Germline Variants in the POT1-Gene in High-Risk Melanoma Patients in Austria

Christoph Müller,* Milica Krunic,† Judith Wendt,* Arndt von Haeseler,†,‡ and Ichiro Okamoto*,†

*Department of Dermatology, Medical University of Vienna, Austria, †Center for Integrative Bioinformatics Vienna, Max F. Perutz Laboratories, University of Vienna, Medical University Vienna, Austria, and ‡Bioinformatics and Computational Biology, Faculty of Computer Science, University of Vienna, Austria

ORCID IDs: 0000-0002-5031-7255 (C.M.); 0000-0003-2385-7122 (I.O.)

ABSTRACT Risk of melanoma is in part determined by genetic factors. Currently the only established high penetrance familial melanoma genes are CDKN2A and CDK4. Recent studies reported germline variants in POT1 in melanoma families. In the present study, we sequenced the entire POT1 gene in 694 patients from the M3-study. Patients with multiple primary melanomas (n = 163) or with a positive family history (n = 133) were classified as high-risk melanoma patients. Additionally, 200 single primary melanoma patients and 198 non-melanoma controls were sequenced. For prediction analysis 10 different tools were used.

In total 53 different variants were found, of which 8 were detected in high-risk melanoma patients, only. Two out of these 8 variants were located in exons and were non-synonymous: g.124510982 G > A (p.R80C) and g.124491977 T > G (p.N300H). While g.124491977 T > G was predicted to be neutral, 80% of the prediction tools classified g.124510982 G > A as deleterious. The variant, g.124467236 T > C, which possibly causes a change in the splice site was identified in a case with a positive family history in the present study. Another variant in the 5-UTR, g.124537261 A > G, was found in 2 high-risk patients. So, in conclusion, melanoma associated POT1 germline variants seem to be rare. Further studies are required to evaluate the role of POT1 for genetic counseling.

KEYWORDS POT1 melanoma familial genetics Austria

Approximately 10–15% of all melanoma patients report a positive family history, multiple primary melanomas or early onset of melanoma diagnosis (Müller et al. 2016). The most important high penetrance gene is the cyclin-dependent kinase Inhibitor 2A (CDKN2A), responsible for about 30% of all familial melanoma cases. Melanoma associated mutations in cyclin-dependent kinase 4 (CDK4), which were also classified as high penetrance mutations, seem to be very rare as only a few families were reported since the initial report in 1996 (Zuo et al. 1996). Only recently, a mutation in the telomerase reverse transcriptase gene (TERT) was described in melanoma patients, adding further data to the already existing evidence that stability of telomeres is important in melanoma biology.

Shelterin, a protein complex composed of six subunits, is involved in the protection of the chromosome ends and in the regulation of the telomerase activity (Aoude et al. 2015). Recently this complex gained particular interest in melanoma genetics as germline variants were found in 3 shelterin genes in melanoma prone families (Robles-Espinoza et al. 2014; Shi et al. 2014; Aoude et al. 2015): POT1, ACD and TERF2IP. The human POT1 gene is located at 7q31.33 and has 19 transcripts. The isoform 1 of the protein, where the variants were originally found, consists of 19 exons and of 634 amino acids. Since the initial description of POT1 as a predisposition gene for hereditary melanoma (Robles-Espinoza et al. 2014; Shi et al. 2014), no further variants associated with melanoma has been described except for one in a single melanoma prone family in the U.S.A. (Wilson et al. 2017). Therefore, the frequency of these variants in other populations remains unclear. This information is crucial to decide whether high-risk patients should be tested for POT1 in a routine genetic counseling of melanoma families (Goldstein et al. 2007). Here we present for the first time data of POT1 variants in high-risk melanoma patients in Austria.
PATIENTS AND METHODS

