The wide spectrum of *POT1* gene mutations correlates with multiple cancer types.

Oriol Calvete[1,2]; Pablo Garcia-Pavia[3,x,y]; Kunze K.[4] Conxi Lázaro[5]; Teresa Ramón y Cajal[6]; Gaelle Bougeard[7]; Gemma Llort[8]; Victoria Fernandez[1]; Miguel Urioste[2,9]; Javier Benítez[1,2].

[1] Human Genetics Group, Spanish National Cancer Research Center (CNIO), Madrid, Spain.

[2] Center for Biomedical Network Research on Rare Diseases (CIBERER), Madrid, Spain.

[3] Department of Cardiology. Hospital Universitario Puerta de Hierro, Mahadahonda, Madrid, Spain.

[x] CIBER on Cardiovascular Diseases (CIBERCV) Madrid, Spain.

[y] Francisco de Vitoria University, Madrid, Spain.

[4] Department of Pathology, Justus-Liebig-University Giessen, Giessen, Germany.

[5] Hereditary Cancer Program. Catalan Institute of Oncology (ICO-IDIBELL), Hospital Duran i Reynals. L'Hospitalet de Llobregat, Spain.

[6] Medical Oncology Service, Hospital Sant Pau, Barcelona, Spain.

[7] Genetics Department, Rouen University Hospital, France.

[8] Genetic Counseling Unit, Corporació Sanitària Parc Taulí, Barcelona, Spain.

[9] Familial Cancer Clinical Unit, Spanish National Cancer Research Center (CNIO), Madrid, Spain.

Corresponding author

Javier Benitez. Spanish National Cancer Research Center, (CNIO). Melchor Fernández Almagro, 3. 28029, Madrid, Spain. E-mail: jbenitez@cnio.es. Tel.: +34917328057. FAX:

ABSTRACT

The POT1 protein forms part of the shelterin complex, which binds and protects telomeres. Germline mutations in the *POT1* gene have recently been shown to be involved in tumors in different tissues such as familial colorectal, glioma and melanoma tumors, which demonstrate the importance of this gene. Recently, we uncovered a mutation in the *POT1* gene (p.R117C) as causative of cardiac angiosarcomas in families with multiple tumors. Our *in silico* studies predicted that the POT1 p.R117C protein had lost the ability to interact with TPP1 and ssDNA. *In vitro* studies corroborated this prediction, and showed that this lack of function leads to abnormally long telomeres with increased fragility.

In order to better understand the spectrum of mutations in the *POT1* gene and its relation with tumorigenesis, we extended the study to families with multiple tumors (with and without angiosarcomas) and sporadic angiosarcomas and cardiac sarcomas.

We found four new mutations that were not described previously and another patient carrying the previously described p.R117C mutation. *In silico* studies predicted that these new mutations were damaging in the same manner as previously described for the POT1 p.R117C mutation. These mutations were present in both, families and sporadic cases with angiosarcomas and sarcomas, although the major part was involved in families with AS and in cardiac tumors. The wide spectrum of mutations in the *POT1* gene leading to different tumorigenesis processes demonstrates the general importance of this gene.

INTRODUCTION

POT1 (protection of telomeres 1) is a component of the so-called shelterin complex, which binds and protects telomeres (Palm & de Lange 2008). POT1 binds TPP1 from the shelterin complex through a conserved domain located at the C-terminus of POT1; in turn, TPP1 binds the TRF1/2 proteins (Figure 1a). Two other conserved domains of the POT1 protein (oligonucleotide/oligosaccharide-binding 1 and 2; OB1 and OB2) are located at the N-terminus and directly interact with the telomere and the residues p.146 to p.152 of the POT1 protein compose the OB-fold. The stacking residues T1; T2; A3; G4; G5; G6; T7; T8; A9 and G10 interact with the single-stranded (ss) telomeric DNA sequence TTAGGGTTAG (Lei et al. 2004) (Figure 1b).

Mutations in the *POT1* gene have recently been shown to be involved in tumors in different tissues. Germline mutations in the *POT1* gene were described to be responsible for familial glioma (Bainbridge et al. 2014), melanoma (Robles-Espinoza et al. 2014; Shi et al. 2014) and colorectal cancer (Chubb et al. 2016). Somatic mutations in the *POT1* gene were also described to be involved in chronic lymphocytic leukemia (Ramsay et al. 2013). Recently, we identified a deleterious missense germline mutation (rs780936436) in the *POT1* gene (p.R117C) which caused cancer in three families with multiple tumors including cardiac angiosarcomas (CAS) and in one family with breast angiosarcomas (AS) (Calvete et al. 2015) (Figure 1b). A constitutional mutation in the *POT1* gene (R432*) was also found in one out of five sporadic CAS tumors (Kunze et al. 2014) (Figure 1b).

