 doi:10.20944/preprints202103.0121.v1

Article

A novel low-risk germline variant in the SH2 domain of the SRC gene affects multiple pathways in familial colorectal cancer

Diamanto Skopelitou [1,2,3,4,*], Beiping Miao [1,2,3,*], Aayushi Srivastava [1,2,3,4], Abhishek Kumar [5], Magdalena Kuswik [6], Dagmara Dymerska [6], Nagarajan Paramasivam [7], Matthias Schlesner [8], Jan Lubinski [6], Kari Hemminki [1,9,10,§], Asta Försti [1,2,3,§] and Obul Reddy Bandapalli [1,2,3,4,§,#]

[1] Molecular Genetic Epidemiology, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany; mando.skopelitou@yahoo.de (D.S.); b.miao@kitz-heidelberg.de (B.M.); srivastava.aayushi97@gmail.com (A.S.); k.hemminki@dkfz.de (K.H.); a.foersti@kitz-heidelberg.de (A.F.)
[2] Hopp Children's Cancer Center (KiTZ), 69120 Heidelberg, Germany
[3] Division of Pediatric Neurooncology, German Cancer Research Center (DKFZ) and German Cancer Consortium (DKTK), 69120 Heidelberg, Germany
[4] Medical Faculty Heidelberg, Heidelberg University, 69120 Heidelberg, Germany
[5] Institute of Bioinformatics, International Technology Park, Bangalore, India; abhishek@ibioinformatics.org (A.K.)
[6] Department of Genetics and Pathology, Pomeranian Medical University, Szczecin, Poland; magdalenakuswik@gmail.com (M.K.); dymerska@pum.edu.pl (D.D.); lubinski@pum.edu.pl (J.L.)
[7] Computational Oncology, Molecular Diagnostics Program, National Center for Tumor Diseases (NCT), Germany; n.paramasivam@dkfz.de
[8] Bioinformatics and Omics Data Analytics, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany; m.schlesner@dkfz.de
[9] Cancer Epidemiology, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany
[10] Faculty of Medicine and Biomedical Center in Pilsen, Charles University in Prague, 30605 Pilsen, Czech Republic
* Equal contribution
§ Shared senior authorship
# Correspondence: o.bandapalli@kitz-heidelberg.de; Tel.: +49-6221-421809

Abstract: Colorectal cancer (CRC) shows one of the largest proportions of familial cases among different malignancies, but only 5-10% of all CRC cases are linked to mutations in established predisposition genes. Thus, familial CRC constitutes a promising target for the identification of novel, high- to moderate-penetrance germline variants underlying cancer susceptibility by next generation sequencing. In this study, we performed whole genome sequencing on 3 members of a family with CRC aggregation. Subsequent integrative *in silico* analysis using our in-house developed variant prioritization pipeline resulted in the identification of a novel germline missense variant in *SRC* gene (V177M), a proto-oncogene highly upregulated in CRC.

Functional validation experiments in HT-29 cells showed that introduction of SRC^{V177M} resulted in increased cell proliferation and enhanced protein expression of phospho-SRC (Y419), a potential marker for SRC activity. Upregulation of *paxillin*, *β-Catenin* and *STAT3* mRNA levels, increased levels of phospho-ERK, CREB and CCND1 proteins and downregulation of the tumor suppressor p53 further proposed the activation of several pathways due to the SRC^{V177M} variant.

The findings of our pedigree-based study contribute to the exploration of the genetic background of familial CRC and bring insights into the molecular basis of upregulated SRC activity and downstream pathways in colorectal carcinogenesis.

Keywords: Familial colorectal cancer; SRC; germline variant; whole genome sequencing

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 March 2021 doi:10.20944/preprints202103.0121.v1

1. Introduction

Colorectal cancer (CRC) shows one of the largest proportions of familial cases among different malignancies, and thus it constitutes a promising target for next generation sequencing (NGS) as a tool for unravelling the underlying genetic alterations [1]. Besides the established cancer predisposing genes, including mismatch repair genes *MLH1*, *MSH2*, and *PMS2* as well as *APC*, *MUTYH* and *SMAD4*/*BMPR1A*, recent sequencing studies have identified the *NTHL1*, *RNF43*, *POLE* and *POLD1* genes as novel susceptibility genes underlying CRC inheritance [2-8]. Since germline variants in the described genes are considered to contribute to only 5-10% of all CRC cases, the remaining proportion of familial CRC, not linked to the discovered cancer predisposing genes, has to be further investigated [8-10].

In order to bring insight into the genetic background of unexplored familial CRC, we performed whole genome sequencing (WGS) in combination with integrative *in silico* analysis on a family presenting CRCs in three generations. Sequencing data were analyzed using the Familial Cancer Variant Prioritization Pipeline version 2 (FCVPPv2), developed by us and implemented in the previous analysis of various familial malignancies, such as Hodgkin lymphoma and thyroid cancer [11-13]. The results converged to a few candidate genes which were further evaluated by additional *in silico* analysis.

2. Materials and Methods

2.1. Patient samples & ethical permissions

The CRC affected family of this study was recruited from Poland. The respective pedigree is shown in, representing the CRC affected members III-1 and IV-8 and the unaffected family member IV-7 that were included in our WGS analysis. Collection of blood samples and clinical information from subjects was undertaken with informed written consent in accordance with the tenets of the declaration of Helsinki.

The study was approved by the Bioethics Committee of the Pomeranian Medical Academy in Szczecin No: BN-001/174/05.

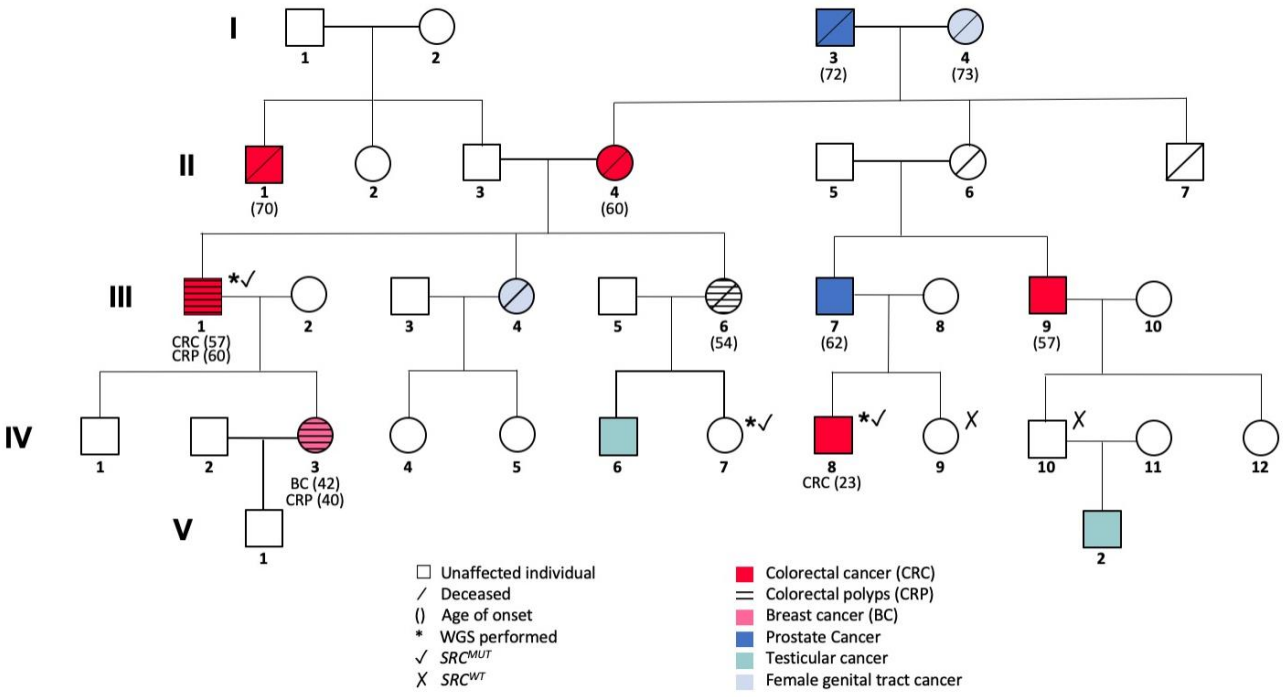


Figure 1. Pedigree of the studied CRC affected family with SRCV177M mutation.