Study participants
In total, DNA of 694 participants was analyzed. All participants were Caucasians with European ancestry and were recruited in Austria as described elsewhere (BURGSTALLER-MUEHLBACHER et al. 2015). High-risk melanoma patients (n = 296), included patients with multiple primary melanomas (n = 163) and patients with a positive family history (n = 133) and were compared to a reference group of single melanoma (n = 200) and non melanoma patients (n = 198). Descriptive data were shown for gender, age at diagnosis, Breslow index, tumor localization and histological subtype in Table 1. In multiple primary melanoma patients, data (date of surgery, localization, histological description such as histological subtype and Breslow index) refers to the patient, data (date of surgery, localization, histological description such as histological subtype and Breslow index) refers to the first primary melanoma. Informed consent was obtained from all individual participants included in the study. The study was approved by the ethics committee of the Medical University of Vienna.

Genotyping
The DNA was purified from whole blood as described previously (BURGSTALLER-MUEHLBACHER et al. 2015). Next generation-sequencing of POT1 was performed at the Genome Centre, Queen Mary, University of London (http://www.smd.qmul.ac.uk/gec/). For the preparation of DNA libraries 0.5 μg of genomic DNA was used. Amplicon libraries were created with the Fluidigm Access Array according to the manufacturer’s protocol. The 150-bp paired-end sequencing was done on the Illumina MiSeq v2 platform.

The datasets generated during the current study are available in the NCBI Sequence Read Archive (https://www.ncbi.nlm.nih.gov/sra).

Data analysis
The reads were mapped against human genome reference (hg19) using NextGenMap (SIDLACEK et al. 2013) (v0.5.0) with default parameters plus several additional options: identity (-i) was set to 0.85, maximum number of consecutive indels allowed (-C) was set to 120 and we used alignment algorithms that support affine gap costs (-affine). Read groups in aligned reads (BAM files) were replaced using Picard tools (http://broadinstitute.github.io/picard) option AddOrReplaceReadGroups. The aligned reads were then indexed using SAMtools (Li et al. 2009) (v1.1). Local realignment around insertions and deletions and quality base score recalibration were performed using the Genome Analysis Tool Kit (MCKENNA et al. 2010) (GATK, v2.6). To call variants (SNPs and indels) in aligned reads files, we used UnifiedGenotyper from GATK with parameters: -dcov set to 2000, -standard_min_confidence_threshold_for_calling set to 30.0, -standard_min_confidence_threshold_for_emitting set to 10, -glm set to BOTH and for option -dbSNP we used human_9606 variants from dbSNP database (SHERRY et al. 2001). GATK called variants were first divided into SNPs and indels using SelectVariant. SNPs to be filtered out were labeled using VariantFiltration with the following filter expressions: "cluster WindowSize = 10, "MQ0 >= 4 && ((MQ0 / (1.0 * DP)) > 0.1), "DP < 5", "QUAL < 30.0 QUAL > 30.0 && QUAL < 50.0", "QD < 0.8", "FS > 60.0" and -missingValuesInExpressionsShouldEvaluateAsFailing. Indels to be filtered out were labeled using VariantFiltration with the filter expressions: "QD < 2.5 || ReadPosRankSum < 20.0 || FS > 200.0", "-missingValuesInExpressionsShouldEvaluateAsFailing". The variants were then combined by GATK CombineVariants tool. Rearranging results was done using our in-house developed python and R scripts.

Prediction analysis of non-synonymous POT1 variants
Two non-synonymous POT1 variants, found in high risk melanoma patients only, were analyzed using 10 prediction tools as described previously (BURGSTALLER-MUEHLBACHER et al. 2015; Müller et al. 2016): MutationTaster2 (SCHWARZ et al. 2014), PolyPhen-2 (Polymorphism Phenotyping-v2, HumDiv and HumVar) (ADZHUBEI et al. 2010), PROVEAN (Protein Variation Effect Analyzer) (CHOI et al. 2012), SIFT (sorts intolerant from tolerant substitutions) (NG and HENKES 2001), SNAP2 (screening for non-acceptable polymorphisms-2) (BROMBERG and ROST 2007), PANTHER (Protein ANalysis THrough Evolutionary Relationships) (Xu et al. 2013), CADD (Combined Annotation Dependent Depletion) (KIRCHER et al. 2014), GERP++ (DAVYDOV et al. 2010) and phyloP (POLLARD et al. 2010). For the latter 2, the tables of the UCSC Genome browser th_allHg19RS_BW and phyloP46wayPlacental were used. Most of those tools provide information about the effect of an amino acid exchange on the protein function. GERP++ and phyloP give a score depending on the conservation by comparing different species. The cut-off score for PROVEAN was -2.5, values below indicate the prediction as deleterious. In SIFT, values have a range from 0 to 1, whereas a score below 0.05 means that the variant is predicted to

