

Original Paper

Dicer Sequencing, Whole Genome Methylation Profiling, mRNA and smallRNA Sequencing Analysis in Basal Cell Carcinoma

Michael Sand^{a,b} Annabelle Bromba^a Daniel Sand^c Thilo Gambichler^a
Schapoor Hessam^a Jürgen C. Becker^{d,e,f} Eggert Stockfleth^a Thomas Meyer^a
Falk G. Bechara^a

^aDepartment of Dermatology, Venereology and Allergology, Ruhr-University Bochum, Bochum, Germany, ^bDepartment of Plastic Surgery, St. Josef Hospital, Catholic Clinics of the Ruhr Peninsula, Essen, Germany, ^cDepartment of Physiological Science, University of California Los Angeles (UCLA), Los Angeles, CA, USA, ^dDepartment of Translational Skin Cancer Research, University Hospital Essen, Essen, Germany, ^eGerman Cancer Consortium (DKTK), Essen, Germany, ^fGerman Cancer Research Center (DKFZ), Heidelberg, Germany

Key Words

Basal cell carcinoma • Epithelial skin cancer • smallRNA • microRNA • Dicer • Methylation • Microarray

Abstract

Background/Aims: Perturbations in the expression of microRNAs (miRNAs) and their maturing machinery components such as Dicer have been previously described for basal cell carcinoma (BCC). However, the mutational status of Dicer in BCC is unclear. Further, the sclerodermiform subtype of BCC (sBCC) has not been previously investigated regarding its methylation profile or its smallRNA expression profile via RNA sequencing. We conducted this study to investigate the mutational status of Dicer in BCC. **Methods:** Dicer sequencing was performed on the Illumina MiSeq System in a total of 16 BCC samples (8 nodular BCCs, 8 sBCCs) and mapped against the human reference genome (i.e., hg19). Dicer sequencing was performed in all 16 BCC samples. We performed whole genome methylation profiling with Infinium MethylationEPIC BeadChips as well as mRNA and smallRNA sequencing in 5 sBCCs with the Illumina NextSeq500 next-generation sequencing system. **Results:** Compared to the wildtype Dicer sequence, we found 5 to 7 variants per sBCC sample including insertion, deletion, and multiple nucleotide variants. Global methylation profiles were highly similar between groups. mRNA sequencing revealed S100A9, KRT14, KRT10, S100A8, S100A7, COX1, KRT1, COX3, and smallRNA sequencing analysis miR-21, miR-99a, miR26-a-2, let-7f, let-7g,

let-7i, miR-100, and miR-205 were the most strongly expressed in sBCCs. **Conclusion:** We identified a variety of Dicer mutations that could play a role in aberrant miRNA expression in BCC. The noted RNA sequences should be further evaluated in functional studies to explore their potential pathogenetic role in sBCC.

© 2019 The Author(s). Published by
Cell Physiol Biochem Press GmbH&Co. KG

Introduction

Basal cell carcinoma (BCC) is the most frequent type of skin tumor, and the incidence rate of BCC is increasing worldwide [1, 2]. Despite being the most common human cancer, the molecular pathogenesis of BCC remains elusive. The invention of *noncoding RNAs (ncRNAs)*, including *microRNAs (miRNAs)*, brought a new perspective on cancer pathogenesis and regulation. For BCC, several authors have shown perturbations in the expression of *ncRNAs* such as *miRNAs*, *circularRNAs*, and *long noncoding RNAs* [3-8]. The same is true for *miRNA* maturing machinery components such as Dicer in BCC [9, 10]. Dicer is one of the two most important *miRNA* maturing enzymes responsible for cytosolic *miRNA* maturation [11]. Its expression is significantly lower in BCC compared to that in healthy control skin, indicating a possible mutation of the Dicer gene [10]. Therefore, we sequenced Dicer, the major *miRNA* maturing enzyme which has previously shown mutations in a variety of other cancers [12]. Dicer sequencing was performed in nodular and sclerodermiform BCC (nBCC, sBCC) screening for possible mutations compared to the human reference genome hg19 [13]. Further, in a subset of sBCC, which is clinically the most aggressive BCC subtype associated with higher rates of recurrence, whole genome methylation profiling, and sequencing analysis of *messenger RNA (mRNA)* and *smallRNA* was performed to further characterize the epigenetic landscape of sBCC.

Materials and Methods

Ethics

The present study was performed in accordance with the approval of the Ethical Review Board of the Ruhr-University Bochum, Germany (registration number: 3265-08) and the Ethics Committee of the regional Medical Association (Nordrhein; registration number: 2017015). Further, it was conducted within the framework of the declaration of Helsinki, and informed consent was signed by all study subjects.

This study conforms to the applicable local requirements regarding ethical and investigational committee review, informed consent, and other statutes or regulations regarding the protection of the rights and welfare of human subjects participating in medical research (Ethical Review Board of the Ruhr-University Bochum, registration number: 4749-13 and the Ethics Committee of the regional Medical Association (Nordrhein; registration number: 2017015).

Subjects

In 16 patients (11 women, 5 men; median age: 71.5 years), biopsy specimens (4-mm punch biopsies) of 8 nBCCs and 8 sBCCs were obtained during surgical tumor removal. See Table 1 for details. Specimens were immediately stored in RNAlater (Qiagen, Hilden, Germany) at -80°C (fresh frozen tissue). The tumors were additionally archived as Formalin-Fixed Paraffin-Embedded (FFPE) Blocks.

Table 1. Details of basal cell carcinoma specimen. Abbreviations: nBCC, nodular basal cell carcinoma; sBCC, sclerodermiform BCC; M, male; F, female

Sample ID Number	Sex	Age	Histology	Body area
1	M	66	nBCC	face
2	F	55	nBCC	shoulder
3	M	72	nBCC	pectoral
4	F	72	nBCC	auricular
5	M	80	nBCC	auricular
6	F	89	nBCC	face
7	M	65	nBCC	face
8	F	66	nBCC	nose
9	F	92	sBCC	auricular
10	F	85	sBCC	face
11	F	74	sBCC	lip
12	F	80	sBCC	nose
13	F	68	sBCC	head
14	F	71	sBCC	face
15	M	65	sBCC	ear
16	F	66	sBCC	face

Genomic DNA and RNA isolation

Genomic DNA (gDNA) and total RNA were isolated in parallel from the same sample using the AllPrep DNA/RNA Mini Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. An on-column DNase digestion step was included in the RNA isolation protocol. The final volume of the eluate was 50 μ L for gDNA and 30 μ L for total RNA.

The total RNA samples were isolated for RNA sequencing analysis. We used an aliquot of each total RNA sample to determine RNA concentration and purity on the NanoDrop ND-1000 spectral photometer (Peqlab, Erlangen, Germany). All samples were analyzed on the 2100 Bioanalyzer (Agilent Technologies, Santa Clara, USA) using RNA 6000 Nano LabChip Kits (Agilent Technologies, Santa Clara, USA) to determine the RIN values.

