-exome sequencing identifies new candidate genes involved in familial predisposition to serrated polyposis syndrome. *Cancers* 2021;13:929–21.
 - 25 Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009;25:1754–60.
 - 26 Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain W1118; Iso-2; Iso-3. *Fly* 2012;6:80–92.
 - 27 Rodan AR, Jenny A. WNK kinases in development and disease. *Curr Top Dev Biol* 2017;123:1–47.
 - 28 Dhillon AS, Hagan S, Rath O, Kolch W. Map kinase signalling pathways in cancer. *Oncogene* 2007;26:3279–90.
 - 29 Leggett B, Whitehall V. Role of the serrated pathway in colorectal cancer pathogenesis. *Gastroenterology* 2010;138:2088–100.
 - 30 Costa AM, Pinto F, Martinho O, Oliveira MJ, Jordan P, Reis RM. Silencing of the tumor suppressor gene WNK2 is associated with upregulation of MMP2 and JNK in gliomas. *Oncotarget* 2015;6:1422–34.
 - 31 Peña-Chilet M, Roldán G, Perez-Florido J, Ortuño FM, Carmona R, Aquino V, Lopez-Lopez D, Loucera C, Fernandez-Rueda JL, Gallego A, García-García F, González-Neira A, Pita G, Núñez-Torres R, Santoyo-López J, Ayuso C, Minguez P, Avila-Fernandez A, Corton M, Moreno-Pelayo Miguel Ángel, Morin M, Gallego-Martínez A, Lopez-Escamez JA, Borrego S, Antiñolo G, Amigo J, Salgado-Garrido J, Pasalodos-Sanchez S, Morte B, Carracedo Ángel, Alonso Ángel, Dopazo J, Spanish Exome Crowdsourcing Consortium. CSVS, a crowdsourcing database of the Spanish population genetic variability. *Nucleic Acids Res* 2021;49:D1130–7.
 - 32 Moniz S, Jordan P. Emerging roles for WNK kinases in cancer. *Cell Mol Life Sci* 2010;67:1265–76.
 - 33 Murillo-de-Ozores AR, Chávez-Canales M, de Los Heros P, Gamba G, Castañeda-Bueno M. Physiological processes modulated by the chloride-sensitive WNK-SPAK/OSR1 kinase signaling pathway and the cation-coupled chloride cotransporters. *Front Physiol* 2020;11:585907.
 - 34 Wiesner GL, Daley D, Lewis S, Ticknor C, Platzter P, Lutterbaugh J, MacMillen M, Baliner B, Willis J, Elston RC, Markowitz SD. A subset of familial colorectal neoplasia kindreds linked to chromosome 9q22.2-31.2. *Proc Natl Acad Sci U S A* 2003;100:12961–5.
 - 35 Greenman C, Stephens S, Smith R, Dalgleish GL, Hunter C, Bignell G, Davies H, Teague J, Butler A, Stevens C, Edkins S, O'Meara S, Vastrik I, Schmidt EE, Avis T, Barthorpe S, Bhamra G, Buck G, Choudhury B, Clements J, Cole J, Dicks E, Forbes S, Gray K, Halliday K, Harrison R, Hills K, Hinton J, Jenkinson A, Jones D, Menzies A, Mironenko T, Perry J, Raine K, Richardson D, Shepherd R, Small A, Tofts C, Varian J, Webb T, West S, Widaa S, Yates A, Cahill DP, Louis DN, Goldstraw P, Nicholson AG, Brasseur F, Looijenga L, Weber BL, Chiew Y-E, DeFazio A, Greaves MF, Green AR, Campbell P, Birney E, Easton DF, Chenevix-Trench G, Tan M-H, Khoo SK, Teh BT, Yuen ST, Leung SY, Wooster R, Futreal PA, Stratton MR. Patterns of somatic mutation in human cancer genomes. *Nature* 2007;446:153–8.
 - 36 Martínez-Jiménez F, Muiños F, Sentís I, Deu-Pons J, Reyes-Salazar I, Arnedo-Pac C, Mularoni L, Pich O, Bonet J, Kranas H, Gonzalez-Perez A, Lopez-Bigas N. A compendium of mutational cancer driver genes. *Nat Rev Cancer* 2020;20:555–72.
 - 37 Moniz S, Martinho O, Pinto F, Sousa B, Loureiro C, Oliveira MJ, Moita LF, Honavar M, Pinheiro C, Pires M, Lopes JM, Jones C, Costello JF, Paredes J, Reis RM, Jordan P. Loss of WNK2 expression by promoter gene methylation occurs in adult gliomas and triggers Rac1-mediated tumour cell invasiveness. *Hum Mol Genet* 2013;22:84–95.
 - 38 Dutruel C, Bergmann F, Rooman I, Zucknick M, Weichenhan D, Geiselhart L, Kaffenberger T, Rachakonda PS, Bauer A, Giese N, Hong C, Xie H, Costello JF, Hoheisel J, Kumar R, Rehli M, Schirmacher P, Werner J, Plass C, Popanda O, Schmezer P. Early epigenetic downregulation of WNK2 kinase during pancreatic ductal adenocarcinoma development. *Oncogene* 2014;33:3401–10.
 - 39 Delghanizadeh S, Khoddami V, Mosbrugger TL, Hammoud SS, Edes K, Berry TS, Done M, Samowitz WS, DiSario JA, Luba DG, Burt RW, Jones DA. Active BRAF-V600E is the key player in generation of a sessile serrated polyp-specific DNA methylation profile. *PLoS One* 2018;13:e0192499.
 - 40 Moniz S, Verissimo F, Matos P, Brazão R, Silva E, Kotelevets L, Kotelevets L, Chastre E, Gaspach C, Jordan P. Protein kinase WNK2 inhibits cell proliferation by negatively modulating the activation of MEK1/ERK1/2. *Oncogene* 2007;26:6071–81.
 - 41 Moniz S, Matos P, Jordan P. WNK2 modulates MEK1 activity through the Rho GTPase pathway. *Cell Signal* 2008;20:1762–8.
 - 42 Fang JY, Richardson BC. The MAPK signalling pathways and colorectal cancer. *Lancet Oncol* 2005;6:322–7.
 - 43 Zhou S-L, Zhou Z-J, Hu Z-Q, Song C-L, Luo Y-J, Luo C-B, Xin H-Y, Yang X-R, Shi Y-H, Wang Z, Huang X-W, Cao Y, Fan J, Zhou J. Genomic sequencing identifies WNK2 as a driver in hepatocellular carcinoma and a risk factor for early recurrence. *J Hepatol* 2019;71:1152–63.
 - 44 Sherr CJ. The pezcoller Lecture: cancer cell cycles revisited. *Cancer Res* 2000;60:3689–95.
 - 45 Liu H, Liu K, Dong Z. The role of P21-activated kinases in cancer and beyond: where are we heading? *Front Cell Dev Biol* 2021;9.
 - 46 Wu J, Meng X, Gao R, Jia Y, Chai J, Zhou Y, Wang J, Xue X, Dang T. Long non-coding RNA LINC00858 inhibits colon cancer cell apoptosis, autophagy, and senescence by activating WNK2 promoter methylation. *Exp Cell Res* 2020;396:112214.
 - 47 Alves ALV, Costa AM, Martinho O, da Silva VD, Jordan P, Silva VAO, Reis RM. WNK2 inhibits autophagic flux in human glioblastoma cell line. *Cells* 2020;9.
 - 48 Slack-Davis JK, Eblen ST, Zecevic M, Boerner SA, Tarsafalvi A, Diaz HB, Marshall MS, Weber MJ, Parsons JT, Catling AD. PAK1 phosphorylation of MEK1 regulates fibronectin-stimulated MAPK activation. *J Cell Biol* 2003;162:281–91.
 - 49 Jia Y, Wang Y, Zhang C, Chen MY. Upregulated CBX8 promotes cancer metastasis via the WNK2/MMP2 pathway. *Mol Ther Oncolytics* 2020;19:188–96.
 - 50 Kumar M, Gouw M, Michael S, Sámano-Sánchez H, Pancsa R, Glavina J, Diakogianni A, Valverde JA, Bukirova D, Čalyševa J, Palopoli N, Davey NE, Chemes LB, Gibson TJ. ELM-the eukaryotic linear motif resource in 2020. *Nucleic Acids Res* 2020;48:D296–306.
 - 51 Rinehart J, Vázquez N, Kahle KT, Hodson CA, Ring AM, Gulcicek EE, Louvi A, Bobadilla NA, Gamba G, Lifton RP. WNK2 kinase is a novel regulator of essential neuronal cation-chloride cotransporters. *J Biol Chem* 2011;286:30171–80.
 - 52 Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, Tukiainen T, Birnbaum DP, Kosmicki JA, Duncan LE, Estrada K, Zhao F, Zou J, Pierce-Hoffman E, Berghout J, Cooper DN, DePristo M, Do R, Flannick J, Fromer M, Gauthier L, Goldstein J, Gupta N, Howrigan D, Kiezun A, Kurki MI, Moonshine AL, Natarajan P, Orozco L, Peloso GM, Poplin R, Rivas MA, Ruano-Rubio V, Rose SA, Ruderfer DM, Shakir K, Stenson PD, Stevens C, Thomas BP, Tiao G, Tusie-Luna MT, Weisburd B, Won H-H, Yu D, Altshuler DM, Ardissino D, Boehnke M, Danesh J, Donnelly S, Elosua R, Florez JC, Gabriel SB, Getz G, Glatt SJ, Hultman CM, Kathiresan S, Laakso M, McCarroll S, McCarthy MI, McGovern D, McPherson R, Neale BM, Palotie A, Purcell SM, Saleheen D, Scharf JM, Sklar P, Sullivan PF, Tuomilehto J, Tsuang MT, Watkins HC, Wilson JG, Daly MJ, MacArthur DG, Consortium EA, Exome Aggregation Consortium. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 2016;536:285–91.

