Frequent *DPH3* **promoter mutations in skin cancers**

Supplementary Material

DNA and RNA extraction

DNA and RNA were extracted from fresh frozen tumor and surrounding tissues using the QIAGEN AllPrep DNA/RNA/miRNA Universal Kit. Tissues were homogenized in a Tissuelyser LT (Qiagen, Hilden, Germany) with 5 mm stainless steel beads in 600 ml RLT buffer and were further processed. Concentrations of total DNA and RNA were measured by using an ultraviolet–visible spectrophotometer (NanoDrop Technologies, Wilmington, USA) and absorption ratio at 260/280 nm was determined. RNA consistency was examined on a Bioanalyzer 2100 System (Agilent Technologies, Palo Alto, CA) using the corresponding RNA nanochips for a representative number of samples. Haematoxylin and eosin-stained sections of formalin fixed paraffin embedded (FFPE) tissues were examined for tumor content by a pathologist. Neoplastic areas were selected and dissected by tru-cut (2-3x0.6 mm) from the paraffin blocks with a manual tissue arrayer (Breecher MTA-I) to avoid stromal cells and to increase the specificity of melanoma cells to more than 95%. DNA was extracted using Qiagen DNeasy Tissue Kit (Qiagen) with additional proteinase K digestion at 55^oC for 48 h.

PCR and Sanger sequencing

DNA, 25 mM MgCl₂, 0.11 mM dNTP, 5% DMSO and 0.11 mM of each primer. Temperatures $\frac{1}{1}$ Determination of the mutational status of the TERT core promoter region (from position -27 to -286 from ATG start site) by PCR and Sanger sequencing was performed as described previously [1]. The mutational status of the DPH3 promoter was also determined by PCR and Sanger sequencing using one primer pair (F 5'CGAAGGGGTAACGCCCCAG 3', R 5'GGTCCCAGACGTGACGTAGC 3'). PCR was carried out in a 10-µl volume containing 10 ng

for PCR were set at 95°C for 45sec, annealing at 56°C and polymerization at 72°C for 30sec each for 35 cycles. Amplified products were purified with ExoSAP (GE Healthcare, Buckinghamshire, UK) to remove unused primer and were subjected to 35 cycles of sequencing reaction with a dideoxy terminator kit and forward and reverse primers in separate reactions (BigDyeTerminator v3.1 Cycle Sequencing Kit, Applied Biosystems, Austin, TX, USA). Sequencing reaction products were precipitated with ethanol and analysed on a capillary sequencer (AbiPrism 3130xl Genetic Analyzer). The sequencing data were analysed using Geneious Pro 5.6.5 software with reference to the sequences from the NCBI gene database, TERT (chr5: 1,295,071 - 1,295,521) and DPH3 (chr3: 16,306,256-16,306,755) (hg19 coordinates). The DPH3 promoter sequences are deposited in the GenBank database under accession numbers BankIt1784645: KP299043 - KP299143.

Cloning of DPH3 and OXNAD1 promoter region

For reporter luciferase assays a 414 bp region covering the core bidirectional promoter region (chr3: 16,306,400-16,306,813, hg19 coordinates) was amplified using genomic DNA From asinglehealthy individual. The amplicon was cloned into a T-overhang vector (TOPO PCR2.1, Invitrogen). The C>T (G>A on the negative strand) mutations at -8, -9 bp positions and CC>TT tandem mutation at -8/-9 bp positions from DPH3 transcription start site were generated using QuikChange Site Directed Mutagenesis kit (Agilent Technologies). A total of 4 reporter constructs carrying promoter region with wild-type (Wt) sequence, with -8T mutant, with -9T mutant and with -8/-9TT mutants were prepared for both orientations. All the plasmid constructs were sequenced by Sanger sequencing to confirm the respective mutations. For the generation of DPH3 reporter constructs the inserts from TOPO PCR2.1 were restriction digested at XhoI and BamHI sites of TOPO PCR 2.1 and subcloned into XhoI

2

and BglII sites of pGL4.10[luc2]. The OXNAD1 promoter inserts were restriction digested at SacI and XhoI sites and subcloned into the same sites of pGL4.10[luc2] promoter-less vector (Promega) (Figure 3).

