A novel germline mutation in the POT1 gene predisposes to familial non-medullary thyroid cancer

Aayushi Srivastava1,2,3,4*, Beiping Miao2,3*, Diamanto Skopelitou1,2,3,4, Varun Kumar5, Abhishek Kumar1,6,7, Nagarajan Paramasivam8, Elena Bonora9, Kari Hemminki1,10,§, Asta Försti1,2,3,§, Obul Reddy Bandapalli1,2,3,4,9,#

1Division of Molecular Genetic Epidemiology, German Cancer Research Center (DKFZ), D69120 Heidelberg, Germany
2Hopp Children’s Cancer Center (KiTZ), Heidelberg, Germany
3Division of Pediatric Neurooncology, German Cancer Research Center (DKFZ), German Cancer Consortium (DKTK), Heidelberg, Germany
4Heidelberg University, Medical Faculty, Heidelberg, Germany
5Department of Medicine I and Clinical Chemistry, University Hospital of Heidelberg, Germany
6Institute of Bioinformatics, International Technology Park, Bangalore, India
7Manipal Academy of Higher Education (MAHE), Manipal 576104, Karnataka, India
8Computational Oncology, Molecular Diagnostics Program, National Center for Tumor Diseases (NCT), Germany.
9Unit of Medical Genetics, Department of Medical and Surgical Sciences, S.Orsola-Malpighi Hospital, University of Bologna, 40138 Bologna, Italy
10Faculty of Medicine and Biomedical Center in Pilsen, Charles University in Prague, 30605 Pilsen, Czech Republic.

* Equal contribution
§ Shared senior authorship
# Correspondence

Obul Reddy Bandapalli
Running title: POT1 germline mutation in thyroid cancer
Abstract

Non-medullary thyroid cancer (NMTC) is a common endocrine malignancy with a genetic basis that has yet to be unequivocally established. In a recent whole genome sequencing study of five families with recurrence of NMTCs, we shortlisted promising variants with the help of bioinformatics tools. Here, we report in silico analyses and in vitro experiments on a novel germline variant (p.V29L) in the highly conserved oligonucleotide/oligosaccharide binding domain of the Protection of Telomeres 1 (POT1) gene in one of the families. The results showed that the variant demonstrates a reduction in telomere-bound POT1 levels in the mutant protein as compared to its wild-type counterpart. HEK293T cells carrying POT1\textsuperscript{V29L} showed increased telomere length in comparison to wild type cells, strongly suggesting that the mutation causes telomere dysfunction and may play a role in predisposition to NMTC in this family. This study reports the first germline POT1 mutation in a family with a predominance of thyroid cancer, thereby expanding the spectrum of cancers associated with mutations in the shelterin complex.

Introduction

Thyroid cancer is the most frequently diagnosed malignant endocrine tumor with a world average age-standardized incidence rate of 6.7/100,000 persons per year (Bray, Ferlay et al., 2018). Non-medullary thyroid carcinoma (NMTC) accounts for up to 95% of all thyroid cancers (Hincza, Kowalik et al., 2019, Peiling Yang & Ngeow, 2016). Based on the population-based registers of the Nordic countries the risk of NMTC is about three-fold higher when a first-degree relative is diagnosed with NMTC compared to those without affected family members (Fallah, Pukkala et al., 2013). Apart from the rare syndromic forms of familial NMTC (FNMT), including familial adenomatous polyposis, Gardner syndrome, Cowden syndrome, Carney complex type 1, Werner syndrome, and Dicer1 syndrome, the genetic basis of FNMT is largely unknown (Hincza et al., 2019, Peiling Yang & Ngeow, 2016). FNMT has been associated with an earlier age of onset, a higher incidence of multifocality and more aggressive disease compared to its sporadic counterpart (El Lakis, Giannakou et al., 2019, Fallah et al., 2013). Thus, it is important to identify genetic factors behind the familial disease to facilitate genetic counseling and clinical management of the patients.

Various approaches, including genome-wide association studies, linkage analyses, targeted sequencing, and whole exome sequencing, have been employed to gain understanding into the genetic basis of FNMT. Several genes and loci, including mainly low-penetrance variants near or in FOXE1, SRGAP1, TITF-1/NKX2-1, DIRC3, and CHEK2, have been suggested to affect non-syndromic FNMT susceptibility (Hincza et al., 2019, Peiling Yang & Ngeow, 2016). In addition, an imbalance of the telomere-telomerase complex has been demonstrated in the peripheral blood of familial papillary thyroid cancer patients (Capezzzone, Cantara et al., 2008).

Recently, we performed whole genome sequencing (WGS) on five families with documented recurrence of NMTC and analyzed these samples with our in-house developed variant prioritization pipeline (FCVPPv2) along with other in silico tools (Srivastava, Kumar et al., 2019). This allowed us to identify a novel missense variant (p.V29L) in the protection of telomeres 1 (POT1) gene in one of the families.
POT1 is a critical component of the shelterin complex, which binds and protects telomeres by modulating telomere capping, replication, and extension by telomerase (de Lange, 2018). Structurally, it is the only member of the shelterin complex that contains two N-terminal oligonucleotide/oligosaccharide binding (OB) domains, which can bind the single-stranded TTAGGG repeats as well as a C terminus, which can bind to TPP1, anchoring it to the shelterin complex composed of four other components: TRF1, TRF2, TIN2 and RAP1 (de Lange, 2018). Germline variants in POT1 have been described in familial melanoma (Potrony, Puig-Butille et al., 2019, Robles-Espinoza, Harland et al., 2014, Shi, Yang et al., 2014, Wong, Robles-Espinoza et al., 2019), glioma (Bainbridge, Armstrong et al., 2015), Li-Fraumeni-like syndrome (Calvete, Martinez et al., 2015b), colorectal cancer (Chubb, Broderick et al., 2016), chronic lymphocytic leukemia (Speedy, Kinnersley et al., 2016) and Hodgkin lymphoma (McMaster, Sun et al., 2018).