For small RNA sequencing analysis, total RNA including small RNAs was isolated from the 5 BCC FFPE tissue samples using the miRNeasy FFPE Kit (Qiagen) following the manufacturer's instructions. An on-column DNase digestion step was included in the RNA isolation protocol. The final volume of the eluate was 20 μ L. An aliquot of each gDNA sample was used to determine purity and approximate DNA concentration on the NanoDrop ND-1000 spectral photometer (Peqlab, Erlangen, Germany). The gDNA samples were additionally quantified using the highly sensitive fluorescent dye-based Qubit® dsDNA HS Assay Kit (Thermo Fisher Scientific, Karlsruhe, Germany). The TruSeq Custom Amplicon Low Input library preparation protocol was followed.

Dicer sequencing

The *microRNA* maturing enzyme Dicer 1 (ribonuclease III, Gene ID 23405 on Chromosome 14) was defined as the target gene. The online tool Illumina Design Studio (<http://designstudio.illumina.com/>) was used to design an amplicon-based enrichment panel (TruSeq Custom Amplicon Low Input) that targets the DICER1 gene. The design strategy was based on human reference genome hg19 including the coding sequence of DICER1 with a targeted amplicon size of 175 bp [13].

Library preparation was performed with the TruSeq Custom Amplicon Low Input Library Kit (Illumina, San Diego, USA) using a dual pool approach according to the manufacturer's protocol.

Hybridization of oligos to genomic DNA was performed in a 96-well plate, followed by extension and ligation to form DNA templates consisting of the regions of interest flanked by universal primer sequences. Using indexed primers, DNA templates were amplified by polymerase chain reaction (PCR) and elongated with barcodes and sequencing adapters. In addition to the 16 gDNA samples, a control DNA (2800M) and a non-template control were processed in parallel to control the complete library preparation process, resulting in 32 DNA-oligo-pool samples. The identification of each sample was ensured by a unique index sequence combination. The DNA 1000 and High Sensitivity LabChip Kit (Agilent Technologies, Santa Clara, USA) were used on the 2100 Bioanalyzer (Agilent Technologies, Santa Clara, USA) to analyze the quality and fragment distribution of the generated libraries. The expected PCR product size, including adapter sequences, was approximately 280 bp for the targeted amplicon size of 175 bp.

Furthermore, all libraries were quantified using the highly sensitive fluorescent dye-based Qubit® dsDNA HS Assay Kit (ThermoFisherScientific, Karlsruhe, Germany) according to the manufacturer. Afterward, the single TruSeq Custom Amplicon libraries were pooled together in equimolar ratios into a final sequencing library.

The final pooled sequencing library was quantified again using the Qubit® dsDNA HS Assay Kit. After quantification, the final sequencing library pool was diluted to 2 nM, followed by denaturation with NaOH.

For cluster generation and sequencing, we used the MiSeq® Reagent Kit v2 (300 cycles; Illumina, San Diego, USA). Before sequencing, cluster generation was performed. Clusters represent areas on the MiSeq® flow cell with clonally amplified DNA library fragments. The cluster generation of loaded fragments was carried out by a two-dimensional amplification of bound fragments (i.e., bridge amplification). Fragments were immobilized on the flow cell involving 2 adapters attached at opposing ends of the fragment and complementary slide-anchored primers. By several cycles of flipping and binding the immobilized fragments to the surrounding primer lawn followed by amplification, approximately 1,000 copies of the original fragments were produced and localized in a tight cluster. After cluster generation, sequencing primers were hybridized to the adapter sequences at the end of the fragments, and sequencing was carried out.

We performed bidirectional sequencing, starting at the end of the sense strand (read 1), then at the end of the complementary strand (read 2). Both reads had a length of 150 bases, finally producing 300 bases of

sequence information in 2 x 150 bp paired-end reads. Sequencing was operated by the MiSeq Control Software v2.5.0.5 (Illumina, San Diego, USA). Primary data analysis including signal processing, de-multiplexing, and trimming of adapter sequences of the two read files as well as read merging was performed using the MiSeq® inherited MiSeq® Reporter 2.5.1 software (Illumina, San Diego, USA). Image and signal processing was conducted by applying the Targeted Re-sequencing processing pipeline. The Illumina Sequence Analysis Viewer (SAV) 1.9.1 was used for imaging and evaluation of the sequencing run performance. After completion of the sequencing run, technical quality parameters were evaluated using the SAV software. The CLC Genomics Workbench v9.5.1 (Qiagen, Hilden, Germany) and the MiSeq® Reporter v2.5.1.3 were used for in-depth data analysis, particularly consisting of sequencing quality control, mapping, variant calling, and annotation. The quality filtered variant lists were annotated with known variants from the human SNP database Build 147 as found at <ftp://ftp.ncbi.nlm.nih.gov/snp/>. Afterward, they were further annotated and filtered based on overlap information including gene symbol, exon numbers, and amino acid changes based on the mRNA and coding sequence in the applied hg19 reference file. Where available, possible splice site effects were added to the lists by a functional consequence analysis. Finally, clinical annotations were added. Therefore, data from the Exome Aggregation Consortium (<http://exac.broadinstitute.org/>) were accessed and added as an annotation.

Furthermore, annotations were added to identify known pathogenic variants or explore novel variants that may have a deleterious effect by virtue of a change in amino acid, disruption of a regulatory motif or the disease-association or pathway-association of the affected gene. Therefore, the databases HGMD (Human Gene Mutation Database), EVS (Exome Variant Server) and ClinVar (public archive of relationships among sequence variation and human phenotype) were accessed. The initially generated quality filtered variant lists were finally filtered to reduce the number of reported variants to those with possible clinical relevant or pathological effects. See Table 2 for details. In addition, we conducted a second variant analysis of all 16 samples starting from the Binary Alignment Map files. Variants were

Table 2. Summary of variants of Dicer compared to the human reference gene after filtering to reduce the number of reported variants to those with possible clinical or pathological effects after considering the information given in the databases Human Genome Mutation Database (HGMD), Exome Variant Server (EVS) and ClinVar (public archive of relationships among sequence variation and human phenotype). Abbreviation: SNV, single nucleotide variant