SUPPLEMENTARY MATERIAL

SUPPLEMENTARY MATERIALS AND METHODS

Variant Identification

DNA extraction

Germline DNA samples were extracted using the QIAamp DNA Blood kit (Qiagen, Hilden, Germany) following the manufacturer's instructions.

WES and variant prioritization in the discovery cohort

WES was performed using the HiSeq2000 Platform (Illumina, San Diego, USA).

All sequencing and quality control details were previously described.[1]

Variant prioritization

Variant prioritization was carried out considering several aspects. First, we only took into consideration those variants present in the canonical transcripts. Also, heterozygous variants were prioritized over homozygous variants since a dominant inheritance to SPS is mostly hypothesized. In addition, a minimum allele frequency of 0.1% was required for variant filtering and only nonsynonymous and/or truncating variants were prioritized. The missense variants had to fulfill at least 3 out of 6 pathogenic predictions used for analysis (PhyloP, SIFT, Polyphen, MutationTaster, CADD and LRT). The next crucial step of variant prioritization considered data integration with the first cohort results. We prioritized genes that presented germline variants in both cohorts and conducted an extensive literature research over possible connections

between candidate genes and SPS. Only candidate genes with putative pathogenic genetic variants detected in the discovery and the validation cohort were further considered. Among them, only those with a function compatible with SPS, CRC or cancer were selected.

Variant validation

Integrative Genome Viewer (IGV, <http://software.broadinstitute.org/software/igv/>) and Sanger sequencing (Eurofins Genomics, Luxembourg) were performed to do a final validation of the identified variants. Segregation was also carried out in patients with family history, when possible. We used the *WNK2* NM_001282394 transcript as reference.