2.2. Whole genome sequencing and variant calling, annotation and filtering

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 March 2021 doi:10.20944/preprints202103.0121.v1

Peripheral blood samples were collected from affected and unaffected family members who agreed to participate in the study as well as from the validation cohort. Genomic DNA was isolated using a modified Lahiri and Schnabel method [14]. WGS was performed using Illumina-based small read sequencing. After mapping to the reference human genome (assembly version Hs37d5) with BWA [15], duplicates were removed using Picard (http://broadinstitute.github.io/picard/).

Applying SAM tools [16] and Platypus [17], single nucleotide variants (SNVs) and small indels were detected, respectively. ANNOVAR [18], 1000 Genomes [19], dbSNP [20] and Exome Aggregation Consortium (ExAC) [21] were used for variant annotation. Variants with a quality score of ≥20 and a coverage score of ≥5x, SNVs passing the strand bias filter (a minimum one read support from both forward and reverse strand) and indels passing all the Platypus internal filters were further checked for minor allele frequencies (MAFs). With respect to the 1000 Genomes Project Phase 3, non-TCGA ExAC data [21], NHLBI-ESP6500 and local data sets, variants with a MAF ≤0.1% in the European population were selected for further analysis. A pairwise comparison of shared rare variants among cohort was performed to check for sample swaps and family relatedness.

2.3. Familial segregation of the cancer predisposing variant

The studied family shows aggregation of CRC and multiple other malignancies such as prostate, female genital tract, testicular and breast cancer. In order to define familial segregation criteria for the pathogenic variant predisposing for cancer development in this family, the hereditary line of malignant diseases was retraced, assigning to each analyzed family member a probability of being a Mendelian case and carrier of the mutation (Figure 1).

The first case sequenced in this family (III-1) developed CRC as well as colorectal polyps (CRP) at the age of 57 and 60 years, respectively, and was thus considered as a carrier of the mutation. Tracing genetic cancer predisposition back to his CRC affected mother (II-4) and further to his cancer-affected grandparents (I-3, I-4; prostate, female genital tract cancer, respectively), the cancer predisposing mutation might have been further inherited to his first cousin once removed (IV-8; via II-6 and III-7). Since this family member (IV-8) developed CRC at the young age of 23 years, he was regarded as the second case of the family and thus carrier of the mutation. On the other hand, the CRC unaffected family member included in this study (IV-7) was 39 years at the time of recruitment. Her first-degree relatives were affected by cancer or CRP (IV-6, III-6, respectively), suggesting that she might show the genotype without expressing the disease phenotype yet. Thus, she was considered as a possible carrier of the mutation.
The identified variants were filtered according to the described definitions of III-1 and IV-8 as cases and IV-7 as a possible carrier of the family, respectively summarized in Supplementary Table S1.

2.4. Evaluation of the pathogenicity of identified variants using FCVPPv2

Applying our in-house developed FCVPPv2, the cancer predisposing potential of coding variants was evaluated, including non-synonymous, stop-gain, small indels and exonic variants of unknown classification.

Ranking all variants using the combined annotation dependent depletion tool (CADD) v1.3, only the top 10% of potentially deleterious variants represented by a PHRED-like CADD score ≥10 were deduced for further analysis [23]. Since evolutionary conservation is regarded to correlate with the functional importance of a genomic position, conservational screening of variants was performed using the following scoring tools with respective cutoff values given in brackets: Genomic Evolutionary Rate Profiling (GERP ≥2.0), PhastCons (>0.3) and PhyloP score (≥3.0) [24, 25]. In order to further assess the intolerance of genes against functional genetic variation, three intolerance scores (<0) based on allele frequency data from our in-house datasets, from ESP [22] and ExAC [26] were applied. Furthermore, intolerance screening of variants included the application of the Z-Score (>0) and pLI score (≥0.9), developed from ExAC consortium

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 March 2021 doi:10.20944/preprints202103.0121.v1

specifically for missense and loss-of-function variants, respectively. Next, the deleteriousness of non-synonymous and splice site SNVs was evaluated, using 10 different scoring systems and 2 meta-prediction tools derived from dbNSFP v3.0 (database for nonsynonymous SNPs' functional predictions) [27]. In order to be further considered in the analysis, the variants should fulfill following filtering criteria: PHRED-like CADD-score of ≥10, ≥2 out of 3 conservational scores, ≥60% of 4 intolerance scores and ≥60% of 12 deleteriousness scores. The remaining top exonic candidates were assessed for allele frequencies in the non-Finnish European population using the latest version of gnomAD browser (https://gnomad.broadinstitute.org/) [28], for predicted cancer drivers by means of the Cancer Genome Interpreter (CGI, https://www.cancergenomeinterpreter.org/) [29] and for predicted functional effects of respective amino acid substitutions by Snap² [30]. Conclusively, recent literature was checked for reported gene-cancer relations and potentially cancer-related protein functions of the top candidates.

2.5. Confirmation of familial segregation by Sanger sequencing

Polymerase Chain Reaction (PCR) was performed with HotStarTaq DNA Polymerase (Qiagen, #203205) at an annealing temperature of 56°C in order to amplify exon 5 of the *SRC* gene (ENST00000358208.4) from DNA of family members III-1, IV-7, IV-8, IV-9 and IV-10. The primers were designed with Primer3 v.0.4.0 (http://bioinfo.ut.ee/primer3-0.4.0/) as followed: *SRC* forward 5′-GGCTACATCCCCAGCAACTA-3′, reverse 5′-CCTCCCTACTCCACAAACCA-3′. PCR amplicons were validated by gel electrophoresis and purified with ExoSAP purification kit according to the manufacturer´s instruction. Sequencing reaction was performed using the BigDye Terminator v3.1 Ready Reaction Cycle Sequencing kit (Thermo Fisher Scientific, #4337455), followed by manual analysis of the electrophoretic profiles of *SRC* sequences.

2.6. Screening of familial CRC index cases and healthy individuals by Taqman assay

The *SRC* variant was screened in 1690 familial CRC cases not related to the studied family and 1676 healthy elderly individuals, both from Poland, using a custom-made Taqman assay.

2.7. Plasmid preparation and cell culture

pcDNA3-MTS-WT-c-Src-FLAG (#44652) was purchased from Addgene and used in functional experiments as the wild type *SRC* plasmid (SRC^{WT}). The mutant *SRC* plasmid (SRC^{V177M}) was created by using QuikChange II XL Site-Directed Mutagenesis Kit (Agilent, #200521) and following primers designed based on Agilent QuikChange Primer Design (https://www.agilent.com/store/primerDesignProgram.jsp): forward 5'-gtctcactttctcgcatgaggaaggtccctctc-3', reverse 5'-gagagggaccttcctcatgcgagaaagtgagac-3'. After confirmation by Sanger sequencing, both plasmids were transformed into XL10-Gold Ultracompetent Cells (Agilent, #200314) and plasmid extraction was performed using PureLink™ HiPure Plasmid Midiprep Kit (Thermo Fisher Scientific, #K210004) according to manufacturer's instructions.

Human colon cancer cell line HT-29 was a kind gift from Peter Krammer´s lab (DKFZ). HT-29 cells were cultured in RPMI and used for transfection of *pcDNA3* (HT29-*pcDNA3*), SRC^{WT} (HT29-SRC^{WT}) and SRC^{V177M} (HT29-SRC^{V177M}). Stably transfected pools of cells were selected by using G418.