Number	Region	Type	Number of Variants
1	95571635-95571636	Insertion	5
	95571636	Deletion	
	95583035-95583036	Insertion	
	95583036	SNV	
2	95571635-95571636	Deletion	6
	95571636	Insertion	
	95583035-95583036	Insertion	
	95583036	SNV	
3	95571635-95571636	Deletion	5
	95571636	Insertion	
	95583035-95583036	Insertion	
	95583036	SNV	
4	95566108-95566109	Insertion	6
	95571635-95571636	Deletion	
	95571636	Insertion	
	95583035-95583036	Insertion	
5	95583036	SNV	5
	95571635-95571636	Deletion	
	95571636	Insertion	
	95583035-95583036	Insertion	
6	95583036	Deletion	5
	95571635-95571636	Insertion	
	95571636	Deletion	
	95583035-95583036	Insertion	
7	95583036	SNV	5
	95571635-95571636	Deletion	
	95571636	Insertion	
	95583035-95583036	Insertion	
8	95571635-95571636	Deletion	7
	95571636	Insertion	
	95583032	SNV	
	95583032	SNV	
9	95583035-95583036	Insertion	5
	95583036	SNV	
	95571635-95571636	Deletion	
	95571636	Insertion	
10	95583035-95583036	Insertion	5
	95583036	SNV	
	95571635-95571636	Deletion	
	95571636	Insertion	
11	95583035-95583036	Deletion	5
	95571636	Deletion	
	95583035-95583036	Insertion	
	95583036	SNV	
12	95571635-95571636	Deletion	5
	95571636	Insertion	
	95583035-95583036	Insertion	
	95583036	SNV	
13	95571635-95571636	Deletion	5
	95571636	Insertion	
	95583035-95583036	Deletion	
	95583036	Insertion	
14	95571635-95571636	Deletion	7
	95571636	Deletion	
	95579534	SNV	
	95579534	SNV	
15	95583035-95583036	Insertion	5
	95583036	SNV	
	95571635-95571636	Deletion	
	95571636	Insertion	
16	95583035-95583036	Deletion	5
	95571636	Deletion	
	95583035-95583036	Insertion	
	95583036	SNV	
	95583036	Deletion	

identified using the FreeBayes algorithm [14]. We used both SNPs and InDel variants reaching as low as a minimal detection rate of 1%. Resulting variant candidates have been Phred-quality filtered to remove sites with a high chance of being polymorphic. Samples for further smallRNA sequencing analysis were chosen by comparing the results from the mentioned pipeline using the FreeBayes algorithm and the initial analysis performed with the CLC Genomics Workbench v9.5.1 (Qiagen, Hilden, Germany) and the MiSeq® Reporter v2.5.1.3. We selected samples for further analysis based on a high similarity between the results of both methods as well as the number of variants and the occurrence of unique variants compared to the mutational profile. Data mining was completed by reviewing the results and manually checking the HGMD via <http://www.hgmd.cf.ac.uk/ac/index.php> for possible corresponding pre-described mutations [15].

Whole genome methylation profiling

To detect and quantify DNA methylation in sBCC samples 10–14, Infinium MethylationEPIC BeadChips (Illumina, San Diego, USA) with over 850,000 methylation sites were used in combination with bisulfite conversion of DNA samples and the Infinium HD Assay for Methylation. First, gDNA was modified by bisulfite conversion. Using the EZ DNA Methylation™ Kit (Zymo Research, Irvine, USA), 500 ng of methylated DNA was treated with bisulfite according to the manufacturer's instructions, thereby converting unmethylated cytosine into uracil. After the on-column desulfonation step, bisulfite-converted DNA samples were then further processed using the Illumina Infinium HD Assay for Methylation (Illumina, San Diego, USA) according to the manufacturer's instructions as follows. First, bisulfite-converted DNA was denatured and neutralized. Second, the DNA was amplified by three orders of magnitude in a whole-genome amplification step. After enzymatic fragmentation, the amplified DNA was precipitated using isopropanol. After resuspension, the DNA was hybridized onto the microarrays. The microarray hybridization was performed according to the Infinium HD Assay Methylation Protocol Guide (Illumina, San Diego, USA). The bisulfite-converted DNA samples were hybridized at 48°C for 16 hours on separate Illumina Infinium MethylationEPIC Arrays (Illumina, San Diego, USA). DNA fragments anneal to locus-specific 50-nucleotide oligomers which are covalently linked to different bead types.

After hybridization and washing of the microarrays, the primers on the BeadChips were extended complementary to the binding DNA sample. This reaction incorporates labeled nucleotides. Incorporation of a G corresponds to the methylated state, and incorporation of an A corresponds to the unmethylated state on the complementary strand. Allele-specific single base extension of the primer incorporates a biotin nucleotide or a dinitrophenyl-labeled nucleotide (C and G nucleotides are biotin labeled, A and T nucleotides are dinitrophenyl labeled). Hybridized DNA was then removed, and the labeled extended primers were stained and dried. Fluorescent signal intensities of the labeled and stained nucleotides on the BeadChips were detected with a dual color channel approach and the Illumina iScan™ System (Illumina, San Diego, USA). Intensity files were generated by the iScan System for the green and the red channel (Illumina, San Diego, USA). The software tool GenomeStudio v2011.1 (Illumina, San Diego, USA) was used for quality control and visualization of the performance of microarray analysis as well as for methylation analysis. Raw microarray data were imported into GenomeStudio v2011.1 and analyzed. For the Infinium Methylation Assay, the β -value was calculated as: $\beta = \text{Max}(\text{Signal B}, 0) / \text{Max}(\text{Signal A}, 0) + \text{Max}(\text{Signal B}, 0) + 100$. For Infinium I assays, signal A and signal B were produced by two different bead types and reported in the same color.

mRNA and smallRNA sequencing analysis

For mRNA sequencing analysis, one RNA sequencing library was prepared from 5 total RNA samples (sBCC samples 10–14) from fresh frozen tissue with the Illumina TruSeq® Stranded mRNA HT technology (Illumina, San Diego, USA), according to the manufacturer's protocol.

For smallRNA sequencing, one smallRNA sequencing library was prepared from 5 total RNA samples derived from FFPE tissue samples (sBCC samples 10–14) using the NEBNext® Small RNA Library Prep Kit for Illumina (New England BioLabs, Frankfurt, Germany). The kit targets microRNAs and other smallRNAs that have a 3' hydroxyl group resulting from enzymatic cleavage by Dicer or other RNA processing enzymes. Sequencing of all libraries (mRNA and smallRNA) was performed at a final concentration of 1.8 pM and with a 1% PhiX v3 control library spike-in on the NextSeq500 sequencing system (Illumina, San Diego, USA). For cluster generation and sequencing of all samples, high-output single-end 75-cycle (1x75bp SE) runs were performed for the mRNA and smallRNA libraries. Sequencing was controlled by the NextSeq Control

Software v2.1.0 (Illumina, San Diego, USA). The resulting *mRNA* reads were quality checked and mapped against the human reference genome. The resulting *smallRNA* reads were quality checked, mapped against miRBase 21, and the human GRCh38 noncoding *RNAs* [16, 17]. In sum, the achieved quality values of the NextSeq500 sequencing runs met the manufacturer's given specifications (Illumina, San Diego, USA). All data could be used for in-depth data analysis.