Table 1 Participant characteristics

| Gender | Controls | SPM | PFH | >1 PM |
|--------|----------|-----|-----|-------|
| female | 74       | 80  | 60  | 55    |
| male   | 124      | 120 | 73  | 108   |
| Mean age (SD) | 53.8 (15.8) | 52.6 (16.4) | 49.7 (15.9) | 53.9 (15.1) |
| Missing | 0        | 2   | 1   | 0     |
| Mean Breslow in mm (SD) | — | 1.4 (1.9) | 1.1 (1.3) | 1.1 (1.3) |
| Localization | — | 10/4 | 4/3 | 9/0 |
| Head and Neck | — | 15 | 16 | 24 |
| Upper Extremity | — | 21 | 12 | 13 |
| Trunk | — | 115 | 72 | 85 |
| Lower Extremity | — | 43 | 30 | 40 |
| Missing/Occult | — | 24/0 | 3/0 | 1/0 |
| Histological Subtype | — | 8 | 9 | 19 |
| LM/LMM | — | 38 | 19 | 25 |
| NMM | — | 87 | 74 | 74 |
| SSM | — | 67 | 31 | 45 |

SD: standard deviation; LM: lentigo maligna; LMM: lentigo maligna melanoma; NMM: nodular melanoma; SSM: superficial spreading melanoma; SPM: single primary melanoma; PFH: patients with a positive family history; >1 PM: patients with multiple primary melanomas.
be deleterious. In CADD values above 15 were classified as deleterious. The range Polyphen2 scores is from values of 0 to 1; higher scores are more likely to be found in deleterious variants with a cut-off score of 0.5. SNAP 2 has output scores between -100 (strong neutral prediction) to 100 (strong effect prediction). PANTHER calculates the preservation time to give a prediction. Longer times indicate a more likely functional impact.

As protein sequence for the data input, the POT1 isoform 1 (ENST00000357628) was used.

**Data availability**

All raw sequencing data are deposited in the NCBI Sequence Read Archiv (SRA) under the BioProject ID PRJNA400454.
### Table 3 High risk patients and melanoma characteristics