2.8. Cell proliferation assays

HT-29 cells were seeded in 24-well plates and 24h later transfected with either 150ng of SRC^{WT}, SRC^{V177M} or *pcDNA3* vector as a negative control. After washing with PBS and trypsinizing the cells, viable cells were selected with trypan blue exclusion of dead cells and quantified by cell counting with the haemocytometer under a 10x objective at six

 doi:10.20944/preprints202103.0121.v1

different time points: day 0, 1, 2, 3, 4 and 5. Numbers of viable cells and respective proliferation curves were compared between HT29-*SRC*^{WT} and HT29-*SRC*^{V177M} cells.

2.9. Quantitative Polymerase Chain Reaction

RNA extraction from cells (HT29-*SRC*^{WT}, HT29-*SRC*^{V177M} and HT29-*pcDNA3*) was performed with Trizol and subsequent RNA purification with sodium acetate. ProtoScript First Strand cDNA Synthesis kit (NEB, #E6300S) was used for cDNA synthesis according to the manufacturer's instructions. Quantitative Polymerase Chain Reaction (qPCR) was performed by means of QuantiFast® SYBR® Green PCR (Qiagen, #204054). The utilized primer pairs for SRC downstream targets (*paxillin, PXN; β-Catenin, CTNNB1; signal transducer and activator of transcription 3, STAT3; AKT*) and the housekeeping gene *HPRT* (*hypoxanthine phosphoribosyltransferase*) as a reference are summarized with respective primer sequences in Supplementary Table S2. Relative gene expression was calculated with the 2ΔCT method and compared between HT29-*SRC*^{WT} and HT29-*SRC*^{V177M} cells.

2.10. Western Blot

Protein lysates from HT29-*SRC*^{WT}, HT29-*SRC*^{V177M} and HT29-*pcDNA3* cells were prepared and quantified by means of Pierce™ BCA Protein Assay Kit (Thermo Fisher Scientific, #23225). NuPAGE™ 4-12% Bis-Tris Protein Gels and the respective running buffer (Thermo Fisher Scientific; #NP0321PK2, #NP0001) were used for separation of 20µg of each protein sample. Blotted membranes were blocked with 2% milk for 1 hour, incubated overnight at 4℃ with primary antibody dilutions and subsequently for 1 hour at room temperature with the respective HRP-conjugated secondary antibody, diluted in 5% milk. Blots were developed using Amersham ECL Western Blotting Detection Kit (GE Healthcare Life Sciences, #RPN2108). Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) and β-Actin proteins were used for loading quantity control. All probed antibodies are summarized in Supplementary Table S3 with respective product details, dilution buffers and dilution factors.

3. Results

3.1. Familial cancer variant prioritization pipeline identifies a novel germline variant in SRC gene

In order to screen WGS data of the analyzed family members for cancer predisposing variants, we applied our in-house developed FCVPPv2 pipeline (Figure 2a). Filtering with a MAF ≤0.1% revealed a total number of 107,917 variants. By considering the familial segregation of the potentially cancer-causing mutation, 4,550 variants were deduced for genomic location-based filtering. Most of these variants were annotated to affect intronic or intergenic regions, leaving 38 coding variants that were further analyzed. Removal of synonymous variants due to their potentially less deleterious nature resulted in 22 non-synonymous, stop-gain and variants of unknown classification. Application of the PHRED-like CADD score further narrowed down this number to 16 variants. Additionally, screening for evolutionary conservation, intolerance of the genes against functional genetic variation as well as predicted deleteriousness reduced the number of variants to 12, to 5 and ultimately to 3 final candidates, respectively: the non-synonymous variants in OGFOD2 (R11Q) and SRC gene (V177M) and the stop-gain variant in ZNF408 gene (Q460X), which could only be annotated by 2 out of totally 12 deleteriousness scores due to its impact as a nonsense mutation (Table 1).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 March 2021 doi:10.20944/preprints202103.0121.v1

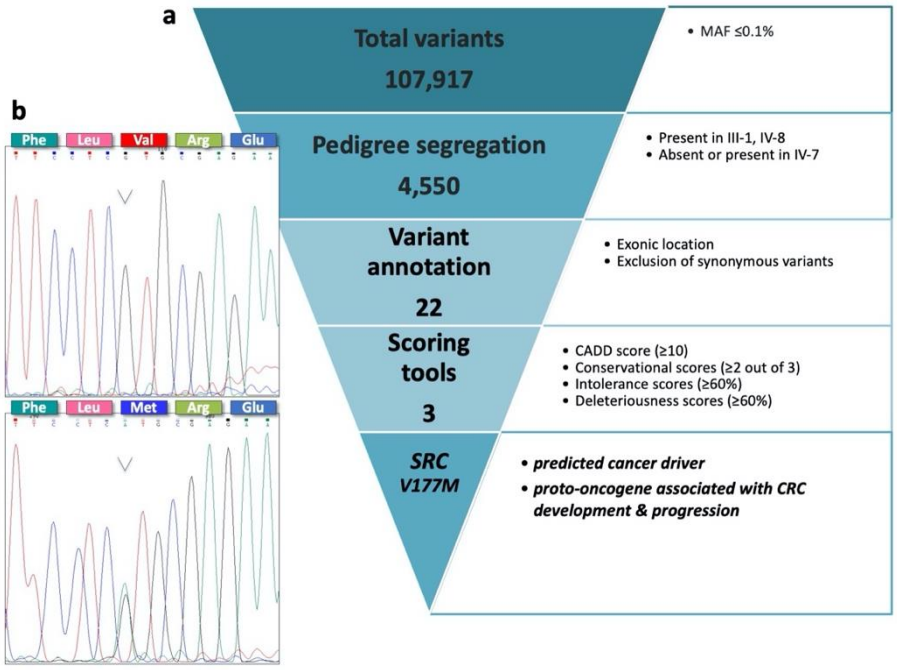


Figure 2. Prioritization of the missense variant in the *SRC* gene (V177M). **(a)** Flow chart depicting the filtering process of exonic variants according to the FCVPPv2. **(b)** Electropherograms representing the wild-type *SRC* sequence (upper panel) identified in family members IV-9 and IV-10 and the heterozygous SRC^{V177M} variant (lower panel) identified in family members III-1, IV-7 and IV-8. The respective substitution Val → Met is displayed in the amino acid sequences.

Table 1. Exonic germline variants prioritized in the studied CRC family. Chromosomal positions, classifications, pedigree segregation, allele frequencies, PHRED-like CADD scores, conservational scores and the percentage of positive intolerance and deleteriousness scores are summarized. CGI results, respective protein functions derived from GeneCards are included [31]. non-syn - non-synonymous; NFE - Non-Finnish European population; PP – predicted passenger; PD – predicted driver; OG – oncogene.

Gene name	Chromosomal position	Exonic classification	Pedigree segregation	Allele frequency		CADD SCORE	Conservational scores			Deleteriousness scores* (%)	Intolerance scores (%)	Amino acid change	CGI	Protein function
				ExAC	gnomAD NFE		GERP++	PhyloP	PhastCons					
OGFOD2	12-123461223-G-A	nonsyn SNV	III1, IV8	6.2×10^{-4}	7.1×10^{-4}	24.8	5.67	6.69	1	91.67	60	R11Q	PP	Iron ion binding, oxidoreductase activity
SRC	20-36022656-G-A	nonsyn SNV	III1, IV7, IV8	1.8×10^{-5}	3.9×10^{-5}	26.8	4.88	6.64	1	91.67	100	V177M	PD, OG	Embryonic development, gene transcription, cell cycle progression, cell growth, adhesion, migration, transformation apoptosis, immune response
ZNF408	11-46726628-C-T	stop-gain SNV	III1, IV7, IV8	1.1×10^{-4}	7.2×10^{-5}	36	5.15	1.53	0.81	100**	60	Q460X	PP	DNA binding protein, highly expressed in retina

* Deleteriousness scores: Following predictions were considered as favorable: SIFT – damaging; Polyphen2_HumDiv/HumVar – probably/possibly damaging; LRT – deleterious; MutationTaster – disease causing (automatic); MutationAssesor – high/medium; FATHMM – damaging; MetaSVM – damaging; MetaLR – damaging; VEST3 ⩾0.5; PROVEAN – damaging; Reliability Index ⩾5.