Genevestigator

In-depth differential expression analysis was performed with Genevestigator (Nebion AG, Zurich, Switzerland). For comparison and differential expression analysis with healthy controls, we used publicly available content on the human mRNASeq_HUMAN_GL platform in Genevestigator. The following samples were available and used as control samples: 99 healthy, untreated and polyA+ enriched skin tissue samples from experiments GSE63980 and E-MTAB-5678 (control group 1) as well as 12 epidermal keratinocyte samples from healthy subjects from experiment GSE96601 (control group 2). In a second step, the first 100 filtered genes with the highest log-ratios were analyzed in Genevestigator. In a third step, only those candidate genes with obvious upregulation in cancer subtypes were noted. The obtained results were verified by an analysis of the differentially expressed genes obtained from control group 1 with microarray data on the human AFFY_U133PLUS_2 Array in Genevestigator. For that purpose, the public experiment GSE103439 (research area: BCC) was curated and integrated to Genevestigator. Control samples included healthy, untreated samples from skin, HF keratinocytes, dermal fibroblasts, dermal papillae, epidermal keratinocytes, epidermis and foreskin-derived neonatal fibroblasts.

Results

Quality control

RNA integrity including the *RNA* integrity number (RIN) and DV200 values were assessed on the Agilent 2100 Bioanalyzer (Agilent, Santa Clara, CA). Raw data from this study were deposited in the National Center for Biotechnology Information's Gene Expression Omnibus and is accessible through the Gene Expression Omnibus Series accession number GSE128795. All quality criteria predefined by the manufacturer for successful microarray and sequencing analysis were fulfilled.

Dicer sequencing

Variants compared to the human reference sequence were analyzed in all samples. After quality filtering and complete annotation and before filtering for clinically relevant variants, we found 5 to 54 variants for the 16 BCC samples. After filtering for clinically relevant variants, we found 5 to 7 variants. The results are compiled as supplementary material (Supplementary Table 1 – for all supplemental material see www.cellphysiolbiochem.com) and Circos® plots shown in Fig. 1 and Fig. 2 depicting insertion, deletion, and multiple nucleotide variants (MNVs). The most frequently observed variants were insertion between position 95583035 and 95583036, insertion between position 95571635 and 95571636, single nucleotide variant deletion at position 95583036, and deletion at position 95571636. Comparing the Dicer sequencing data in nBCC to sBCC, we found no statistically significant difference between the two BCC subtypes ($p > 0.05$). The full analysis of the Dicer sequencing data is available as supplementary data compiled in Supplementary Table 1.

Whole genome methylation profiling

A total of 863904 CpG sites, 2932 nonCpG loci, and 59 single nucleotide polymorphisms (SNPs) were analyzed. The number of detected sites per array spot for all samples were in a very small range (between 866118 and 866222), indicating very similar data quality among the samples. The average methylation level β is shown in the box-plot in Fig. 3. We noted equal data distribution in all samples with a median average β value of 0.3 to 0.9, which indicates a high degree of similarity between the samples. We used R v3.0.2 software to generate heatmaps with the packages ggplot package v2 1.0.1, reshape2 package v1.4.1 and

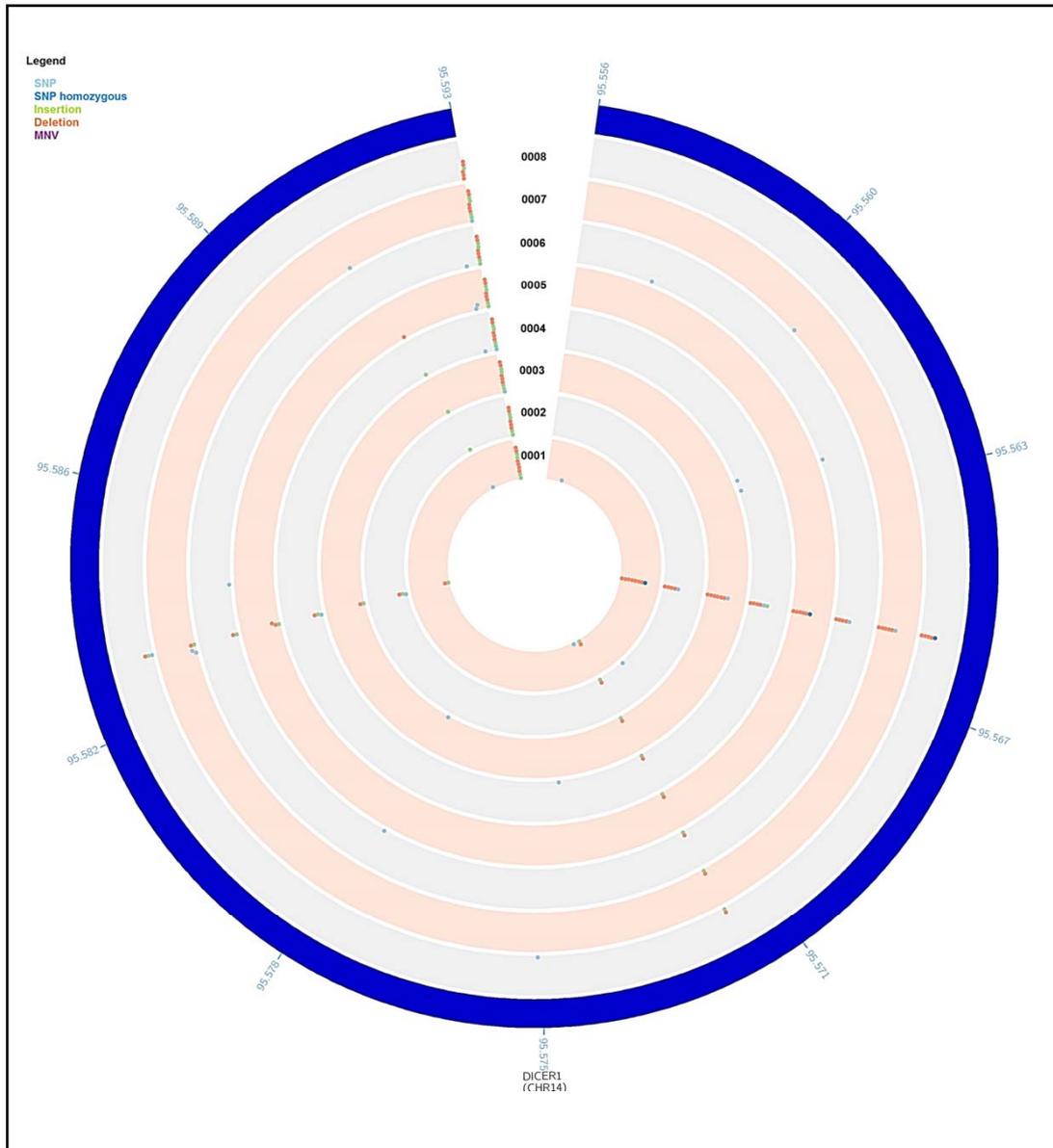


Fig. 1. Color-coded Circos plot with one track per sample was generated showing variants of the DICER1 gene in nodular basal cell carcinoma samples (samples 1 to 8). Tracks are labeled with the sample number. Within one track, dots are stacked if they are at the same or adjacent coordinates. Variants in the input files are symbolized in the plot with a colored dot in the DICER gene at the affected coordinate. The color of the dot is coding for the type of the variant. Variations include single nucleotide polymorphisms (SNPs; light blue), homozygous SNPs (turquoise), insertions (green), deletions (orange), and multiple nucleotide variation (MNV; purple).

