Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Lee Y-R, Yehia L, Kishikawa T, et al. WWP1 gain-of-function inactivation of PTEN in cancer predisposition. N Engl J Med 2020;382:2103-16. DOI: 10.1056/NEJMoa1914919

Supplementary Appendix

WWP1 Gain-of-Function Inactivates PTEN to Drive Cancer Predisposition

Yu-Ru Lee, Lamis Yehia, Takahiro Kishikawa, Ying Ni, Brandie Leach, Jinfang Zhang, Nivedita Panch, Jing Liu, Wenyi Wei, Charis Eng & Pier Paolo Pandolfi

Table of Contents

Supplementary Methods

Patients for Validation Studies

Eligible patients met at least the relaxed International Cowden Consortium (ICC) operational diagnostic criteria (Table S1). For all eligible patients, we reviewed the Cleveland Clinic (CC) score,[1] a semi-quantitative score based on weighting clinical features, and that estimates the pretest probability of finding a germline *PTEN* mutation (http://www.lerner.ccf.org/gmi/ccscore). Since all individuals are *PTEN* wildtype, we used the CC score as a surrogate for phenotypic burden. Scoring criteria, i.e., the clinical phenotype variables, are evaluated by the Center for Personalized Genetic Healthcare (CPGH, Cleveland Clinic, Cleveland, OH) medical geneticists and genetic counselors based on medical records, or concurrently during physical exams. We selected 83 probands based on clinical manifestations, high phenotypic burden (CC score ≥10), and/or pathognomonic features such as Lhermitte-Duclos disease.

For further analysis of *WWP1* germline variants in an extended series of patients, based on our findings from the WES discovery series, we prioritised patients with oligopolyposis, a Bannayan-Riley-Ruvalcaba (BRRS) diagnosis, and those having a high CC score (CC score ≥15). For the oligopolyposis series (n=126), eligible patients had ≥5 gastrointestinal polyps, including at least 1 hamartomatous or hyperplastic/serrated polyp, and tested wildtype for genes known to be associated with a polyposis phenotype (*PTEN*, *BMPR1A*, *SMAD4*, *ENG*, *APC*, *STK11*).[2,3] For the BRRS series (n=123), classic phenotypic manifestations include macrocephaly in combination with intestinal hamartomatous polyposis, vascular malformations, lipomas, and genital lentiginosis.[4] In the absence of formal diagnostic criteria for BRRS, BRRS-like individuals included those meeting at least the relaxed ICC criteria, combined with

phenotypic enrichment of features overlapping with BRRS, namely macrocephaly, lipomatosis, hemangiomas, vascular malformations, intestinal polyposis, autism/developmental delay, and thyroid involvement (Hashimoto's thyroiditis or nodules). The high CC score series of patients (*n*=99) included individuals with high phenotypic burden (CC score ≥15) regardless of phenotype. Since the CC score is a predictor of germline *PTEN* mutation status, and these patients are wildtype for any *PTEN* alterations, the high CC score series is valuable for gene discovery, particularly of genes that could impact PTEN function.

We reviewed medical records, including pathology reports, for each research participant and extracted family history from clinical notes associated with cancer genetics and/or genetic counseling visits, where applicable and with the individuals' consent. Cleveland Clinic Institutional Review Board approval (protocol *PTEN*-8458) and written informed consents from all research participants were obtained for this study.

Whole-exome sequencing and bioinformatic analysis

Genomic DNA was extracted from patient peripheral blood leukocytes by the Genomic Medicine Biorepository (GMB) of the Cleveland Clinic Genomic Medicine Institute (GMI, Cleveland, OH) following standard procedures (https://www.lerner.ccf.org/gmi/gmb/). We subjected germline genomic DNA of eligible probands to whole-exome sequencing. Exome enrichment was performed with the TruSeq SBS v.3 or Nextera Rapid Capture Exome (Illumina Inc., San Diego, California), and subsequent 100 bp paired-end sequencing was performed with Illumina HiSeq 2000 or 2500 platforms. Sequencing was performed at an Illumina Sequencing Service

Center (Illumina Inc.), Personalis Inc. (Menlo Park, CA), and the Cleveland Clinic Genomics Core (Cleveland, OH).

Raw sequencing reads were mapped to the human reference haploid genome sequence (Genome Reference Consortium human genome build 37, hg19) using the Burrows-Wheeler Aligner (BWA version 0.6.1).[5] Indel realignment, base and quality score recalibrations, and removal of PCR duplicates from the resultant Binary Alignment Map (BAM) files were performed using the Genome Analysis Toolkit (GATK), Sequence Alignment/Map (SAMtools) and Picard.[6,7] Variant discovery and genotype calling of single nucleotide variations (SNVs) and small insertions and deletions (indels, <50 bp) were performed using the GATK Haplotype Caller.

Variant annotation and prioritization

Following alignment and variant calling, we annotated resultant variants using ANNOVAR.[8] We excluded synonymous variants and intronic variants that do not affect splicing. We prioritized variants with a minor allele frequency (MAF) ≤ 0.0005 (0.05%). To predict the potential impact of missense variants on protein function, we used a combination of SIFT, PolyPhen-2, and MutationTaster.[9-11] A variant is predicted to be deleterious if it is damaging according to at least two of the three computational prediction programs. All resultant variants were inspected through the Integrative Genomics Viewer (IGV) to exclude false positive calls.[12]

Sanger sequencing validation

Prioritized gene variants were validated by PCR-based region-specific mutation analysis using Sanger sequencing. Gene-specific primers encompassing exonic regions harboring the particular

variations were designed. Sequencing was performed in the forward and reverse directions at the Eurofins Genomics DNA sequencing facility (Louisville, KY). All primer sequences are listed in Table S10 and resultant chromatograms in Figs. S1, S2, and S4. We analyzed the chromatograms using the Mutation Surveyor DNA Variant Analysis Software (SoftGenetics, State College, PA) and mutations reported according to the Human Genome Variation Society (HGVS).

High resolution melting (HRM) analysis

Patient-derived germline genomic DNA is first diluted to 50 ng/µl using UltraPure™ DNase/RNase-free distilled water (Thermo Fisher Scientific, Waltham, MA). First, we perform polymerase chain reaction (PCR) in a 96-well plate format using a thermal cycler. Each reaction consists of 4 µl of the LightScanner Master Mix (BioFire Defense, Salt Lake City, Utah), 1 µl of each the forward primer and reverse primer diluted to 4 µM, 4 µl of distilled water, and 1 µl of 50 ng/µl genomic DNA. The reaction mixture is overlaid over 10-20 µl of mineral oil (Sigma Aldrich, St. Louis, MO) per well. PCR cycling conditions are 95°C for 2 min, followed by 37 cycles consisting of 95°C for 30 sec and the primer-specific annealing temperature for 30 sec, and finally 72°C for 1 min. Primer pairs and optimized annealing temperatures are listed in Table S10. Following PCR, we performed high resolution melting (HRM) curve analysis using the LightScanner instrument (Idaho Technology, Salt Lake City, UT). Suspected variants were validated using Sanger sequencing, as described above. We utilized the cBioPortal MutationMapper tool (https://www.cbioportal.org/mutation_mapper) to depict the identified *WWP1* germline variants relative to the WWP1 protein structure.[13]

The Cancer Genome Atlas (TCGA) pan-cancer germline variant analysis

TCGA dataset was utilized as a representative of apparently sporadic cancers. We obtained TCGA germline genomic high-confidence variant calls from the Institute for Systems Biology Cancer Genomics Cloud (ISB-CGC, https://isb-cgc.appspot.com).[14] Clinical data were downloaded from the Genomic Data Commons (GDC, https://portal.gdc.cancer.gov). We analyzed 10,389 samples from 33 cancer types (Table S3). Similar to CS, CS-like, and BRRS patient samples, we annotated germline variants from TCGA using ANNOVAR[8] and implemented identical variant prioritization criteria. We prioritized exonic and splicing variants that passed quality control filters. MAF $\leq$ 0.0014 or 0.1%, representing the allele frequency of the most common variant we identified in our patients, was used as a cut-off for filtration. To predict the potential impact of missense variants on protein function, we used a combination of SIFT, PolyPhen-2, and MutationTaster.[9-11] A variant is predicted to be deleterious if it is damaging according to at least two of the three computational prediction programs. We utilized cBioPortal MutationMapper to depict the identified *WWP1* germline variants relative to the WWP1 protein structure (https://www.cbioportal.org/mutation_mapper).[13]

ExAC and gnomAD data analysis

Exome Aggregation Consortium (ExAC) germline genomic data excluding TCGA (non-TCGA ExAC) were downloaded from the ExAC browser (http://exac.broadinstitute.org/downloads), last accessed on August 16, 2019. This dataset includes 53,105 population controls without a reported cancer diagnosis. We also downloaded *WWP1* germline variants from the non-cancer whole-exome sequencing (WES) dataset in the Genome Aggregation Database (gnomAD v2.1.1, https://gnomad.broadinstitute.org/gene/ENSG00000123124?dataset=gnomad_r2_1_non_cancer), last accessed on October 25, 2019. The gnomAD dataset comprises of 141,456 unrelated

individuals (125,748 with WES data). We annotated germline variants from non-TCGA ExAC and non-cancer gnomAD using ANNOVAR[8] and implemented identical variant prioritization criteria as described above for TCGA.