Gene panel sequencing

Gene-panel sequencing was completed using the Illumina Miseq platform (Illumina, San Diego, USA) in germline DNA from SPS patients of the validation cohort. Library quality control was accomplished with Bioanalyzer 2000 (Agilent, Santa Clara, CA, USA). Samples were indexed with adapters containing different barcodes and pooled together, followed by massively parallel sequencing using an AmpliSeq DNA panel of 575 amplicons divided into two pools, with an amplicon length of 125-375 bp (Illumina, San Diego, USA).

Development of a cellular model for variant characterization

Cell lines

HT-29 (Cat No. HTB-38) and HEK293T (Cat No. CRL-3216) cell lines were purchased from American Type Culture Collection (ATCC, Manassas, VA). HT-

29 was cultured with McCoy 5A (Modified) (Gibco, Waltham, MA) and the later with DMEM (Gibco), both supplemented with 10% fetal bovine serum (FBS) at 37°C in 5%CO₂. All functional assays were performed with medium supplemented with 5% FBS.

WNK2 knock-out generation

The CRISPR-Cas9 system was used to perform the *WNK2* knock-out. The sgRNA was designed using The Benchling CRISPR guide design tool (<http://benchling.com>). Bottom and top strands of sgRNA were purchased from IDT (Coralville, IA) and cloned into the BbsI site of the PX458 plasmid following the protocol by Ran F. A. et al.[2] The plasmid was transiently transfected into HT-29 cells using X-tremeGENE HP DNA transfection reagent (Roche, Basel, Switzerland). After 3-5 days, cells were sorted for GFP+ cells. Sorted cells were let to rest for 2 days and then seeded at 1 cell/well density in 96-well plates with HT-29 conditioned media. After two weeks, several clones were analyzed for *WNK2* gene editing. DNA, RNA and protein samples were extracted and analyzed. Genomic DNA was analyzed by Sanger sequencing (Eurofins Genomics) and screened for indel insertions manually and by using the TIDE web tool (<https://tide.nki.nl/>).[3] Predicted exonic off-target regions with high homology with the target region were identified by using the CRISPOR tool (<http://crispor.tefor.net/>) and excluded by PCR and Sanger sequencing.[4] Gene expression was accessed by quantitative real-time PCR, and protein expression was evaluated by Western-Blot.

Site-directed mutagenesis

To generate *WNK2* variants by site-directed mutagenesis, the pUC57-*WNK2* vector (**supplementary table 1**) was digested into smaller segments because of the length of *WNK2* and polymerase restrictions. We used three different restriction enzymes: XbaI, NcoI and SalI. Both 4083 bp (5' XbaI-3' NcoI) and 2868 bp (5' NcoI-3' SalI) fragments, which were part of the coding sequence of *WNK2*, were recovered and subcloned separately into the pUC19 plasmid. The generation of *WNK2* variants was carried out using the Q5 Site-Directed Mutagenesis Kit (NEB, Ipswich, MA) according to the manufacturer's instructions. Primers were designed using the NEBaseChanger tool and purchased from IDT (Coralville, IA). Mutated fragments were inserted back into the pUC57-*WNK2* backbone. Variant insertion was verified by Sanger sequencing.

Variant reintroduction into the *WNK2* knock-out model

To achieve a stable, reproducible *WNK2* expression, the *WNK2* ORF was transferred from pUC57 into the lentiviral plasmid pLVX-TetOne-Puro (TakaraBio, Kusatsu, Japan). Each mutated *WNK2* ORF with its variant and the *WNK2* WT sequence were cloned independently in the pLVX-TetOne-Puro vector thanks to the additional EcoRI and AgeI restriction sites flanking the *WNK2* ORF.

Lentiviral particles were generated for each *WNK2* variant. Briefly, HEK293T cells were co-transfected with the lentiviral vectors (VSVG2, psPAX2) and pLVX-TetOne-Puro containing either *WNK2* ORF WT or the mutated *WNK2* ORF using the CalPhos Mammalian Transfection kit (TakaraBio, Kusatsu, Japan). The supernatant was recovered both after 24h and 48h after

transfection. Viral particles were pooled and then concentrated in a 10% sucrose cushion at 20,000xg for 3 hours at 4°C. The pellet was resuspended in 5mL of McCoy medium and prepared for infection. HT-29 cells were plated at 600,000 cells/well in a 6-well plate and let grow for two days before infection. Lentiviral transduction was performed in the presence of 8 µg/mL of polybrene. The plate was centrifuged at 1600 xg for 2h at 32 °C and then, culture media was removed and replaced with fresh complete media.

Functional titer by limiting dilution was performed to establish an optimal multiplicity of infection (MOI) of 1, which was further validated by Real-Time PCR. After obtaining genomic DNA samples from infected cells, the viral *WPRE* region was amplified by RT-PCR with RealQ Plus 2x Master Mix (Ampliqon, Odense, Denmark) and specific primers. The assessment of lentiviral vector copies was inferred from a standard curve using known amounts of pLVX-TetOne. The final integrated lentivirus copy number was calculated as total viral copies/number of cells assayed. The number of cells was estimated from the amount of genomic DNA loaded in the RT-PCR, considering that the mass of the human genome (haploid) is 3,181 pg.

Protein extraction and Western Blot

To obtain whole-cell protein extracts, cells were lysed with RIPA buffer solution (Sigma-Aldrich, MA, USA) supplemented with cOmplete Protease Inhibitor Cocktail and PhosSTOP (Roche, Basel, Switzerland) and recovered using a cell-scrapper. Protein concentration was measured using Pierce BCA Protein Assay kit (ThermoFisher, Waltham, MA). Western blot for WNK2 detection was carried out running 40ug of protein extract of each sample in NuPAGE™ 3-8%

Tris-Acetate gels, according to the manufacturer's protocol (ThermoFisher, Waltham, MA). Protein transfer into PVDF membranes (Millipore, Bedford, MA) was left to perform overnight at 4°C. Proteins were blotted with anti-WNK2 primary antibody an anti-rabbit secondary Dylight 800 antibody. Protein detection was carried out using Odyssey Imaging System (LI-COR, Lincoln, NE).