In silico studies suggested that the POT1 p.R117C protein had lost the ability to interact with ssDNA and TPP1 (Figure 1a) (Calvete et al. 2015). *In vitro* studies confirmed these *in silico* predictions and indicated that carriers of this mutation had reduced levels of POT1 bound to the telomere and to the TPP1 protein, which correlated with abnormally long telomeres with increased fragility (Calvete et al. 2015) (Figure 1a). Longer telomeres were also found in carriers of mutations in the *POT1* gene described in melanoma tumors (Robles-Espinoza et al. 2014; Shi et al. 2014).

In order to better understand the role of POT1 not only as telomere protector, but also as one of the main genes responsible for the development of different familial cancer types, we extended our study to 34 families with multiple tumors (10 with and 24 without angiosarcomas) and 30 cases of sporadic angiosarcomas and cardiac sarcomas.

MATERIALS AND METHODS

Patients

A total of 64 patients with different tumors were selected for the whole *POT1* gene study. They were two CAS and 8 AS from 10 families with multiple tumors, 24 families with multiple tumors but without angiosarcomas, 26 sporadic cases of CAS, 2 cardiac sarcomas, and 2 sporadic angiosarcomas that were not cardiac (Supplementary Table 1). Patients were selected from Spain (Spanish National Cancer Research Center, Hospital Puerta de Hierro, Hospital de Sant Pau, Corporació Sanitària Parc Taulí and Institut Català d'Oncologia), Germany (Liebig-University) and France (Rouen University Hospital). The ethics committees of the various institutions approved this study, and written informed consent was obtained from all participants prior to inclusion in the study.

DNA samples

Genomic DNA was isolated from peripheral blood lymphocytes using the FlexiGene DNA Kit (QIAGEN). DNA from normal tissue was extracted from paraffin-embedded tissue samples using the DNeasy® Blood & Tissue Kit (Qiagen, Cat. No. 69504) following the manufacturer's instructions.

POT1 gene study

Sanger sequencing of the entire *POT1* gene was performed in the full series of selected patients. The primers used for Sanger sequencing of the 15 exons are listed in Supplementary Table 2. PCR was performed using standard conditions.

In silico studies

The heat map representation of predicted tolerance to independent amino acid substitutions was made using SNAP2 software (implemented in PredictProtein), which assesses the potential functional impact of the variants (Yachdav et al. 2014). Potential splice site alterations was calculated with Alamut®-Mutation Interpretation Software version 2.3.2 (available at www.interactive-biosoftware.com) and SplicePort (Dogan et al. 2007). Solvent accessibility of amino acids (PACC score) was predicted using the PROFAcc algorithm (Schlessinger & Rost 2005) implemented in ProteinPredict (Yachdav et al. 2014). Protein-binding regions were predicted using the ISIS algorithm (Ofran & Rost 2007).

RESULTS

Sequencing of the entire *POT1* gene in the different series uncovered four new mutations (one nonsense and three missense mutations) that were not described before, and the same mutation described in Calvete et al. (Calvete et al. 2015) (Table 1). The missense mutations were all within functional domains (Figure 1b) and they were considered damaging by the functional predictors and the tolerance to amino acid change score (Table 1).

Regarding the families with multiple tumors including angiosarcomas, two mutations were found in the *POT1* gene. However, no mutations in the *POT1* gene were found in any individual of the 24 studied families with multiple tumors without angiosarcomas. A missense mutation (p.T497L) was found in a breast AS from a French family with multiple tumors (Table 1). Putative protein-protein binding sites were calculated for the putative protein containing the mutation. The protein-protein binding site at position p.499 was lost in this putative protein, as previously observed for the POT1 p.R117C protein. This resulted in the loss of capability of the mutant protein to interact with the TPP1 protein (Supplementary Table 3) (Figure 1b). This prediction was confirmed by *in vitro* assays (Calvete et al. 2015). A truncating mutation (Gln301*) was found in an individual with CAS from a second family (Table 1). The putative POT1 Gln301* protein was not included in the *in silico* study because the entire TPP1 domain is deleted and it is therefore expected to have lost its ability to interact with TPP1.

