This is a repository copy of Germline CDKN2A/P16INK4A mutations contribute to genetic determinism of sarcoma.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/115743/

Version: Accepted Version

Article:
Jouenne, F, Chauvot de Beauchene, I, Bollaert, E et al. (33 more authors) (2017) Germline CDKN2A/P16INK4A mutations contribute to genetic determinism of sarcoma. Journal of Medical Genetics, 54 (9). pp. 607-612. ISSN 0022-2593

https://doi.org/10.1136/jmedgenet-2016-104402

© 2017, Article author(s) (or their employer(s) unless otherwise stated in the text of the article). All rights reserved. No commercial use is permitted unless otherwise expressly granted. This is an author produced version of a paper published in Journal of Medical Genetics. Uploaded in accordance with the publisher's self-archiving policy.

Reuse
Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
Germline CDKN2A/P16INK4A mutations contribute to genetic determinism of sarcoma

Fanélie Jouenne¹,², Isaure Chauvat de Beauchene³, Emeline Bollaert⁴, Marie-Françoise Avril⁵, Olivier Caron⁶, Olivier Ingster⁷, Axel Lecesne⁶, Patrick Benusiglio⁶, Philippe Terrier¹, Vincent Caumette¹, Daniel Pissaloux⁴, Arnaud de la Fouchardièrè⁶, Odile Cabaret¹, Birama N'Diaye¹, Amélie Velghe⁴, Gaëlle Bougeard⁶, Graham J. Mann¹⁰, Serge Koscielny¹¹, Jennifer H Barrett¹², Mark Harland¹², Julia Newton-Bishop¹², Nelleke Gruis¹³, Remco Van Doorn¹³, Marion Gauthier-Villars¹⁴, Gaëlle Pierron¹⁴, Dominique Stoppa-Lyonnet¹⁴, Isabelle Coupier¹⁵, Rosine Guimbaud¹⁶, Capucine Delnatte¹⁷, Jean-Yves Scoazec¹, Alexander M. Eggermont¹⁸, Jean Feunteun¹⁹, Luba Tchetanov²⁰, Jean-Baptiste Demoulin⁴, Thierry Frebourg⁹, Brigitte Bressac- de Paillerets¹,²

Authors’ Affiliations:

¹Gustave Roussy, Université Paris-Saclay, Département de Biologie et Pathologie Médicales, Villejuif, F-94805, France
²INSERM, U1186, Université Paris-Saclay, Villejuif, F-94805, France
³Physics Department T38, Technical University of Munich, James-Franck-Str. 1, 85748 Garching, Germany
⁴De Duve Institute, Université Catholique de Louvain, 1200 Brussels, Belgium
⁵Department of Dermatology, Assistance Publique-Hôpitaux de Paris, Hôpital Cochin Tarnier, Paris, France; Faculté de Médecine Paris 5 Descartes, Paris, France
⁶Gustave Roussy, Université Paris-Saclay, Département de Médecine Oncologique, Villejuif, F-94805, France
⁷CHU Angers, Service de Génétique médicale, Angers, France
⁸Centre Leon Berard, Department of Pathology, Lyon, France
⁹INSERM, U1079, Faculty of Medecine, Normandy University and Departement of Genetics, Rouen University Hospital, Normandy Centre for Genomic and personalized Medicine
Corresponding Author : Brigitte Bressac- de Paillerets, Département de Biopathologie, Gustave Roussy, 114 rue Edouard Vaillant, 94805 Villejuif. Phone: 33-1-42115490; Fax : 33-1-42115267 ; Email : brigitte.bressac@gustaveroussy.fr
Abstract

Background Sarcomas are rare mesenchymal malignancies whose pathogenesis is poorly understood; both environmental and genetic risk factors could contribute to their etiology.

