



OPEN

Concomitant DNA methylation and transcriptome signatures define epidermal responses to acute solar UV radiation

Nicholas Holzschek^{1,2}✉, Jörn Söhle¹, Torsten Schläger¹, Cassandra Falckenhayn¹, Elke Grönniger¹, Ludger Kolbe¹, Horst Wenck¹, Lara Terstegen¹, Lars Kaderali², Marc Winnefeld¹ & Katharina Gorges¹✉

The simultaneous analysis of different regulatory levels of biological phenomena by means of multi-omics data integration has proven an invaluable tool in modern precision medicine, yet many processes ultimately paving the way towards disease manifestation remain elusive and have not been studied in this regard. Here we investigated the early molecular events following repetitive UV irradiation of in vivo healthy human skin in depth on transcriptomic and epigenetic level. Our results provide first hints towards an immediate acquisition of epigenetic memories related to aging and cancer and demonstrate significantly correlated epigenetic and transcriptomic responses to irradiation stress. The data allowed the precise prediction of inter-individual UV sensitivity, and molecular subtyping on the integrated post-irradiation multi-omics data established the existence of three latent molecular phototypes. Importantly, further analysis suggested a form of melanin-independent DNA damage protection in subjects with higher innate UV resilience. This work establishes a high-resolution molecular landscape of the acute epidermal UV response and demonstrates the potential of integrative analyses to untangle complex and heterogeneous biological responses.

Solar UV irradiation has complex and ambivalent effects on the human organism. Beneficial effects of sun exposure are thought to be mainly mediated by vitamin D, which is synthesized in the skin through a photosynthetic reaction triggered by exposure to UVB. Vitamin D was primarily acknowledged for its importance in bone formation, increasing evidence however points to its influence on the proper functioning of nearly every tissue in our bodies¹. In contrast to this however, solar UV irradiation is also the most abundant risk factor for skin cancer and other extrinsically influenced skin disorders^{2,3}. It is well established that UV irradiation both directly and indirectly induces DNA damage. Direct damage is mainly a result of UVB and to lesser extent UVA irradiation, causing dimerization of adjacent pyrimidine bases, a frequent cause of mutations during replication⁴. Indirect DNA damage results mainly from oxidative stress, caused by free radicals and cellular reactive oxygen species, which increase after UV irradiation⁵. Damaged DNA, if not properly repaired, interferes with many cellular mechanisms such as transcription, the cell cycle and replication and can give rise to mutations and epigenetic alterations, driving genomic instability and ultimately carcinogenesis.

Human skin has developed several defense systems to guard against the damaging effects of UV: Prominently these include structural changes to the tissue such as epidermal thickening and the synthesis of melanin, but they also comprise quick molecular adaptations like the suspension of cell cycle and gene transcription, as well as the activation of DNA repair pathways. The extent of protection afforded by these mechanisms however is characterized by high inter-individual variation⁶. The stratification of individual UV response is thus highly important for risk assessment in cancer prevention (UV-protection), therapeutic dose determination (PUVA therapy) and in the understanding of the biological processes leading to malignancies (e.g. squamous skin cancers). Fitzpatrick skin type categories⁷ have been widely used as an indicator and predictor of sun sensitivity in epidemiology and experimental photobiology. However, this categorization is hampered by subjectivity and is prone to recall

¹Beiersdorf AG, Research and Development, Tropelwitzstraße 17, 22529 Hamburg, Germany. ²Institute for Bioinformatics, University Medicine Greifswald, Greifswald, Germany. ✉email: nicholas.holzschek@beiersdorf.com; katharina.roeck@beiersdorf.com

error⁸. In a study assessing the reliability of Fitzpatrick skin type classifications for instance, only ~60% of all study participants self-identified as the same skin type after repeated questioning a few months later⁹. Several authors have investigated the relationship between Fitzpatrick skin type and minimal erythema dose (MED)^{10–13}, a more objective measure of UV sensitivity frequently used in clinical or research settings, showing increasing MED with higher Fitzpatrick classification in general, but with considerable intergroup variation.

DNA methylation is a covalent epigenetic modification of cytosine to 5-methylcytosine, occurring within CpG dinucleotides^{14,15}. Although methylations of adenine have been reported as well, these have so far received considerably less attention. Methylation of CpG sites in the human genome is an important regulatory mechanism that can lead to the activation or repression of gene transcription. Modifications are established and maintained by a set of specific enzymes called DNA methyltransferases. DNA methylation is generally considered to represent a regulatory interface between environmental cues and the genome and might cause or allow long-lasting changes in gene transcriptional activity¹⁶. Our current knowledge about epigenetic changes associated with acute UV irradiation, its contribution to transcriptomic alterations and implication in skin photobiology, remains very limited. Previous studies have shown however, that chronic solar UV gives rise to large hypomethylated blocks of DNA in the healthy epidermis and that these blocks are conserved in cutaneous squamous skin carcinomas¹⁷, underlining the importance of studying DNA methylation in the context of solar irradiation. In addition, a multi-omics analysis of UV irradiated keratinocytes recently identified several new UV target genes including CYP24A1, GJA5, SLAMF7 and ETV1¹⁸, demonstrating the value of multi-layered omics analyses in unraveling biological phenomena and enabling more reliable biomarker detection, as it has similarly been shown in cancer research, allowing molecular diagnosis and prognosis, often utilizing DNA methylation markers^{19,20}.