** Only 2 out of 12 deleteriousness scores were available for this variant.

Checking the top-listed variants with the latest version of gnomAD, revealed allele frequencies of ≤0.1% in the Non-Finnish European population for all the variants [28]. On the other hand, CGI reported SRC as the only predicted cancer driver with an oncogenic function, whereas OGFOD2 and ZNF408 were annotated as passenger mutations [29].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 March 2021 doi:10.20944/preprints202103.0121.v1

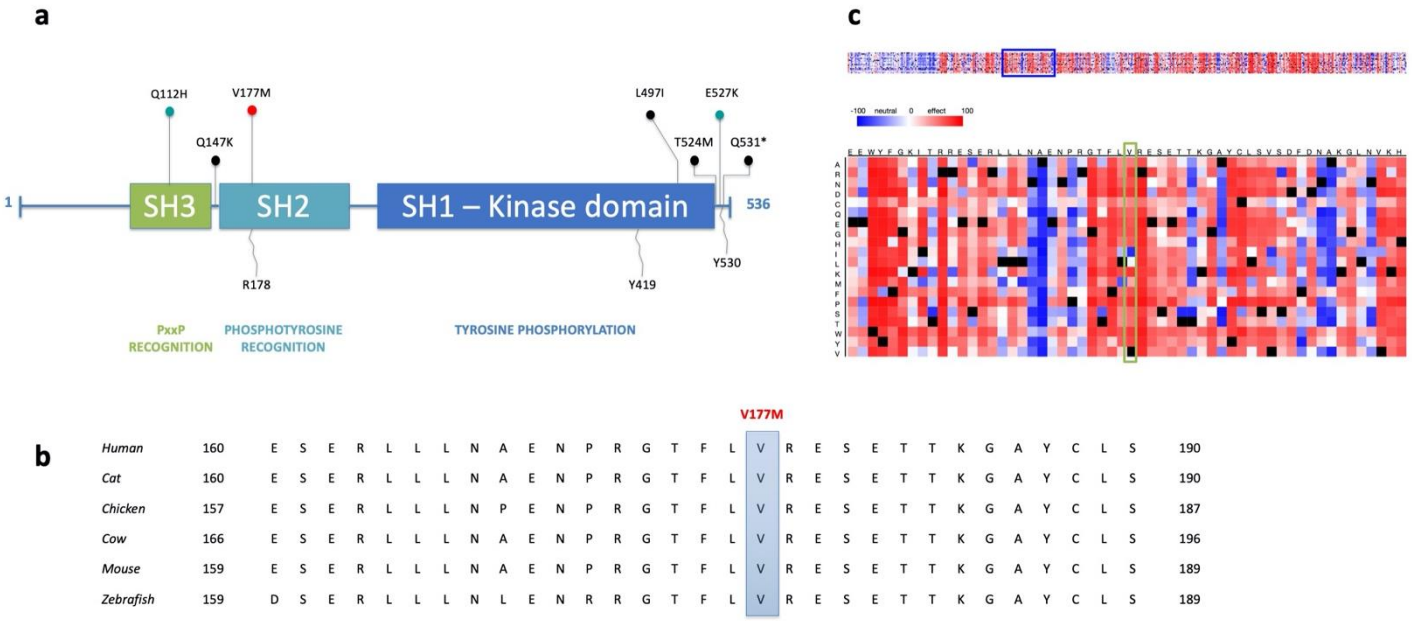
These predictions were confirmed by the results of literature search stressing the carcinogenic potential of SRC.

3.2. Confirmation of familial segregation & screening of a large cohort of familial CRC index patients and healthy individuals

Targeted Sanger sequencing for exon 5 of the *SRC* gene confirmed pedigree segregation of the prioritized variant, showing the heterozygous mutation SRC^{V177M} in 2 family members (III-1, IV-8) with CRC and in the possible carrier (IV-7) and the wild-type sequence in 2 family members without CRC, for whom the DNA samples were available and tested by Sanger sequencing (IV-9, IV-10, Figure 2b). Furthermore, targeted genotyping of 1690 unrelated familial CRC cases and 1676 healthy elderly individuals, both from Poland, using custom-made Taqman assay identified the SRC^{V177M} variant in 4 additional index cases diagnosed at the ages of 48, 50, 60 and 65 years, respectively, and in 3 healthy individuals aged 63, 65 and 89 years, respectively (*OR 1.65, 95%; CI 0.39-6.93, p=0.49*).

3.3. The identified variant affects the highly conserved SH2 domain of the SRC protein

Analysis of the SRC protein sequence proposed a high functional impact of the affected position: First of all, the identified missense variant (V177M) alters an amino acid residue within the SH2 domain (p.151–248), a protein domain enabling physical interactions with phosphotyrosine-containing target peptides in the course of intracellular signaling cascades (Figure 3a). As part of several proteins including the Src, Fps and Abl families, the SH2 domain shows high conservation, being identical in approximately 35% of all SH2 domains [34]. In particular, the universally conserved arginine residue R178 within the SH2 domain has been reported to play a central role in phosphotyrosine recognition and formation of electrostatic interactions [35]. Since the amino acid residue affected by the variant (V177M) is located directly adjacent to R178, the identified variant may have an impact on protein function and further protein-protein interactions. Alignment of SRC protein sequences of multiple species extracted from Ensembl (GRCh37/hg19), further revealed a high conservation of the whole protein (Supplementary Figure S1) and in particular of the affected region among all concerned species(Figure 3b) [33]. Similar results were obtained by Snap² indicating an overall relatively high impact of potential substitutions at the respective position of the predicted amino acid change (Figure 3c).



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 March 2021 doi:10.20944/preprints202103.0121.v1

Figure 3. SRC domain structure and protein sequence highlighting the conserved and functionally important region affected by the missense variant V177M. **(a)** SRC protein domains are represented with respective domain functions: SH3 (84 – 145), SH2 (151 – 248), SH1 (270 – 523). Phosphorylation sites required for activation (R178 and Y419) and autoinhibition (Y530) are included. The identified germline variant affects the amino acid residue (V177M, red pin) directly adjacent to R178 within the SH2 domain, crucial for phosphotyrosine recognition. Additional *SRC* germline variants are indicated by blue pins (Q112H in CRC; E527K in thrombocytopenia, myelofibrosis, bleeding, bone pathologies [32]). Somatic mutations identified in CRC are represented by black pins: truncating mutation Q531*, missense mutations extracted from cBioPortal (www.cbioportal.org) on 20th of March 2020 using TCGA PanCancer data. **(b)** Extract of SRC protein sequence alignment downloaded from Ensembl (GRCh37/hg19) [33] for the following species: human (ENST00000373578.2), cow (ENSBTAT00000011767.3), mouse (ENSMUST00000029175.7), chicken (ENSGALT00000006127.2), cat (ENSFCAT00000006993.2), zebrafish (ENSDART00000102843.4). As highlighted, the variant affects an amino acid residue identical in all sequences and thus highly conserved across the concerned species. **(c)** Predicted functional effects of amino acid substitutions are represented by the heat map extracted from Snap², whereby the color red indicates a strong predicted effect, white an inconclusive prediction and blue a weak predicted effect. The position of the amino acid residue affected by the identified missense mutation (V177M) is highlighted with horizontal yellow box, showing an overall relatively high impact of potential substitutions at the respective position.

Based on the established oncogenic role of SRC in general cancer development and in particular in CRC and on the described analysis results of the FCVPPv2, the identified SRC^{V177M} variant was considered to bear pathogenic potential leading to its prioritization for functional validation.