Genevestigator analysis

After filtering for candidate genes with a log-ratio ≥ 2 , a control mean expression ≥ 1 and a cancer mean expression > 0 , 107 genes were described as differentially expressed when compared to control group 1 and 196 genes when compared to control group 2. In the next step, only those candidate genes with obvious upregulation in cancer subtypes were noted; we found a total of 15 genes for control group 1 and 63 genes for control group 2. Based on the 15 genes which were differentially expressed compared to control group 1, verification with microarray data based on the human AFFY_U133PLUS_2 array in Genevestigator

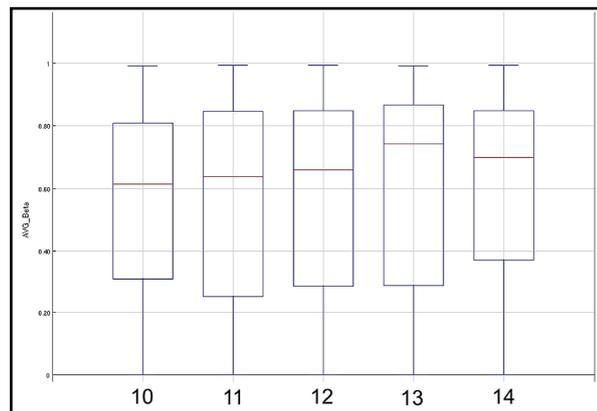


Fig. 3. A box plot for the 5 sclerodermiform basal cell carcinoma samples 10 to 14 showing the average methylation level β over all CpG sites. Equal data distribution in all these samples with a median average β -value of 0.3 to 0.9 is found for this data set indicating a high degree of similarity between the samples.

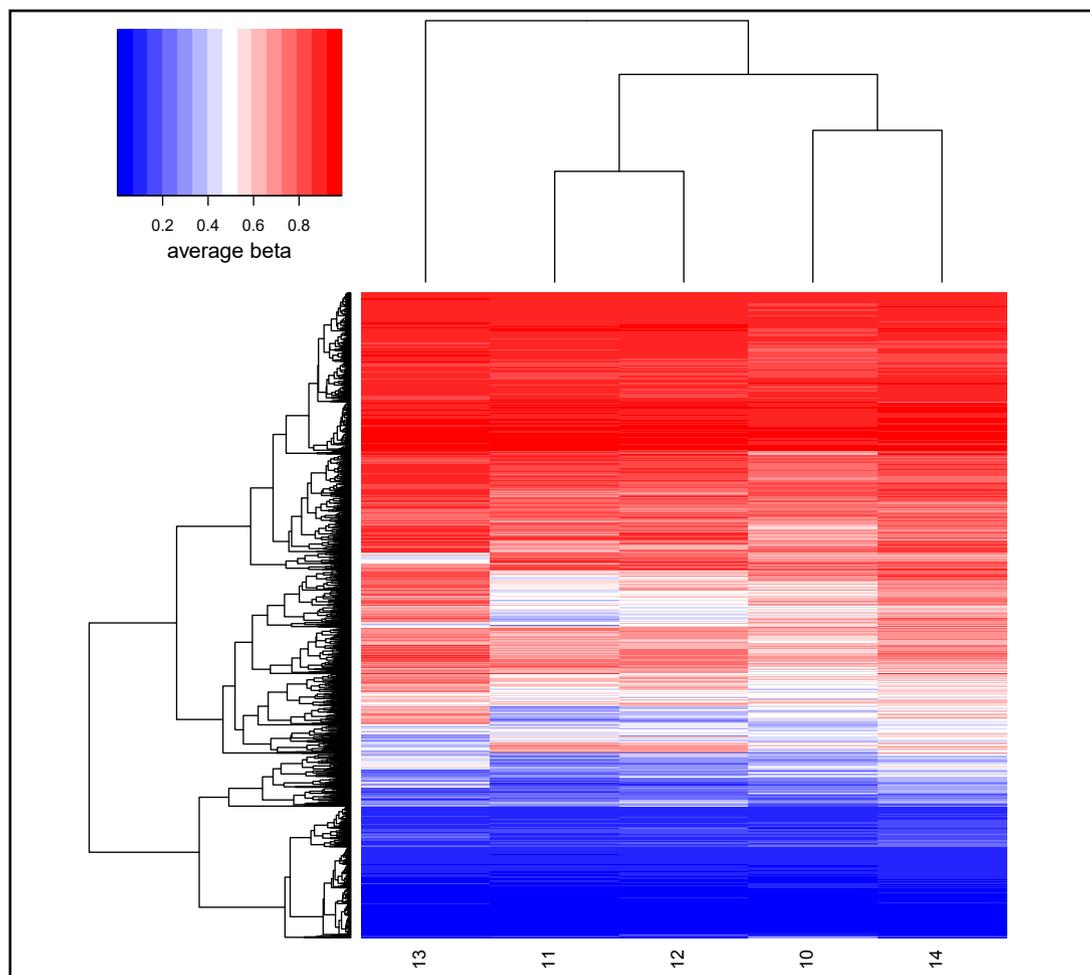
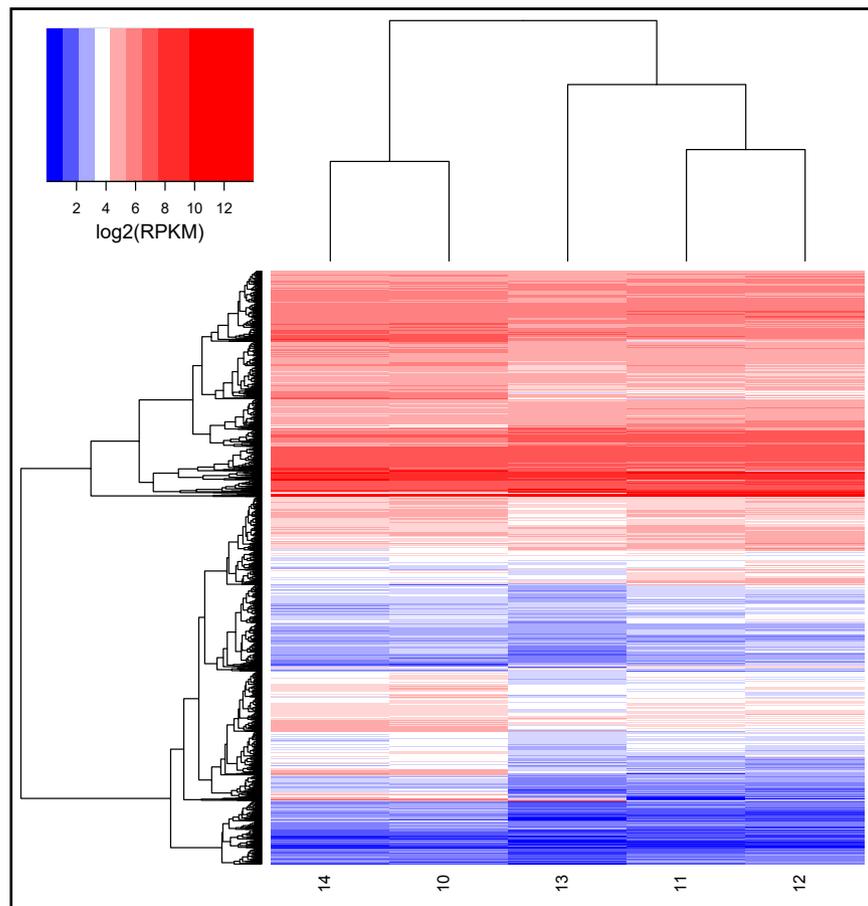