Gene variant enrichment analysis in TCGA and non-TCGA ExAC

We investigated the frequencies of germline variants in classical genes known to be associated with cancer predisposition syndromes. These include *PTEN* (OMIM 601728), *TP53* (OMIM 191170), *BMPR1A* (OMIM 601299), and *STK11* (OMIM 602216). We annotated all variants using ANNOVAR.[8] All variants reported as pathogenic or likely pathogenic in ClinVar were retained (https://www.ncbi.nlm.nih.gov/clinvar/). For all remaining variants, we implemented identical variant prioritization criteria as described above for TCGA and non-TCGA ExAC. Of note, this strategy is likely to overestimate the frequency of predicted deleterious germline variants in the known cancer predisposition genes, conservatively implemented to have identical filtration criteria as *WWP1*.

Statistical analyses

OpenEpi software (http://openepi.com/Menu/OE_Menu.htm) was used to calculate odds ratios (OR) for *WWP1* variant enrichment. For analyses between different population groups, 2x2 tables were used to calculate OR. The 95% confidence intervals (CI) and corresponding P values were calculated using the mid-P exact test. OR > 1.0 implies that the incidence rate is greater for the population of interest compared with the standard population. P values of < 0.05 were considered to be statistically significant.

Cell culture, cell transfection and establishment of stable cell lines

All human cell lines were obtained from the American Type Culture Collection (ATCC; http://www.atcc.org) and checked for mycoplasma using the MycoAlert Mycoplasma Detection Kit (Lonza). 293T, HCT116, DLD-1 and mouse embryonic fibroblast (MEF) cells were maintained in DMEM supplemented with 10% fetal bovine serum, 2 mM glutamine, 100 U/ml penicillin and streptomycin (Invitrogen). Transfections were performed using Lipofectamine 2000 reagent (Invitrogen) according to the manufacturer's instruction. In brief, $5x10^5$ cells were transfected with 5 μg of DNA plasmids in a 6-well dish. Cells were recovered in the completed media for 12 hours and then harvested at the indicated times. Stable cell lines were generated by lentivirus transduction.

Xenotransplantation

For assaying tumor growth in the xenograft model, 7-week-old male *FOXN*^{nu} nude mice housed in specific pathogen-free environments were injected s.c. with either 1.5 x 10^6 HCT116 or 1.0 x 10^6 DLD-1 derivatives mixed with RPMI medium and Matrigel (vol/vol, 1:1). The mice were implanted with tumors, on both sides, with different genotypes. The care and treatment of animals were approved by the Beth Israel Deaconess Medical Center IACUC Committee on Animal Research.

Plasmids, reagents and antibodies

pLenti-HA-WWP1 WT-Puro, Flag-WWP1, and Myc-WWP1 were gifts from Dr. Wei's lab. All mutant constructs of WWP1 were generated using a QuikChange Lightning Site-Directed Mutagenesis (Agilent Technologies). All mutations were confirmed by Sanger sequencing.

Lipofectamine 2000, DMEM, Opti-MEM reduced serum media and fetal bovine serum (FBS) were purchased from Invitrogen. Anti-Flag-M2 affinity gel and puromycin were purchased from Sigma Aldrich. Polybrene was purchased from Santa Cruz Biotechnology, Inc. For Western blotting: anti-PTEN (9559), anti-EGFR (4267), anti-Ubiquitin (3936), anti-Phospho-AKT (pSer473, 9271), anti-AKT (pan AKT, 4685), anti-pS6 (2211), anti-S6 (2217) antibodies were all purchased from Cell Signaling Technology; Mouse anti-PTEN antibody (6H2.1) was purchased from Cascade BioScience; anti-WWP1 (human) (H00011059-M01) for Western blotting and immunoprecipitation was purchased from Novus Biologicals; anti-WWP1 (A302-950) was purchased from Bethyl; anti-Actin (A3853) and anti-Flag-M2 were purchased from Sigma Aldrich; anti-HSP90 was purchased from BD Biosciences (610419).

Lentivirus production and infection

To generate recombinant lentivirus, 293T cells were co-transfected with VSVG, PMDL, REV, and indicated lentivirus-based constructs. The virus-containing supernatant was harvested. For infection, the viral stock was supplemented with 10 μg/ml of polybrene and the infected cells were selected by 2 μg/ml of puromycin for at least two days.

***In vivo* ubiquitination assay**

To analyze *in vivo* ubiquitination of PTEN, cells were transfected with various constructs, together with His-Ubiquitin and Myc-PTEN. Cells were lysed by buffer A (6 M guanidine-HCl, Na_2HPO_4/NaH_2PO_4 [pH 8.0], and 10 mM imidazole), and lysates were incubated with Ni-NTA agarose for 1.5 hours at 4°C. The beads were washed once with buffer A, twice with buffer A/TI (1 vol buffer A: 3 vol buffer TI [25 mM Tris-HCl, pH 6.8, and 20 mM imidazole]), and three

times with buffer TI, and then analyzed by Western blot. In all experiments, an equal amount of His-Ubiquitin expression was verified by Western blot analysis.

Western blotting and immunoprecipitation

For Western blotting, cells were lysed in RIPA buffer (Boston BioProducts) supplemented with protease (Roche) and phosphatase (Roche) inhibitors. Proteins were separated on NuPAGE 4-12% Bis-Tris gradient gels (Invitrogen), transferred to polyvinylidene difluoride (PVDF) membranes (Immobilon P, Millipore) and the blots were probed with the indicated antibodies. For immunoprecipitation, 293T cells were transfected with the indicated expression vectors by using Lipofectamine 2000 (Life Technologies). 24 hours after transfection, cells were lysed in RIPA buffer with protease (Roche) and phosphatase (Roche) inhibitors. 500 μg of total lysates were pre-cleared for 30 minutes at 4°C, and then immunoprecipitated with anti-PTEN (Cell Signaling Technology 9559, 1:500) antibody overnight at 4°C. The Protein-A or Protein-G sepharose beads (GE Healthcare) were then added and incubated for another 2 hours. The immunoprecipitates were washed with RIPA buffer three times. In denaturing conditions, standard Laemmli-Buffer with 5% final concentration of β-mercaptoethanol was added to the samples, which were then boiled and separated on NuPAGE 4-12% Bis-Tris gradient gels (Invitrogen).

Native gel analysis

To determine if PTEN dimerization occurred, cells were lysed in lysis buffer containing 20 mM Tris-HCl pH7.5, 150 mM NaCl, 1% NP40, 1 mM EDTA, 1 mM protease (Roche) and phosphatase inhibitor (Roche) for further immunoprecipitation. For the native elution, pre-chilled

0.1 M glycine pH 2.5 was used to elute immunocomplexes for 10 minutes at 4°C, further neutralized with 1 M Tris-HCl pH 8.0. Tris-Glycine-Native Sample buffer (Boston BioProducts, BP-120) was added and samples were immediately run on NativePAGE Gel (Cat. No. BN1004BOX, Thermo Fisher Scientific).

Soft agar colony formation assay

In 6-well plates, 2 ml bottom layer medium (DMEM and 10% FBS with 0.6% agarose) was added to each well and cooled down for 30 min at room temperature. Then 2 ml top layer medium (DMEM and 10% FBS with 0.3% agarose) mixed with cells was added to each well and cooled down for 10 min at 4°C. The 5 x 10^4 of HCT116 or 1.5 x 10^4 DLD-1 cells were cultured at 37°C for 21 days. Then 1 ml 0.005% Crystal Violet was added to each well for 1h at room temperature. Colonies were counted with ImageJ software. Each experiment was performed in triplicate.

Cellular fractionation

Membrane versus cytosolic fractionation of HCT116 or DLD-1 cells transduced with the indicated constructs was performed using the ProteoExtract Native Membrane Protein Extraction Kit (Calbiochem), and according to the manufacture's procedures.

Establishment of CRISPR knock-in cells

In order to generate *WWP1*$^{K740N/+}$ mutant cells, Alt-R CRISPR-Cas9 System was performed in accordance with manufacturer's protocol with slight modification (IDT). 125 pmol of crRNA and tracrRNA (IDT) were incubated with 100 pmol of Cas9 enzyme protein for 5 min at room

temperature to make RNP complex. Subsequently, RNA complex and 100 pmol of ssODN were reverse transcribed into 4.0 x 10^5 of trypsinized HCT116 cells with Lipofectamine CRISPRMAX Transfection Reagent (Thermo Fisher). Cells were cultured in 12 well plates for 2 days and reseeded into 96-well plates for single cell cloning. After 2 weeks, genomic DNA was extracted from each clone and subjected to PCR amplification. PCR products were digested with restriction enzyme MboI (New England Bio Labs) for 30min at 37 °C and the intensity of cut bands were separated by gel electrophoresis to screen for positive clones that showed homologous recombination. Genomic DNA isolated from positive clones were subjected to Sanger sequencing for confirmation. The sequences of guide RNA and ssODN were designed using Benchling CRISPR designing tool (https://www.benchling.com/). The sequences used:

WWP1 K740N crRNA: UCACAUGACCUGAAGUUGGGGUUUUAGAGCUAUGCU

WWP1 K740N ssODN:

ATATTTATTTCACCCAGAGATAACAACATTGAAGAATGTGGCTTAGAAATGTAC

TTTTCTGTTGACATGGAGATTTTGGGAAAAGTTACTTCACATGATCTGAACTTGG

GAGGTTCCAATATTCTGGTGACTGAGGAGAACAAAGATGAATATATTGGGTAA

GGTGATATACCTTATTAAGCTTAATTTCTAGAACACTT

WWP1 KI PCR Fw: 5'- TGCAGGCACTATTTCATGGA -3'

WWP1 KI PCR Rv: 5'- GGCAGTGGGAAAATGAAGCA -3'.