RNA extraction and Quantitative Real-time PCR

Total RNA was extracted using the RNeasyMini Kit according to the manufacturer's protocol (Qiagen). RNA was quantified using Nanodrop (ThermoFisher, Waltham, MA) and retrotranscribed with the High-Capacity cDNA reverse Transcription kit (Applied Biosystems). Quantitative PCR was performed in the QuantStudio 1 System (Applied Biosystems) by using Taqman probes for *WNK2*, *CCDN1* and *GAPDH* as a housekeeping gene reference. *MMP2* gene expression was measured by using RealQ Plus 2x Master Mix (Ampliqon, Odense, Denmark) and specific primers for *MMP2* and *GAPDH* detection. The relative quantification was performed with the $-\Delta\Delta C_t$ method.

Functional Characterization of Genetic Variants

MAPK pathway activity: ERK1/2

To detect total phospho-ERK1/2, 600,000 cells were plated in a P60 dish and cultured in McCoy 5A media supplemented with 5% FBS and 1µg/mL of doxycycline for two days. After 24h starvation, cells were induced with 1 ng/mL hEGF (Miltenyi Biotec, Bergisch Gladbach, Germany) for 10 minutes. To preserve phosphorylated proteins, cells were lysed using Lysis buffer 6 (Bio-

Techne, Minnesota, USA) supplemented with Pepstatin A, Leupeptin and Aprotinin (10µg/mL) (Merck, Darmstadt, Germany) and recovered using a cell-scrapper. The whole-cell extract was incubated for 15 minutes on ice and centrifuged at 2000xg for 5 minutes at 4°C. Total protein extract was quantified using Pierce BCA Protein Assay kit (ThermoFisher Waltham, MA), and 15 µg of total protein extracts were assayed. Detection of phosphorylated ERK1/2 was performed using the Phospho-ERK1 (T202/Y204)/ERK2 (T185/Y187) DuoSet IC ELISA kit according to the manufacturer's protocol (Bio-Techne, Minnesota, USA). Phospho-ERK1/2 levels between EGF-stimulated and non-stimulated conditions were quantified.

MAPK pathway activity: PAK1/2 phosphorylation

To detect phosphorylated PAK1, an In-Cell ELISA (ICE) assay was performed. 96-well plates were coated with 50 µg/mL Poly-L-Lysine before cell seeding. 25,000 cells were plated per well and cultured in McCoy 5A media supplemented with 5% FBS and 1µg/mL of doxycycline. After 48 hours, cells were stimulated with 10 ng/mL of hEGF for 5 minutes. To immediately fix phosphorylated proteins and avoid cell detachment, cells were treated with a 2X 8% formaldehyde solution, supplemented with 100 mg/L $\text{Ca}^{2+}/\text{Mg}^{2+}$ and PhosSTOP for 20 min. After fixation, cells were permeabilized with 0.05% Triton X100 for 20 min and blocked with the Intercept Blocking Buffer (LI-COR, Lincoln, NE) for 1h. Cells were incubated with both phospho-PAK1 (rabbit) and β-actin (mouse) antibodies overnight at 4 °C. The multiplexed detection of both targets was performed with the anti-mouse Dylight 800 (ThermoFisher) and anti-rabbit IRDye 680RD (LICOR) antibodies. Plates were scanned with

Odyssey (LI-COR) and analysed with Image Studio 4.0 software. Phospho-PAK1 levels between EGF-stimulated and non-stimulated conditions were quantified.

References

1. Soares de Lima Y, Arnau-Collell C, Díaz-Gay M, Bonjoch L, Franch-Expósito S, Muñoz J, Moreira L, Ocaña T, Cuatrecasas M, Herrera-Pariente C, Carballal S, Moreno L, Díaz de Bustamante A, Castells A, Bujanda L, Cubiella J, Rodríguez-Alcalde D, Balaguer F, Castellví-Bel S. Germline and Somatic Whole-exome Sequencing Identifies New Candidate Genes Involved in Familial Predisposition to Serrated Polyposis Syndrome. *Cancers* 2021;13(4):1–21.
2. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* 2013;8(11):2281–2308.
3. Brinkman EK, Chen T, Amendola M, van Steensel B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res* 2014; 42(22):e168–e168.
4. Concordet JP, Haeussler M. CRISPOR: Intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Res* 2018;46(W1):W242-W245.

Supplementary table 1. List of reagents (plasmids, restriction enzymes and Taqman probes) used in this study.