3.4. Functional validation of the prioritized variant in SRC gene

3.4.1. Enhanced cell proliferation of SRC^{V177M} expressing CRC cells in vitro

In order to investigate the proliferative impact of the prioritized SRC^{V177M} variant, cell proliferation assays were conducted at 6 different time points using HT-29 cells. Cells transfected with SRC^{V177M} showed a significant increase in cell numbers compared to HT29-SRC^{WT} cells starting from day 1 ($p \leq 0.0001$). Cells transfected with *pcDNA3* showed the lowest cell numbers compared to both, HT29-SRC^{V177M} and HT29-SRC^{WT} cells, at all time points (Figure 4a).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 March 2021 doi:10.20944/preprints202103.0121.v1

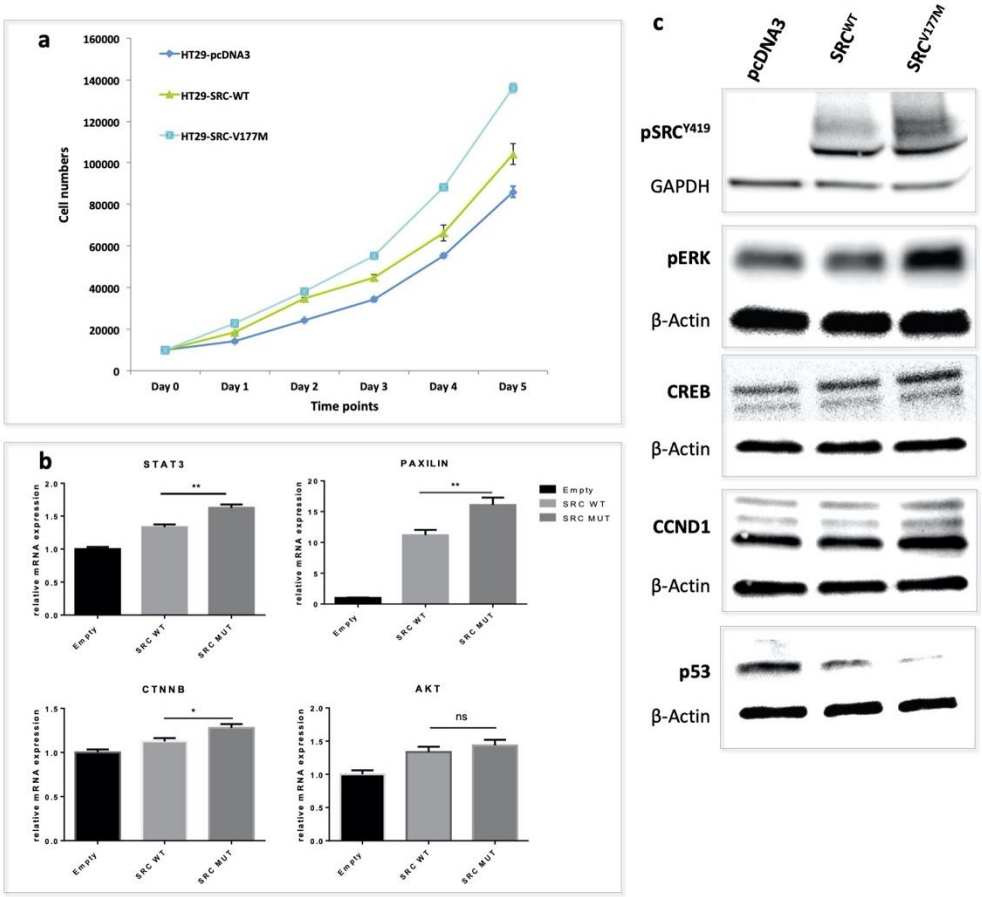


Figure 4. Impact of the *SRC*^{V177M} variant on cell proliferation and key components of SRC signaling pathways. **(a)** Cell proliferation assays show significantly increased cell numbers of HT29-*SRC*^{V177M} compared to HT29-*SRC*^{WT} cells and the control. **(b)** qPCR results represent significantly increased mRNA levels of *PXN, CTNNB* and *STAT3* in HT29-*SRC*^{V177M} cells. * - *p<0.05*; ** - *p<0.01*, ns – no significance. **(c)** Western blot results indicate enhanced protein expression of pSRC^{Y419}, pERK, CREB and CCND1 as well as decreased p53 protein in HT29-*SRC*^{V177M} cells.

3.4.2. Enhanced STAT3, CTNNB and PXN gene transcription induced by SRCV177M variant

In order to investigate the impact of the identified variant on pre-translational level, mRNA levels of potential target genes were quantified. Results of qPCR experiments showed significant upregulation of *CTNNB*, *STAT3* and *PXN* mRNA levels in HT29-*SRC*^{V177M} compared to HT29-*SRC*^{WT} cells (*CTNNB*: *p<0.05*; *STAT3*, *PXN: p<0.01)*, whereas no significant difference could be observed for *AKT* mRNA levels. Thus, our experiments propose the involvement of the mutated SRC protein in pre-translational regulation of *CTNNB*, *STAT3* and *PXN* genes being associated with cell proliferation, invasion and metastasis (Figure 4b).

3.4.3. SRC^{V177M} variant leads to increased SRC phosphorylation at Y419, a potential marker for SRC activity

In order to investigate the effect of the prioritized *SRC*^{V177M} variant on SRC protein conformation and intrinsic kinase activity *in vitro*, HT-29 cells were transfected with the mutated plasmid and checked for phospho-SRC (pSRC) protein levels. Phosphorylation at the tyrosine residue 530 (pSRC^{Y530}) has been reported to induce a closed SRC conformation due to intramolecular binding of the respective phosphotyrosine to the SH2-domain. On the other hand, full activation of SRC requires an open protein conformation enabling autophosphorylation at position 419 (pSRC^{Y419}) within the catalytic domain [36]. Western blot quantification of pSRC^{Y419} as a potential marker for activated

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 March 2021 doi:10.20944/preprints202103.0121.v1

SRC protein resulted in increased pSRC^{Y419} protein expression in HT29-*SRC*^{V177M} cells compared to HT29-*SRC*^{WT} cells. In this way, the *SRC*^{V177M} variant enhanced the auto-phosphorylation and activation of SRC protein by potentially disrupting the pY530 – SH2 domain interaction. Although HT-29 cells were shown to express SRC protein endogenously [37], the included control did not show detectable pSRC^{Y419} protein levels indicating the absence of the open and fully activated SRC protein conformation (Figure 4c).

3.4.4. SRC^{V177M} variant affects pERK, CREB, CCND1 and p53 protein expression

With the aim of further validating the variant-induced upregulation of SRC activity and investigating the respective impact on colorectal carcinogenesis, key components of known SRC signaling pathways were checked for altered protein expression. Western Blot results revealed enhanced protein expression of phospho-ERK (extracellular signal-regulated kinase; pERK), CREB (CAMP responsive element binding protein) and CCND1 (cyclin D1) in HT29-*SRC*^{V177M} compared to HT29-*SRC*^{WT} cells. On the other hand, the tumor suppressor protein p53 showed decreased protein levels in HT29-*SRC*^{WT} and, to an even greater extent, in HT29-*SRC*^{V177M} cells compared to the control HT29-pcDNA3 (Figure 4c).

4. Discussion

By performing WGS and integrative *in silico* analysis on a CRC affected family using our FCVPPv2, we were able to identify a novel germline variant in *SRC* gene (V177M) contributing to cancer predisposition. *SRC* is a commonly known proto-oncogene, the somatic mutations of which promote the development, progression and metastasis of various malignancies including colorectal, breast, prostate, ovarian and testicular cancers [38-40]. However, the present results suggest that the identified *SRC*^{V177M} variant may act as a germline CRC-predisposing variant. In contrast to numerous inactivating mutations in tumor suppressor genes, activating mutations contributing to familial cancer are rare and include the genes *RET, MET , KIT and ALK* [41]. All of these encode kinases, which are activated by the predisposing mutations to different extent, which may be the mode of action of the present kinase, SRC. Non-complete penetrance of cancer and the diversity of cancers in the family may be explained by the observed moderate effect of the *SRC*^{V177M} variant on CRC risk in the Polish population (OR 1.65). This suggests a polygenic mode of inheritance and additional mutations may be needed to express the cancer phenotype.