Generation of *WWP1* truncation mutations

To generate Flag/Myc tagged vectors which express N-terminal or C-terminal fraction of WWP1 protein (Flag-WWP1-C2+WW, Flag-WWP1-C2, and Myc-WWP1-HECT), pFlag-WWP1 and pMyc-WWP1 vectors were subjected to PCR amplification with divergent primers followed by

circularization with ligase in accordance with the manufacturer's protocol of Q5 Site-Directed Mutagenesis kit (New England). Deletion of WW2-WW3 linker and substitution of K740N mutant were generated with the same kit and further sequenced by Sanger sequencing. The primer sequences used:

Flag-WWP1-C2+WW: Fw: 5'- TGAGTCGACTCTAGAGGATC -3',

Rv: 5'- TAATGCCATAATCTGTTGG -3'

Flag-WWP1-C2: Fw: 5'-GAGGGATTTGGACAAGAA-3',

Rv: 5'-TTCAGATTCCAATTCTGC-3',

Myc-WWP1-HECT: Fw: 5'-ATCAATGTGTCCCGG-3',

Rv: 5'-TGGTGAAGCAGTGGC-3'

Flag-WWP1-C2+WW ΔLinker: Fw: 5'-TCGGCTTCAATGTTAGCTGCAGAAAATG-3',

Rv: 5'-AGGCCGCTGCCACGTTGT-3'

Myc-WWP1-HECT K740N: Fw: 5'-ATGACCTGAACTTGGGAGGTTC-3',

Rv: 5'-GTGAAGTAACTTTTCCCAAAATC-3'

Establishment of WWP1 mutant mouse embryonic fibroblast (MEF) cells

All experimental procedures were approved by the Institutional Animal Care and Use Committee (IACUC, RN150D) at the Beth Israel Deaconess Medical Center (protocol no. 076-2017). In order to generate *WWP1*$^{K736N/+}$ knock-in mice, mutant allele was incorporated with Easi-CRISPR system in the Beth Israel Deaconess Transgenic Facility. The sequences of guide RNA and ssODN were designed using Benchling CRISPR designing tool. One-cell stage fertilized mouse embryos were microinjected with 200 ng/μl Cas9 protein (PNA bio Company), 100 ng/μl of sgRNAs (synthesized by PNA bio Company), and 50 ng/μl of ssDNA (synthesized by IDT),

subsequently transferred into pseudopregnant C57BL6/J6 female mice. Genomic DNA isolated from F0 newborn tails was subjected to PCR followed by digestion with StyI (New England Bio Labs) which specifically digests the mutant allele. The positive founders were further validated with sanger sequenceing. *Wwp1*$^{K736N/+}$ and *Wwp1*$^{+/+}$ MEF cells were established from the embryos of wildtype C57BL6/J female (The Jackson Laboratory) which was crossed with *Wwp1*$^{K736N/+}$ male mouse.

The sequences used:

Wwp1 K736N sgRNA #1: ACUUCACAUGAUUUAAAGUU

Wwp1 K736N sgRNA #2: UAUAACGAUCAAUAGAUGUG

Wwp1 K736N ssODN:

CCAGAGATAACAATATTGAAGAATGTGGCTTAGAAATGTACTTTTCTGTAGACA

TGGAGATTCTGGGAAAAGTTACTTCACATGATTTA**AACCTT**GGAGGTTCCAATA

TCCTGGTGACGGAAGAAAACAAAGATGAATATATTGGGTAGGTTGATAGAGCT

GATTGAATTCATTCTGTAGAAGGTATGGGATATGATTAGAAATATCCTATATTT

ACACTCTTAACCTTAGTAGTTTTTATGAAACACATGGAAAACTCTCGTCAGGTT

CTTTCTCCGCTGGCTCAGGTCCTTGTGCTCTGTGCTGATCCAGATCCTGCTCGAG

TTTGCTCGAACAATGCATATCGTTAGTCACTTTCCCATGTGCATTCGCAATAGGC

AGTCTCGCTTCTGACCAAGTCAGAAGCCATGACAACATGCTGCTGAAAGCGATT

TCCAGTCTTCAAAATCTCAGAAAGAAAAGTTGTATGTTCTCCTAGCCTATTACA

CATATACATATACAATGCAATGTGTGTGTGTGTGTAAAAGTGCACTTCGGGATA

GTATAACCTGAAGAAGAGAACAGGAAAGAGTAACTTATCAGAGATGATAGTGA

AAGACTGAATGAATGTAGGTAACTGAAACTTGATTCATAAAAGAGGATATAAA

GTATATTTTTTCTAGTTTTAAATCTGCCATTTTTTGGTGATGGATAAATGCATGC

AAATATTCAATAGTGAAAATCCAGTGTGAAGCTTCACAGCTTCAGAATGTTTAT

AACGATCAATAGATGTGGAGTCTGGTTAATTTTGGTGTATGAATTTTTTATTTTT

AAGTTCTTTATGATAAAAATTTAACGTAAAGTATTTTGCAAAGGCGATGTAGCC

TCCTCG G (the bold part indicates mutant nucleotides that substitute K736 to N736, and concurrently generate a new StyI recognition site.)

Wwp1 K736N PCR Fw: 5'- CCATTCTACAAGCGTATGCT -3'

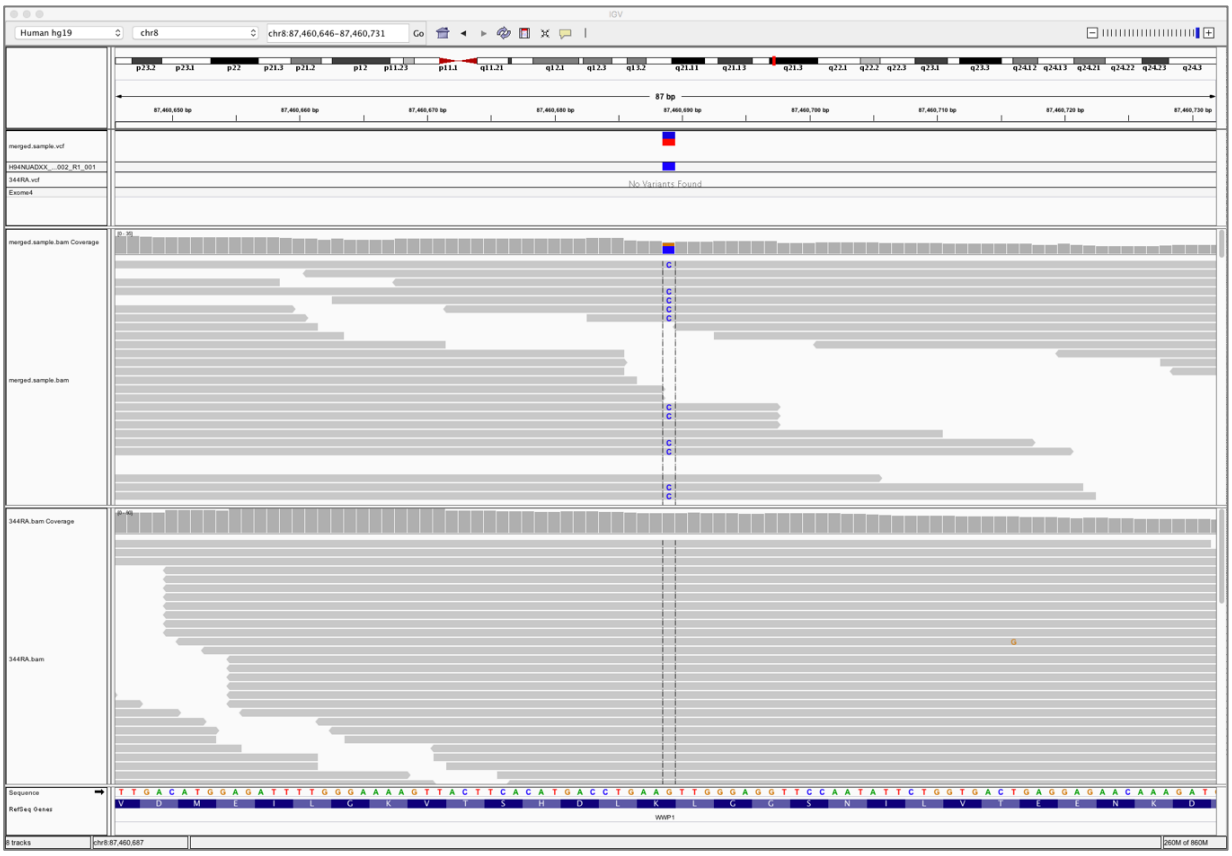
Wwp1 K736N PCR Rv: 5'- GCATGCATTTATCCATCACC -3'

Seqence primer for genotyping: 5'- ACTGGAAATCGCTTTCAGCAG -3'.