Methods and results We performed whole-exome sequencing (WES) in a familial aggregation of 3 individuals affected with soft-tissue sarcoma (STS) without TP53 mutation (Li-Fraumeni-Like, LFL) and found a shared pathogenic mutation in CDKN2A tumor suppressor gene. We searched for individuals with sarcoma among 474 melanoma-prone families with a CDKN2A-/+ genotype and for CDKN2A mutations in 190 TP53-negative LFL families where the index case was a sarcoma. Including the initial family, 8 independent sarcoma cases carried a germline mutation in the CDKN2A/p16INK4A gene. In 5/7 formalin-fixed paraffin-embedded (FFPE) sarcomas, heterozygosity was lost at germline CDKN2A mutations sites demonstrating complete loss of function. As sarcomas are rare in CDKN2A/p16INK4A carriers, we searched in constitutional WES of 9 carriers for potential modifying rare variants and identified three in platelet-derived growth factor receptor (PDGFRA) gene. Molecular modeling showed that two never-described variants could impact the PDGFRA extracellular domain structure.

Conclusion Germline mutations in CDKN2A/p16INK4A, a gene known to predispose to hereditary melanoma, pancreatic cancer and tobacco-related cancers accounts also for a subset of hereditary sarcoma. In addition, we identified PDGFRA as a candidate modifier gene.
SHORT REPORT

Sarcomas are a complex group of rare malignant tumors derived from cells that originate from the mesenchyma. These tumors, which can affect both bone and soft tissue, include more than 50 different subtypes. The annual incidence of soft tissue sarcomas (STS) is around 5 new cases per 100,000 population, whereas it is 0.8 for bone sarcomas, in Caucasians. They account for nearly 20% of all pediatric solid malignant cancers, but less than 1% of all adult solid malignant cancers. The pathogenesis of most sarcomas is still poorly understood and both environmental and genetic risk factor could contribute to their etiology. The main environmental factors are carcinogens, viruses, and ionizing radiation, particularly radiation therapy received for a first cancer. The risk of sarcoma is enhanced in several hereditary cancer syndromes, including Li-Fraumeni syndrome (LFS), a rare, dominant Mendelian cancer syndrome linked to TP53 mutations but also possibly to POT1 mutations. Beyond these syndromes, there may be other complex heritable predispositions as well as others, not yet identified.

The potential for intrafamily exome-sequencing approach to identify additional cancer susceptibility genes has been demonstrated. Therefore, we conducted germline whole exome sequencing (WES) in 2 affected members of a three sarcoma-cases family (Patients I-2 and II-1, family 7389, Table 1, figure 1A). We performed data mining applying the classical filtering strategies provided in Ingenuity Variant Analysis (IVA) software (Qiagen). With very stringent frequency filtering (MAF) <0.001%, using a Biological Context of sarcoma, three germline variants shared by both sarcoma-affected relatives (uncle and nephew), were identified in CDKN2A, PDGFRa and SKA3 genes. Because of the loss of function mutation detected in CDKN2A and the well known role of CDKN2A in somatic sarcomagenesis, both in humans and mice, we focused first on this gene. CDKN2A is a known tumor suppressor gene and the first familial melanoma gene identified; it encodes two distinct proteins, p16INK4a and p14ARF, which both function in cell cycle regulation. We
confirmed the germline splice mutation (c.151-2A>G) with Sanger sequencing, and also in DNA extracted from formalin-fixed, paraffin-embedded (FFPE) tumoral tissue from the 3d case, deceased patient I-1 (figure S2A). We had previously identified this specific mutation in three independent, melanoma-prone families. Transcript analysis was performed for a proband, indicating that CDKN2A exon 2 had been skipped in both p16<sup>INK4A</sup> and p14<sup>ARF</sup> transcripts, creating putative frameshifts (figure S3).

Next, we performed Sanger sequencing of CDKN2A for germline mutations in full collection A (190 unrelated families with suspected LFS or Li-Fraumeni like, LFL, whose index case was a sarcoma without detectable TP53 germline mutation). We identified a second carrier of a CDKN2A/P16<sup>INK4A</sup> germline mutation (p.Ile49Ser), a patient diagnosed with a pleomorphic liposarcoma at age 32 years whose brother died of osteosarcoma at age 24 years (Family 18998, Table 1, figure S1A & S2B).