In this study, we elucidated the genetic and functional consequences of the novel POT1 missense variant p.V29L that segregated in an NMTC family. Results from in silico and in vitro analyses suggested a mechanism for the identified POT1 mutations in predisposition to FMNTC.

Results

WGS and variant prioritization

The subject of this study was an Italian family with reported recurrence of NMTC (Fig. 1). Five members of this family were affected by PTC, Hürthle cell cancer, micro-PTC or a combination of two subtypes (II-2, II-3, II-5, II-8, II-9). Three members were possible carriers affected by benign nodules (I-1, II-4, II-6) and two were unaffected (II-1, II-7). WGS was performed on eight of these family members. The variants were filtered based on pedigree data considering family members diagnosed with NMTC or micro-PTC as cases, benign nodules or goiter as potential variant carriers and unaffected members as controls. A total of 101081 variants, with mean allele frequencies less than 0.1%, was reduced by pedigree-based filtering to 2708. We did not identify any deleterious loss-of-function variants, however, six non-synonymous variants in six genes (EPYC, SPOCK1, MYBPC1, ACSS3, NRP1, and POT1) segregated with the disease in the family and passed the filters of the FCVPPv2 (Srivastava et al., 2019). An overview of the process leading to the selection of a candidate variant is outlined in Figure 1B. Given the importance of POT1 in various cancers, we selected it as our candidate variant for further in silico analyses and functional validation. A list of all shortlisted variants and their scores is available in the supplementary data (Table S1).

In silico studies predict the importance of the p.V29L mutation to POT1 protein function

Comparative sequence analysis of the p.V29L position showed it to be highly conserved across selected representative species within the phylogeny (Fig. 2A). The p.V29L variant is located in the OB1 domain of
the protein, as are several other germline and somatic variants reported in a wide spectrum of other human cancers (Fig. 2B, Table S2). It is also evident that the region around the position p.V29L is highly conserved. The tolerance of POT1 protein function to single amino acid substitutions was calculated by SNAP2 and accessed using PredictProtein. The heat map representation of the resulting data shows a highly deleterious effect of almost all substitutions in the position p.V29. An aggregation of highly deleterious effects of any amino acid change can be seen in the selected range (1-72 amino acids; Fig. 2C). These predictions reinforce the biological importance of the OB folds.

We attained the crystal structure of the N-terminal domain of POT1 (aa 1-185) as a complex with ssDNA from the RCSB PDB database (1XJV) (Lei, Podell et al., 2003). This domain binds G-rich telomeric ssDNA with the same specificity and higher affinity than the full-length protein, suggesting that this segment encompasses the entire DNA-binding region of the protein (Lei et al., 2003). The OB-fold shown in this crystal structure consists of a highly curved, five-stranded anti-parallel β-barrel. The interaction of the ssDNA with the concave groove of the OB folds along with the position of our variant (p.V29L) can be seen in Figure 2D. Moreover, we predicted the change in protein stability by the p.V29L substitution using the mutation Cutoff Scanning Matrix approach (mCSM), which relies on graph-based signatures to predict the impact of missense mutations on protein stability. The thermodynamic change in free energy caused by the p.V29L mutation was predicted to be destabilizing (ΔΔG = -0.886 Kcal/mol) (Fig. 2D).

The V29L mutation aggravates DNA dependent functions of POT1

As adverted to in the introduction, both germline and somatic deleterious mutations reported in POT1 tend to be concentrated in its OB folds (Gu, Wang et al., 2017). Therefore, these OB domains are the main target of the mutational events in this protein across different human cancers. Mutations in this region have been reported to affect DNA binding and lead to loss of function of the POT1 protein (Calvete, Martinez et al., 2015a, Ramsay, Quesada et al., 2013). To test whether the missense POT1 variant affects protein function, we performed western blotting with lysates isolated from HEK293T cells transfected with an empty vector, or with the vector carrying cDNA encoding Myc-tagged human wild-type POT1 or Myc-tagged human mutant POT1. We did not detect significant differences in POT1 protein levels between POT1WT and POT1V29L transfected cells (Fig. 3A). We then performed chromatin immunoprecipitation (ChIP) assays to examine the effect of the POT1V29L variant on the binding of POT1 to telomeric chromatin. Our results showed significantly weakened binding of telomeric DNA to POT1V29L as compared to POT1WT (Fig. 3B).

Furthermore, to confirm our findings from the ChIP assay, we performed an electrophoretic mobility shift assay (EMSA) using constructs containing cDNA for wild-type and mutant POT1 that were translated in vitro and incubated with radiolabeled telomeric ssDNA. EMSA results confirmed that the p.V29L alteration affected the ability of POT1 to bind to the 3’ end of the G-rich telomeric overhang, whereas wild-type POT1 was able to efficiently bind to telomeric ssDNA (Fig. 3C). In an attempt to assess the effect of the POT1V29L variant on telomere length, we measured telomere length in the family members using WGS data and found no significant difference. This could be due to naturally occurring
variance in telomere length within the human population as well as due to age differences and the
limited number of samples. This situation is distinct from one that is present in the analysis of cell lines.
We demonstrated this experimentally by passaging 40 times of HEK293T cells transfected with POT1$^{\text{WT}}$
and POT1$^{\text{V29L}}$ and subsequently re-analyzing the telomere lengths. The results showed that the telomere
length was significantly longer in mutated cells compared to the wild-type cells (two-tailed student’s t-
test, P<0.005; Fig. 3D).

Discussion

In our study, we identified a novel germline POT1 missense mutation that segregated with thyroid
cancer in an Italian family. As the scope of personalized therapy and medical genetics advance, the
importance of identifying mutations and pathways affected in different cancers is heightened. Next-
generation sequencing has emerged as the state-of-the-art tool for the identification of driver mutations
in tumors and novel cancer-predisposing genes in Mendelian diseases. The heritability of thyroid cancer
can be attributed to both rare, high-penetrance mutations and common, low-penetrance variants. Our
approach was focused on identifying the former in a familial pedigree.