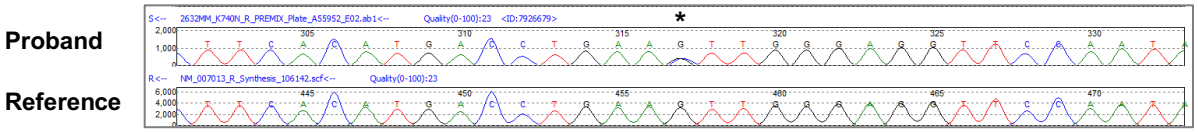
Supplementary Figures

Figure S1. Identification of *WWP1* c.2220G>C, p.K740N variant in a CS-like proband.

A



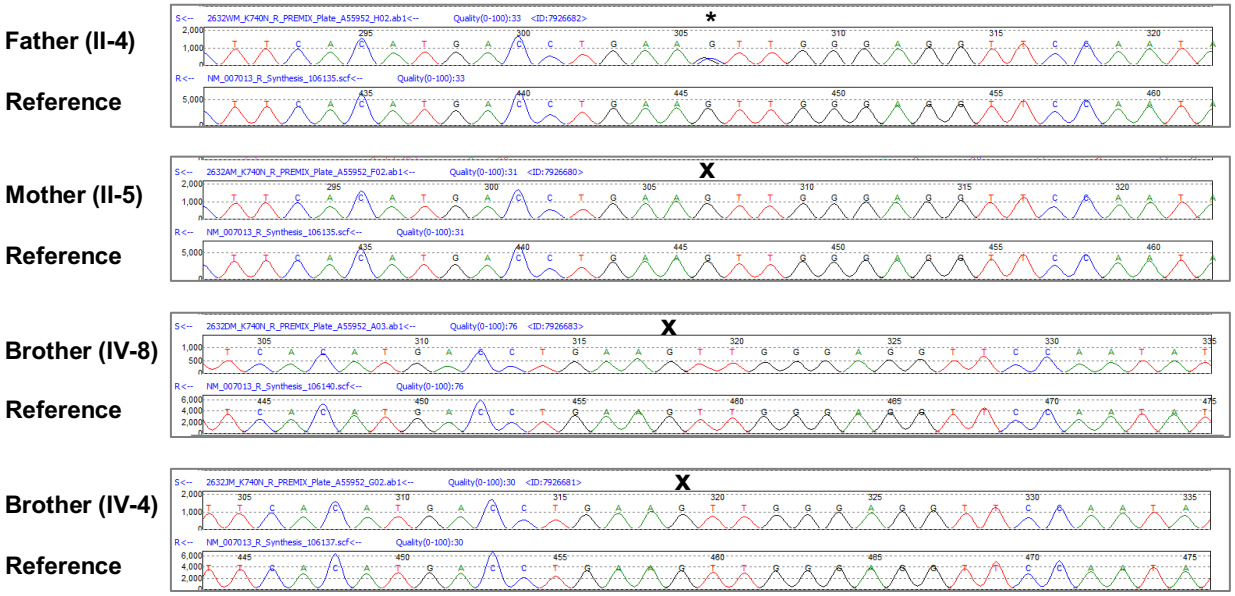
B



(A) Identification of *WWP1* c.2220G>C, p.K740N variant from whole-exome sequencing (WES) data of a CS-like proband (CCF02632-01-001, IV-3 on the pedigree). IGV visualization of the variant (top panel). Lower panel shows a wildtype CS proband as a negative control.

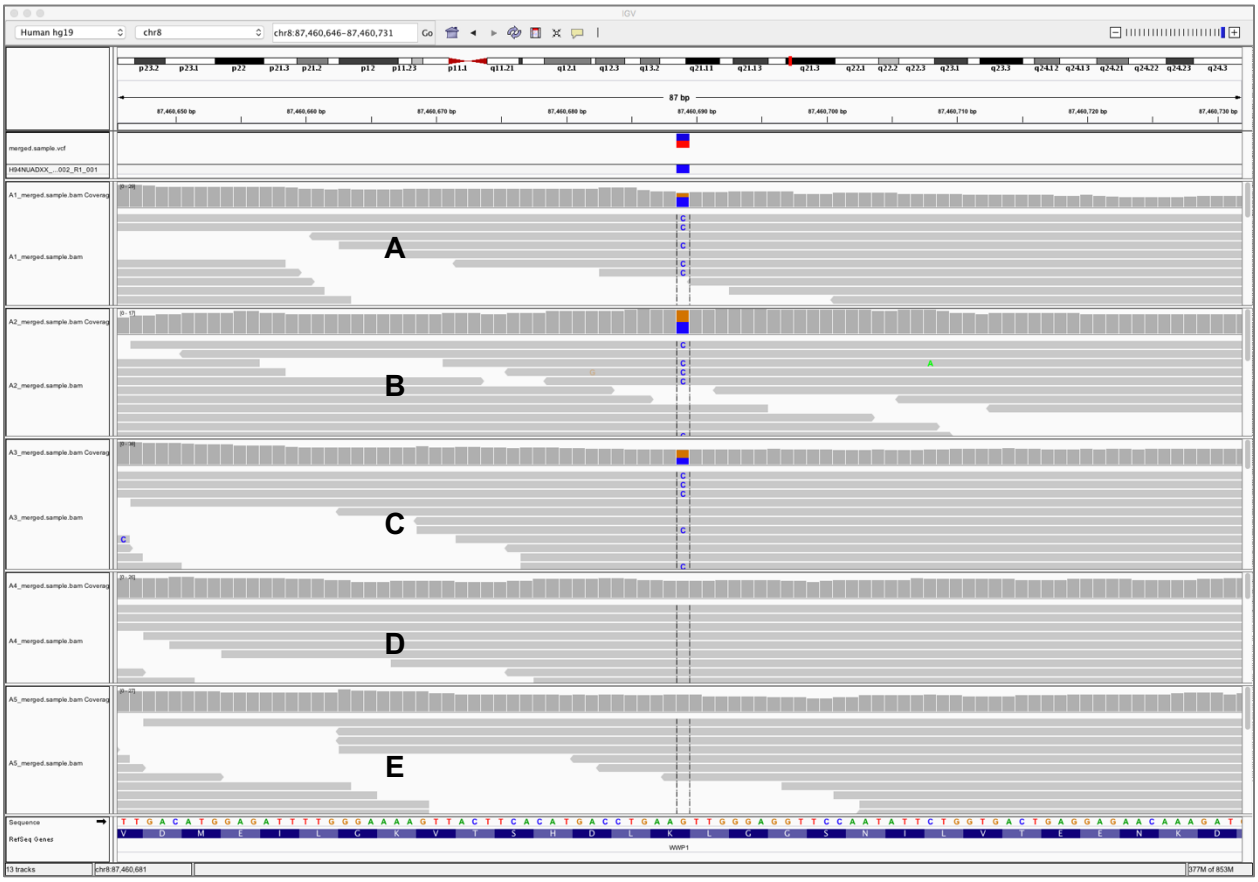
(B) Sanger sequencing validation of *WWP1* c.2220G>C, p.K740N (*). Upper sequence is derived from CCF02632-01-001. Lower sequence is the wildtype reference chromatogram.

Figure S2. Scanning *WWP1* p.K740N variant in family members of proband CCF02632-01-001.