Fig. 4. Heatmap cluster analysis of DNA methylation results for sclerodermiform basal cell carcinoma samples 10 to 14. The color scale reflects the log2 signal intensity and runs from blue (low intensity) to white (medium intensity) to red (strong intensity).

Fig. 5. Heatmap cluster analysis of mRNA-Seq results for sclerodermiform basal cell carcinoma samples 10 to 14. The color scale reflects the \log_2 signal intensity and runs from blue (low intensity) to white (medium intensity) to red (strong intensity).



with public experiments GSE103439 revealed that Chitinase 3 Like 2 (CHI3L2;YKL-39), MYCN Proto-Oncogene (MYCN), and Receptor Activity Modifying Protein 1 (RAMP1) were upregulated in sBCC.

Discussion

Lyle et al. reported that a loss of Dicer in the epidermis induces extensive *DNA* damage and activates the *DNA* damage response and p53-dependent apoptosis [18]. They showed that Dicer and p53 cooperate to suppress mammalian skin carcinogenesis. Further, a lower expression of Dicer in BCC and reports of Dicer mutations in a variety of malignant tumors led to an investigation of the mutation status of the Dicer gene in BCC [10, 12]. We have described a variety of mutations in the Dicer gene; however, their clinical impact remains unclear and needs to be addressed in further validation studies based on the sequencing data acquired.

RNA sequencing showed S100A8 and S100A9 to be overexpressed. Calprotectin (a heterodimer of the two calcium-binding proteins S100A8 and S100A9) plays a role in inflammation-associated cancer [19]. S100A8 has shown a strong expression in hairless mice carrying homozygous mutations in the hairless gene showing rudimentary hair follicles (HFs) with enhanced susceptibility to BCC [20]. Further, the cytokeratin genes KRT1, KRT10, and KRT14 showed strong expression which aligns with immunohistochemistry protein level results [21]. Cytochrome C oxidase subunits I and III (COX1 and COX3) are encoded in the mitochondrial *DNA* and have been previously associated with tumorigenesis (COX-I) and glioblastoma mass size (COX-III) [22, 23]. CHI3L2, part of a group of chitinases

and chitinase-like proteins, has shown an increased expression in glioblastoma [24]. It is closely related to CHI3L1, showing structural similarity with CHI3L1 in size, nucleotide, and amino acid sequences (47% and 51% homology, respectively) with diverse functional activities. While CHI3L1 shows oncogenic properties, CHI3L2 shows a decrease of mitogenesis and proliferation, which may lead to differentiation. Copy number gains of the proto-oncogene MYCN may play a pivotal role in BCC development when BCC is associated with an increase in sonic hedgehog (SHH) expression (i.e., the main player of the SHH signaling pathway) [25].

Mutations in the SHH Patched receptor are highly associated with BCC [26]. Receptor Activity Modifying Protein 1 (RAMP1) is up-regulated in prostate cancer but has not been associated with cutaneous or epidermal tumors such as sBCC [27].

With regards to the *smallRNA* sequencing analysis in sBCC, miR-21 was strongly expressed. Most targets of miR-21 are tumor-suppressors, and miR-21 is strongly upregulated in a variety of cancers [28]. miR-21 controls cell proliferation, migration, and apoptosis in cancer cells by acting as an oncomir. However, its strong expression in sBCC has not been described. A miR-21 antagomir could be one potential form of therapy by inhibiting miR-21 in sBCC—a therapeutic tool for colorectal cancer on an experimental level *in vitro* [29]. Further investigation of this potential therapeutic effect is warranted in future studies. Although miR-99a has been described as a potential tumor suppressor strongly down-regulated in psoriatic skin, it is inversely expressed in sBCC [30]. Other *miRNAs* such as miR-125b, miR-99a, and miR-100 are part of a co-transcribed polycistronic *miRNA* cluster frequently found in cancer [31, 32].

Ultraviolet B (UV-B) radiation is one of the more important causative factors for BCC development. MiR-26a is up-regulated upon UV-B radiation, inducing apoptosis *in vitro* [33]. Interestingly, MiR-26a is strongly down-regulated in cutaneous melanoma; miR-26a replacement represents a potential therapeutic strategy for metastatic cutaneous melanoma [34]. The *let-7 miRNA* cluster has been associated with the proliferation,