| Variant       | dbsSNP         | Carrier  | No. of primaries | 1st melanoma Age/Breslow/Localization | 2nd melanoma Age/Breslow/Localization | 3rd melanoma Age/Breslow/Localization | 4th melanoma Age/Breslow/Localization | CDKN2A status | Family history of melanoma |
|---------------|---------------|----------|------------------|--------------------------------------|--------------------------------------|---------------------------------------|--------------------------------------|---------------|-----------------------------|
| g.124537261 A>G | rs202009081   | PFH      | 1                | 49/0.4mm/Lower Extremity              | 70/in situ/Lower Extremity            | 71/in situ/Back                      | 74/2.4mm/Back                        | wt            | Mother 68 years             |
| g.124537261 A>G | rs202009081   | >1 PM    | 4                | 47/0.3mm/Shoulder                     | 70/in situ/Lower Extremity            | 71/in situ/Back                      | 74/2.4mm/Back                        | wt            | negative                    |
| g.124568913 C>T | n.a.          | >1 PM    | 2                | 66/1mm/Lower Extremity                | 74/5mm/Genital                       | —                                     | —                                     | wt            | negative                    |
| g.124467236 T>C | rs749702835   | PFH      | 1                | 22/0.4mm/Abdomen                      | —                                    | —                                    | —                                     | c.1514G>GC    | Mother 40 years             |
| g.124510982 G>A | rs778692211   | >1 PM    | 2                | 33/Unknown/Lower Extremity            | 68/0.45mm/Back                       | 71/in situ/Back                      | —                                     | wt            | G>GC                        |
| g.124491977 T>G | n.a.          | >1 PM    | 2                | 57/1mm/Back                          | 66/in situ/Lower Extremity            | 53/1.6mm/Back                        | —                                     | wt            | negative                    |
| g.124463018 C>T | rs30211999    | >1 PM    | 3                | 44/0.75mm/Chest                       | 57/0.4mm/Chest                       | —                                     | —                                     | wt            | negative                    |
| g.124463400 T>C | n.a.          | >1 PM    | 2                | 31/0.5mm/Chest                       | 31/in situ/Lower Extremity            | —                                     | —                                     | wt            | negative                    |
| g.124463559 T>C | n.a.          | >1 PM    | 2                | 36/1mm/Abdomen                       | 53/in situ/Lower Extremity            | —                                     | —                                     | p.R24P        | negative                    |

n.a.: not available; PFH: patients with a positive family history; >1 PM: patients with multiple primary melanomas; wt: wild type.

**POT1 variants in high-risk patients**

Descriptive data of the study population is shown in Table 1. In 694 cases and 35 controls, 53 variants were found in POT1 (see Table 2), of which 27 were found in the control group only. Out of 53 variants exclusively found in high-risk melanoma patients (Tables 2A and B), 48 were not listed in the dbsSNP (Table 2B). Four of the variants were exclusively found in high-risk melanoma patients (see Table 2B). Of the variants listed in the public SNP databases (Table 2A), 2 were located in 5’ UTR, 2 in the 3’ UTR, 1 in the coding sequence of exon 7, and 1 in the coding sequence of exon 11.

**Coincidence of CDKN2A mutations**

To exclude coincidence with CDKN2A mutations, we then examined the coincidence of CDKN2A mutations. One of the variants found exclusively in high-risk melanoma patients, predicted to be deleterious by 8 of 10 prediction tools (80%), resulted in an amino acid exchange and was predicted to be deleterious in 8 of 10 prediction tools (80%). Results of all prediction analysis are shown in Table 4.

**Prediction analysis of non-synonymous POT1 variants**

Prediction analysis was performed for non-synonymous variants in the CDKN2A sequence of our cases carrying potential risk variants of POT1. One of the variants found exclusively in high-risk melanoma patients, predicted to be deleterious by 8 of 10 prediction tools (80%), resulted in an amino acid exchange and was predicted to be deleterious in 8 of 10 prediction tools (80%). Results of all prediction analysis are shown in Table 4.
g.124463559 T>C, was associated with an established CDKN2A high-risk mutation, g.21974756 C>G (p.R24P). The carrier of the POT1 variant, g.124467236 T>C, additionally had the CDKN2A variant g.21971211 G>C (c.151-4 G>G-C), which was demonstrated to be non-effective in a previous study (BURGSTÄLLER-MUEHLBACHER et al. 2015).

DISCUSSION

Only recently, novel disease associated germline variants in POT1 were reported in melanoma pedigrees (ROBLES-ESPINOZA et al. 2014; SHI et al. 2014). This finding is of particular interest as the established disease causing mutations in familial melanoma, i.e., mutations in CDKN2A and CDK4 account only for 30–40% of the melanoma pedigrees. Despite this, just one family with a POT1 germline variant associated with melanoma was published so far (WILSON et al. 2014).

In the present study, in which the entire POT1 gene was sequenced in cases at high risk of melanoma and in control patients, a total of 53 variants were found. Despite this, previously published POT1 variants described in melanoma pedigrees (ROBLES-ESPINOZA et al. 2014; SHI et al. 2014) were not detected in our study. However, we found the intronic variant, g.124467236 T>C, in a patient with a positive family history of melanoma which was published so far (WILSON et al. 2017).