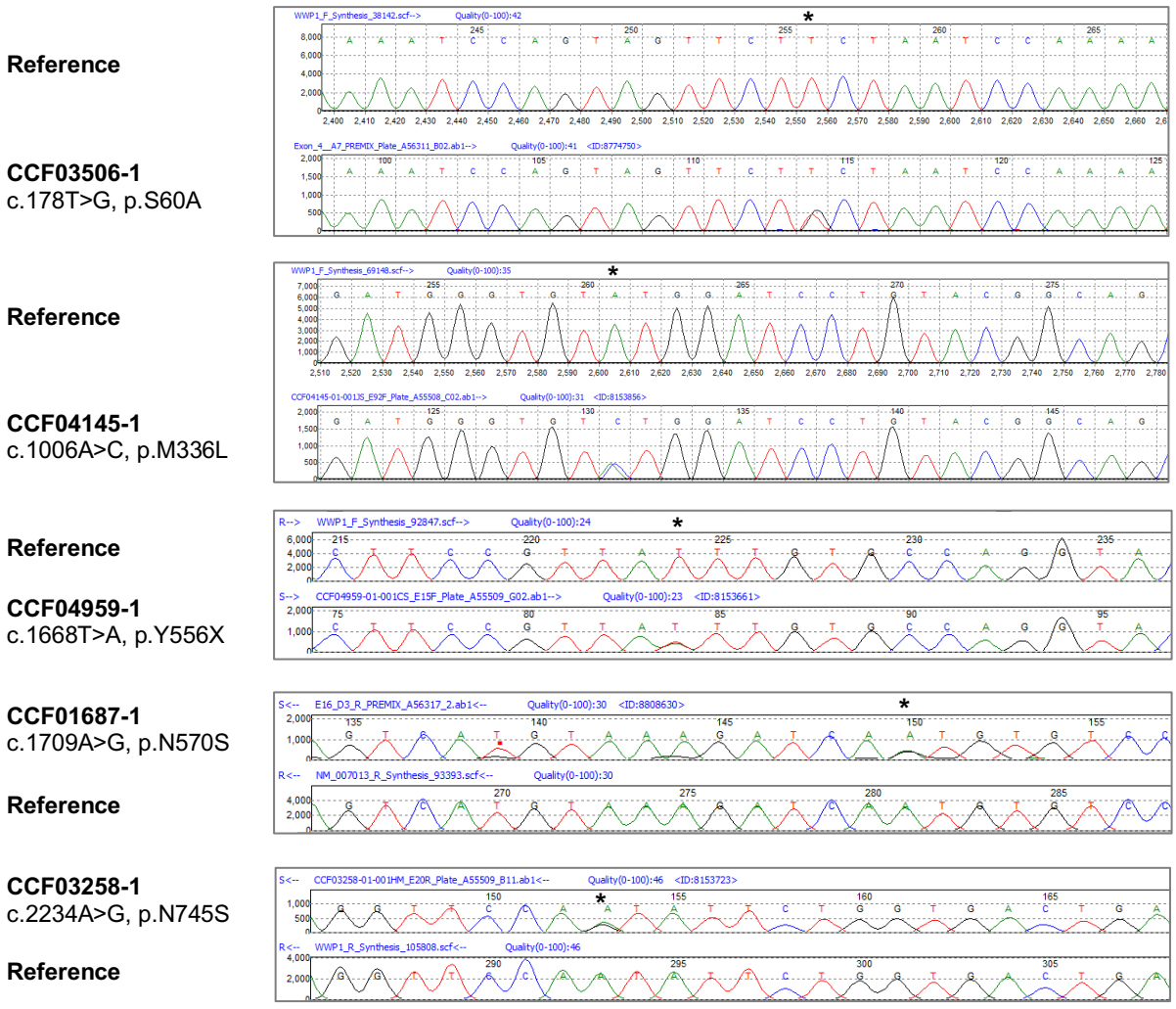
Sanger sequencing validation of *WWP1* c.2220G>C, p.K740N heterozygous variant in family members of proband CCF02632-01-001. The father (II-4) harbors the same *WWP1* germline variant (*), whereas the mother (II-5) and two brothers (IV-4 and IV-8) are wildtype (x). Lower sequence in each chromatogram refers to the wildtype reference sequence.

Figure S3. IGV visualization of *WWP1* c.2220G>C, p.K740N from WES data.



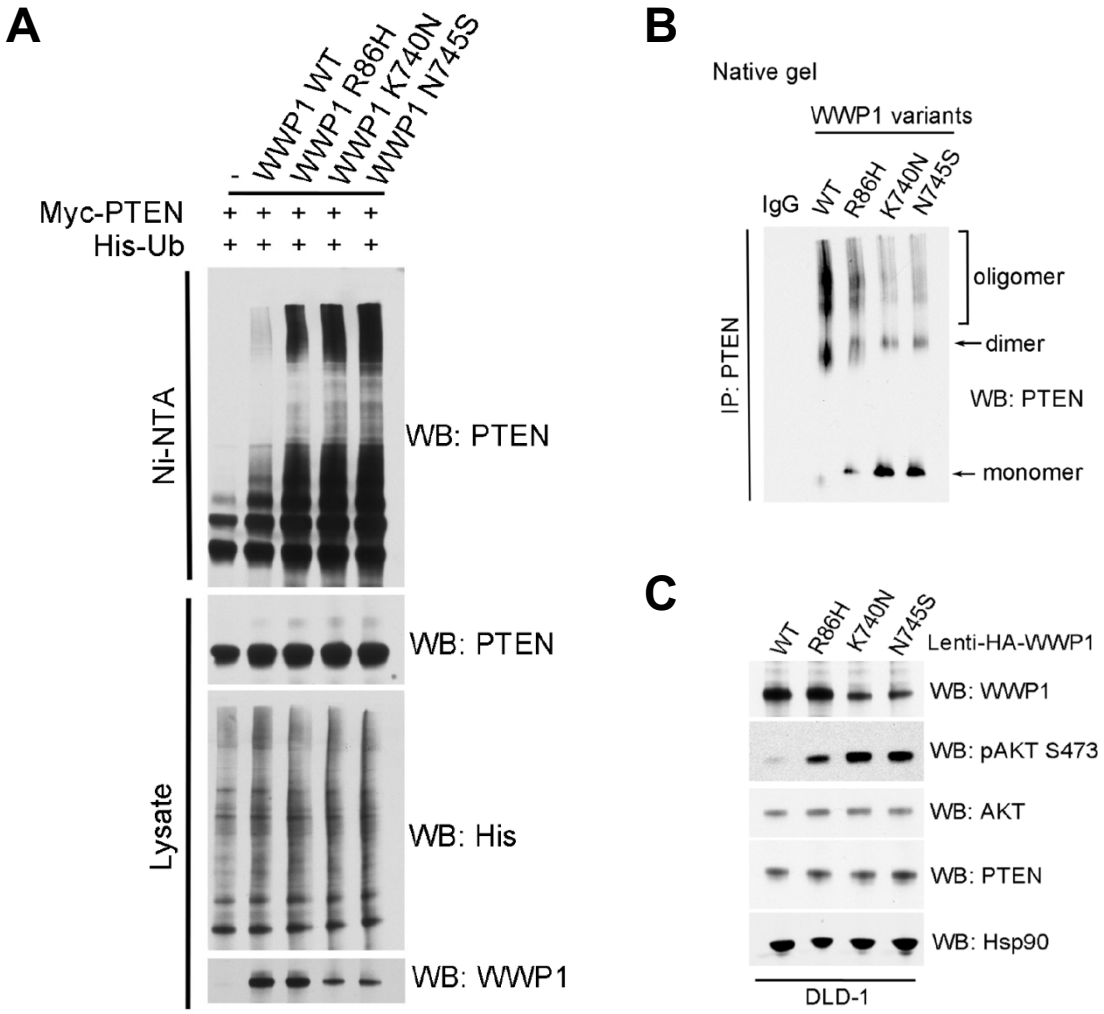
IGV visualization of the *WWP1* c.2220G>C, p.K740N variant from whole-exome sequencing (WES) data of the proband (A, IV-3 on the pedigree, variant positive), brother (B, IV-6 on the pedigree, variant positive), father (C, II-4 on the pedigree, variant positive), mother (D, II-5 on the pedigree, wildtype), and another brother (E , IV-4 on the pedigree, wildtype). Note that sibling IV-6 from the pedigree did not have DNA available for Sanger sequencing validation. The above WES data (B) shows that he is a carrier of the *WWP1* c.2220G>C, p.K740N variant.

Figure S4. Sanger sequencing of *WWP1* variants in unrelated probands with oligopolyposis.



Sanger sequencing validation of heterozygous *WWP1* variants (*) identified in five unrelated probands with oligopolyposis as predominant phenotype. Lower sequence in each chromatogram refers to the wildtype reference sequence.

Figure S5. *WWP1* gain-of-function mutations increase PTEN polyubiquitination, decrease PTEN dimerization/oligomerization resulting in subsequent AKT activation.

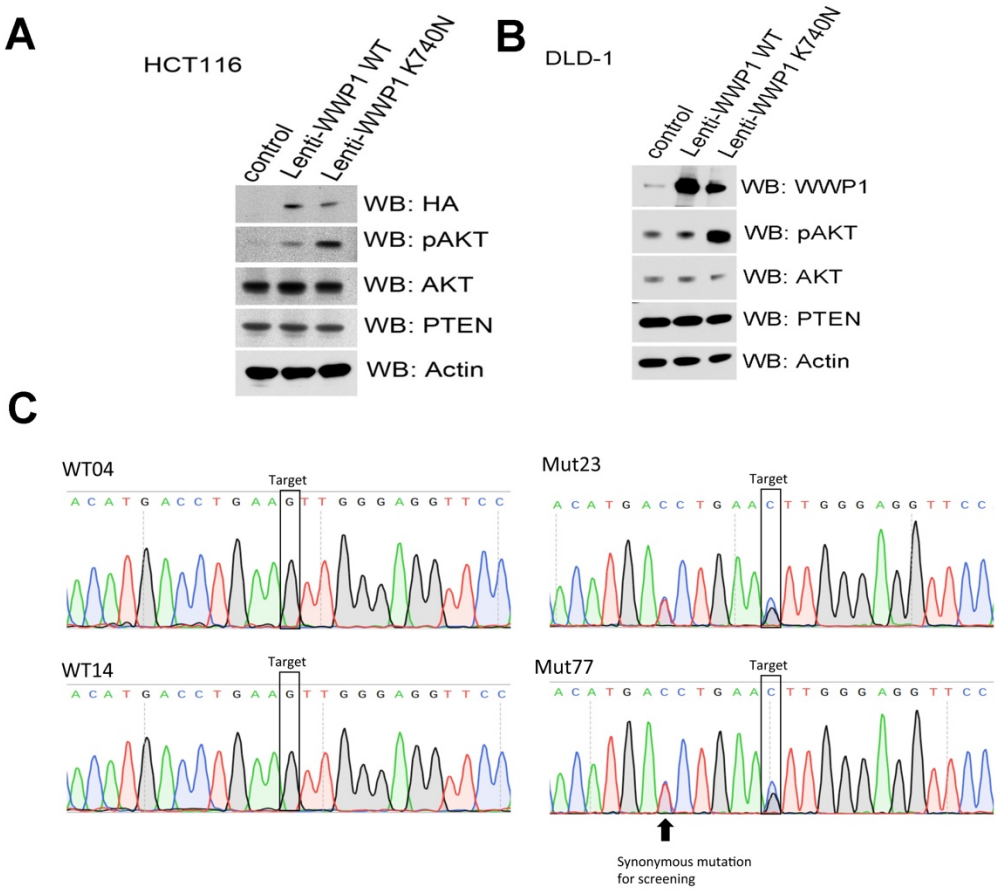


(**A**) Effects of the indicated WWP1 germline variants on WWP1-mediated PTEN poly-ubiquitination. HEK293T cells were transfected with the indicated constructs, and PTEN ubiquitination was analyzed. The ubiquitinated proteins were pulled down under denaturing conditions by nickel-nitrilotriacetic acid (Ni-NTA) agarose and analyzed by Western blot.