The oncogenic role of SRC has been elucidated on molecular basis, referring to cellular functions such as cell migration and invasion. One of the described underlying molecular mechanisms include the focal adhesion-associated adaptor protein PXN: Docking at the phosphorylated tyrosine residue pY397 of Focal adhesion kinase (FAK), SRC can form the active FAK/SRC complex which further phosphorylates and associates with PXN and p130cas. Respective PXN/p130cas phosphotyrosines may then recruit Crk protein, resulting in cellular processes such as actin reorganization, cell spreading and migration [42-44]. In HT-29 cells, SRC mediated increase of FAK, PXN and p130cas tyrosine phosphorylation and resulting cell migration enhancement has been induced by VEGFR-1 stimulation, implicating VEGF signaling upstream of the described molecular mechanisms[45].

In our experiments we showed that introduction of the *SRC*^{V177M} variant resulted not only in upregulated protein expression of pSRC^{Y419}, the fully activated SRC protein in open conformation (Figure 5a), but also in increased *PXN* mRNA levels. Thus, the mutated and activated SRC protein may affect *PXN* expression already at pre-translational level, potentially contributing to the described processes of cell migration. Besides the described upregulation of *PXN* potentially contributing to invasive and migratory cell behavior, we observed increased *CTNNB* mRNA levels as a result of the introduced variant.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 March 2021 doi:10.20944/preprints202103.0121.v1

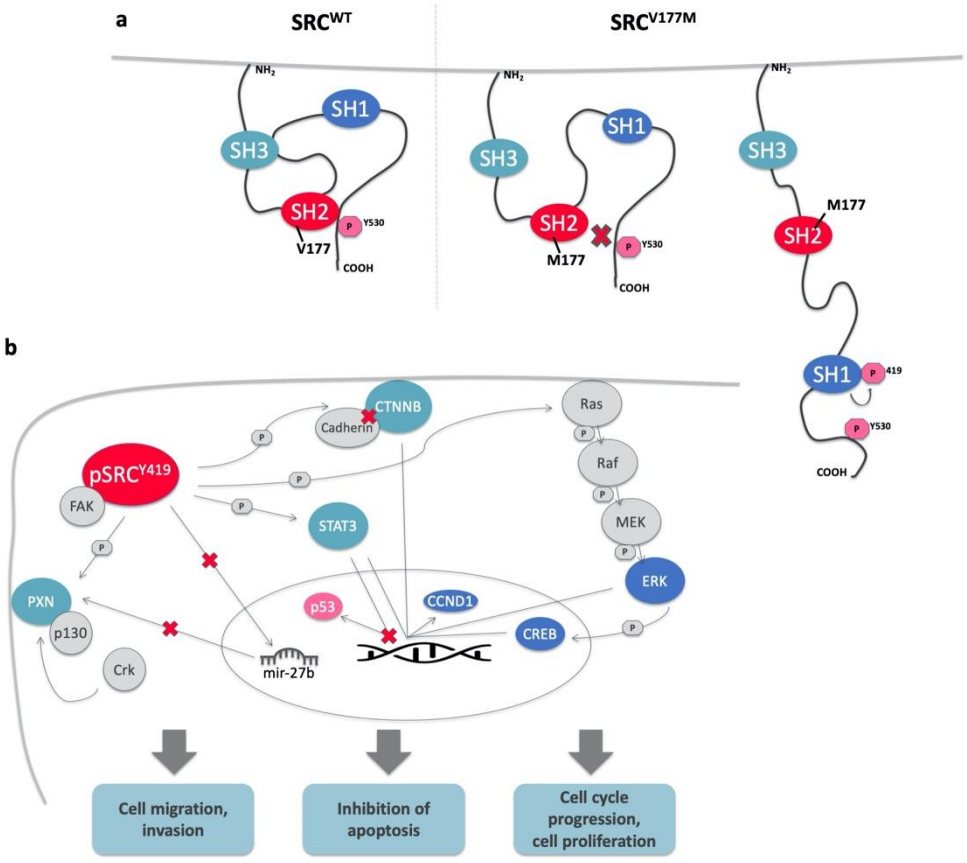


Figure 5. Molecular processes potentially induced by the *SRC*^{V177M} variant. **(a)** The *SRC*^{V177M} variant potentially induces disruption of the intramolecular pY530 – SH2 domain binding leading to autophosphorylation at Y419 and full activation of the SRC protein. **(b)** Activated pSRC^{Y419} enhances cell migration, invasion, proliferation as well as cell cycle progression and inhibits apoptosis due to the illustrated cancer related pathways. The demonstrated activation of *CCND1* gene expression is representative of additional target genes of CTNNB, STAT3, ERK and CREB, not illustrated in this figure. SRC target genes showing increased mRNA levels in HT29-*SRC*^{V177M} cells were colored in light blue, whereas SRC target proteins overexpressed in HT29-*SRC*^{V177M} cells were colored in dark blue. Observed suppression of protein expression due to the variant was represented by pink color. Arrows labeled with (P) are indicating phosphorylation, whereas arrows labeled with (X) are indicating an inhibiting impact.

We also showed that *SRC*^{V177M} upregulates *STAT3* at mRNA level. Although several studies have reported an increase in STAT3 transcriptional activity by SRC phosphorylation leading to gene expression of STAT3 target genes [46-48], little is known about the transcriptional regulation of *STAT3* itself, potentially involving SRC protein. Even though the exact underlying mechanisms remain to be elucidated, our results indicate that the *SRC*^{V177M} variant and thus active pSRC^{Y419} may increase *STAT3* gene expression and may contribute to CRC. Several STAT3 target genes are known to play an important role in cell proliferation and apoptosis, such as *CCND1* [49]. In this study, we observed increased CCND1 protein levels due to the *SRC*^{V177M} variant, which may lead to cell cycle progression and cell proliferation via the known SRC-STAT3-CCND1 association.

Additionally, STAT3 has been reported to mediate SRC-induced transcriptional inhibition of the tumor suppressor p53 [50]. Since we observed decreased p53 protein levels upon *SRC*^{V177M} variant introduction, these findings may also be explained by STAT3 as the mediating factor between activated pSRC^{Y419} and suppressed p53 expression, potentially resulting in inhibition of apoptosis. This may support the established SRC-STAT3-p53 pathway and reciprocally confirm the activating function of the studied *SRC*^{V177M} variant.

Based on the described molecular functions, the affected STAT3 downstream targets CCND1 and p53 could be responsible for the observed increase of viable cell numbers of

 doi:10.20944/preprints202103.0121.v1

HT29-*SRC*^{V177M} cells. Interestingly, overexpression of CCND1 may be traced back to further *SRC*^{V177M} downstream effectors contributing to cell proliferation: 1. CTNNB activating gene transcription of Wnt target genes including CCND1 [51] and 2. the MAPK/ERK pathway being required for CCND1 transcription and assembly with CDK4/6 [52]. Since we observed increased phosphorylation of ERK protein in HT29-*SRC*^{V177M} cells, our results indicate the potential activation of the proposed SRC-Ras-Raf-MEK-ERK1/2 pathway. Mediated by the Ras GTPase, SRC may induce the consecutive phosphorylation of the effector kinase Raf, MAP2K/MEK (Mitogen-activated protein kinase kinase) and MAPK/ERK, which is generally known to result in cell growth and proliferation [53]. Besides *CCND1* transcription, ERK further affects the regulation of gene expression by phosphorylating CREB [53], which we also reported as overexpressed in HT29-*SRC*^{V177M} cells.

The proposed association of SRC with activated PI3/AKT signaling, resulting in cellular processes such as cell growth, proliferation and migration, is considered to rely on increased phosphorylation and activation of AKT protein [54, 55]. Since our experiments investigated only *AKT* mRNA levels and did not reveal any differences between HT29-*SRC*^{V177M} and HT29-*SRC*^{WT} cells, our results indicate the independence of *AKT* gene expression regulation from the investigated mutation, not contradicting the current state of research.

Confirming the functional impact of the studied *SRC*^{V177M} variant on key components of the described cancer related pathways (PXN, Wnt, STAT3, MAPK/ERK signaling, Figure 5b), we aim to confirm the postulated involvement of upregulated SRC activity in colorectal carcinogenesis and further to implicate these molecular mechanisms in cancer development of the studied family.