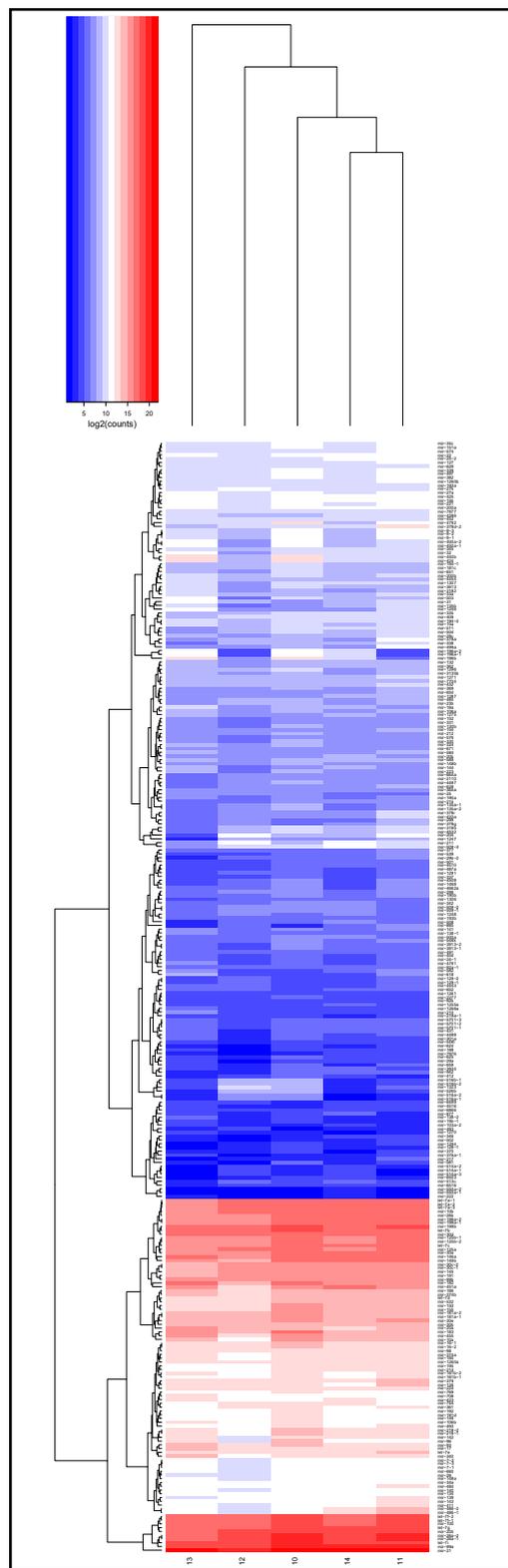


Fig. 6. Heatmap of smallRNA-Seq results for sclerodermiform basal cell carcinoma samples 10 to 14. The color scale reflects the log₂ signal intensity and runs from blue (low intensity) to white (medium intensity) to red (strong intensity).

apoptosis, and invasion potential of cancer cells. While the *let-7 miRNA* cluster has been widely described as a tumor suppressor (i.e., it shows reduced expression in cancer cells), we observed an increased expression of *let-7f*, *let-7g*, and *let-7i* in our group of sBCC [35]. The expression of miR-205 is highly specific for squamous epithelium and has been described as a marker for metastatic head and neck squamous cell carcinoma [36]. It has not been associated with sBCC thus far. MiR-205 expression is prominent in cutaneous stem cells promoting their neonatal expansion [37]. Its regulative potential on the phosphoinositide 3-kinase (PI3K) pathway was discovered by targeting negative regulators of PI3K signaling, thereby restricting the proliferation of cutaneous stem cells. HF stem cells belong to the miR-205 target group of cutaneous stem cells and are the cells of origin for BCC development [38, 39]. At the same time, the PI3K/AKT/mTOR signaling pathway is a downstream target pathway of tazarotene, a retinoid pro-drug that inhibits BCC carcinogenesis [40].

Conclusion

In conclusion, this study provides further evidence for a disturbance in non-protein-coding RNA expression such as *miRNAs* and its maturation machinery in BCC.

Acknowledgements

We are grateful to Dr. Amelie Schreieck and Stefan Kotschote, MS for technical advice.

Funding Information

We would like to thank Brigitte and Dr. Konstanze Wegener-Stiftung, Düsseldorf, Germany for financial support, which made this research possible.

Disclosure Statement

All authors hereby disclose any commercial associations that may pose or create a conflict of interest with the information presented in this manuscript. The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

- 1 Apalla Z, Lallas A, Sotiriou E, Lazaridou E, Ioannides D: Epidemiological trends in skin cancer. *Dermatol Pract Concept* 2017;7:1-6.
- 2 Verkouteren JAC, Ramdas KHR, Wakkee M, Nijsten T: Epidemiology of basal cell carcinoma: scholarly review. *Br J Dermatol* 2017;177:359-372.
- 3 Sand M, Skrygan M, Sand D, Georgas D, Hahn SA, Gambichler T, Altmeyer P, Bechara FG: Expression of microRNAs in basal cell carcinoma. *Br J Dermatol* 2012;167:847-855.
- 4 Sand M, Bechara FG, Sand D, Gambichler T, Hahn SA, Bromba M, Stockfleth E, Hessem S: Circular RNA expression in basal cell carcinoma. *Epigenomics* 2016;8:619-632.
- 5 Sand M, Bechara FG, Sand D, Gambichler T, Hahn SA, Bromba M, Stockfleth E, Hessem S: Long-noncoding RNAs in basal cell carcinoma. *Tumour Biol* 2016;37:10595-10608.
- 6 Balci S, Ayaz L, Gorur A, Yildirim Yaroglu H, Akbayir S, Dogruer Unal N, Bulut B, Tursen U, Tamer L: microRNA profiling for early detection of nonmelanoma skin cancer. *Clin Exp Dermatol* 2016;41:346-351.