The POT1 variant prioritized in this family (p.V29L) underwent several in silico and in vitro studies to
demonstrate the consequence of the amino acid substitution. Since the mutation is located in the OB1-
domain, a thorough literature review aided us in hypothesizing the functional effects of the point
mutation (Gu et al., 2017), as several other germline and somatic variants reported in a wide spectrum
of other human cancers are also clustered around the OB-domains. These predictions were supported
by results from both the in silico studies as well as the in vitro studies. In silico studies showed putative
disruption of POT1 protein function by p.V29L that was later validated by functional studies. The ChIP
assay showed a significant decrease in the mutant POT1 protein’s ability to bind to ssDNA as compared
to its wild-type counterpart (p= 0.01, student’s t-test.). Results from the EMSA supported this inference
by also showing a decreased ability of POT1$^{\text{V29L}}$ in forming a protein-DNA complex. Although we did not
test the capability of the POT1$^{\text{V29L}}$ protein to interact and bind to TPP1, thus allowing it to localize to
double-stranded telomeres, a previous functional study on a variant in the OB-folds of POT1 (p.R117C)
showed disruption in POT1-TPP1 interaction as a result of the mutation (Calvete, Garcia-Pavia et al.,
2017).

Germline deleterious mutations in POT1 have previously been associated with susceptibility to
melanoma (Robles-Espinoza et al., 2014, Shi et al., 2014), glioma (Bainbridge et al., 2015), colorectal
cancer (Chubb et al., 2016), Li-Fraumeni-like syndrome (Calvete et al., 2015b) and chronic lymphocytic
leukemia (Speedy et al., 2016). Orois et al. analyzed seven FNMTC families with the sole aim of
identifying POT1 mutations but were unable to detect any variants in these families (Orois, Badenas et
al., 2020). It is of particular interest to note that predicting the phenotype simply by attaining the
genotype is not possible. Mutations in the same domain of the POT1 protein can lead to an array of
different human cancers.

Although advancements have been made in recent years in the understanding of FNMTC, the hereditary
factors contributing to the susceptibility to and possible unfavorable prognosis of FNMTC have yet to be
adequately explored. On the one hand, over-diagnosis and overtreatment of low-grade disease or benign nodules have to be avoided, and on the other hand, it is imperative to identify aggressive cases with poor prognoses (Jegerlehner, Bulliard et al., 2017). This is only possible if there is a strong understanding of predictive germline variants and their underlying pathways. We acknowledge that one of the limitations of this study is that the proposed disease-causing variant was found in only one family. However, when dealing with rare, high-penetrance variants, it is a challenging task to locate more than one family with a mutation in the same gene. Nonetheless, this draws attention to two aspects. First, it is evident that many other disease-causing loci have yet to be discovered and second, there is a certain ambiguity associated with the selection of one causal variant in a family, as other deleterious variants that are shared amongst patients in the family could also be important in the pathogenesis of the studied phenotype.

In conclusion, the POT1 mutation reported in this study plays a role in NMTC predisposition. While one germline mutation in POT1 has already been reported in a melanoma-prone family with recurrence of thyroid cancers (Wilson, Hattangady et al., 2017), we report the first of such mutations in a family affected solely by NMTCs. Hence, our study expands the spectrum of cancers known to be evoked by inherited mutations in the POT1 gene. Loss of function of this gene may play a role in the pathogenesis of NMTC and may facilitate novel approaches for screening and clinical management of this disease.

Materials and Methods

Patients

The subject of this study was an Italian family with NMTC aggregation recruited at the S. Orsola-Malpighi Hospital, Unit of Medical Genetics in Bologna, Italy. Samples from four affected members, one unaffected member, and three possible carriers were available for whole genome sequencing. All blood samples were collected from the participants with informed consent following ethical guidelines approved by “Comitato Etico Indipendente dell’Azienda Ospedaliero-Universitaria di Bologna, Policlinico S. Orsola-Malpighi (Bologna, Italy)” and “comité consultatif de protection des personnes dans la recherche biomédicale, Le centre de lutte contre le cancer Léon-Bérard (Lyon, France)” and DNA was isolated using the QiAMP DNA Blood Mini kit according to the manufacturer’s instructions.

Whole Genome Sequencing

WGS of available DNA samples from the NMTC family members was performed using Illumina-based small read sequencing. Mapping to the human reference genome (assembly GRCh37 version hs37d5) was performed using BWA mem (version 0.7.8) (Li & Durbin, 2009) and duplicates were removed using biobambam (version 0.0.148). Platypus (Rimmer, Phan et al., 2014) was used to call small nucleotide variants (SNVs) and InDels through joint calling on all the samples from the family. Variants were annotated using ANNOVAR, 1000 Genomes, dbSNP and ExAC (Lek, Karczewski et al., 2016, Smigielski, Sirotkin et al., 2000, The Genomes Project (Auton et al., 2015, Wang, Li et al., 2010). Variants with a QUAL score greater than 20 and coverage greater than 5x, and that passed all the Platypus internal filters were evaluated further. Variants with minor allele frequencies (MAFs) greater than 0.1% in the
1000 Genomes Phase 3 and non-TCGA ExAC data were marked as common and removed. A pairwise comparison of shared rare variants was performed to check for sample swaps and family relatedness.

Data Analysis and Variant Prioritization

Prioritization of variants

Variant evaluation was performed using the criteria of our in-house developed variant prioritization pipeline FCVPPv2 (Kumar, Bandapalli et al., 2018). First, all the variants were filtered based on the pedigree data considering cancer patients as cases, individuals with benign nodules as potential mutation carriers and unaffected persons as controls.