To explore further the potential connection between CDKN2A germline mutations and sarcoma, we reviewed the phenotypes in our collection of 296 melanoma-prone French families with CDKN2A/P16<sup>INK4A</sup> germline mutations (collection B; mutations were partially described previously<sup>10</sup>) and found 8 kindreds that contained at least one member with sarcoma. Among them, 5 probands with sarcoma carried the pathogenic familial CDKN2A/P16<sup>INK4A</sup> germline mutation (Table 1, families 14288, 14291, 2225, 15118 and 14289, figure S1B) and three families had incomplete data (two untested index cases; one unconfirmed STS; figure S4A and C; material and methods). Overall, among the 296 families, the difference in sarcoma incidence between CDKN2A mutation carriers (5/593; 0.84%; 95% confidence interval: 0.3%-2%) and non-carriers (1/298; 0.34%; 95%CI 0.02%-2.16%) did not reach statistical significance (p-value =0.67; Fisher’s exact test). Considering the yearly incidence in Caucasians of 5.8 per 100,000<sup>1</sup> and the mean follow-up duration in collection B of 46 years, the probability of observing at least 5 sarcomas in the 593 CDKN2A carriers population was 0.02 (assuming a binomial distribution). In the 298 CDKN2A WT populations, the mean follow-up was 39 years and the probability of observing at least 1 sarcoma was 49%.
Next, we searched for biological arguments. As loss of heterozygosity (LOH) is considered in tumor’s biology as a strong indicator to the causative role of a tumor suppressor, we performed Sanger sequencing in seven FFPE sarcoma blocs available from French patients. We identified LOH at the CDKN2A germline mutation site in 5/7 samples (Table 1; figure S2). These LOHs demonstrate the occurrence of a second genetic hit on CDKN2A and, therefore, complete loss of p16\(^{INK4A}\) function in 5 sarcomas, in accordance with the driver role of CDKN2A tumor suppressor gene in sarcomagenesis\(^8\).

Finally, we interrogated the GenoMEL database containing 178 CDKN2A+ melanoma-prone families (collection C), after removal of 60 French families already included in collection B. We identified three additional independent CDKN2A mutation carriers affected with a sarcoma. One family was from Australia and carried a CDKN2A/p16\(^{INK4A}\) p.Leu32Pro germline mutation (Family 20473, Table 1, figure S1C). The second family was from the UK (21 kb deletion targeting CDKN2A/p14\(^{ARF}\) exon 1b)(data not shown), and the third family was from the Netherlands but the initial diagnosis of fibrosarcoma case was revised to melanoma and therefore, was excluded.

Overall, in collections A, B, and C, ascertained for Li-Fraumeni (A) or multiple cases of melanoma and/or pancreatic cancer (B and C), we identified eight independent families in which at least a CDKN2A/P16\(^{INK4A}\) mutation carrier, had a sarcoma (Table 1). Therefore, based on probabilistic and biological arguments, CDKN2A/P16\(^{INK4A}\) germline mutations can be strongly suspected to increase sarcoma risk. Interestingly, in the literature, two sarcoma cases in CDKN2A /P16\(^{INK4A}\) mutation carriers were identified in families with melanoma/pancreatic cancer and very recently germline CDKN2A mutations were identified in two independent patients presenting with Li-Fraumeni syndrome.\(^{11,12}\) In addition to the well known role of CDKN2A in somatic sarcomagenesis, other observations in animals suggested a germline effect.\(^8\) First, in a mouse model, deletion of the Cdkn2a locus could substitute for mutations in Trp53 to generate soft-tissue sarcomas.\(^{13}\) Second, in a naturally occurring, canine breed-specific histiocytic sarcoma, a genome-wide association study (GWAS) identified a haplotype near CDKN2A.\(^{14}\) In conclusion to our work and published data,
germline mutations in $CDKN2A/P16^{INK4A}$, a gene known to predispose to hereditary melanoma, pancreatic cancer and tobacco-related cancers, accounts also for a subset of hereditary sarcoma.\(^9\)