(**B**) Evaluation of PTEN dimerization and oligomerization in HCT116 cells with stable expression of the indicated constructs by native gel electrophoresis. Total lysates from cells with stable expression of the indicated constructs were immunoprecipitated with a rabbit anti-PTEN antibody, and then the immunocomplexes were natively eluted from the beads. The eluted samples were immediately run on the native gel.

(**C**) Analysis of AKT activation in DLD-1 cells with stable expression of the indicated constructs. Total lysates were resolved by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and then probed with the indicated antibodies.

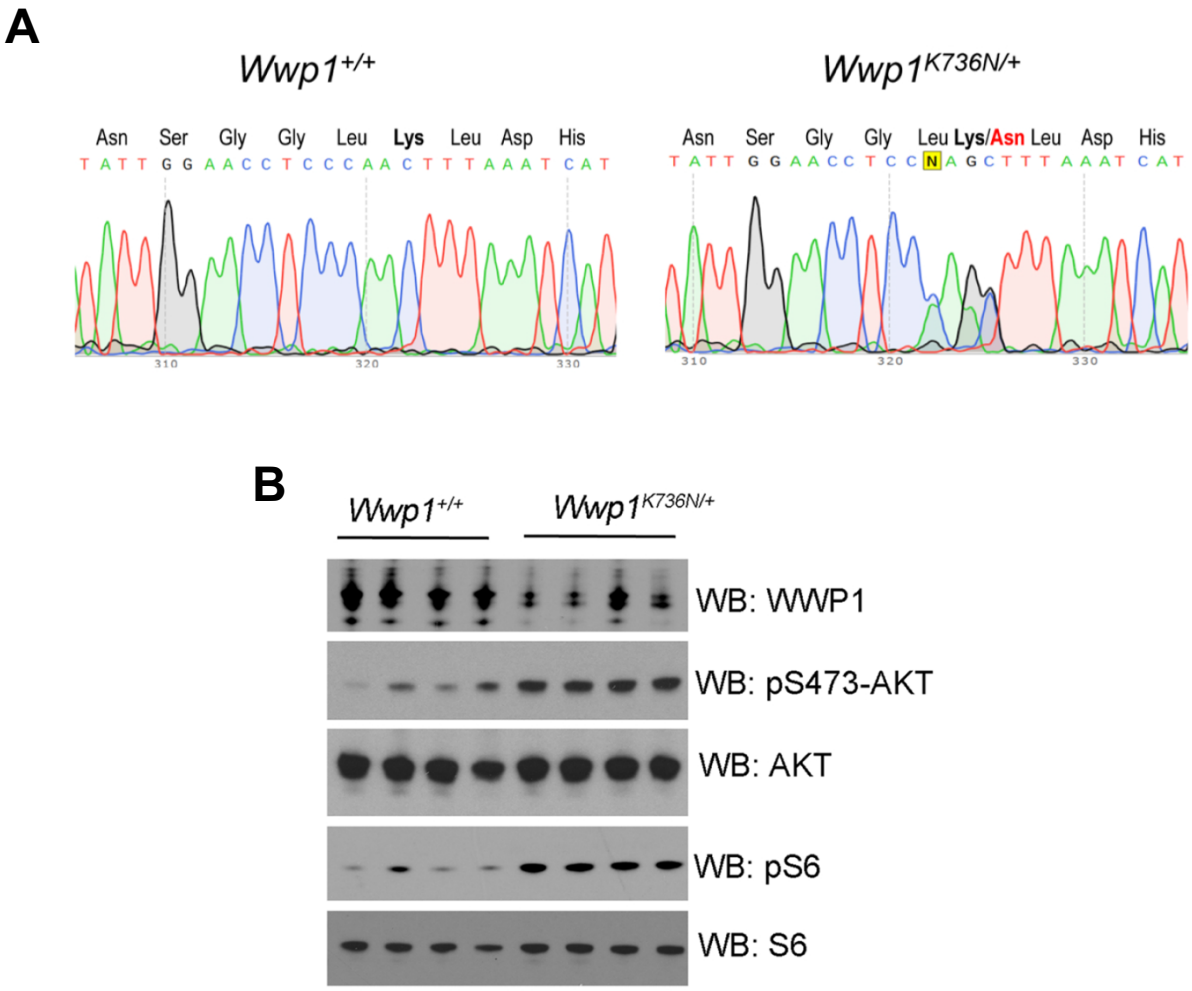
Figure S6. Downstream effects of WWP1 K740N overexpression and generation of heterozygous c.2220G>C, p.K740N CRISPR cell lines.



(**A** and **B**) Analysis of AKT activation in HCT116 or DLD-1 cells stably expressing the indicated constructs.

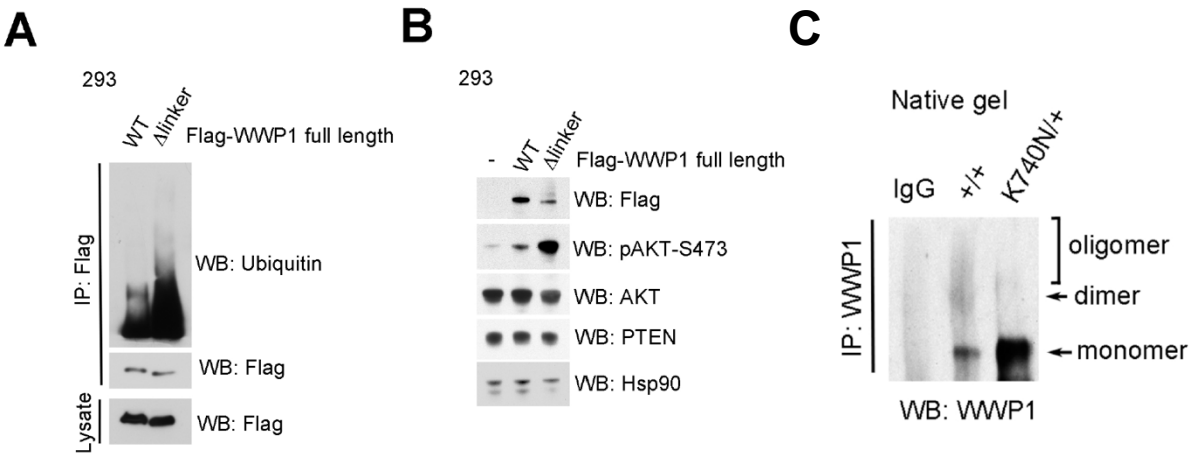
(**C**) Sanger sequencing of genomic DNA isolated from CRISPR clones showing HCT116 cells with either *WWP1*$^{+/+}$ or *WWP1*$^{K740N/+}$ genotype.

Figure S7. Interrogation of mouse embryonic fibroblasts derived from CRISPR knock-in K736N mutant mice.



(**A**) Sanger sequencing of genomic DNA isolated from mice with either *Wwp1*$^{+/+}$ or *Wwp1*$^{K736N/+}$ genotype. (**B**) Analysis of AKT activation in MEFs isolated from mice with either *Wwp1*$^{+/+}$ or *Wwp1*$^{K736N/+}$ genotype. Total lysates were resolved by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and then probed with indicated antibodies.