Variants were then filtered using the Combined Annotation Dependent Depletion (CADD) tool v1.3 (Kircher, Witten et al., 2014), and variants with a scaled PHRED-like CADD score greater than 10, i.e. variants belonging to the top 10% of probable deleterious variants in the human genome, were considered further. Genomic Evolutionary Rate Profiling (GERP) (Cooper, Stone et al., 2005), PhastCons (Siepel, Bejerano et al., 2005) and PhyloP (Pollard, Hubisz et al., 2010) were used to evaluate the evolutionary conservation of the genomic position of a particular variant. GERP scores > 2.0, PhastCons scores > 0.3 and PhyloP scores ≥ 3.0 were indicative of a good level of conservation and were therefore used as thresholds in the selection of potentially causative variants.

Next, all variants were assessed for deleteriousness using 10 tools accessed using dbNSFP (Liu, Wu et al., 2016), namely SIFT, PolyPhen V2-HDV, PolyPhen V2-HVAR, LRT, MutationTaster, Mutation Assessor, FATHMM, MetaSVM, MetLR and PROVEAN and variants predicted to be deleterious by at least 60% of these tools were analyzed further.

Lastly, three different intolerance scores derived from NHLBI-ESP6500 (Petrovski, Wang et al., 2013), ExAC (Lek, Karczewski et al., 2016) and a local dataset, all of which were developed with allele frequency data, were included to evaluate the intolerance of genes to functional mutations. The ExAC consortium has developed two additional scoring systems using large-scale exome sequencing data including intolerance scores (pLI) for loss-of-function variants and Z-scores for missense and synonymous variants. These were used for nonsense and missense variants respectively. However, all the intolerance scores were used to rank and prioritize the genes and not as cut-offs for selection.

After shortlisting variants according to the aforementioned criteria, we performed a literature review on the prioritized candidates and checked if coding variants in important oncogenes, tumor suppressor genes or autosomal dominant familial syndrome genes had been missed by the cut-offs of the pipeline. These variants were handled leniently with regard to conservation and deleteriousness cut-offs and were included in the further analysis.

Candidate variant selection and validation

After filtering the variants based on the FCVPPv2, we visually inspected the WGS data for correctness using the Integrative Genomics Viewer (IGV) (Robinson, Thorvaldsdottir et al., 2017). The final selection was based on a thorough literature review. The selected variant of interest (POT1 p.V29L) was validated.
by Sanger sequencing of DNA samples of all available family members using specific primers for polymerase chain reaction amplification designed with Primer3 (http://bioinfo.ut.ee/primer3-0.4.0/).

Primer details are available on request. Sequencing was performed on a 3500 Dx Genetic Analyzer (Life Technologies, CA, USA) using ABI PRISM 3.1 Big Dye terminator chemistry according to the manufacturer’s instructions. The electrophoretic profiles were analyzed manually. Segregation of the variant with the disease was confirmed.

Further in silico studies

SNAP2 (Hecht, Bromberg et al., 2015), a neural network-based classifier, was accessed via PredictProtein (Yachdav, Kloppmann et al., 2014) to generate a heat map representation of independent substitutions for each position of the protein, based on its tolerance to amino acid substitution. The effect of the p.V29L mutation on the stability of the POT1: DNA interaction was assessed using the mutation Cutoff Scanning Matrix (mCSM) tool (Pires, Ascher et al., 2014).

Protein alignment and structural modeling

Multiple sequence alignments were generated for homologous POT1 sequences to evaluate conservation using T-Coffee (Notredame, Higgins et al., 2000). Alignments for POT1 were generated using the following sequences: NP_056265.2, XP_519345.2, NP_001127526.1, XP_009001386.1, XP_006149256.1, NP_598692.1, XP_002712135.2, XP_010802750.1, XP_005628494.1, XP_001501458.4, XP_006910616.1, XP_010585693.1, XP_004478311.1, XP_007504310.1, XP_001508179.2, NP_996875.1, and NP_001084422.1. Alignments were visualized and formatted manually.

Cell culture

HEK293T cells (RRID: CVCL_0063) were a gift from Andreas Trump (DKFZ, Heidelberg). The cells were maintained in DMEM high glucose supplemented with 10% fetal bovine serum (Gibco), penicillin (50 U/ml, Life Technology), and streptomycin (50 µg/ml, Life Technology). The cells have been authenticated using SNP or STR profiling within the last 3 years and all experiments were performed with mycoplasma-free cells.

Selection and expansion of stable POT1\(^{\text{WT}}\) and POT1\(^{\text{V29L}}\) clones

pLPC myc hPOT1 was a gift from Titia de Lange (Addgene plasmid # 12387; http://n2t.net/addgene:12387 ; RRID:Addgene_12387; (Loayza & De Lange, 2003)). POT1\(^{\text{WT}}\) and POT1\(^{\text{V29L}}\) plasmids were transfected into HEK293T cells at 70–80% confluence using Lipofectamine 2000 (Thermo Fisher Scientific). At 48 hours post-transfection, cells were selected by growth with 1 µg/ml of puromycin for 14 days. The medium was changed every 2-3 days. Surviving colonies were selected and used for further experiments.

Measurement of relative telomere length
Telomere length was measured on DNA extracted from HEK293T POT<sub>WT</sub> and POT<sub>1V29L</sub> cells after 40 passages using real-time PCR as described earlier by others and in our lab (Hosen, Rachakonda et al., 2015). Telomere and albumin primer sequences 5′ to 3′ were:

ACACTAAGGGTTGTTGTTGGGTAGTGT (Telg),
TGTTAGGTATCCCTATCCCTATCCCTATCCCTAACA (Telc),
CGCGGCGGCGGCCTGCGGCTGGGCTGGCCTGGCTGGCGGCCATGCTTTTCAGCTCTGCAAGTC (Albugcr2) and
GCCCGGCCCGCCGCGCCCGTCCCGCCGAGCATTAAGCTCTTTGGCAACGTAGGTTTC (Albdgcr2).