Cell culture and Luciferase reporter assays

A patient derived melanoma cell line (UKRV-Mel-21) was a kind gift from Prof. Stefan Eichmüller of German Cancer Research Center. The cell line was derived from a Caucasian female patient with malignant melanoma. The cell line was cultured in RPMI-1640 media (Gibco) supplemented with 10% FBS (Gibco). Cells were incubated at 37°C in a humidified atmosphere with 5% CO2. The used cell line was authenticated by short tandem repeat profiling [2]. The DNA extracted from the UKRV-Mel-21 cell line was also sequenced and found to have no endogenous mutations in DPH3 promoter.

For reporter assay, cells were seeded in 12-well plates and cotransfected with FuGene-6 (Promega), 1.5µg of reporter construct and 75 ng of pRL-actin expression plasmid for Renilla luciferase in triplicate. pGL3-control plasmid was used for determining transfection efficiency and pRL-actin as an internal control for normalization of luminescence values. pGL3-control and pRL-actin express firefly luciferase and Renilla luciferase genes under the thymidine kinase and actin promoters, respectively.

Promoter-less vector (pGL4.10[luc2]) and non-transfected cells were used as negative controls. Cells were harvested for 24 hours post transfection using 1x passive lysis buffer (Promega) and reporter expression was analyzed using the Dual-Luciferase Reporter assay system (Promega). The relative ratio of firefly luminescence to renilla luminescence was

3

calculated to normalize the variations across samples. Statistical differences were determined using two sided t-test and box-plots were plotted in R.

Statistical methods

We performed a binomial test to evaluate whether the frequency of C>T mutations at 16,306,504 and 16,306,505 is statistically significantly different from the average mutation rate expected per chance for the cancer types included in the study. The reported mutation rates of 11.4, 33.3 and 75.8 mutations/megabase for melanoma, squamous cell carcinoma and basal cell carcinomas respectively were used as a probability (p) of a mutation at a specific base (null hypothesis: p=11.4x10⁻⁶; 33.3x10⁻⁶; 75.8x10⁻⁶) [3-5]. Two-sided P values with 5% significance level was calculated using R statistical software [6].

Transcription factorbinding sitespredictions

TRANSFAC databasewasusedtosearchforpredictedtranscription factorbindingsiteswith threedifferentalgorithmsincluding PROMO,ConSiteandrVista[7-11].

Regulation and conservation data analysis

ChIPseqdatafor161transcription factors(TFs)fromtheEncyclopedia ofDNAElements (ENCODE) consortiumwiththebindingmotifsfromFactorbook repositorywasvisualized usingthe UCSC genome browser and inspected for the DPH3 promoter region [12-14]. Otherconservation andregulation dataincluding H3K27Ac ChIP-seqpeaks,whichisthe histonemark often found near regulatory elements, DNaseI hypersensitivity clusters,

GenomeEvolutionary RateProfiling (GERP)scoresandMultizAlignment of100vertebrates werealsoinspected usingUCSCgenome browser(Suppl.FigureS5).

The GERP algorithm provides a position-specific estimate of evolutionary constraint by

quantifying substitute deficit across species in multiple genome alignments. It was shown

previously that functional mutations tend to occur at the positions with high GERP

evolutionary constraint score [15, 16]. GERP scores for the mutations positions were

obtained from UCSC genome browser [\(http://genome.ucsc.edu/cgi-](http://genome.ucsc.edu/cgi-)

bin/hgTrackUi?g=allHg19RS_BW). According to the algorithm developers, GERP score above

2 is an indication that the position is evolutionary constraint (see the link above). GERP

scores for the recurrently mutated positions "-8", "-9", "-12" and "-13" in the DPH3

promoter regions are 1.21, 4.29, 5.38 and 5.38 respectively, which indicates that the

positions "-9", "-12" and "-13" are highly conserved whereas position "-8" is not.

REFERENCES

1. Heidenreich B, Nagore E, Rachakonda PS, Garcia-Casado Z, Requena C, Traves V, Becker J, Soufir N, Hemminki K and Kumar R. Telomerase reverse transcriptase promoter mutations in primary cutaneous melanoma. Nature communications. 2014; 5:3401.

2. Castro F, Dirks WG, Fahnrich S, Hotz-Wagenblatt A, Pawlita M and Schmitt M. Highthroughput SNP-based authentication of human cell lines. Int J Cancer. 2013; 132(2):308-314. 3. Walia V, Mu EW, Lin JC and Samuels Y. Delving into somatic variation in sporadic melanoma. Pigment cell & melanoma research. 2012; 25(2):155-170.