Here we hypothesized that integrative analysis of UV induced epigenetic and transcriptomic alterations *in vivo* might help to decipher inter-individual responses to environmental challenges and give hints towards early pathogenesis. For this reason, we generated high-resolution multi-omics molecular profiles of the *in vivo* irradiated epidermis. Our results provide evidence that a UV induced epigenetic memory might be established already after short term repetitive UV irradiation. Integrative analyses of methylation and expression data reveal previously unnoted pathways involved in the acute epidermal UV response and allow the precise inter-individual prediction of MED without the need for prior UV irradiation. Finally, analysis of these molecular phototypes indicates the existence of a melanin-independent form of damage protection in individuals with higher innate resilience to UV irradiation.

Results

UV irradiated epidermis shows genome-wide aberrant methylation patterns and substantial transcriptomic reprogramming. Elucidating the complex molecular mechanisms underlying UV-gene interaction might offer new insights into how UV modulates skin homeostasis and disease pathogenesis to help improve the prevention of UV-induced skin aging and related pathologies. In order to obtain a comprehensive picture of the molecular events regulating acute epidermal photobiology, 32 female Caucasian volunteers (Fitzpatrick phototypes 1–4) were irradiated with individually calibrated doses of 0.9 MED using a full spectrum solar simulator on three subsequent days on a sun-protected area on their lower backs. 24 h after the last irradiation, suction blister roofs were extracted from irradiated and control sites of each subject and gene expression profiling (Illumina RNA seq) and concomitant DNA methylation profiling (Illumina EPIC Arrays) were performed. Paired differential expression and methylation analyses between irradiated and control areas revealed that in total 20.5% (FDR < 0.05) of all interrogated CpGs and 32.4% (FDR < 0.05) of all detected gene transcripts were significantly altered in response to irradiation. These considerable changes were spread over the whole genome, with notable exceptions only occurring in the regions around the centromeres and in some constitutively heterochromatic regions e.g. on chromosome 13 (Fig. 1a). In general, a tendency towards hypomethylation was detected with 65.1% of all significant CpGs decreasing in methylation. Notably, the tendency towards hypomethylation increased from open sea regions to CpG-islands (Fig. S1 a).

Large blocks of the genome have previously been shown to be hypomethylated in chronically sun-exposed epidermal samples in comparison to protected skin^{17,21} and have also been associated with clinical measures of photoaging¹⁷. How quickly this epigenetic imprinting in response to UV exposure occurs however, is so far unknown. We thus investigated whether early indications of photoaging were already detectable after acute repetitive UV irradiation and analyzed the methylation status within the previously reported regions¹⁷ in our data. We found that in over a fifth of the originally described genomic blocks (49/224) the observed methylation changes after acute irradiation correlated very well with the reported patterns (Fig. 1b), differing mainly in magnitude in comparison to chronically exposed skin. This delivers evidence that epigenetic alterations in response to extrinsic stimuli can manifest quickly after external stimulation and suggests that even few repetitive sunburns can be sufficient to impact epigenetic imprinting in genomic regions associated with extensive photoaging.

Considering the extent of alterations in methylation patterns in response to acute irradiation and the universal role of DNA methylation in cancer biology, we then also performed a comparison of pan-cancer methylation signatures²² to our data, to establish if any overlap in signatures could be observed. The analysis revealed a small number of genomic regions with methylation changes post-irradiation very much reminiscent of those found aberrantly methylated in cancerous tissue. Most of these showed extensive hypomethylation (Supplementary Fig. S2a,b). Whether these alterations are in fact linked to carcinogenesis or purely a product of stochasticity will remain to be determined, but the overlap and extent of correlation raises concern and might warrant further investigation.