Figure S8. Downstream effects of *WWP1* p.K740N gain-of-function

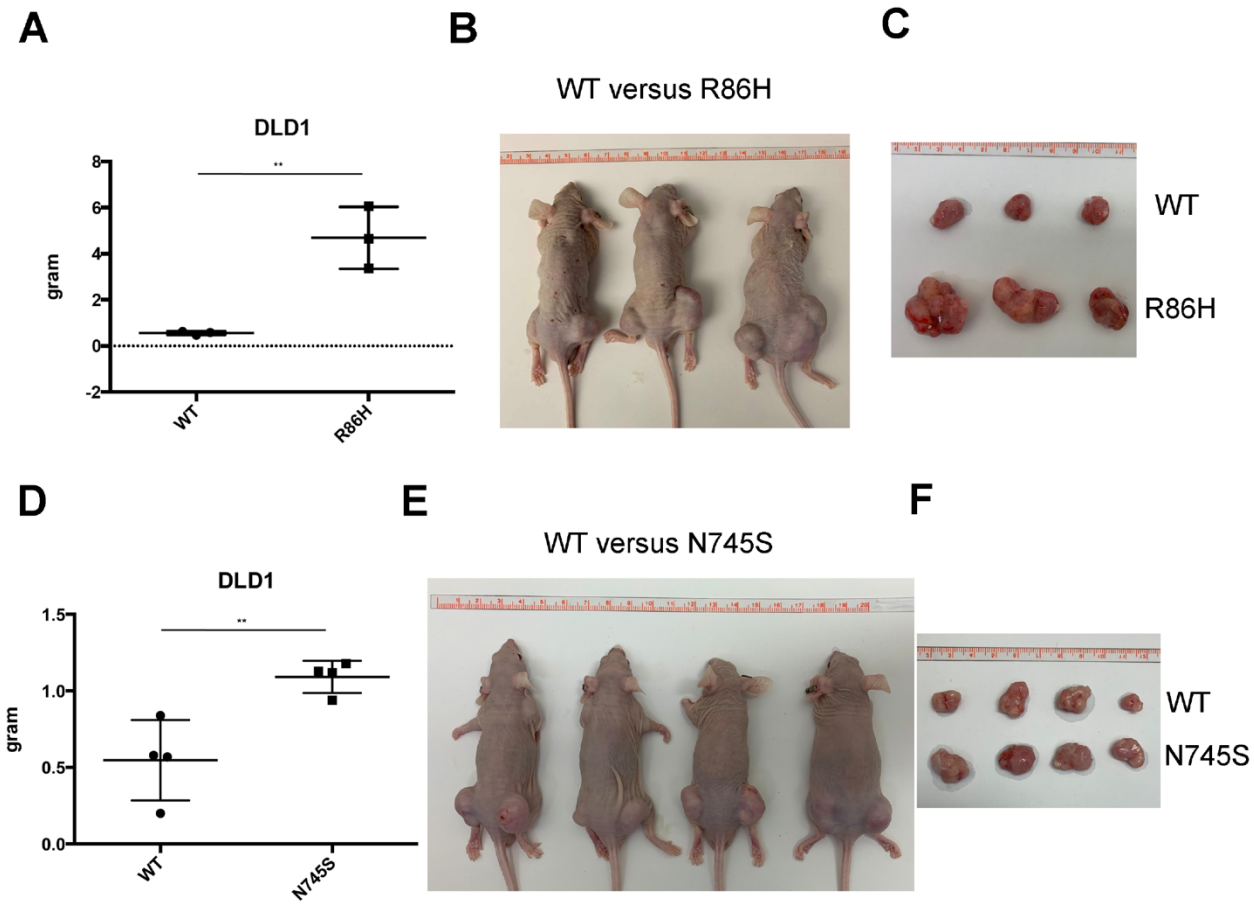


(**A**) Effects of the indicated *WWP1* deletion mutants on WWP1-mediated auto-ubiquitination. HEK293T cells were transfected with the indicated constructs, and WWP1 auto-ubiquitination was analyzed.

(**B**) Analysis of AKT activation in HEK293T cells expressing the indicated constructs. Total lysates were resolved by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and then probed with the indicated antibodies.

(**C**) Evaluation of WWP1 dimerization and oligomerization potential in HCT116 cells with either $WWP1^{+/+}$ or $WWP1^{K740N/+}$ genotype. Total lysates from cells were immunoprecipitated with an anti-WWP1 antibody, and then the immunocomplexes were natively eluted from the beads. The eluted samples were immediately run on the native gel.

Figure S9. Tumor xenograft models of *WWP1* R86H and N745S variants



(**A** to **C**) Tumor xenograft assays were performed by subcutaneously implanting DLD-1 cells stably expressing indicated constructs. The mice were implanted with tumors on both sides. WT tumors were implanted on the left side of mice, while tumors stably expressing R86H were implanted on the right side of mice (n=3 mice per group). Data are shown as mean ± SD (***P < 0.0005, **P < 0.01, P* < 0.05, Student's t test).

(**D** to **F**) Tumor xenograft assays were performed by subcutaneously implanting DLD-1 cells stably expressing indicated constructs. The mice were implanted with tumors on both sides. WT tumors were implanted on the left side of mice, while tumors stably expressing N745S were implanted on the right side of mice (n=4 mice per group). Data are shown as mean ± SD (***P < 0.0005, **P < 0.01, P* < 0.05, Student's t test).

Supplementary Tables

Table S1. International Cowden Consortium operational diagnostic criteria.

Pathognomonic	**Major**	**Minor**
Adult Lhermitte-Duclos disease (LDD) Mucocutaneous lesions Trichilemmomas, facial Acral keratoses Papillomatous papules Mucosal lesions	Breast carcinoma Thyroid carcinoma (non-medullary), especially follicular thyroid carcinoma Macrocephaly (occipital frontal circumference ≥ 97th percentile) Endometrial carcinoma	Other thyroid lesions (e.g., adenoma, multinodular goiter) Mental retardation (i.e., IQ ≤ 75) Gastrointestinal hamartomas Fibrocystic breast disease Lipomas Fibromas Genitourinary tumors (especially renal cell carcinoma) Genitourinary malformations Uterine fibroids
Operational diagnosis in an individual		
Any of following: Mucocutaneous lesions alone, if ≥ six facial papules (three of which must be trichilemmomas) Cutaneous facial papules and oral mucosal papillomatosis Oral mucosal papillomatosis and acral keratoses ≥ Six palmoplantar keratoses ≥ Two major criteria (one of which must be macrocephaly or LDD) One major and ≥ three minor criteria ≥ Four minor criteria		
Operational diagnosis in a family where one individual is diagnostic for CS		
Any one pathognomonic criterion Any one major criterion ± minor criteria Two minor criteria History of Bannayan-Riley-Ruvalcaba syndrome (BRRS)		

Table S2. Demographic and clinical characteristics of unrelated Cowden-like (CS-like) oligopolyposis probands with identified *WWP1* germline variants.

Patient	Clinical phenotypes	*WWP1* (NM_007013)	Population frequencies	Deleteriousness predictions
CCF03506 M, 65 yrs	Macrocephaly (62 cm), gastrointestinal polyps (12 hyperplastic, 3 lipomatous; age at first polyp = 61)	Exon 4: c.178T>G, p.S60A (rs779336984)	1000G: 0 ESP6500: 0 ExAC: 2.84E-05 gnomAD: 4.24E-05	SIFT: 🔵 PP2: 🟤 MT: 🟤 CADD: 🟤
CCF04145 F, 51 yrs	Gastrointestinal polyps (18 adenomatous, 2 hamartomatous, 4 hyperplastic, 1 lipomatous, 1 NOS, 1 other, 1 serrated adenoma; age at first polyp = 48)	Exon 9: c.1006A>C, p.M336L (rs200063533)	1000G: 0 ESP6500: 0 ExAC: 2.84E-05 gnomAD: 6.80E-05	SIFT: 🔵 PP2: 🔵 MT: 🔵 CADD: 🔵
CCF04959 F, 69 yrs	Gastrointestinal polyps (3 adenomatous, 5 hyperplastic; age at first polyp = 59); ascending colon cancer (age 62), ovarian cancer (age 48)	Exon 15: c.1668T>A, p.Y556X	1000G: 0 ESP6500: 0 ExAC: 0 gnomAD: 0	SIFT: 🟤 PP2: 🟤 MT: 🟤 CADD: 🟤
CCF01687 F, 63 yrs	Follicular variant papillary thyroid cancer (age 62), invasive breast cancer (age 63), fibrocystic breast disease, uterine fibroids, gastrointestinal polyps (8 hyperplastic)	Exon 16: c.1709A>G, p.N570S (rs141471813)	1000G: 0 ESP6500: 0.0002 ExAC: 3.77E-05 gnomAD: 2.54E-05	SIFT: 🔵 PP2: 🔵 MT: 🟤 CADD: 🟤
CCF03258 M, 70 yrs	Gastrointestinal polyps (5 hamartomatous, 2 NOS; age at first polyp = 67), bladder cancer (age 70)	Exon 20: c.2234A>G, p.N745S (rs148651938)	1000G: 0.0008 ESP6500: 0.002 ExAC: 0.0014 gnomAD: 0.0016	SIFT: 🔵 PP2: 🟤 MT: 🟤 CADD: 🟤

Patient age refers to the age at consent in years. Results of computational deleteriousness prediction algorithms are depicted as blue (tolerated) or brown (damaging) circles. *Abbreviations*: M, male; F, female; yrs, years; NOS, not otherwise specified; 1000G, 1000 Genomes project; ESP6500, National Heart, Lung, and Blood Institute (NHLBI) Exome Sequencing Project; ExAC, Exome Aggregation Consortium (here, non-TCGA dataset); gnomAD, Genome Aggregation Consortium (here, non-cancer dataset); SIFT, Sorting Intolerant from Tolerant (https://sift.bii.a-star.edu.sg); PP2, PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/); MT, MutationTaster (http://mutationtaster.org); CADD, Combined Annotation Dependent Depletion (https://cadd.gs.washington.edu).