4. Durinck S, Ho C, Wang NJ, Liao W, Jakkula LR, Collisson EA, Pons J, Chan SW, Lam ET, Chu C, Park K, Hong SW, Hur JS, Huh N, Neuhaus IM, Yu SS, et al. Temporal dissection of tumorigenesis in primary cancers. Cancer discovery. 2011; 1(2):137-143.

5. Jayaraman SS, Rayhan DJ, Hazany S and Kolodney MS. Mutational landscape of basal cell carcinomas by whole-exome sequencing. J Invest Dermatol. 2014; 134(1):213-220.

6. Team RC. (2014). R: A Language and Environment for Statistical Computing: R Foundation for Statistical Computing).

7. Matys V, Kel-Margoulis OV, Fricke E, Liebich I, Land S, Barre-Dirrie A, Reuter I, Chekmenev D, Krull M, Hornischer K, Voss N, Stegmaier P, Lewicki-Potapov B, Saxel H, Kel AE and Wingender E. TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. Nucleic acids research. 2006; 34(Database issue):D108-110.

8. Messeguer X, Escudero R, Farre D, Nunez O, Martinez J and Alba MM. PROMO: detection of known transcription regulatory elements using species-tailored searches. Bioinformatics (Oxford, England). 2002; 18(2):333-334.

9. Farre D, Roset R, Huerta M, Adsuara JE, Rosello L, Alba MM and Messeguer X. Identification

of patterns in biological sequences at the ALGGEN server: PROMO and MALGEN. Nucleic acids research. 2003; 31(13):3651-3653.

10. Sandelin A, Wasserman WW and Lenhard B. ConSite: web-based prediction of regulatory elements using cross-species comparison. Nucleic acids research. 2004; 32(Web Server issue):W249-

252.

11. Loots GG and Ovcharenko I. rVISTA 2.0: evolutionary analysis of transcription factor binding sites. Nucleic acids research. 2004; 32(Web Server issue):W217-221.

12. ENCODE project consortium. An Integrated Encyclopedia of DNA Elements in the Human Genome. Nature. 2012; 489(7414):57-74.

13. Wang J, Zhuang J, Iyer S, Lin XY, Greven MC, Kim BH, Moore J, Pierce BG, Dong X, Virgil D, Birney E, Hung JH and Weng Z. Factorbook.org: a Wiki-based database for transcription factor- binding data generated by the ENCODE consortium. Nucleic acids research. 2013; 41(Database issue):D171-176.

14. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM and Haussler D. The human

genome browser at UCSC. Genome research. 2002; 12(6):996-1006.

15. Cooper GM, Stone EA, Asimenos G, Green ED, Batzoglou S and Sidow A. Distribution and intensity of constraint in mammalian genomic sequence. Genome research. 2005; 15(7):901-913.

16. Cooper GM, Goode DL, Ng SB, Sidow A, Bamshad MJ, Shendure J and Nickerson DA. Single- nucleotide evolutionary constraint scores highlight disease-causing mutations. Nature methods.

2010; 7(4):250-251.

 $\boldsymbol{\mathsf{A}}$

A

 \overline{C} G G \overline{C} ${\mathsf G}$

> G G G

C

C

 $C C$ \overline{C} G G т G \subset C \subset G т С. С C ſ., G

16.306,510 bp

 $\overline{\epsilon}$

 $\overline{}$

в

 $\mathbb T$

TA G

G

Supplementary Figure 1

Manual review of *DPH3* **promoter mutations using Integrative Genomics Viewer and validation by Sanger sequencing.**

Representative Integrative Genomics Viewer screenshots (**A**) and matched Sanger sequencing chromatograms (**B**) of *DPH3* promoter mutations at chr3: 16,306,505 C>T (-9C>T) and chr3: 16,306,504-16,306,505 CC>TT (-8/-9CC>TT). Average depth of coverage at the chr3: 16,306,505 was x60 for both tumor and normal samples.