Telomere/single-copy gene (T/S) values were calculated by 2<sup>−ΔCt</sup> and relative T/S values (i.e., RTL values) were generated by dividing sample T/S values with the T/S value of reference DNA sample (genomic DNA pooled from 10 healthy individuals). All the experiments were done in triplicates and repeated twice.

Western Blot

Protein lysates were prepared and quantified using the BCA protein assay kit (Pierce, Darmstadt, Germany). 20 µg of the proteins were then blotted onto 0.2 µM nitrocellulose membranes and blocked with 5% milk. Membranes were incubated overnight at 4 °C with the target Anti-Myc tag antibody [9E10] - ChIP Grade (ab32). Immune complexes were detected with the corresponding HRP-conjugated secondary antibody (Anti-rabbit IgG, HRP-linked Antibody, cell signaling, 7074). The loading quantity control was incubated with the Anti-beta-Actin antibody [AC-15] (HRP) (ab49900) overnight at 4°C. Blots were developed by using ECL Western blot substrate (EMD Millipore, Darmstadt, Germany).

Chromatin Immunoprecipitation (ChIP) assay and telomere dot-blots

The ChIP assay was performed following the protocol by Liu F. et al. with minor modifications (Liu, Feng et al., 2017). Briefly, the cells were cultured in 15 cm² dishes with 70% confluence and fixed for 10 min at 25°C with a working solution of 1% (v/v) formaldehyde on a shaking platform. The cross-linking reaction was quenched by adding glycine to a final concentration of 0.125 M. ChIP was performed using the Chromatin immunoprecipitation assay kit (EMD Millipore, #17-295) following the manufacturer’s instructions using 5 µg of a mouse monoclonal to Myc tag – ChIP Grade antibody (Anti-Myc tag antibody [9E10] - ChIP Grade ab32). The lysates were sonicated with 5 × 5 min with 5s on/off intervals (Bioruptor, Diagenode) to get the DNA lengths between 200 and 1000 bp. The immunoprecipitated DNA was purified with the iPure kit (Diagenode, c03010014). Purified DNA was slot blotted onto a Hybond N+ membrane with the help of a dot-blot apparatus (Bio-Rad, 170-6545) and subsequently hybridized with a biotin-labeled (TTAGGG)<sub>3</sub>probe synthesized by Sigma. The North2South<sup>®</sup> Chemiluminescent Hybridization and Detection Kit (Thermo Fisher: 17097) was used to detect the biotin signal with the help of a CCD camera. Signals were then quantified by Image J, and the fold of enrichment was calculated. The amount of telomeric DNA after ChIP was normalized to the total input telomeric DNA.

Electrophoretic mobility shift assay (EMSA)

The gel shift assay of POT<sub>1WT</sub> and POT<sub>1V29L</sub> was performed as described previously (Ramsay et al., 2013). In brief, 20 µl reaction was prepared in EMSA buffer (25 mM HEPES-NaOH (pH 7.5), 100 mM NaCl, 1 mM
EDTA and 5% glycerol) supplemented with 1 µg of poly(dI-dC) and around 30-40 ng of γP32 labeled ds telomere Probe (GGTTAGGGTTAGGGTTAGGG) per reaction. The reactions were incubated for 30 min. at 25°C. POT1wt and POT1V29L were immunoprecipitated from HEK293T cells, ectopically expressing respective POT1 proteins, by lysing them in 20 mM Tris pH 7.5, 40 mM NaCl, 2 mM MgCl2, 0.5% NP40, 50U/ml Benzonase, supplemented with protease and phosphatase inhibitors. After 15 min. of incubation on ice, the NaCl concentration was adjusted to 450 mM and the incubation was continued for another 15 minutes. Lysates were clarified by centrifugation (13200 rpm, 20 min, 4°C) and 1.0 mg of total protein was used per immunoprecipitation in IP buffer (25 mM Tris-Cl (pH 7.5), 150 mM NaCl, 1.5 mM DTT, 10% glycerol, 0.5% NP40) supplemented with protease and phosphatase inhibitors.

Endogenous proteins were captured onto protein G-magnetic beads (NEB; #S1430S), washed extensively in IP buffer and used for POT1wt and POT1V29L source. After gel shift incubation, the reaction contents were loaded onto a pre-electrophoresed 5% acrylamide/bis (37.5:1) gel in 0.5xTBE and run at 100V at 25°C. The gels were dried and analyzed by autoradiography. The labeled probe consensus alone served as a negative control of the EMSA.

Acknowledgments

The authors thank the members of the families for participating in this study, Genomics and Proteomics Core Facility (GPCF) of the German Cancer Research Center (DKFZ) for providing excellent library preparation and sequencing services and the Omics IT and Data Management Core Facility (ODCF) of the DKFZ for the whole genome sequencing data management. KH was supported by the EU Horizon 2020 program grant No. 856620.

Author contributions

K.H., A.F. and O.R.B. conceived and designed the study. E.B. provided the NMTC family samples. N.P. ran the WGS pipeline. A.K., A.S., B.M., N.P., and O.R.B., analyzed the data. A.S., B.M., D.S. and V.K. performed the experiments. A.S. and O.R.B. wrote the first draft of the manuscript. All authors read, commented on and approved the manuscript.

Conflict of interest

The authors declare no conflict of interest

The Paper Explained

Problem

Familial non-medullary thyroid cancer (FNMTCT) accounts for 3–7% of all thyroid cancers with the transmission pattern being an autosomal dominant mode of inheritance. The risk of developing cancer for first-degree relatives is increased 3-5-fold compared to the general population and the genetic alterations responsible for this disease are hardly known. Several low-penetrance predisposing loci have
been identified using GWAS studies in recent years, but no high-penetrance mutations have been confirmed so far.