As melanoma risk in $CDKN2A$ mutation carriers is clearly associated with $MC1R$ frequent alleles acting as modifiers, \(^{15}\) we formulated the hypothesis that the very low frequency of sarcoma cases observed in $CDKN2A/P16^{INK4A}$ positive melanoma-prone families could be explained by rare modifiers alleles. In a model of oligogenic inheritance, it is challenging to identify rare germline variants that act in synergy to initiate cancer and GWAS are unable to identify rare disease-predisposing variants.\(^{16}\) Candidate pathogenic variants for sarcoma risk in $ATM$, $ATR$, $BRCA2$ and $ERCC2$ genes were identified recently in a large sarcoma case control study, as well as $POT1$ variants in cardiac angiosarcoma, but other genes not yet identified could also play a role.\(^5\)\(^6\) To explore this hypothesis, we considered the two additional germline variants identified in $PDGFRA$ and $SKA3$ genes in the WES data.\(^6\) In $SKA3$ gene, an insertion of 2T was supposed to have occurred in a stretch of 12 T but was unconfirmed by Sanger sequencing (figure S5). The platelet-derived growth factor receptor alpha gene ($PDGFRA$) harbored a germline missense mutation, c.335T>G, p.Leu112Arg, located in the extracellular receptor domain and predicted deleterious by 2 computational methods (GVGD and SIFT). This mutation was verified by Sanger sequencing and was also found in DNA extracted from FFPE-sarcoma tissue from the third family member, patient I-1, therefore being present in the 3 sarcoma-affected patients (Family 7389, Figure 1A & S7A).

Next, we performed additional WES analyses in blood-extracted DNA from 7 probands affected with sarcoma that carried germline $CDKN2A$ mutations (14288-II.1, 14289-I.1, 2225-II.1, 14291-II.1 and 15118-II.1 in collection B; family 20473-I.4 in collection C; and 18998-II.1 in collection A). Subsequently, we data mined the WES available for a total of 9 $CDKN2A/P16^{INK4A}$ carriers affected with sarcoma, including 2 relatives. We applied the classical filtering strategies provided in IVA software (Qiagen) (figure S6).\(^7\) For variant frequency, we defined rare variants as those with a minor
allele frequency (MAF) <0.5%. The outcome of our filtering strategy was the selection of 82 variants spanning 76 genes. Among previously published sarcoma susceptibility genes, we found no mutations in TP53, ATR, BRCA2 and ERCC2. We found a c.8584+1G>A putative splice site mutation in ATM gene in patient 7389-I.2, but this variant was absent in the sarcoma affected relative, II.1. We also found, in patient 14291-II.1, a POT1 c.1127A>G, p.Gln376Arg missense variant, present at a frequency of 0.07% in Eur-Am ESP, and predicted deleterious by 4 prediction methods (SIFT, MutationTaster, Polyphen 2 and Condel). This variant was also present in the unaffected mother. More interestingly, we detected 2 other germline missense mutations (verified by Sanger sequencing, figure S7B & C) located in the extracellular receptor domain of the platelet-derived growth factor receptor alpha gene (PDGFRA), including one absent in public databases. The PDGFRA missense variant c.227A>G, p.Asn76Ser, predicted deleterious by 4 computational methods (GVGD, SIFT, Mutation Taster and Polyphen 2) was not present in unaffected mother that carried the CDKN2A p.Gly101Trp mutation (Family 14291, figure S1B). In the sarcoma-proband I-4 of family 20473 (figure S1C), we identified another germline PDGFRA variant, c.1388C>G, p.Thr463Ser, described with an allelic frequency of 0.02%, and predicted deleterious by 2 computational methods (Mutation Taster and Condel). Co-segregation analysis was not informative (figure S1C).

The PDGFRα, composed of extracellular, trans-membrane, and intracellular domains (figure 1C) is activated by the binding of its ligand, which induces dimerization, followed by kinase domain activation. Germline oncogenic gain-of-function mutations in PDGFRA cause familial gastrointestinal stromal tumors (GIST) associated with other tumors. Accordingly, the variants described above were not oncogenic in classical cell transformation assays (data not shown). Nevertheless, these variants could favor sarcomagenesis by interfering with various PDGFRA molecular functions, either canonical or not. To study the impact of PDGFRA germline variants on the 3D receptor structure, we performed molecular modeling of 3 PDGFRA missense variants identified in CDKN2A carriers with sarcoma, the two variants absent from public databases, p.Asn76Ser (N76S), p.Leu112Arg (L112R), and the rare variant, p.Thr463Ser (T463S) (ESP Eur. Am.
0.02%). We added as a control, a frequent SNP, p.Ser478Pro (S478P) described with an allelic frequency of 10.26% (ESP Eur. Am.) and predicted neutral by 5 computational methods (GVGD, SIFT, Mutation Taster, Polyphen 2 and Condel), identified in patient 14288-II.1 and 14289-I.1 (figure S7D).