Although the *SRC* variant was reported in only 6/125,748 exomes of the gnomAD database, it was found in 4 additional index cases among 1690 tested Polish CRC families and in 3 out of 1676 controls, implying that it may be a moderate risk allele, explaining the non-complete penetrance of cancer in the family and also the diverse pattern of cancers which may be due to additional mutation(s).

Even though post-translational modifications of the SRC protein have been widely studied as the underlying cause of high SRC activity in cancer, less is known about activating variation of the *SRC* gene in human CRC besides the somatic truncating mutation *Src*^{531} [56, 57]. By identifying a germline mutation of activating function, we were able to bring insight into the understanding of genetically determined upregulation of SRC activity in colorectal malignancies and to implement genetic *SRC* variation in familial CRC inheritance.

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1, Supplementary Table S1: Summary of family members WGS was performed on, including personal data and the consideration of being a carrier of the cancer-causing mutation, Supplementary Table S2: Alphabetical list of qPCR primers with respective forward and reverse sequences, Supplementary Table S3: Alphabetical list of primary and secondary antibodies with respective product details and dilution conditions, Supplementary Figure S1: Alignment of multiple SRC protein sequences.

Author Contributions: Conceptualization, K.H., A.F. and O.R.B.; methodology, D.S., B.M., A.S., D.D., J.L. and O.R.B.; software, A.K., N.P., M.S. and O.R.B.; validation, D.S., B.M., A.S., M.K., D.D., J.L. and O.R.B.; formal analysis, D.S., A.K., N.P. and O.R.B.; investigation, D.S., B.M., A.S., M.K., A.F. and O.R.B.; resources, D.D. and J.L.; data curation, D.S., A.K., N.P., M.S. and O.R.B.; writing—original draft preparation, D.S. and O.R.B.; writing—review and editing, D.S., A.S., K.H., A.F. and O.R.B.; visualization, D.S., B.M. and O.R.B.; supervision, K.H., A.F. and O.R.B.; project administration, K.H., A.F. and O.R.B.; funding acquisition, K.H. and A.F. All authors have read and agreed to the published version of the manuscript.

Funding: This article is based upon work from COSTAction CA17118, supported by COST (European Cooperation in Science and Technology) and Transcan ERA-NET funding from the German Federal Ministry of Education and Research (BMBF). KH was supported by the EU Horizon 2020 program grant No. 856620.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Bioethics Committee of the Pomeranian Medical Academy in Szczecin (No: BN-001/174/05, 25 October 2005).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study. Written informed consent has been obtained from the patient(s) to publish this paper.

Data Availability Statement: Unfortunately, for reasons of ethics and patient confidentiality, we are not able to provide the sequencing data into a public database. The data underlying the results presented in the study are available from the corresponding author or from Asta Försti (Email: a.foersti@kitz-heidelberg.de).

Acknowledgments: We thank the DKFZ Genomics and Proteomics Core Facility for Illumina Sequencing Services and Omics IT and Data Management Core Facility (ODCF), DKFZ for managing the NGS data.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Frank C, Fallah M, Ji J, Sundquist J, Hemminki K: **The population impact of familial cancer, a major cause of cancer**. *International journal of cancer* 2014, **134**:1899-1906.
2. Weren RD, Ligtenberg MJ, Kets CM, de Voer RM, Verwiel ET, Spruijt L, van Zelst-Stams WA, Jongmans MC, Gilissen C, Hehir-Kwa JY *et al*: **A germline homozygous mutation in the base-excision repair gene NTHL1 causes adenomatous polyposis and colorectal cancer**. *Nat Genet* 2015, **47**(6):668-671.
3. Kuiper RP, Hoogerbrugge N: **NTHL1 defines novel cancer syndrome**. *Oncotarget* 2015, **6**(33):34069-34070.
4. Yan HHN, Lai JCW, Ho SL, Leung WK, Law WL, Lee JFY, Chan AKW, Tsui WY, Chan ASY, Lee BCH *et al*: **RNF43 germline and somatic mutation in serrated neoplasia pathway and its association with BRAF mutation**. *Gut* 2017, **66**(9):1645-1656.
5. 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**(2):520-529.
6. Briggs S, Tomlinson I: **Germline and somatic polymerase epsilon and delta mutations define a new class of hypermutated colorectal and endometrial cancers**. *J Pathol* 2013, **230**(2):148-153.
7. Palles C, Cazier JB, Howarth KM, Domingo E, Jones AM, Broderick P, Kemp Z, Spain SL, Guarino E, Salguero I *et al*: **Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas**. *Nat Genet* 2013, **45**(2):136-144.
8. Valle L, de Voer RM, Goldberg Y, Sjursen W, Forsti A, Ruiz-Ponte C, Caldes T, Garre P, Olsen MF, Nordling M *et al*: **Update on genetic predisposition to colorectal cancer and polyposis**. *Molecular aspects of medicine* 2019, **69**:10-26.
9. Jasperson KW, Tuohy TM, Neklason DW, Burt RW: **Hereditary and familial colon cancer**. *Gastroenterology* 2010, **138**(6):2044-2058.
10. Lorans M, Dow E, Macrae FA, Winship IM, Buchanan DD: **Update on Hereditary Colorectal Cancer: Improving the Clinical Utility of Multigene Panel Testing**. *Clin Colorectal Cancer* 2018, **17**(2):e293-e305.

11. Bandapalli OR, Paramasivam N, Giangiobbe S, Kumar A, Benisch W, Engert A, Witzens-Harig M, Schlesner M, Hemminki K, Forsti A: **Whole genome sequencing reveals DICER1 as a candidate predisposing gene in familial Hodgkin lymphoma**. *International journal of cancer* 2018, **143**(8):2076-2078.
12. Kumar A, Bandapalli OR, Paramasivam N, Giangiobbe S, Diquigiovanni C, Bonora E, Eils R, Schlesner M, Hemminki K, Forsti A: **Familial Cancer Variant Prioritization Pipeline version 2 (FCVPPv2) applied to a papillary thyroid cancer family**. *Scientific reports* 2018, **8**(1):11635.
13. Srivastava A, Kumar A, Giangiobbe S, Bonora E, Hemminki K, Forsti A, Bandapalli OR: **Whole Genome Sequencing of Familial Non-Medullary Thyroid Cancer Identifies Germline Alterations in MAPK/ERK and PI3K/AKT Signaling Pathways**. *Biomolecules* 2019, **9**(10).
14. Lahiri DK, Schnabel B: **DNA isolation by a rapid method from human blood samples: effects of MgCl2, EDTA, storage time, and temperature on DNA yield and quality**. *Biochem Genet* 1993, **31**(7-8):321-328.
15. Li H, Durbin R: **Fast and accurate short read alignment with Burrows-Wheeler transform**. *Bioinformatics* 2009, **25**(14):1754-1760.
16. Li H: **A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data**. *Bioinformatics* 2011, **27**(21):2987-2993.
17. Rimmer A, Phan H, Mathieson I, Iqbal Z, Twigg SR, Consortium WGS, Wilkie AO, McVean G, Lunter G: **Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications**. *Nat Genet* 2014, **46**(8):912-918.
18. Wang K, Li M, Hakonarson H: **ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data**. *Nucleic Acids Res* 2010, **38**(16):e164.
19. Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA *et al*: **A global reference for human genetic variation**. *Nature* 2015, **526**(7571):68-74.
20. Smigielski EM, Sirotkin K, Ward M, Sherry ST: **dbSNP: a database of single nucleotide polymorphisms**. *Nucleic Acids Res* 2000, **28**(1):352-355.
21. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB *et al*: **Analysis of protein-coding genetic variation in 60,706 humans**. *Nature* 2016, **536**(7616):285-291.
22. Petrovski S, Wang Q, Heinzen EL, Allen AS, Goldstein DB: **Genic intolerance to functional variation and the interpretation of personal genomes**. *PLoS Genet* 2013, **9**(8):e1003709.
23. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J: **A general framework for estimating the relative pathogenicity of human genetic variants**. *Nat Genet* 2014, **46**(3):310-315.
24. Cooper GM, Stone EA, Asimenos G, Program NCS, Green ED, Batzoglou S, Sidow A: **Distribution and intensity of constraint in mammalian genomic sequence**. *Genome Res* 2005, **15**(7):901-913.
25. Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, Clawson H, Spieth J, Hillier LW, Richards S *et al*: **Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes**. *Genome Res* 2005, **15**(8):1034-1050.
26. Ward LD, Kellis M: **HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants**. *Nucleic Acids Res* 2012, **40**(Database issue):D930-934.
27. Liu X, Wu C, Li C, Boerwinkle E: **dbNSFP v3.0: A One-Stop Database of Functional Predictions and Annotations for Human Nonsynonymous and Splice-Site SNVs**. *Hum Mutat* 2016, **37**(3):235-241.
28. Karczewski K, Francioli L, Tiao G, Cummings B, Alföldi J, Wang Q, Collins R, Laricchia K, Ganna A, Birnbaum D *et al*: **Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes**. In.: bioRxiv; 2019.
29. Tamborero D, Rubio-Perez C, Deu-Pons J, Schroeder MP, Vivancos A, Rovira A, Tusquets I, Albanell J, Rodon J, Tabernero J *et al*: **Cancer Genome Interpreter annotates the biological and clinical relevance of tumor alterations**. *bioRxiv* 2017:140475.