- 7 Sand M, Bechara FG, Gambichler T, Sand D, Friedlander MR, Bromba M, Schnabel R, Hessam S: Next-generation sequencing of the basal cell carcinoma miRNome and a description of novel microRNA candidates under neoadjuvant vismodegib therapy: an integrative molecular and surgical case study. *Ann Oncol* 2016;27:332-338.
- 8 Sand M, Hessam S, Amur S, Skrygan M, Bromba M, Stockfleth E, Gambichler T, Bechara FG: Expression of oncogenic miR-17-92 and tumor suppressive miR-143-145 clusters in basal cell carcinoma and cutaneous squamous cell carcinoma. *J Dermatol Sci* 2017;86:142-148.
- 9 Sand M, Skrygan M, Georgas D, Arenz C, Gambichler T, Sand D, Altmeyer P, Bechara FG: Expression levels of the microRNA maturing microprocessor complex component DGCR8 and the RNA-induced silencing complex (RISC) components argonaute-1, argonaute-2, PACT, TARBP1, and TARBP2 in epithelial skin cancer. *Mol Carcinog* 2012;51:916-922.
- 10 Sand M, Gambichler T, Skrygan M, Sand D, Scola N, Altmeyer P, Bechara FG: Expression levels of the microRNA processing enzymes Drosha and Dicer in epithelial skin cancer. *Cancer Invest* 2010;28:649-653.
- 11 Sand M, Gambichler T, Sand D, Skrygan M, Altmeyer P, Bechara FG: MicroRNAs and the skin: tiny players in the body's largest organ. *J Dermatol Sci* 2009;53:169-175.
- 12 Foulkes WD, Priest JR, Duchaine TF: DICER1: mutations, microRNAs and mechanisms. *Nat Rev Cancer* 2014;14:662-672.
- 13 Haeussler M, Zweig AS, Tyner C, Speir ML, Rosenbloom KR, Raney BJ, Lee CM, Lee BT, Hinrichs AS, Gonzalez JN, Gibson D, Diekhans M, Clawson H, Casper J, Barber GP, Haussler D, Kuhn RM, Kent WJ: The UCSC Genome Browser database: 2019 update. *Nucleic Acids Res* 2019;47:D853-D858.
- 14 Garrison E, Marth G: Haplotype-based variant detection from short-read sequencing. arXiv:12073907 [q-bioGN], 2012. URL: <https://arxiv.org/abs/1207.3907v2>.
- 15 Stenson PD, Ball EV, Mort M, Phillips AD, Shiel JA, Thomas NS, Abeysinghe S, Krawczak M, Cooper DN: Human Gene Mutation Database (HGMD): 2003 update. *Hum Mutat* 2003;21:577-581.
- 16 Frankish A, Diekhans M, Ferreira AM, Johnson R, Jungreis I, Loveland J, Mudge JM, Sisu C, Wright J, Armstrong J, Barnes I, Berry A, Bignell A, Carbonell Sala S, Chrast J, Cunningham F, Di Domenico T, Donaldson S, Fiddes IT, Garcia Giron C, et al.: GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res* 2019;47:D766-D773.
- 17 Kozomara A, Griffiths-Jones S: miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res* 2014;42:D68-73.
- 18 Lyle S, Hoover K, Colpan C, Zhu Z, Matijasevic Z, Jones SN: Dicer cooperates with p53 to suppress DNA damage and skin carcinogenesis in mice. *PLoS One* 2014;9:e100920.
- 19 Gebhardt C, Nemeth J, Angel P, Hess J: S100A8 and S100A9 in inflammation and cancer. *Biochem Pharmacol* 2006;72:1622-1631.
- 20 Xu J, Weng Z, Arumugam A, Tang X, Chaudhary SC, Li C, Christiano AM, Elmets CA, Bickers DR, Athar M: Hair follicle disruption facilitates pathogenesis to UVB-induced cutaneous inflammation and basal cell carcinoma development in Ptch(+/-) mice. *Am J Pathol* 2014;184:1529-1540.
- 21 Yamamoto O, Asahi M: Cytokeratin expression in trichoblastic fibroma (small nodular type trichoblastoma), trichoepithelioma and basal cell carcinoma. *Br J Dermatol* 1999;140:8-16.
- 22 Pannunzio A, Coluccia M: Cyclooxygenase-1 (COX-1) and COX-1 Inhibitors in Cancer: A Review of Oncology and Medicinal Chemistry Literature. *Pharmaceuticals (Basel)* 2018;11:101-120.
- 23 Oksuz E, Atalar F, Tanirverdi G, Bilir A, Shahzadi A, Yazici Z: Therapeutic potential of cyclooxygenase-3 inhibitors in the management of glioblastoma. *J Neurooncol* 2016;126:271-278.
- 24 Areshkov PO, Avdieiev SS, Balynska OV, Leroith D, Kavsan VM: Two closely related human members of chitinase-like family, CHI3L1 and CHI3L2, activate ERK1/2 in 293 and U373 cells but have the different influence on cell proliferation. *Int J Biol Sci* 2012;8:39-48.
- 25 Ruiz-Perez MV, Henley AB, Arsenian-Henriksson M: The MYCN Protein in Health and Disease. *Genes (Basel)* 2017;8:113.
- 26 Lupi O: Correlations between the Sonic Hedgehog pathway and basal cell carcinoma. *Int J Dermatol* 2007;46:1113-1117.
- 27 Logan M, Anderson PD, Saab ST, Hameed O, Abdulkadir SA: RAMP1 is a direct NKX3.1 target gene up-regulated in prostate cancer that promotes tumorigenesis. *Am J Pathol* 2013;183:951-963.
- 28 Pfeffer SR, Yang CH, Pfeffer LM: The Role of miR-21 in Cancer. *Drug Dev Res* 2015;76:270-277.

- 29 Song MS, Rossi JJ: The anti-miR21 antagomir, a therapeutic tool for colorectal cancer, has a potential synergistic effect by perturbing an angiogenesis-associated miR30. *Front Genet* 2014;4:301.
- 30 Hawkes JE, Nguyen GH, Fujita M, Florell SR, Callis Duffin K, Krueger GG, O'Connell RM: microRNAs in Psoriasis. *J Invest Dermatol* 2016;136:365-371.
- 31 Zhang L, Ge Y, Fuchs E: miR-125b can enhance skin tumor initiation and promote malignant progression by repressing differentiation and prolonging cell survival. *Genes Dev* 2014;28:2532-2546.
- 32 Emmrich S, Rasche M, Schoning J, Reimer C, Keihani S, Maroz A, Xie Y, Li Z, Schambach A, Reinhardt D, Klusmann JH: miR-99a/100~125b tricistrons regulate hematopoietic stem and progenitor cell homeostasis by shifting the balance between TGFbeta and Wnt signaling. *Genes Dev* 2014;28:858-874.
- 33 Zhang T, Qian H, Hu C, Lu H, Li JB, Wu YF, Li W: miR-26a Mediates Ultraviolet B-Induced Apoptosis by Targeting Histone Methyltransferase EZH2 Depending on Myc Expression. *Cell Physiol Biochem* 2017;43:1188-1197.
- 34 Reuland SN, Smith SM, Bemis LT, Goldstein NB, Almeida AR, Partyka KA, Marquez VE, Zhang Q, Norris DA, Shellman YG: MicroRNA-26a is strongly downregulated in melanoma and induces cell death through repression of silencer of death domains (SODD). *J Invest Dermatol* 2013;133:1286-1293.
- 35 Balzeau J, Menezes MR, Cao S, Hagan JP: The LIN28/let-7 Pathway in Cancer. *Front Genet* 2017;8:31.
- 36 Fletcher AM, Heaford AC, Trask DK: Detection of metastatic head and neck squamous cell carcinoma using the relative expression of tissue-specific mir-205. *Transl Oncol* 2008;1:202-208.
- 37 Wang D, Zhang Z, O'Loughlin E, Wang L, Fan X, Lai EC, Yi R: MicroRNA-205 controls neonatal expansion of skin stem cells by modulating the PI(3)K pathway. *Nat Cell Biol* 2013;15:1153-1163.
- 38 Peterson SC, Eberl M, Vagnozzi AN, Belkadi A, Veniaminova NA, Verhaegen ME, Bichakjian CK, Ward NL, Dlugosz AA, Wong SY: Basal cell carcinoma preferentially arises from stem cells within hair follicle and mechanosensory niches. *Cell Stem Cell* 2015;16:400-412.
- 39 Wang GY, Wang J, Mancianti ML, Epstein EH Jr.: Basal cell carcinomas arise from hair follicle stem cells in Ptch1(+/-) mice. *Cancer Cell* 2011;19:114-124.
- 40 So PL, Wang GY, Wang K, Chuang M, Chiueh VC, Kenny PA, Epstein EH, Jr.: PI3K-AKT signaling is a downstream effector of retinoid prevention of murine basal cell carcinogenesis. *Cancer Prev Res (Phila)* 2014;7:407-417.