As the PDGFRα signaling complex has remained uncharacterized at the structural level, we modeled two extracellular immunoglobulin (Ig)-like domains (D1 and D5; figure 1C) containing these variants by homology with related domains. Structurally, all these domains feature five to eight β-strands that form two β-sheets (a β-sandwich). Figure 1C illustrates how the variants N76S and L112R affect the structure of D1. In particular, N76S promoted larger β-strands fold (β3 and β4) prior and after the mutation site, contributing to stabilization of a perfect antiparallel β-sheet, constituted with β1, β3, and β4 strands and maintained by a regular, stable H-bond network that contrasted with the fluctuating network in the native protein. Moreover, this variant promoted destabilization of two small β-strands (β2 and β5) that were present in the native protein. Variant L112R induced β-strand (β5) formation in place of the random coil rather observed in the native protein and increased β-folding in segments more distant from the mutation point (β strands β1, β2, and β4). Our analysis of the impact of T463S and S478P variants in the D5 domain suggested only a slight increase in residual flexibility, but all its structural features were well-preserved with respect to the native protein. It should be noted that a comprehensive characterization of PDGFRα variants located in the extracellular domains may require detailed analysis of the full-length protein structure in the native and mutated states.

Overall, our data identified PDGFRα as a new sarcoma candidate modifier gene. Unfortunately, PDGFRα was not included in the 72 genes panel studied in the recent study of 1162 patients with sarcoma.² PDGFRα belongs to the large family of membrane RTKs and play primary roles in mesenchymal tissue development. Recent whole-genome or whole-exome analyses have revealed numerous somatic mutations localized in the RTK-III extracellular domain which could have transforming potential, based on their structural and physicochemical effects on the receptor.²¹ These mutations in PDGFRα extracellular domains could affect non-canonical RTK functions. Upon
ligand activation, RTKs are internalized and translocated into endosomal compartments for signaling.

Overall, our genetic and molecular modeling results suggested that PDGFRA germline variants that affect the extracellular domain could play a role in sarcomagenesis, but the functional mechanism remains unknown.

Acknowledgements The authors would like to thank the following individuals for useful discussions: Sebastien Forget, Nabila Bouatia-Naji, Soto Romuald Kiando, Corine Bertolotto, Alicia Goldstein, Peter Kanetsky and Céline Lefebvre. The authors are grateful to Dr Stéphanie Baert-Desurmont, Dr Aurore Coulomb for providing sarcoma specimens, to Elisabeth Holland for providing DNA samples. The authors thank Florence Demenais and her collaborators, Hamida Mohamdi and Eve Maubec for the establishment of melanoma-prone families (Collection B). This collection constitutes also 20 years of contributions from a French Network of Dermatologists and Oncogeneticists, including, in particular for this study: Dr Caroline Abadie, Dr Pascale Andry-Benzaken, Pr François Aubin, Dr Séverine Audebert, Pr Philippe Bahadoran, Dr Emmanuelle Barouk-Simonet, Dr Pascaline Berthet, Dr Françoise Boitier, Dr Valérie Bonadona, Pr Jean-Marie Bonnetblanc, Dr Marie Noëlle Bonnet-Dupeyron, Dr Virginie Bubien, Dr Jean Chiesa, Dr Marie-Agnès Collonge-Rame, Pr Stephane Dalle, Pr François Eisinger, Dr Sandra Fert-Ferrer, Dr Jean-Pierre Fricker, Dr Paul Gesta, Dr Damien Giacchero, Pr Brigitte Gilbert-Dussardier, Dr Sophie Giraud, Pr Florent Grange, Pr Jean-Jacques Grob, Pr Bernard Guillot, Dr Ewa Hainault-Wierzbicka, Pr Pascal Joly, Dr Christine Lasset, Pr Pierre Laurent Puig, Dr Marine Lebrun, Dr Sophie Lejeune, Pr Dominique Leroux, Dr Jean Marc Limacher, Pr Dan Lipsker, Dr Michel Longy, Dr Alain Lortholary, Dr Sandrine Mansard, Dr Ludovic Mansuy, Dr Véronique Mari, Dr Ludovic Martin, Dr Tanguy Martin Denavit, Dr Christina Mateus, Dr Michèle Mathieu, Pr Eve Maubec, Dr Christine Maugard, Pr Nicolas Meyer, Dr Gwénaël Nadeau Pr Laurence Olivier-Faivre, Dr Philippe Parent, Dr Jean-Luc Perrot, Dr Gabriella Pichert, Dr Nicolas Poulalhon, Pr Caroline Robert, Pr Hagay Sobol, Pr Luc Thomas, Pr Pierre Vabres and Dr Hélène Zattara. Finally, we thank Peter Kanetsky for management of GenoMEL melanoma-prone families’ database set up and curation.
Contributors BB-deP, J-BD, LT and TF designed the study; TF, OC, OI, M-FA, AL, PB, PT, DP, AdlaF, OC, IC, GJM, MH, JN-B, NG, RVD, RG, CD, MG-V, GP, DS-L provided patients clinical data and samples. FJ, EB, VC, IC-de B, GB, AV performed the experiments. FJ, IC-deB, BN’D, BB-deP, J-BD, LT, JB were involved with data analyses and interpretation. FJ, BB-deP, JF, LT, JB-D, AE, PB, M-FA, TF, J-YS wrote, reviewed the manuscript.