Table S3. TCGA pan-cancer dataset.

TCGA code	Study Name	Sample size
LAML	Acute Myeloid Leukemia	142
ACC	Adrenocortical carcinoma	92
BLCA	Bladder Urothelial Carcinoma	412
LGG	Brain Lower Grade Glioma	515
BRCA	Breast invasive carcinoma	1076
CESC	Cervical squamous cell carcinoma and endocervical adenocarcinoma	305
CHOL	Cholangiocarcinoma	45
COAD	Colon adenocarcinoma	419
ESCA	Esophageal carcinoma	184
GBM	Glioblastoma multiforme	393
HNSC	Head and Neck squamous cell carcinoma	526
KICH	Kidney Chromophobe	66
KIRC	Kidney renal clear cell carcinoma	387
KIRP	Kidney renal papillary cell carcinoma	289
LIHC	Liver hepatocellular carcinoma	375
LUAD	Lung adenocarcinoma	518
LUSC	Lung squamous cell carcinoma	499
DLBC	Lymphoid Neoplasm Diffuse Large B-cell Lymphoma	41
MESO	Mesothelioma	82
OV	Ovarian serous cystadenocarcinoma	412
PAAD	Pancreatic adenocarcinoma	185
PCPG	Pheochromocytoma and Paraganglioma	179
PRAD	Prostate adenocarcinoma	498
READ	Rectum adenocarcinoma	145
SARC	Sarcoma	255
SKCM	Skin Cutaneous Melanoma	470
STAD	Stomach adenocarcinoma	443
TGCT	Testicular Germ Cell Tumors	134
THYM	Thymoma	123
THCA	Thyroid carcinoma	499
UCS	Uterine Carcinosarcoma	57
UCEC	Uterine Corpus Endometrial Carcinoma	543
UVM	Uveal Melanoma	80
TCGA	All studies	10389
TCGA-CSCC	Cowden syndrome component sporadic cancers represented in TCGA	3683

Tables provided in the Excel spreadsheet:

Table S4. Germline *WWP1* variants identified in The Cancer Genome Atlas (TCGA)

Table S5. Germline *WWP1* variants identified in the Exome Aggregation Consortium minus TCGA (non-TCGA ExAC)

Table S6. Germline *WWP1* variants identified in the gnomAD non-cancer population

Table S7. Germline variants in representative genes classically known to be associated with polyposis and/or hereditary colon cancer predisposition in the TCGA and ExAC (non-TCGA) datasets

Table S8. Germline variants in representative classical cancer susceptibility genes identified in the TCGA and ExAC (non-TCGA) datasets

Table S9. Enrichment analysis of germline variants in representative classical cancer susceptibility genes in the TCGA and ExAC (non-TCGA) datasets

Table S10. Primer pair sequences for PCR-based amplification of *WWP1* exons 1-25 prior to high resolution melt (HRM) analysis.

Primer	Sequence
WWP1 exon 1 F	CCGCGTGCGGGTT
WWP1 exon 1 R	GCGCCGCGCCGTC
WWP1 exon 2 F	TAGGAATTATCACACTGAAAGCTATTTAT
WWP1 exon 2 R	ACTATCAAAATGTATGAAGTCTGACTATTA
WWP1 exon 3 F	ACAAACTCTTATCTAACACATGAATAAAT
WWP1 exon 3 R	AGTTCATCAGGGAATCACCATATTA
WWP1 exon 4 F	GCAACACATACTCACTAGTGAAT
WWP1 exon 4 R	TTCCGTCCTTTACACAAGGT
WWP1 exon 5 F	GAATGACAACGTCTACTCCTG
WWP1 exon 5 R	ATAAACTTATAAGGAAATAAAGGTCAGAAC
WWP1 exon 6 F	TGTAGGTTTGATTTCTGAAGCATT
WWP1 exon 6 R	TTTCCTTGAAATTTTACAAACATATAAAGT
WWP1 exon 7 F	GAATAAGCAAGAAGTATATAAAAACCTAGT
WWP1 exon 7 R	GTCGCCACCCATATTCAAA
WWP1 exon 8 F	AATTCTCATGAAATGTTGTTGAATTATACT
WWP1 exon 8 R	TCCTTTAAACAAGGTGTCATAAATAGC
WWP1 exon 9.1 F	GTCAATTGATTAGGTAATTGATTGCTTT
WWP1 exon 9.1 R	CAATTCTGCACTGGTAGAAGGAATA
WWP1 exon 9.2 F	CAGTTCAAGAAATACTGACTTCCT
WWP1 exon 9.2 R	TGCTACTTAAATGCAAGTCAGAC
WWP1 exon 10 F	ACTGTTATTTTATTTCTCTCCCTAATCTT
WWP1 exon 10 R	CTCCCTTATAGATCATTAGCAAAAAC
WWP1 exon 11 F	CTGCTTTTCATTTTCTTATATTGAGAGAT
WWP1 exon 11 R	CACCTGAGGGAGGAAGT
WWP1 exon 12 F	ACTGTTATTTTATTTCTCTCCCTAATCTT
WWP1 exon 12 R	CTTATAGATCATTAGCAAAAACAGCTTAC
WWP1 exon 13 F	AGATCCTAAATAGTTATTAAATATTATACTCAC
WWP1 exon 13 R	CTACTGATATATGCAAGACGAC

WWP1 exon 14 F	CCTAGAAAAGGTTCATCTTGTGATT
WWP1 exon 14 R	ACTAAGCAATATTTAAAGGAGAACTTCAG
WWP1 exon 15 F	ACTAAATATTCAACTTTATTATGAGTTAATTTGGT
WWP1 exon 15 R	GTATAAAGAGAAGACTAAATTATTGAGTATAAGAT
WWP1 exon 16 F	GCCAGGTACTATAGTGGTAAATAAATC
WWP1 exon 16 R	CAGAATACAATGAGACAAATGGTG
WWP1 exon 17 F	CTTGTACTGCATGTGAGTATTTAATATTC
WWP1 exon 17 R	CAATGAGATATACTTATTTCAAGGCAGA
WWP1 exon 18 F	ACATATTAATGCTAAAAAATAACCTTGACT
WWP1 exon 18 R	GCCTTCTCTAATACATTTGTAACAAAC
WWP1 exon 19 F	TTTCTTTTATCATTGTTCTCTAAGAAATTG
WWP1 exon 19 R	AAATGACTTATTGCAAAACAAAATCTTC
WWP1 exon 20 F	CCCATTTTGTTTCATTTTGTTAATTAGTG
WWP1 exon 20 R	GAGAGAGGATTTAGTTCTGAAGTG
WWP1 exon 21 F	AGAACTATAGAAGTTCTTGGTGTTT
WWP1 exon 21 R	CACTAACATTCCCTACAATAGCAT
WWP1 exon 22 F	GTGGCCTGATTCTGTGC
WWP1 exon 22 R	CCATTTCATTAGAAAATATGTTGAGTACTT
WWP1 exon 23 F	GCCCAACCTTTAGCACATCTT
WWP1 exon 23 R	TCCCTTCACAGTCACCATAGG
WWP1 exon 24 F	AGGACCAGATAAACTTTGTCTTATT
WWP1 exon 24 R	GCTTTCACCTTCGTATTTAGGAT
WWP1 exon 25 F	AACATCTGATTTTGTTTTTGTTTCTG
WWP1 exon 25 R	TTAAATGCAAGAGCTCCTCCA

Optimal annealing temperature for exon 1 is 65°C, whereas all other exons (2-25) have an optimized annealing temperature of 60°C. F, forward primer; R, reverse primer.

References

1. Tan MH, Mester J, Peterson C, et al. A clinical scoring system for selection of patients for PTEN mutation testing is proposed on the basis of a prospective study of 3042 probands. Am J Hum Genet 2011;88:42-56.
2. Heald B, Mester J, Rybicki L, Orloff MS, Burke CA, Eng C. Frequent gastrointestinal polyps and colorectal adenocarcinomas in a prospective series of PTEN mutation carriers. Gastroenterology 2010;139:1927-33.
3. Ngeow J, Heald B, Rybicki LA, et al. Prevalence of germline PTEN, BMPR1A, SMAD4, STK11, and ENG mutations in patients with moderate-load colorectal polyps. Gastroenterology 2013;144:1402-9, 9 e1-5.
4. Gorlin RJ, Cohen MM, Jr., Condon LM, Burke BA. Bannayan-Riley-Ruvalcaba syndrome. Am J Med Genet 1992;44:307-14.
5. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;25:1754-60.
6. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 2010;20:1297-303.
7. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009;25:2078-9.
8. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 2010;38:e164.
9. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res 2003;31:3812-4.

10. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. Nat Methods 2010;7:248-9.
11. Schwarz JM, Rodelsperger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. Nat Methods 2010;7:575-6.
12. Robinson JT, Thorvaldsdottir H, Winckler W, et al. Integrative genomics viewer. Nat Biotechnol 2011;29:24-6.
13. Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2012;2:401-4.
14. Huang KL, Mashl RJ, Wu Y, et al. Pathogenic Germline Variants in 10,389 Adult Cancers. Cell 2018;173:355-70 e14.