Results

We identified a novel germline variant in the POT1 gene, a component of the shelterin complex. Functional *in vitro* studies showed that the mutation weakens the binding of POT1 to telomere DNA, leading to abnormal telomere elongation. This is the first POT1 mutation identified in a family with a predominance of thyroid cancers, thereby expanding the spectrum of cancers associated with mutations in the shelterin complex.

Impact

Identification of POT1 mutations associated with familial non-medullary thyroid cancers is important, as it provides the possibility of early diagnosis and screening of individuals at risk of developing NMTCs and enables the development of personalized therapy.

For More Information

Supplementary tables

Supplementary table 1: Short-listed variants with scores

Supplementary table 2: Known POT1 germline variants

References

Bainbridge MN, Armstrong GN, Gramatges MM, Bertuch AA, Jhangiani SN, Doddapaneni H, Lewis L, Tombrello J, Tsavachidis S, Liu Y, Jalali A, Plon SE, Lau CC, Parsons DW, Claus EB, Barnholtz-Sloan J, Il'yasova D, Schildkraut J, Ali-Osman F, Sadetzki S et al. (2015) Germline mutations in shelterin complex genes are associated with familial glioma. *J Natl Cancer Inst* 107: 384

Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 68: 394-424

Calvete O, Garcia-Pavia P, Dominguez F, Bougeard G, Kunze K, Braeuninger A, Teule A, Lasa A, Ramon YCT, Llort G, Fernandez V, Lazaro C, Urioste M, Benitez J (2017) The wide spectrum of POT1 gene variants correlates with multiple cancer types. *Eur J Hum Genet* 25: 1278-1281

Calvete O, Martinez P, Garcia-Pavia P, Benitez-Buelga C, Paumard-Hernandez B, Fernandez V, Dominguez F, Salas C, Romero-Laorden N, Garcia-Donas J, Carrillo J, Perona R, Trivino JC, Andres R, Cano JM, Rivera B, Alonso-Pulpon L, Setien F, Esteller M, Rodriguez-Perales S et al. (2015a) A mutation in the
POT1 gene is responsible for cardiac angiosarcoma in TP53-negative Li-Fraumeni-like families. *Nat Commun* 6: 8383

Calvete O, Martinez P, Garcia-Pavia P, Benitez-Buelga C, Paumard-Hernández B, Fernandez V, Dominguez F, Salas C, Romero-Laorden N, García-Donas J, Carrillo J, Perona R, Triviño JC, Andrés R, Cano JM, Rivera B, Alonso-Pulpon L, Setien F, Esteller M, Rodriguez-Perales S et al. (2015b) A mutation in the POT1 gene is responsible for cardiac angiosarcoma in TP53-negative Li-Fraumeni-like families. *Nature Communications* 6: 8383

Capezzone M, Cantara S, Marchisotta S, Filetti S, De Santi MM, Rossi B, Ronga G, Durante C, Pacini F (2008) Short telomeres, telomerase reverse transcriptase gene amplification, and increased telomerase activity in the blood of familial papillary thyroid cancer patients. *J Clin Endocrinol Metab* 93: 3950-7

Chubb D, Broderick P, Dobbins SE, Frampton M, Kinnersley B, Penegar S, Price A, Ma YP, Sherborne AL, Palles C, Timofeeva MN, Bishop DT, Dunlop MG, Tomlinson I, Houlston RS (2016) Rare disruptive mutations and their contribution to the heritable risk of colorectal cancer. *Nat Commun* 7: 11883

Cooper GM, Stone EA, Asimenos G, Program NCS, Green ED, Batzoglou S, Sidow A (2005) Distribution and intensity of constraint in mammalian genomic sequence. *Genome Res* 15: 901-13

de Lange T (2018) Shelterin-Mediated Telomere Protection. *Annu Rev Genet* 52: 223-247

El Lakis M, Giannakou A, Nockel PJ, Wiseman D, Gara SK, Patel D, Sater ZA, Kushchayeva YY, Klubo-Gwiezdzinska J, Nilubol N, Merino MJ, Kebebew E (2019) Do patients with familial nonmedullary thyroid cancer present with more aggressive disease? Implications for initial surgical treatment. *Surgery* 165: 50-57

Fallah M, Pukkala E, Tryggvadottir L, Olsen JH, Tretli S, Sundquist K, Hemminki K (2013) Risk of thyroid cancer in first-degree relatives of patients with non-medullary thyroid cancer by histology type and age at diagnosis: a joint study from five Nordic countries. *J Med Genet* 50: 373-82

Gu P, Wang Y, Bisht KK, Wu L, Kukova L, Smith EM, Xiao Y, Bailey SM, Lei M, Nandakumar J, Chang S (2017) Pot1 OB-fold mutations unleash telomere instability to initiate tumorigenesis. *Oncogene* 36: 1939-1951

Hecht M, Bromberg Y, Rost B (2015) Better prediction of functional effects for sequence variants. *BMC Genomics* 16: S1

Hincza K, Kowalik A, Kowalska A (2019) Current Knowledge of Germline Genetic Risk Factors for the Development of Non-Medullary Thyroid Cancer. *Genes (Basel)* 10

Hosen I, Rachakonda PS, Heidenreich B, Sitaram RT, Ljungberg B, Roos G, Hemminki K, Kumar R (2015) TERT promoter mutations in clear cell renal cell carcinoma. *Int J Cancer* 136: 2448-52
Jegerlehner S, Bulliard JL, Aujesky D, Rodondi N, Germann S, Konzelmann I, Chiolero A, Group NW (2017) Overdiagnosis and overtreatment of thyroid cancer: A population-based temporal trend study. *PLoS One* 12: e0179387

Kircher M, Witten DM, Jain P, O’Roak BJ, Cooper GM, Shendure J (2014) A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* 46: 310-5

Kumar A, Bandapalli OR, Paramasivam N, Giangiobbe S, Diquigiovanni C, Bonora E, Eils R, Schlesner M, Hemminki K, Forsti A (2018) Familial Cancer Variant Prioritization Pipeline version 2 (FCVPPv2) applied to a papillary thyroid cancer family. *Sci Rep* 8: 11635