Funding This work was supported by an INCA grant 2013-1-MELA-05 and personal donations from C. and N. de Paillerets and M.-H. Wagner, awarded to B.B.- d.P. The work of R.v.D and N.A.G. was supported by the Dutch Cancer Society (UL 2012-5489) and grant # R01 CA83115 from the National Cancer Institute to GenoMEL international consortium.

Competing interests B. Bressac-de Paillerets is an inventor on the MITF patent which is not licensed. No potential conflicts of interest were disclosed by the other authors.
Table 1: CDKN2A/P16INK4A germline mutations identified in eight families with members affected by sarcoma and candidate modifiers

<table>
<thead>
<tr>
<th>Collections</th>
<th>Family ID</th>
<th>Patient ID</th>
<th>Cases Clinical context</th>
<th>P16 exon</th>
<th>p16&lt;sup&gt;INK4A&lt;/sup&gt; mutation&lt;sup&gt;a&lt;/sup&gt;</th>
<th>p14&lt;sup&gt;APC&lt;/sup&gt; AA change</th>
<th>Sarcoma LOH at CDKN2A/P16&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Melanoma/ pancreatic cancer reports</th>
<th>P16INK4A Loss of function</th>
<th>PDGFRA variants</th>
<th>Other candidate modifiers&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>7389</td>
<td>I-1</td>
<td>LS (57)</td>
<td>IVS 1</td>
<td>c.151-2A&gt;G</td>
<td>c.194-2A&gt;G</td>
<td>No</td>
<td>Yes (unpublished data)</td>
<td>Splice</td>
<td>p.Leu112Arg</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I-2</td>
<td>AS (67)</td>
<td>IVS 1</td>
<td>c.151-2A&gt;G</td>
<td>c.194-2A&gt;G</td>
<td>Yes</td>
<td></td>
<td>p.Leu112Arg</td>
<td></td>
<td>None</td>
</tr>
<tr>
<td></td>
<td></td>
<td>II-1</td>
<td>STS (33)</td>
<td>IVS 1</td>
<td>c.151-2A&gt;G</td>
<td>c.194-2A&gt;G</td>
<td>Yes</td>
<td></td>
<td>p.Leu112Arg</td>
<td></td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>II-2</td>
<td>OS (24)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>None</td>
</tr>
<tr>
<td>B</td>
<td>1428</td>
<td>II-1</td>
<td>OS (16)</td>
<td>2</td>
<td>c.194T&gt;C, p.Leu65Pro</td>
<td>p.Ala79Ala</td>
<td>Yes</td>
<td>Yes (Landi MT., 2004)</td>
<td>Yes (Landi MT., 2004)</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>None</td>
</tr>
<tr>
<td>B</td>
<td>2225</td>
<td>II-1</td>
<td>UCS (43)</td>
<td>2</td>
<td>c.225_243del, p.Ala76fs</td>
<td>p.Arg90fs</td>
<td>Yes</td>
<td></td>
<td>ND</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>B</td>
<td>1428</td>
<td>I-1</td>
<td>FS (69)</td>
<td>2</td>
<td>c.301G&gt;T, p.Gly101Trp</td>
<td>p.Arg115L</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>I-1</td>
<td>FS (69)</td>
<td>2</td>
<td>c.301G&gt;T, p.Gly101Trp</td>
<td>p.Arg115L</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>II-1</td>
<td>OS (13)+ MM (26,27)</td>
<td>2</td>
<td>c.301G&gt;T, p.Gly101Trp</td>
<td>p.Arg115L</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td>None</td>
</tr>
<tr>
<td>B</td>
<td>1511</td>
<td>II-1</td>
<td>SVS (57)</td>
<td>2</td>
<td>c.301G&gt;T, p.Gly101Trp</td>
<td>p.Arg115L</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>I-2</td>
<td>OS (7)+ 8MM RMS (25)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>None</td>
</tr>
</tbody>
</table>