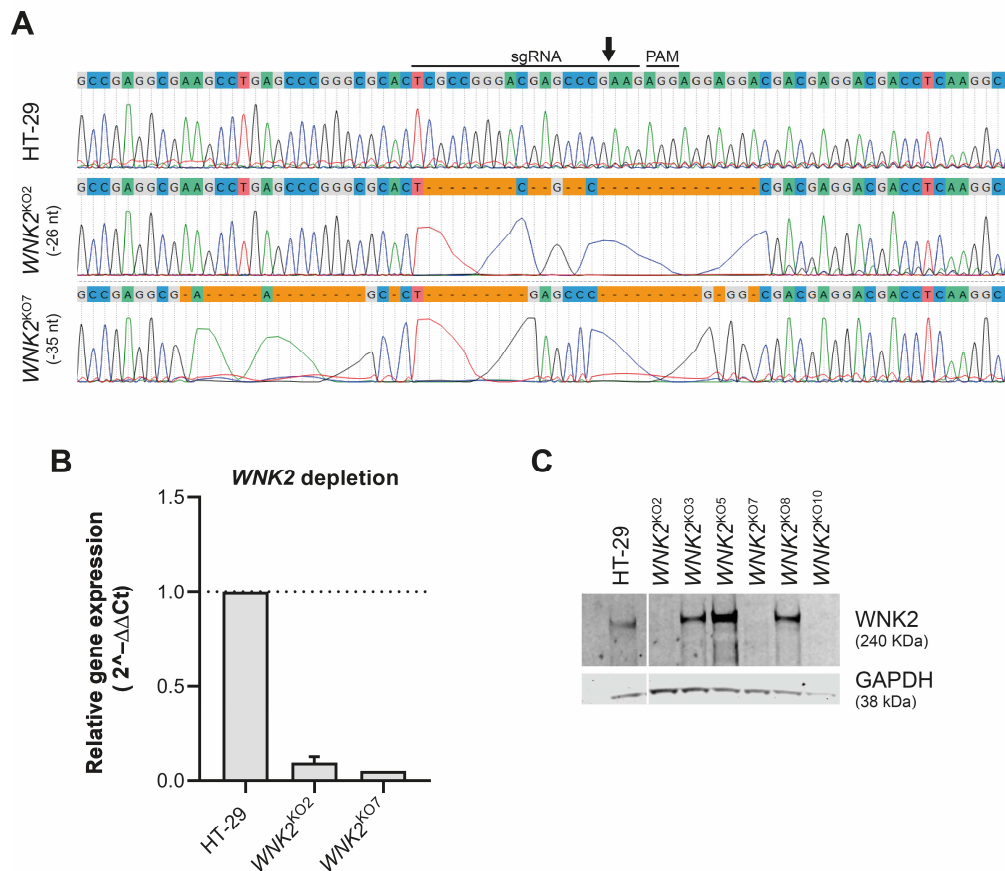
Reagent	Source	Identifier
Plasmids		
pSpCas9(BB)-2A-GFP (PX458)	Addgene (gift from Feng Zhang)	Cat No. 48138
pUC19	Invitrogen (Waltham, MA)	Cat No. 18265017
pLVX-TetOne-Puro inducible system	TakaraBio, provided by Dr. Eulàlia de Nadal (IRB, Barcelona)	Cat No. 631847
Second generation packaging vectors pCMV-VSV-G2 and psPAX2	provided by Dr. Ester Sánchez-Tilló (IDIBAPS, Barcelona)	
WNK2 custom gene synthesis (pUC57-WNK2)	Bio Basic Inc. (Toronto, Canada).	
Restriction enzymes		
BbsI	New England Biolabs (Ipswich, MA)	Cat No. R0539S
NcoI-HF	New England Biolabs (Ipswich, MA)	Cat No. R3193S
Sall-HF	New England Biolabs (Ipswich, MA)	Cat No. R3138S
EcoRI-HF	New England Biolabs (Ipswich, MA)	Cat No. R3101S
AgeI-HF	New England Biolabs (Ipswich, MA)	Cat No. R3552S
Antibodies		
rabbit polyclonal anti-WNK2	Invitrogen (Waltham, MA)	Cat No. PA5-53440
rabbit monoclonal anti-GAPDH (clone 14C10)	Cell Signaling (Danvers, MA)	Cat No. 2118
rabbit monoclonal anti-vinculin (7F9)	Santa Cruz Biotechnology	Cat No. sc-73614
rabbit polyclonal phospho-PAK1(Ser199/204)/PAK2 (Ser192/197)	Cell Signaling (Danvers, MA)	Cat No. 2605
mouse monoclonal anti- β -actin	Sigma Aldrich (St Louis, MO)	Cat No. A2228
DyLight 800 goat anti-mouse IgG	Invitrogen (Waltham, MA)	Cat No. SA5-10176
IRDye 680RD goat anti-rabbit IgG	LI-COR (Lincoln, NE)	Cat No. 926-68071

DyLight 800 goat anti-rabbit IgG	Invitrogen (Waltham, MA)	Cat No. SA5-10036
Taqman probes		
GAPDH-VIC/MGB	Applied Biosystems (Foster City, CA)	4326317E
WNK2-FAM/MGB	Applied Biosystems (Foster City, CA)	hs00396601_m1
CCND1-FAM/MGB	Applied Biosystems (Foster City, CA)	hs00765553_m1

