Original research

Germline mutations in \textit{WNK2} could be associated with serrated polyposis syndrome


\textbf{ABSTRACT}

\textbf{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.

\textbf{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 \textit{WNK2} gene was disrupted in HT-29 cells by gene editing, and \textit{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 progress, survival and adhesion.

\textbf{Results} We identified 2 rare germline variants in the \textit{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.

\textbf{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 \textit{WNK2} gene. Functional studies suggested germline \textit{WNK2} variants affect protein function in the context of the MAPK pathway, a molecular hallmark in this disease.

\textbf{INTRODUCTION}

Colorectal cancer (CRC) is one of the most common cancers worldwide with a significant associated mortality. Aside from lung cancer, with an avoidable environmental cause, CRC is responsible for more deaths than any other malignancy in Western countries. The vast majority of CRC cases develop through an adenoma-carcinoma sequence. 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.

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. 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. 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. Although its prevalence in the population is unknown, it could be higher than expected according to data from CRC screening. It is also probably underrecognised due to 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. Smoking, body mass index and alcohol intake have been highlighted as important environmental risk factors for serrated polyps. Importantly, twin studies also showed that around 13%–30% of the variation in CRC susceptibility involves inherited genetic factors. 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. 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.

First reported in 2014, germline loss-of-function variants in RNF43 were associated with the development of multiple sessile serrated adenomas. 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. 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. Other genes involved in polyposis predisposition such as BMPRIA, SMAD4, PTEN and GREM1 have been also screened with negative results. 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 SPS 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. 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. The updated 2019 WHO criteria were not available when this study was initiated and developed for the discovery cohort. A complete clinical characterisation of this discovery cohort was previously published. 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 Clinic de Barcelona, Institut Català d’Oncologia-IDIBELL and Fundación Pública Galicia 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 fulfill 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. 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) and RNFL43, 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). 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/). 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 plasmin, 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 (Biorad, Hercules, California, USA). Phospho-ERK1/2 levels between EGFP-stimulated and unstimulated 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 antiozyme Dylight 800 (ThermoFisher, Waltham, Massachusetts, USA) and anti-rabbit 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 60,000 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-scapper, 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 Elisspot 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. 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. 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. 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 SBS1 and SBS5, considered clock-like mutational signatures, were the most represented in both samples, and no other distinctive signature was apparent.

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, and c.3341C>T (p.Val1114Met) 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...
Cancer genetics

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.
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 (WNK2KO2 and WNK2KO7) with no expression of WNK2 at both mRNA and protein levels (online supplemental figure 1).

Afterwards, we evaluated the phenotype of the selected WNK2KO2 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 WNK2KO2, which showed the highest ERK1/2 and PAK1/2 phosphorylation levels at all tested doses.

We further characterised WNK2KO2 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 WNK2KO2 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. 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. WNK2KO2 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 germine 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 WNK2KO2 and WNK2KO7 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). WNK2KO2 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 p.Pro702Leu and p.Ala1607Val variants (figure 3F), compared with the rescued WT phenotype.

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

<table>
<thead>
<tr>
<th>Variant</th>
<th>Exon</th>
<th>Prediction tools</th>
<th>gnomAD</th>
<th>Allele count</th>
<th>Number of homozygotes</th>
<th>Family</th>
<th>SPS criteria</th>
<th>CRC index case</th>
<th>CRC family history</th>
<th>Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>c.2105C&gt;T (p.Pro702Leu)</td>
<td>9</td>
<td>5</td>
<td>0.000125</td>
<td>19/152,210</td>
<td>0</td>
<td>PanSPS_044</td>
<td>3</td>
<td>Y</td>
<td>N</td>
<td>Validation</td>
</tr>
<tr>
<td>c.3341C&gt;T (p.Thr1114Met)</td>
<td>12</td>
<td>3</td>
<td>0.0000526</td>
<td>8/152,182</td>
<td>0</td>
<td>PanSPS_078</td>
<td>1</td>
<td>N</td>
<td>Y</td>
<td>Validation</td>
</tr>
<tr>
<td>c.4820C&gt;T (p.Ala1607Val)</td>
<td>20</td>
<td>5</td>
<td>0.000138</td>
<td>21/152,228</td>
<td>0</td>
<td>CAR_SP54.4</td>
<td>1,3</td>
<td>N</td>
<td>Y</td>
<td>Discovery</td>
</tr>
<tr>
<td>c.5588T&gt;C (p.Leu1863Pro)</td>
<td>23</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>PanSPS_055</td>
<td>3</td>
<td>N</td>
<td>N</td>
<td>Validation</td>
</tr>
<tr>
<td>c.6157G&gt;A (p.Val2053Ile)</td>
<td>25</td>
<td>5</td>
<td>0.0000657</td>
<td>10/152,200</td>
<td>0</td>
<td>CAR_SP6.3</td>
<td>2,3</td>
<td>N</td>
<td>Y</td>
<td>Discovery</td>
</tr>
</tbody>
</table>

Prediction tools: number of pathogenicity positive predictions according to missense bioinformatic prediction tools (PhylD, 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.