Supplementary Figure S2

ChIPseq peaks from ENCODE data in the *DPH3* **promoter region.** ChIPseq data for 161 transcription factors (TFs) from ENCODE, visualized using the UCSC genome browser. Each box encloses peak cluster of TF occupancy, with the darkness of the box being proportional to the maximum signal strength observed in any cell line contributing to the cluster (UCSC track description: [http://ucscbrowser.genap.ca/cgi-](http://ucscbrowser.genap.ca/cgi-bin/hgTables?db=hg19&hgta_group=regulation&hgta_track=wgEncodeRegTfbsClusteredV3&hgta_table=wgEncodeRegTfbsClusteredV3&hgta_doSchema=describe+table+schema)

[bin/hgTables?db=hg19&hgta_group=regulation&hgta_track=wgEncodeRegTfbsClusteredV3&hgta_table=wgEncodeRegTfbsClustered](http://ucscbrowser.genap.ca/cgi-bin/hgTables?db=hg19&hgta_group=regulation&hgta_track=wgEncodeRegTfbsClusteredV3&hgta_table=wgEncodeRegTfbsClusteredV3&hgta_doSchema=describe+table+schema) [V3&hgta_doSchema=describe+table+schema\)](http://ucscbrowser.genap.ca/cgi-bin/hgTables?db=hg19&hgta_group=regulation&hgta_track=wgEncodeRegTfbsClusteredV3&hgta_table=wgEncodeRegTfbsClusteredV3&hgta_doSchema=describe+table+schema) . Within a cluster, green color highlight indicates the highest scoring site of a Factorbookidentified canonical motif for the corresponding factor (Factorbook repository of ENCODE data analysis:

http://www.factorbook.org/mediawiki/index.php/Welcome_to_factorbook) [12]. In particular Ets factors ELK1, ELK4, ELF1 and

GABPA show high signal in the mutated region (chr3: 16,306,504-16,306,509, marked with yellow frame).

Supplementary Figure S3

DPH3 **mRNA expression in different skin lesions.** *DPH3* mRNA expression measured by quantitative real-time PCR in a different skin lesions. (**A**) Differences in the levels of expression in BCC tumors with three different mutations and without mutations (**B**) differences in expression based on presence or absence of the mutations in BCC. (**C**) Comparison of *DPH3* expression in melanoma and nevi tissues with and without mutations. (**D**) Measurement of *DPH3* expression in squamous cell carcinomas.

OXNAD1 Expression in Basal Cell Carcinomas according to presence of DPH3 promoter mutations

Supplementary Figure S4

OXNAD1 **mRNA expression in skin BCC.** *OXNAD1* mRNA expression measured by quantitative real-time PCR in BCC tumors with different *DPH3* mutations in the left panel. Differences in expression based on presence or absence of mutations are shown in the right panel.

Supplementary Figure S5

Regulation and conservation data analysis in the *DPH3* **promoter region.** Conservation and regulation data for the *DPH3* promoter region including H3K27Ac ChIP-seq peaks (the histone mark often found near regulatory elements), DNaseI hypersensitivity clusters, CpG islands, Genome Evolutionary Rate Profiling (GERP) scores and Multiz Alignment of 100 vertebrates, visualized using UCSC genome browser. The region carries typical promoter marks and the motif containing the mutations, with exception of the -8 position, is highly conserved across mammalian lineage.

Supplementary Table S1. Clinical characteristics of 21 melanomas with whole exomes sequenced

n.a.=not available

^a Fitzpatrick phototyping scale

^b LMM, Lentigo Maligna Melanoma; SSM, Superficial Spreading Melanoma; NM, Nodular Melanoma; ALM, Acral Lentiginous Melanoma

Supplementary Table S2. DPH3 promoter mutations in melanoma whole-exome dataset $(n=21)$

^aWith base quality of >20 at the given position

^bMutations confirmed by Sanger sequencing

^cTandem CC>TT mutation

^dOne read supports tandem CC>TT mutation

Supplementary Table S3. Distribution and association of DPH3 promoter mutations according to patient and tumor characteristics in melanoma and non-melanoma skin cancers

^a Fitzpatrick phototyping scale
^b LMM, Lentigo Maligna Melanoma; SSM, Superficial Spreading Melanoma; NM, Nodular Melanoma; ALM, Acral Lentiginous Melanoma

Two-sided P-values were derived from χ^2 -test and considered statistically significant if < 0.05.

Wt = wildtype; Mut = mutation

Supplementary Table S4. Exome Coverage Statistics