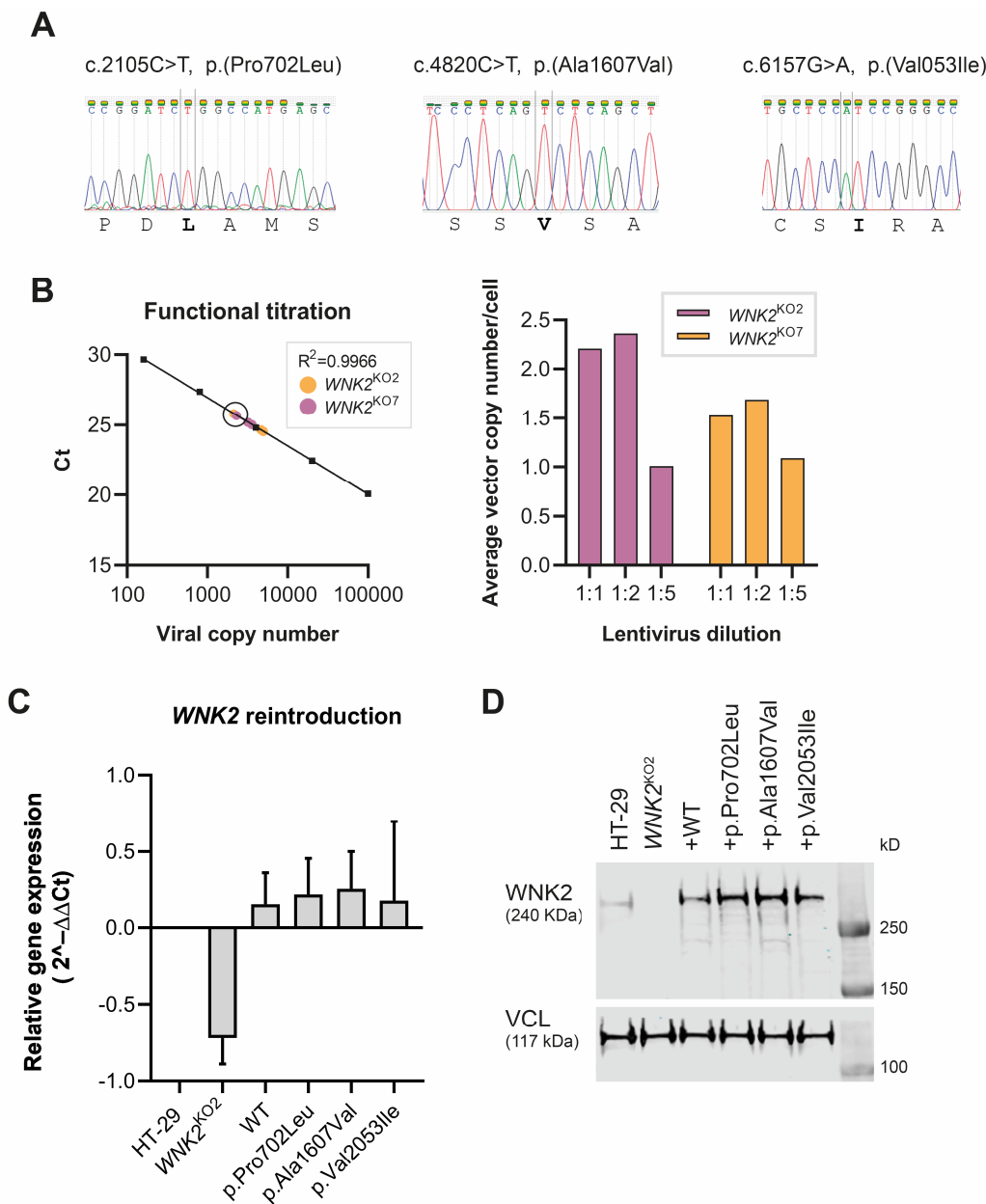
Supplementary table 2. List of primers used in this study.

	Forward	Reverse
Validation cohort validation		
<i>WNK2</i> c.2105C>T	5'- AGAGCTGCTGCTGAGTGTG- 3'	5'-ACAGTGGCCACAGGAGGT- 3'
<i>WNK2</i> c.3341C>T	5'- CCAGACTGCCACACTTCTG- 3'	5'- CTCTAGGACTAAGACCCCAG G-3'
<i>WNK2</i> c.5588T>C	5'- AGCGACTCTGGGGACGAG- 3'	5'- CATTGTCGCTGCTGATGTAG- 3'
WNK2 knock-out generation		
<i>WNK2</i> sgRNA	5'- caccGTCGCCGGGACGAGCC CGAAG-3'	5'- aaacCTTCGGGCTCGTCCCGG CGAC-3'
<i>WNK2</i> sgRNA cut validation	5'- GCTTCTGCTCTGCAAGACG- 3'	5'- CAGAGGGTGACACACATGG- 3'
Off-target validation		
<i>CENPB</i>	5'- GAGCCGCTTTGTCTCGGG- 3'	5'- TGCGCAGGTCCGGATTCTC- 3'
<i>MTX1</i>	5'- GAGTTCCCGTCACCTAAGC G-3'	5'- CTCTTGCCAAACTGCACGTG- 3'
<i>ZNF865</i>	5'- CTTCCGATCCTCCACGCC- 3'	5'- CTGACTTGCTAGGGGTTGGG -3'
Site directed mutagenesis		
<i>WNK2</i> c.2105C>T	5'- TTCCCGGATCtGGCCATGAG C-3'	5'- GTGCTGCTGGAGGGACGG-3'
<i>WNK2</i> c.4820C>T	5'- ACCTCCTCAGtCTCAGCTGG G-3'	5'-GGGCACGTGCTCCTGGTA- 3'
<i>WNK2</i> c.6157G>A	5'- GCCCTGCTCCaTCCGGGCC TC-3'	5'- TGCTGGGTCTGCACTGCCTT C-3'
c.2105C>T validation	5'- TCTACCGTGTACTCAGACTC G-3'	5'-ATCTGGAGGGGCTTCAGC- 3'
c.4820C>T validation	5'- GTGGACAGCACCATCAAGA- 3'	5'-CCTGAGACCGCCTCCTTA- 3'
c.6157G>A validation	5'- CAAGCTGCTAAATCCCCTG G-3'	5'- ATAGGTCACTCTCTCAGACGT T-3'
Real-Time PCR		

<i>WPRE</i>	5'- CCGTTGTCAGGCAACGTG-3'	5'- AGCTGACAGGTGGTGGCAAT -3'
<i>MMP2</i>	5'- AGCGAGTGGATGCCGCCTT TAA-3'	5'- CATTCCAGGCATCTGCGATG AG-3'
<i>GAPDH</i>	5'- GTCTCCTCTGACTTCAACAG CG-3'	5'- ACCACCCTGTTGCTGTAGCC AA-3'