Regarding the sporadic tumors series, a mutation (p.P116L) was found in an individual with CAS (Table 1). The same protein-protein binding site at position p.499 for TPP1 binding is lost in the putative POT1 p.P116L protein (Supplementary Table 3). In addition, the putative POT1 p.P116L protein changed the orientation from exposed (wt) to buried (PACC score for solvent accessibility) for two residues in the OB-fold (p.152 and p.266), as previously described for the POT1 p.R117C protein (Figure 1b) (Supplementary Table 4). Therefore, the POT1 p.P116L protein is also predicted to have lost its capacity to bind ssDNA. The sporadic tumor series also contained an individual with a cardiac sarcoma carrying the previously described p.R117C mutation and another one with an intronic variant (c.547-1) (Table 1). The intronic variant was located in the splice acceptor site of the 4^{th} intron, and might lead to skipping the acceptor site for splicing out intron 4. The splice acceptor score was calculated for the wild type *POT1* DNA (score: 0.61; acceptance threshold: 0.45) (Supplementary Figure 1). *POT1* c.547-1G>A putatively lost the acceptor site (score: 0.00; acceptance threshold: 0.45), which would result in the skipping of the splice acceptor site. Therefore, removal of 4th intron of *POT1* c.547-1G>A putatively leads to removal of

exon 5 as well (Supplementary Figure 1). The *in silico* studies with the putative POT1 c.547-1G>A protein were performed assuming that the entire 5th exon was lost. In the putative protein containing the *POT1* c.547-1G>A mutation, the site at p.499 for binding TPP1 was also predicted to be lost (Figure 1b) (Supplementary Table 3) (no RNA was available)

In summary, two different mutations in the *POT1* gene were found in families with multiple tumors including angiosarcomas but no *POT1* mutations were found in families without angiosarcomas. In addition, another three mutations in the *POT1* gene were found in three sporadic tumors, one in a sporadic CAS patient and two in sporadic cardiac sarcoma individuals. All putative proteins are expected to lose their ability to interact with TPP1, as previously demonstrated *in vitro* for the POT1 R117C protein (Calvete et al. 2015). Moreover, the putative POT1 p.P116L protein is predicted to have lost its capacity to form the OB-fold as well. No significant accessibility score changes were found for the other putative proteins (Supplementary Table 4).

DISCUSSION

Mutations in the *POT1* gene

Four new predicted damaging mutations (p.R116C p.Gln301*, p.T497L and c.547-1G>A) (Table 1) that appear to be involved in tumorigenesis have been added to the mutation spectrum of the *POT1* gene. *In silico* studies predicted that these new mutations have the same effect on the POT1 protein as described for the POT1 p.R117C mutation (Calvete et al. 2015) (Supplementary Tables 3 and 4). Therefore, these mutant proteins are expected to be unable to bind to TPP1. In addition, the mutation p.P116L is predicted to be defective in its interaction with ssDNA. *In vitro* studies of the POT1 p.R117C protein substantiated the *in silico* prediction (Calvete et al. 2015); therefore, the new mutations described in the current study are expected to deregulate POT1 function in the same manner and to increase telomere length, making them unstable. Carriers of the mutations described in this work are expected to have abnormally long telomeres with increased damage and fragility. Variants in telomere structure and maintenance genes lead to telomere fragility, which is commonly associated with different cancer types (Karami et al. 2016).

POT1 and angiosarcomas

In the present work, mutations in the *POT1* gene were found in 20% of the 10 studied families with multiple tumors including angiosarcomas (cardiac and breast), but no mutations were found in 24 families with multiple tumors without angiosarcomas. Mutations in the *POT1* gene were also found in 3 out of the 30 sporadic tumors (2 in cardiac sarcomas and 1 in CAS) (Table 2a). The mutations in the sporadic cardiac sarcoma individuals demonstrate that the mutations in the *POT1* gene described in this work are not limited to angiosarcomas.

Because cardiac tumors are rare, we added previously published cases (Kunze et al, 2014; Calvete et al, 2015). Therefore, Table 2b includes 12 additional families with multiple tumors with angiosarcomas, four of them had the POT1 p.R117C mutation. Ten additional families with multiple tumors without angiosarcomas did not present *POT1* mutations (Calvete et al. 2015). Regarding sporadic CAS patients, one out of five additional patients presented a truncating mutation (p.R432*) in the *POT1* gene (Kunze et al. 2014). Based on these data, we can conclude that mutations in the *POT1* gene are present in 27.3% and 11.4% of families with multiple tumors including angiosarcomas and sporadic CAS tumors, respectively (Table 2b).

POT1 mutations in other diseases

Other described *POT1* mutations involved in familial glioma and familial melanoma tumors also led to abnormally long telomeres (Robles-Espinoza et al. 2014; Shi et al. 2014; Bainbridge et al. 2014). No information regarding telomere length was available for the colorectal cancer patients carrying *POT1* mutations (Chubb et al. 2016).

All these mutations described in different types of tumors appear randomly distributed along the gene and the conserved domains, independently of the cancer type (Figure 1b). Table 2c summarizes the frequency of alterations in these diseases. Although the incidence of *POT1* mutations is lower than in AS cases, they should be considered especially in familial melanoma (2.4%).