WNK2, colorectal cancer; gnomAD, Genome Aggregation Database; N, no; SPS, serrated polyposis syndrome; Y, yes.

**Cancer genetics**

The functional characterisation of WNK2 variants was replicated in clone WNK2KO (online supplemental figure 3), focusing on the assays more directed to the MAPK pathway itself (ERK1/2 and PAK1/2 phosphorylation levels in both WNK2KO and WNK2KO 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±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±SD (n=4). (F) Fibronectin (FN)-mediated cell adhesion assay using crystal violet staining. Data represent mean±SD (n=3). (G) MMP2 mRNA relative expression after treatment with 1 ng/mL of hEGF. Data represent mean±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.

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 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.babelomics.org). By applying the previous filters of frequency and pathogenicity, we identified two rare, protein-altering

Figure 3

Variants in WNK2 activate the MAPK pathway and contribute to cell cycle deregulation, cell proliferation and altered cellular adhesion. WNK2 KO cells 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 KO 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.
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 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 p21Rac and consequent activation of ERK1/2 (figure 4).\(^40\)\(^41\)

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 tissue\(^38\) and in hepatocarcinoma cell lines.\(^43\) All tested WNK2 variants showed the same tendency, with a prominent WNK2 malfunction detected for the p.Pro702Leu variant, which displayed the most significant phosphorylation levels of both PAK1/2 and ERK1/2 in WNK2\(^{KO}\) and WNK2\(^{KO}\) clones.

---

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

<table>
<thead>
<tr>
<th>Variant</th>
<th>Exon</th>
<th>gnomAD</th>
<th>Allele number</th>
<th>Number of homozygotes</th>
<th>CADD</th>
<th>Patient</th>
<th>SPS criteria</th>
<th>Onset age</th>
<th>CRC index case</th>
<th>Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>c.106_107insG (p.Pro36Argfs*121)</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
<td>BN-174</td>
<td>1.3</td>
<td>22</td>
<td>N</td>
<td>DE</td>
</tr>
<tr>
<td>c.1853G&gt;A (p.Ser618Asn)</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
<td>BN-174</td>
<td>1.3</td>
<td>22</td>
<td>N</td>
<td>DE</td>
</tr>
<tr>
<td>c.2758G&gt;A (p.Ala920Thr)</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td>GCP208001</td>
<td>1.3</td>
<td>27</td>
<td>N</td>
</tr>
<tr>
<td>c.3418G&gt;A (p.Gly1140Ser)</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>23.9</td>
<td>MUC-6</td>
<td>1.3</td>
<td>42</td>
<td>NA</td>
</tr>
<tr>
<td>c.3823C&gt;T (p.Thr1208Met)</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>26.6</td>
<td>MUC-4</td>
<td>*</td>
<td>30</td>
<td>N</td>
</tr>
<tr>
<td>c.5476C&gt;T (p.Arg1886Tnp)</td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>25.1</td>
<td>BN-210</td>
<td>3</td>
<td>55</td>
<td>N</td>
</tr>
<tr>
<td>c.5656C&gt;T (p.Arg1886Tnp)</td>
<td>23</td>
<td>0.02233</td>
<td>219/152</td>
<td>194</td>
<td>1</td>
<td>21.2</td>
<td>BN-145</td>
<td>3</td>
<td>29</td>
<td>NA</td>
</tr>
<tr>
<td>c.5906C&gt;G (p.Pro1969Arg)</td>
<td>24</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>25.3</td>
<td>GCP308001</td>
<td>3</td>
<td>59</td>
<td>N</td>
</tr>
<tr>
<td>c.6080C&gt;G (p.Ala2027Gly)</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>25.1</td>
<td>BN-83</td>
<td>3</td>
<td>18</td>
<td>NA</td>
</tr>
<tr>
<td>c.6512G&gt;A (p.Glu2171Asp)</td>
<td>28</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>23</td>
<td>BN-104</td>
<td>1.3</td>
<td>21</td>
<td>NA</td>
</tr>
</tbody>
</table>

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.

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.

---

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.

---

Cancer genetics

---

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.\textsuperscript{34} 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.\textsuperscript{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.\textsuperscript{35} 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.\textsuperscript{46,47}

Cell adhesion is an important feature of the cell malignancy process and is closely regulated by PAK1 activation.\textsuperscript{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.\textsuperscript{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.\textsuperscript{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.\textsuperscript{30} 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.\textsuperscript{31} 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).\textsuperscript{32} 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 $\approx$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.
Cancer genetics

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

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

Twitter Laura Valle @VlalleResearch

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


Funding YSdL was awarded an INPhINIT fellowship (LCF/BD/0118/160058) from la Caixa Foundation JDI 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 CIBEREH. MD-G was supported by a contract from Agència de Gestió d’Ajuts Universitaris i de Recerca – AGAUR (Generalitat de Catalunya, 2019FI_2B, 00203). LB was supported by a Juan de la Cierva postdoctoral fellowship (FCI-2017-32593). IBARW and AKS are supported by European Union’s Horizon 2020 research and innovation programme under agreement no. 772957 (Solve).

References