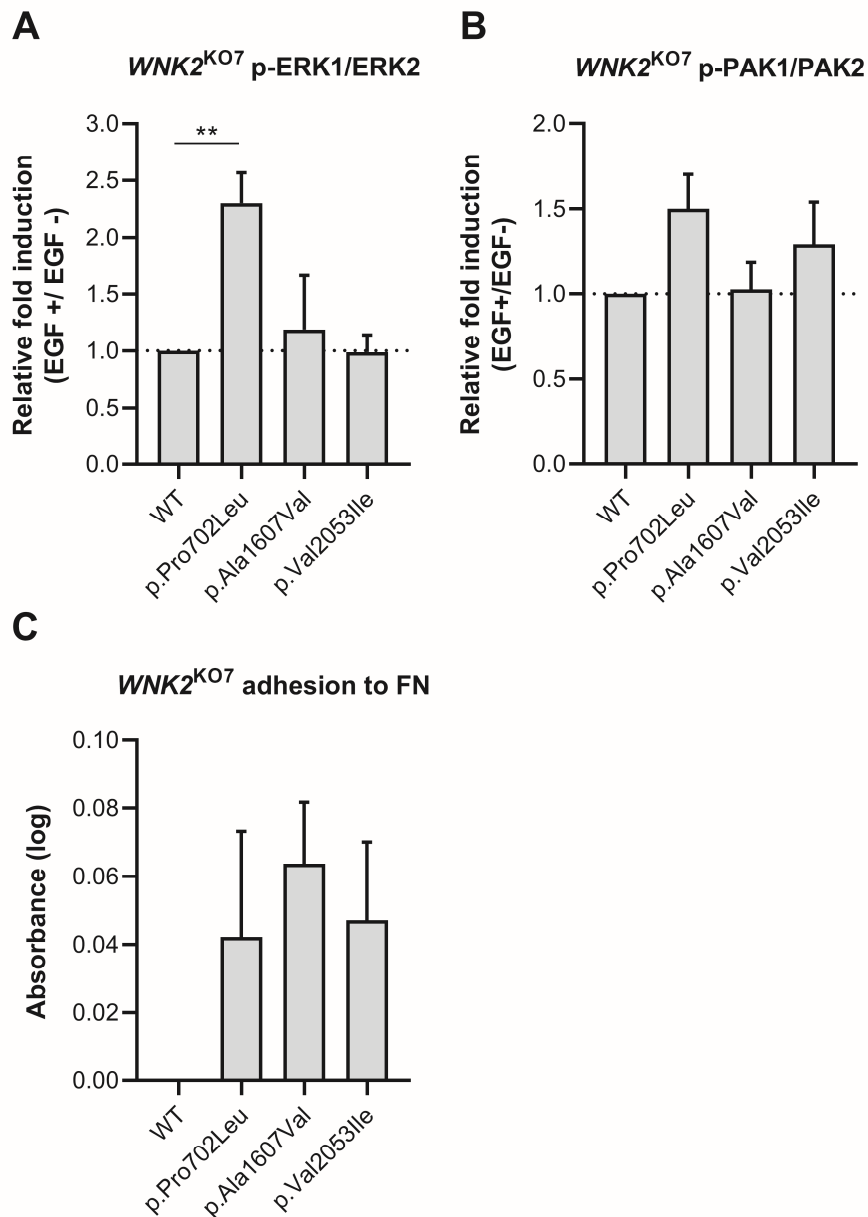


Supplementary figure 1. CRISPR/Cas9-mediated *WNK2* gene inactivation in HT-29 cells. (A) Sanger sequencing profile of both clones aligned with the wild-type sequence, showing indel mutations near the break site (arrow). The 20-nt sgRNA and the 3-nt PAM sequences are depicted. (B) *WNK2* mRNA relative expression of both clones assessed by Real Time PCR. Data represent mean \pm SD (n = 2). The experiment was performed in triplicate and repeated two times (n = 2). (C) Representative Western Blot analysis of several *WNK2*^{KO} clones, confirming *WNK2* depletion on clones *WNK2*^{KO2} and *WNK2*^{KO7}. Uncropped Western can be consulted further down in this same document.



Supplementary figure 2. Development of a cellular model for *WNK2* variant characterization. (A) Sanger sequencing of site-directed mutagenesis products confirmed the correct generation of *WNK2* p.(Pro702Leu), p.(Ala1607Val) and p.(Val2053Ile) variants. Changes in the nucleotide sequence are marked, and the altered amino acid sequence is highlighted in bold. (B) Representative functional titration of the lentiviral particles stock to ensure that a single lentiviral

copy is inserted per cell. The example corresponds to the pLVX-TetOne-WNK2-WT plasmid reintroduced into both clones to rescue the expression of wild-type *WNK2*. On the left panel, the WPRE lentiviral sequence is detected by SYBR green-based Real-Time PCR by using a standard curve with known lentiviral vector copy numbers. Samples infected with different dilutions of the lentiviral particles stock were assayed, and the selected dilution (one vector copy number/cell) is circled. On the right panel, the extrapolated viral copy numbers were corrected taking into account the amount of DNA assayed, determining the integrated viral copies per cell. Samples were assayed in triplicate. (C) *WNK2* mRNA relative expression in *WNK2*^{KO2} cells expressing either the WT sequence (rescued phenotype) or each of the selected variants. The dose of 1 ug/mL of doxycycline induced near wild-type *WNK2* expression levels. Data represent mean \pm SD The experiment was performed in triplicate and repeated two times (n = 2). (D) Representative Western Blot analysis of *WNK2*^{KO2} cells expressing either the WT sequence or each of the selected variants after 1 ug/mL doxycycline treatment. VCL, vinculin.