In summary, we observed that mutations in the *POT1* gene are not limited to familial angiosarcomas, but also occur in sporadic angiosarcomas and cardiac sarcomas. *POT1* mutations are described mainly in cardiac tissue pathologies and a putative relation between cardiac tumors and malfunction of telomere biology might exist. The wide spectrum of mutations in the *POT1* gene leading to tumorigenesis in different tissues demonstrates the general importance of this gene. However, the molecular landscape that leads to tumorigenesis is still not well understood.

FIGURE LEGENDS

Figure 1: Shelterin and telomere regulation. a) Above: POT1, which belongs to the shelterin complex, binds telomeres to regulate telomerase activity. Below: Carriers of the POT1 p.R117C mutation show reduced levels of POT1 bound to telomeres and to TPP1, which correlates with abnormally long and fragile telomeres (Calvete et al. 2015). b) POT1 binds TPP1 through the conserved domain located at the C-terminus. The POT1 R117C protein has lost the ability to bind to TPP1 due to the lost of the protein-binding site at position p.499 (white arrow). The OB1 and OB2 domains at the N-terminus directly interact with the telomere. The OB-fold (residues p.146 to p.152) is shown (white rectangle). The stacking residues T1; T2; A3; G4; G5; G6; T7; T8; A9; G10 (grey arrowheads) interact with the ssDNA. In the putative POT1 R117C protein, the PACC solvent accessibility scores of positions p.152 and p.266 (T8; A9) have changed from exposed (wt) to buried, which results in loss of its ability to bind to ssDNA. Mutations in the *POT1* gene found in this work (blue arrowheads), in a previous sporadic CAS (Kunze et al. 2014) (green arrowhead), in melanoma tumors (Robles-Espinoza et al. 2014; Shi et al. 2014) (white arrowheads), in glioma tumors (Bainbridge et al. 2014) (black arrowhead) and in colorectal cancer (Chubb et al. 2016) (red arrowhead) are shown.

Supplementary Figure 1: *In silico* analysis of splice site skipping in the *POT1* c.547-1G>A variant. Above: Splice acceptor site (black triangle) of the 4th intron of wild type (wt) POT1 DNA (score: 0.61; acceptance threshold: 0.45). Below: Skipping of the splice acceptor site (black triangle) of the 4th intron in *POT1* c.547-1G>A DNA (score: 0.00; acceptance threshold: 0.45) that would remove the 5th exon.

REFERENCES

Bainbridge, M.N. et al., 2014. Germline Mutations in Shelterin Complex Genes Are Associated With Familial Glioma. *JNCI Journal of the National Cancer Institute*, 107(1), pp.dju384–dju384.

Calvete, O. et al., 2015. A mutation in the POT1 gene is responsible for cardiac angiosarcoma in TP53-negative Li–Fraumeni-like families. *Nature Communications*, 6, p.8383.

Chubb, D. et al., 2016. Rare disruptive mutations and their contribution to the heritable risk of colorectal cancer. *Nature communications*, 7, p.11883.

Dogan, R.I. et al., 2007. SplicePort-An interactive splice-site analysis tool. *Nucleic Acids Research*, 35.

Karami, S. et al., 2016. Telomere structure and maintenance gene variants and risk of five cancer types. *International Journal of Cancer*, 139(12), pp.2655–2670.

Kunze, K. et al., 2014. A recurrent activating PLCG1 mutation in cardiac angiosarcomas increases apoptosis resistance and invasiveness of endothelial cells. *Cancer Research*, 74, pp.6173–6183.

Lei, M., Podell, E.R. & Cech, T.R., 2004. Structure of human POT1 bound to telomeric single-stranded DNA provides a model for chromosome end-protection. *Nature structural & molecular biology*, 11, pp.1223–1229.

Ofran, Y. & Rost, B., 2007. ISIS: Interaction sites identified from sequence. In *Bioinformatics*.

Palm, W. & de Lange, T., 2008. How shelterin protects mammalian telomeres. *Annual review of genetics*, 42, pp.301–334.

Ramsay, A.J. et al., 2013. POT1 mutations cause telomere dysfunction in chronic lymphocytic leukemia. *Nat Genet*, 45, pp.526–530.

Robles-Espinoza, C.D. et al., 2014. POT1 loss-of-function variants predispose to familial melanoma. *Nature Genetics*, (May 2013).

Schlessinger, A. & Rost, B., 2005. Protein flexibility and rigidity predicted from sequence. *Proteins: Structure, Function and Genetics*, 61, pp.115–126.

Shi, J. et al., 2014. Rare missense variants in POT1 predispose to familial cutaneous malignant melanoma. *Nature Genetics*, (September 2013).

Yachdav, G. et al., 2014. PredictProtein-an open resource for online prediction of protein structural and functional features. *Nucleic acids research*, 42, pp.W337–43.

Acknowledges

Funding