In conclusion, melanoma driving POT1 germline variants might be rare. However, further studies are required to assemble comprehensive information on the frequency and the role of POT1 in familial melanoma. It is also important to note that germline variants in POT1 were reported to be associated with other types of cancer such as colorectal cancer (CHUBB et al. 2016), glioma (BAINBRIDGE et al. 2015) and chronic lymphatic lymphoma (CALVETE et al. 2015; KARAMI et al. 2016; SPEEDY et al. 2016). As none of the variants described were found in melanoma cases, further studies might reveal that POT1 variants are specific to specific cancer types.

ACKNOWLEDGMENTS

We thank all participants of the M3 study for their contribution. This project was funded by the Anniversary Fund of the Austrian National Bank (grant number 15079) and the Medical Scientific Fund of the Mayor of the City of Vienna (grant number 10077). Study design, data collection, data analysis, manuscript preparation and/or publication decisions were not influenced by the funding sources. Conflicts of interest: None.

LITERATURE CITED

Adzhubei, I. A., S. Schmidt, L. Peshkin, V. E. Ramensky, A. Gerasimova et al., 2010 A method and server for predicting damaging missense mutations. Nat. Methods 7(4): 248–249. https://doi.org/10.1038/nmeth0410-248

Aoude, L. G., A. L. Pritchard, C. D. Robles-Espinoza, K. Wadt, M. Harland et al., 2015 Nonsense mutations in the shelterin complex genes ACD and TERF2IP in familial melanoma. J. Natl. Cancer Inst. 107: dju408. https://doi.org/10.1093/jnci/dju408

Table 4 Prediction of the variant g.124510982 G>A and g.124491977 T>G

| Prediction tools | g.124510982 G>A Prediction | g.124491977 T>G Prediction |
|------------------|-----------------------------|-----------------------------|
| **Mutation**     | **Disease causing**         | **Polymorphism**            |
| Polyphen2        | Probably Damaging           | Benign                      |
| HumDiv Score     | 0.987                       | 0.148                       |
| HumVar Score     | Possibly Damaging           | Benign                      |
| Score            | 0.791                       | 0.048                       |
| Provean Prediction | Deleterious                | Neutral                     |
| Score            | –5.503                      | –0.623                      |
| Sift Effect      | Tolerated                   | Tolerated                   |
| Score            | 0.16                        | 0.11                        |
| CADD PHRED 12 score | 31                        | 0.014                       |
| SNAP2 Prediction | Neutral                     | Neutral                     |
| Score            | -15                         | -89                         |
| Panther Preservation time | 57%                     | 93%                         |
| Message          | Probably damaging           | Probably benign             |
| GERP++ Score     | 5.57                        | -4.9                        |
| PhyLOP Score     | 2.77                        | -0.470331                   |
| Sum deleterious Total in % | 8                         | 0                           |

Of the 53 genetic variants found, 8 were exclusive in high-risk melanoma patients. Two of them, g.124491977 T>G and g.124510982 G>A, both non-synonymous variants, were tested for their alleged functionality. While g.124491977 T>G was predicted to be neutral by all 10 tools, g.124510982 G>A was predicted to be damaging by 80% of the prediction tools and is therefore very likely to be biologically functional. Comparing the wild type amino acid arginine with the resulting cysteine, there are differences in some amino acid features. The mutant residue is smaller and charged neutral, compared to the negatively charged wild type amino acid. Consequently, the correct folding of the protein could be influenced due to the more hydrophobic nature of the resulting amino acid (VENSKLAAR et al. 2010).

One potential limitation of this study is the fact that family history was largely reported and histopathologic reports confirming the diagnosis of relatives were not available for all cases. In the current study, the potential effect of the variants was assessed by computational analyses. Naturally, functional analyses are required to determine the exact role of these variants in melanoma development.