30. Hecht M, Bromberg Y, Rost B: **Better prediction of functional effects for sequence variants**. *BMC Genomics* 2015, **16 Suppl 8**:S1.
31. Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, Stein TI, Nudel R, Lieder I, Mazor Y *et al*: **The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses**. *Curr Protoc Bioinformatics* 2016, **54**:1 30 31-31 30 33.
32. Turro E, Greene D, Wijgaerts A, Thys C, Lentaigne C, Bariana TK, Westbury SK, Kelly AM, Selleslag D, Stephens JC *et al*: **A dominant gain-of-function mutation in universal tyrosine kinase SRC causes thrombocytopenia, myelofibrosis, bleeding, and bone pathologies**. *Sci Transl Med* 2016, **8**(328):328ra330.
33. Hunt SE, McLaren W, Gil L, Thormann A, Schuilenburg H, Sheppard D, Parton A, Armean IM, Trevanion SJ, Flicek P *et al*: **Ensembl variation resources**. *Database (Oxford)* 2018, **2018**.
34. Marengere LE, Pawson T: **Structure and function of SH2 domains**. *J Cell Sci Suppl* 1994, **18**:97-104.
35. Waksman G, Kominos D, Robertson SC, Pant N, Baltimore D, Birge RB, Cowburn D, Hanafusa H, Mayer BJ, Overduin M *et al*: **Crystal structure of the phosphotyrosine recognition domain SH2 of v-src complexed with tyrosine-phosphorylated peptides**. *Nature* 1992, **358**(6388):646-653.
36. Gargalionis AN, Karamouzis MV, Papavassiliou AG: **The molecular rationale of Src inhibition in colorectal carcinomas**. *Int J Cancer* 2014, **134**(9):2019-2029.
37. Barraclough J, Hodgkinson C, Hogg A, Dive C, Welman A: **Increases in c-Yes expression level and activity promote motility but not proliferation of human colorectal carcinoma cells**. *Neoplasia* 2007, **9**(9):745-754.
38. Wiener JR, Windham TC, Estrella VC, Parikh NU, Thall PF, Deavers MT, Bast RC, Mills GB, Gallick GE: **Activated SRC protein tyrosine kinase is overexpressed in late-stage human ovarian cancers**. *Gynecol Oncol* 2003, **88**(1):73-79.
39. Wheeler DL, Iida M, Dunn EF: **The role of Src in solid tumors**. *Oncologist* 2009, **14**(7):667-678.
40. Leonetti E, Gesualdi L, Scheri KC, Dinicola S, Fattore L, Masiello MG, Cucina A, Mancini R, Bizzarri M, Ricci G *et al*: **c-Src Recruitment is Involved in c-MET-Mediated Malignant Behaviour of NT2D1 Non-Seminoma Cells**. *Int J Mol Sci* 2019, **20**(2).
41. Rahman N: **Realizing the promise of cancer predisposition genes**. *Nature* 2014, **505**:302-308.
42. Schaller MD, Parsons JT: **pp125FAK-dependent tyrosine phosphorylation of paxillin creates a high-affinity binding site for Crk**. *Mol Cell Biol* 1995, **15**(5):2635-2645.
43. Feller SM: **Crk family adaptors-signalling complex formation and biological roles**. *Oncogene* 2001, **20**(44):6348-6371.
44. Lamorte L, Rodrigues S, Sangwan V, Turner CE, Park M: **Crk associates with a multimolecular Paxillin/GIT2/beta-PIX complex and promotes Rac-dependent relocalization of Paxillin to focal contacts**. *Mol Biol Cell* 2003, **14**(7):2818-2831.
45. Lesslie DP, Summy JM, Parikh NU, Fan F, Trevino JG, Sawyer TK, Metcalf CA, Shakespeare WC, Hicklin DJ, Ellis LM *et al*: **Vascular endothelial growth factor receptor-1 mediates migration of human colorectal carcinoma cells by activation of Src family kinases**. *Br J Cancer* 2006, **94**(11):1710-1717.
46. Cao X, Tay A, Guy GR, Tan YH: **Activation and association of Stat3 with Src in v-Src-transformed cell lines**. *Mol Cell Biol* 1996, **16**(4):1595-1603.
47. Bowman T, Broome MA, Sinibaldi D, Wharton W, Pledger WJ, Sedivy JM, Irby R, Yeatman T, Courtneidge SA, Jove R: **Stat3-mediated Myc expression is required for Src transformation and PDGF-induced mitogenesis**. *Proc Natl Acad Sci U S A* 2001, **98**(13):7319-7324.
48. Yu CL, Meyer DJ, Campbell GS, Larner AC, Carter-Su C, Schwartz J, Jove R: **Enhanced DNA-binding activity of a Stat3-related protein in cells transformed by the Src oncoprotein**. *Science* 1995, **269**(5220):81-83.
49. Carpenter RL, Lo HW: **STAT3 Target Genes Relevant to Human Cancers**. *Cancers (Basel)* 2014, **6**(2):897-925.
50. Niu G, Wright KL, Ma Y, Wright GM, Huang M, Irby R, Briggs J, Karras J, Cress WD, Pardoll D *et al*: **Role of Stat3 in regulating p53 expression and function**. *Mol Cell Biol* 2005, **25**(17):7432-7440.
51. Oving IM, Clevers HC: **Molecular causes of colon cancer**. *Eur J Clin Invest* 2002, **32**(6):448-457.

52. Diehl JA: **Cycling to cancer with cyclin D1**. *Cancer Biol Ther* 2002, **1**(3):226-231.

53. Chen J, Elfiky A, Han M, Chen C, Saif MW: **The role of Src in colon cancer and its therapeutic implications**. *Clin Colorectal Cancer* 2014, **13**(1):5-13.

54. Chen R, Kim O, Yang J, Sato K, Eisenmann KM, McCarthy J, Chen H, Qiu Y: **Regulation of Akt/PKB activation by tyrosine phosphorylation**. *J Biol Chem* 2001, **276**(34):31858-31862.

55. Datta K, Bellacosa A, Chan TO, Tsichlis PN: **Akt is a direct target of the phosphatidylinositol 3-kinase. Activation by growth factors, v-src and v-Ha-ras, in Sf9 and mammalian cells**. *J Biol Chem* 1996, **271**(48):30835-30839.

56. Irby RB, Mao W, Coppola D, Kang J, Loubeau JM, Trudeau W, Karl R, Fujita DJ, Jove R, Yeatman TJ: **Activating SRC mutation in a subset of advanced human colon cancers**. *Nat Genet* 1999, **21**(2):187-190.

57. Irby RB, Yeatman TJ: **Role of Src expression and activation in human cancer**. *Oncogene* 2000, **19**(49):5636-5642.