Lei M, Podell ER, Baumann P, Cech TR (2003) DNA self-recognition in the structure of Pot1 bound to telomeric single-stranded DNA. *Nature* 426: 198-203

Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, Tukiainen T, Birnbaum DP, Kosmicki JA, Duncan LE, Estrada K, Zhao F, Zou J, Pierce-Hoffman E, Berghout J, Cooper DN et al. (2016) Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536: 285-91

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25: 1754-60

Liu F, Feng X, Ma W (2017) Analysis of Telomere Proteins by Chromatin Immunoprecipitation (ChIP). *Methods Mol Biol* 1587: 205-214

Liu X, Wu C, Li C, Boerwinkle E (2016) dbNSFP v3.0: A One-Stop Database of Functional Predictions and Annotations for Human Nonsynonymous and Splice-Site SNVs. *Hum Mutat* 37: 235-41

Loayza D, De Lange T (2003) POT1 as a terminal transducer of TRF1 telomere length control. *Nature* 423: 1013-8

McMaster ML, Sun C, Landi MT, Savage SA, Rotunno M, Yang XR, Jones K, Vogt A, Hutchinson A, Zhu B, Wang M, Hicks B, Thirunavukarason A, Stewart DR, Koutros S, Goldstein AM, Chanock SJ, Caporaso NE, Tucker MA, Goldin LR et al. (2018) Germline mutations in Protection of Telomeres 1 in two families with Hodgkin lymphoma. *Br J Haematol* 181: 372-377

Notredame C, Higgins DG, Heringa J (2000) T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* 302: 205-17

Orois A, Badenas C, Reverter JL, Lopez V, Potrony M, Mora M, Halperin I, Oriola J (2020) Lack of Mutations in POT1 Gene in Selected Families with Familial Non-Medullary Thyroid Cancer. *Horm Cancer*

Peiling Yang S, Ngeow J (2016) Familial non-medullary thyroid cancer: unraveling the genetic maze. *Endocr Relat Cancer* 23: R577-R595
Petrovski S, Wang Q, Heinzen EL, Allen AS, Goldstein DB (2013) Genic intolerance to functional variation and the interpretation of personal genomes. PLoS Genet 9: e1003709

Pires DEV, Ascher DB, Blundell TL (2014) mCSM: predicting the effects of mutations in proteins using graph-based signatures. Bioinformatics (Oxford, England) 30: 335-342

Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A (2010) Detection of nonneutral substitution rates on mammalian phylogenies. Genome Res 20: 110-21

Potrony M, Puig-Butille JA, Ribera-Sola M, Iyer V, Robles-Espinoza CD, Aguilera P, Carrera C, Malvehy J, Badenas C, Landi MT, Adams DJ, Puig S (2019) POT1 germline mutations but not TERT promoter mutations are implicated in melanoma susceptibility in a large cohort of Spanish melanoma families. Br J Dermatol 181: 105-113

Ramsay AJ, Quesada V, Foronda M, Conde L, Martinez-Trillos A, Villamor N, Rodriguez D, Kwarciai A, Garabaya C, Gallardo M, Lopez-Guerra M, Lopez-Guillermo A, Puente XS, Blasco MA, Campo E, Lopez-Otin C (2013) POT1 mutations cause telomere dysfunction in chronic lymphocytic leukemia. Nat Genet 45: 526-30

Rimmer A, Phan H, Mathieson I, Iqbal Z, Twigg SRF, Consortium WGS, Wilkie AOM, McVean G, Lunten G (2014) Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications. Nat Genet 46: 912-918

Robinson JT, Thorvaldsdottir H, Wenger AM, Zehir A, Mesirov JP (2017) Variant Review with the Integrative Genomics Viewer. Cancer Res 77: e31-e34

Robles-Espinoza CD, Harland M, Ramsay AJ, Aoude LG, Quesada V, Ding Z, Pooley KA, Pritchard AL, Tiffen JC, Petljak M, Palmer JM, Symmons J, Johansson P, Stark MS, Gartsie MG, Snowden H, Montgomery GW, Martin NG, Liu JZ, Choi J et al. (2014) POT1 loss-of-function variants predispose to familial melanoma. Nat Genet 46: 478-481

Shi J, Yang XR, Ballew B, Rotunno M, Calista D, Fargnoli MC, Ghiorzo P, Bressac-de Paillerets B, Nagore E, Avril MF, Caporaso NE, McMaster ML, Cullen M, Wang Z, Zhang X, Bruno W, Pastirino L, Queirolo P, Banuls-Roca J, Garcia-Casado Z et al. (2014) Rare missense variants in POT1 predispose to familial cutaneous malignant melanoma. Nat Genet 46: 482-6

Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, Clawson H, Spieth J, Hillier LW, Richards S, Weinstock GM, Wilson RK, Gibbs RA, Kent WJ, Miller W, Haussler D (2005) Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res 15: 1034-50

Speedy HE, Kinnersley B, Chubb D, Broderick P, Law PJ, Litchfield K, Jayne S, Dyer MJS, Dearden C, Follows GA, Catovsky D, Houlston RS (2016) Germ line mutations in shelterin complex genes are associated with familial chronic lymphocytic leukemia. Blood 128: 2319-2326
Srivastava A, Kumar A, Giangiobbe S, Bonora E, Hemminki K, Forsti A, Bandapalli OR (2019) Whole Genome Sequencing of Familial Non-Medullary Thyroid Cancer Identifies Germline Alterations in MAPK/ERK and PI3K/AKT Signaling Pathways. *Biomolecules* 9

Wilson TL, Hattangady N, Lerario AM, Williams C, Koepp E, Quinonez S, Osborne J, Cha KB, Else T (2017) A new POT1 germline mutation-expanding the spectrum of POT1-associated cancers. *Fam Cancer* 16: 561-566