<sup>a</sup> All subjects provided written informed consent for participation in these oncogenetic research studies which were approved by local research ethics committees. Collection (A) comprised 190 families with suspected Li-Fraumeni syndrome that included at least one member with sarcoma. Collection (B) comprised 300 melanoma-prone French families CDKN2A/p16INK4A+ Collection (C) comprised 250 CDKN2A/p16INK4A+ melanoma-prone families from the international GenoMEL database.

<sup>b</sup> Clinical context: LS, liposarcoma; AS, angiosarcoma; STS, soft-tissue sarcoma; OS, osteosarcoma; UCS, uterine carcinosarcoma; FS, fibrosarcoma; SVS, synovialosarcoma, MM, cutaneous melanoma; ages (y) at diagnosis appears in parentheses.

<sup>c</sup> CDKN2A/P16<sup>INK4A</sup> exons 1alpha, 2, and 3 were sequenced by Sanger method based on transcript NM_000077.4 to screen for mutations in collections A, B, and C. None were present in public controls databases such as 1000 genomes or ExAC.

<sup>d</sup> LOH analyses were performed by Sanger sequencing at germline CDKN2A mutations sites in DNA extracted from sarcoma FFPE samples; NA, not analyzed.
REFERENCES


Figures legends

Figure 1. Whole exome sequencing in a 3-sarcoma cases family without TP53 germline mutation: identification of CDKN2A and PDGFRA germline mutations, co-segregating with sarcomas

A, Pedigree of the Li-Fraumeni-like family. Cancer diagnosis and age at onset is indicated for affected members; hatched circles/squares indicate sarcoma: AGS, angiosarcoma; LPS, liposarcoma; STS, soft-tissue sarcoma. Genotypes of CDKN2A and PDGFRA for all samples available for testing are shown. Patients with WES data are indicated with a black star.

B, WES germline SNV filtering and interpretation, for 2 patients of Family 7389. We used Ingenuity Variant Analysis software (v.2.1.20130711, IVA, Qiagen) and predetermined filters (see Bioinformatics analysis, supplementary online). Starting with 307 690 variants spanning 17 673 genes, successive filters lead to 3 variants spanning 3 genes (CDKN2A, PDGFRA and SKA3).

C, Structural properties of PDGFRα wild-type and variants. (Upper row) the PDGFRα protein has a modular structure composed of five Ig-like domains (D1, D2, D3, D4, and D5), a trans-membrane domain (TMD), and a cytoplasmic region. The cytoplasmic region consists of a regulatory juxtamembrane region (JMR) and a catalytic kinase domain, with a N-lobe and a C-lobe, which harbors a kinase insert domain (KID). (Lower row) The X-ray analysis structures are represented as ribbon diagrams, based on the KIT structural data. D1 and D5 are denoted as ovals. (Middle row) schematic representations of D1 and D5 topologies. (Bottom figures) superimposed conformations of wild-type PDGFRα (blue) and PDGFRα germline variants (pink), obtained from molecular dynamics (MD) simulations. Representative conformations were selected by RMSDs clustering and are presented as ribbon diagrams.