Supplementary figure 3. Functional characterization of *WNK2*^{K07} cellular model. *WNK2*^{K07} cells expressing either the WT sequence (rescued phenotype) or each of the selected variants were functionally characterized. Data is normalized against the observed values in *WNK2* WT. (A) ERK1/2 phosphorylation levels measured by ELISA in cell samples treated or untreated with 1 ng/mL hEGF. Data is displayed as EGF+/- ratio (n = 3; mean ± SD). (B)

PAK1/2 phosphorylation levels measured by an In-Cell ELISA (ICE) assay in cell samples treated or untreated with 10 ng/mL hEGF. Data is displayed as EGF+/- ratio (n = 3; mean \pm SD) (C) Cells were cultured on fibronectin (FN)-coated plates for 1 hour and the adherent cells were detected by crystal violet staining (mean \pm SD) The experiments were performed in triplicate and repeated three times (n = 3). ** $P < .01$, ANOVA with Fisher's LSD post hoc test.

Supplementary table 3. Predictions according to the eukaryotic linear motif (ELM) resource for the *WNK2* genetic variants identified in additional external SPS cohorts. ELM is a repository of manually curated experimentally validated short linear protein motifs and it is used for investigating functional regions in proteins.

Genetic variant	Exon	Cohort	ELM domain	Positions	Instances (matched sequence)	ELM description
c.106_107insG (p.Pro36Argfs*121)	1	DE	Frameshift			
c.1853G>A (p.Ser618Asn)	8	NL	DEG_SCF_TRCP1_1	617-623	DSGQGST	The DSGxxS phospho-dependent degron binds the F box protein of the SCF-betaTrCP1 complex. The degron is found in various proteins that function in regulation of cell state
c.2105C>T (p.Pro702Leu)	9	ES	DEG_SCF_FBW7_1	712-719	PPSTPMPT	The TPxxS phospho-dependent degron binds the FBW7 F box proteins of the SCF (Skp1_Cullin-Fbox) complex.
			LIG_SH3_3	705-711	SFAPVLP	This is the motif recognized by those SH3 domains with a non-canonical class I recognition specificity
c.2758G>A (p.Ala920Thr)	11	AU	DOC_WW_Pin1_4	924-929	VPPSPH	The Class IV WW domain interaction motif is recognised primarily by the Pin1 phosphorylation-dependent prolyl isomerase

			LIG_SH3_3	922-928	TDVPPSP	This is the motif recognized by those SH3 domains with a non-canonical class I recognition specificity
			MOD_GSK3_1	924-931	VPPSPHHT	GSK3 phosphorylation recognition site
			MOD_ProDKin_1	924-930	VPPSPHH	Proline-Directed Kinase (e.g. MAPK) phosphorylation site in higher eukaryotes
c.3341C>T (p.Thr1114Met)	12	ES	MOD_GSK3_1	1107-1114	PCPTVQLT	GSK3 phosphorylation recognition site
c.3418G>A (p.Gly1140Ser)	14	DE	MOD_CK1_1	1134-1140	SCESYGG	CK1 phosphorylation site
				1141-1147	SDVTSGK	CK1 phosphorylation site
			MOD_CK2_1	1142-1148	DVTSGKE	Casein kinase 2 (CK2) phosphorylation site
			MOD_GSK3_1	1134-1141	SCESYGGG	GSK3 phosphorylation recognition site
			MOD_GSK3_1	1138-1145	YGGSDVTS	GSK3 phosphorylation recognition site
c.3623C>T (p.Thr1208Met)	15	DE	DOC_PP1_RVXF_1	1210-1217	NHKMVTFK	Protein phosphatase 1 catalytic subunit (PP1c) interacting motif binds targeting proteins that dock to the substrate for dephosphorylation. The motif defined is [RK][0,1][V][^P][FW]
c.4820C>T	20	ES	MOD_CK1_1	1605-1611	SSASAGT	CK1 phosphorylation site
			MOD_GSK3_1	1608-1614	HVPTSSAS	GSK3 phosphorylation recognition site
			DOC_WW_Pin1_4	1608-1613	SAGTPV	Pin1 phosphorylation-dependent

(p.Ala1607Val)						prolyl isomerase recognition site
c.5476C>T (p.Arg1826Trp)	23	DE	LIG_14-3-3_CanoR_1	1826-1836	RRAQTASSIEV	Canonical Arg-containing phospho-motif mediating a strong interaction with 14-3-3 proteins
c.5588T>C (p.Leu1863Pro)	23	ES	MOD_GSK3_1	1859-1866	KQASLPVS	GSK3 phosphorylation recognition site
c.5656C>T (p.Arg1886Trp)	23	DE, NL	MOD_GSK3_1	1882-1889	QRPSRAGS	GSK3 phosphorylation recognition site
			MOD_PKA_2	1882-1888	QRPSRAG	Secondary preference for PKA-type AGC kinase phosphorylation
c.5906C>G (p.Pro1969Arg)	24	AU	LIG_SH3_3	1963-1969	LGKPLPP	This is the motif recognized by those SH3 domains with a non-canonical class I recognition specificity
			DOC_MAPK_DCC_7	1965-1973	KPLPPNVGF	A kinase docking motif mediating interaction towards the ERK1/2 and p38 subfamilies of MAP kinases
			DOC_MAPK_MEF2A_6	1965-1973	KPLPPNVGF	A kinase docking motif that mediates interaction towards the ERK1/2 and p38 subfamilies of MAP kinases
c.6080C>G (p.Ala2027Gly)	25	DE	MOD_GSK3_1	2027-2034	AQASVGLT	GSK3 phosphorylation recognition site
c.6157G>A (p.Val2053Ile)	25	ES	LIG_14-3-3_CanoR_1	2054-2058	RASLS	Canonical Arg-containing phospho-motif mediating a strong interaction with 14-3-3

						proteins.
c.6512G>A (p.Ser2171Asn)	28	DE	NA	NA	NA	NA

ELM Eukaryotic Linear Motif, NA, not available. DE, Germany; NL, Netherlands; ES, Spain; AU, Australia.

Uncropped Western Blot – Supplementary figure 1C