Wong K, Robles-Espinoza CD, Rodriguez D, Rudat SS, Puig S, Potrony M, Wong CC, Hewinson J, Aguilera P, Puig-Butille JA, Bressac-de Paillerets B, Zattara H, van der Weyden L, Fletcher CDM, Brenn T, Arends MJ, Quesada V, Newton-Bishop JA, Lopez-Otin C, Bishop DT et al. (2019) Association of the POT1 Germline Missense Variant p.I78T With Familial Melanoma. *JAMA Dermatol* 155: 604-609

Yachdav G, Kloppmann E, Kajan L, Hecht M, Goldberg T, Hamp T, Honigschmid P, Schafferhans A, Roos M, Bernhofer M, Richter L, Ashkenazy H, Punta M, Schlessinger A, Bromberg Y, Schneider R, Vriend G, Sander C, Ben-Tal N, Rost B (2014) PredictProtein—an open resource for online prediction of protein structural and functional features. *Nucleic Acids Res* 42: W337-43
Figures

Figure 1A. Pedigree of the NMTC family with POT1<sup>V29L</sup> mutations.

Figure 1B. Overview of the variant filtering process using the Familial Cancer Variant Prioritization Pipeline version 2 (FCVPPv2). The number of variants passing each step of the pipeline is shown.

Figure 2. In silico studies.

(A). Comparative sequence analysis of POT1 across representative phylogeny. (B). Schematic primary structure of the POT1 protein with known germline mutations identified in various cancers shown relative to the OB domains (blue) and the TPP1 binding region (orange). The variant identified in this study is highlighted in red. (C). Heat map representation of SNAP2 results showing the predicted impact of individual amino acid substitutions (y-axis) for each position (x-axis) on protein function. Dark red indicates a highly deleterious substitution (score = 100), white indicates a minor effect, green indicates a neutral effect or no effect (score = -100), black represents the corresponding wild type residue (upper panel). The section of amino acids belonging to the OB1 domain from the upper panel is expanded and displayed in detail in the lower panel. p.V29L is shown with the blue rectangles (lower panel). (D). Crystal structure of the N-terminal domain of POT1 as a complex with ssDNA (PDB 1XJV). V29L is indicated with an arrow. Stability change caused by the p.V29L substitution as predicted by mCSM.

Figure 3. The POT1<sup>V29L</sup> mutation affects telomeric binding to ssDNA.

(A) Western blot of empty (untagged) vector, Myc-POT1 WT and Myc-POT1 mutant with beta-actin serving as an internal control. (B) Quantification of telomeric DNA bound to POT1 by ChIP analysis. IgG served as a negative control. Results were normalized to input chromatin. Representative ChIP dot blot is shown with input sample dilutions (left panel). Quantification of ChIP resulted in the bar graph (right panel). Green bar: wild-type; red bar: mutant. (C) Results of EMSA showing the decrease in POT1 binding capacity to telomeric ssDNA by the p.V29L substitution. (D) Relative telomere length is significantly longer in POTV29L compared to POT1WT transfected HEK293T cells after 40 passages.
**Figure 1**

A

- **I**
  - 1
  - 2

- **II**
  - 1
  - 2 (42)
  - 3 (44)
  - 4 (48)
  - 5 (43)
  - 6 (47)
  - 7
  - 8 (41)
  - 9 (38)

- **Legend**
  - Red: Papillary thyroid cancer (PTC)
  - Pink: microPTC
  - Black: Hurthle cell cancer
  - Grey: Hurthle cell adenoma
  - Blue: Benign nodules
  - Green: Simple goitre
  - Red +: Mutant (p.V29L)
  - Green +: Wild-type
  - Unaffected individual
  - Deceased
  - WGS performed
  - Age at diagnosis

B

- Total Variants (MAF ≤ 0.1%)
  - 101081

- Pedigree Segregation
  - 2708

- CADD ≥ 10
  - 174

- Coding Variants
  - 9

- Conservational Screening (GERP, PhastCons, PhyloP)
  - 8

- Deleteriousness Screening
  - 6

- Intolerance Scores
  - 6

- Selection of Candidate Variant
  - POT1^{V29L}

- Functional Validation
Figure 2

A

|        | Human | Chimpanzee | Orangutan | Marmoset | Tree shrew | Mouse | Rabbit | Cow | Dog | Horse | Bat | Elephant | Armadillo | Opossum | Platypus | Chicken | Frog |
|--------|-------|------------|-----------|-----------|------------|--------|--------|-----|-----|--------|-----|----------|-----------|---------|----------|---------|------|
|        | G     | T          | I         | V         | N         | V     | Y      | G   | V   | K     | F   | K       | F         | K       | F       | G       | F    |
| p.Val29| V     | I          | V         | N         | Y         | G     | V     |     |     | F     | F   | F       | F         | F       | F       | G       | F    |
|        | F     | F          | F         | P         | Y         | L     | S     | K   | G   | T     |     |         |            |         |         |         |      |
| 20     |       |            |           |           |           |       |       |     |     |       |     |         |            |         |         |         |      |
| 29     |       |            |           |           |           |       |       |     |     |       |     |         |            |         |         |         |      |
| 41     |       |            |           |           |           |       |       |     |     |       |     |         |            |         |         |         |      |

B

C

D

V29L

Predicted Stability Change (ΔG): -0.886 Kcal/mol (Destabilizing)

Mutation:
- Wild-type: V
- Position: 29
- Mutant-type: L
- Chain: A
**Figure 3**

A. Western blot analysis of Myc-POT1 and Beta-actin expression levels under different conditions.

B. ChIP assay showing protein-DNA complex formation and quantification of POT1 expression.

C. Gel electrophoresis results indicating the presence of POT1-DNA complexes with different input concentrations.

D. Box plot comparing telomere length in cells expressing POT1WT and POT1V29L.