Genome-wide correlative analyses of gene expression and methylation reveal coordinated changes in known and novel players of the UV response. Methylation can lead to long-lasting acti-

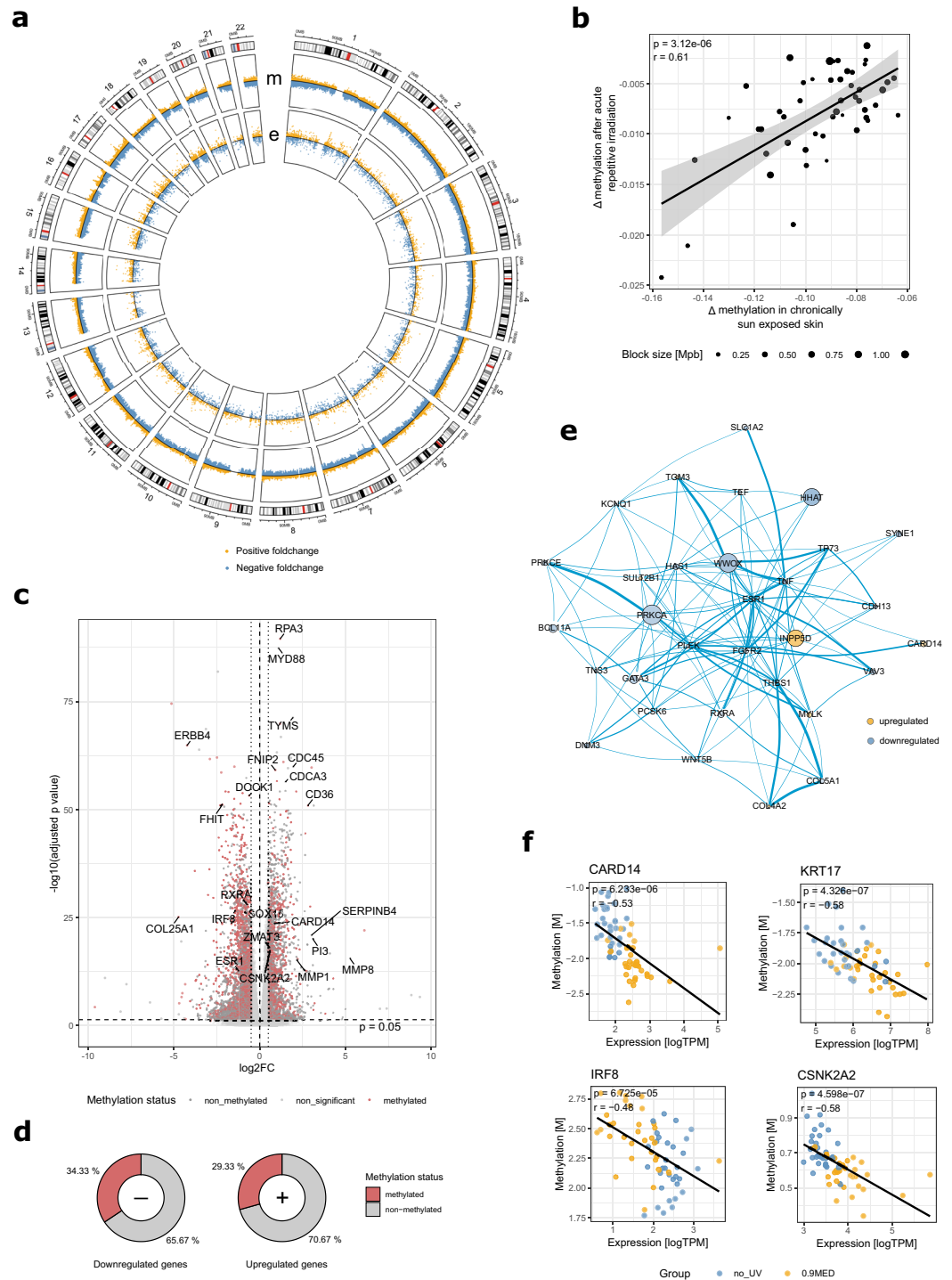


Figure 1. Epigenetic and transcriptomic changes of irradiated samples compared to non-irradiated controls: (a) Circos plot showing differential methylation (m, outer circle) and expression (e, inner circle) in response to irradiation to 0.9 MED in a genomic context (FDR < 0.05). Amplitude of points corresponds to log₂ fold-change with the solid black line representing no change. Hypomethylated CpGs and downregulated genes are colored in blue, hypermethylated CpGs and upregulated genes in yellow. Colored bands in the karyogram mark centromeres (red) and heterochromatin status (grey to black). (b) Differential methylation of 49 genomic regions previously associated with chronic sun-exposure¹⁷ compared to differential methylation after acute repetitive irradiation. (c) Volcano plot of differential gene expression in response to irradiation. Differentially expressed genes with ≥ 3 differentially methylated CpGs are marked in red. (d) Genome-wide ratio of differentially up- and downregulated genes with concomitant change in methylation (≥ 3 CpGs). (e) Protein–protein-interaction network between the most interconnected differentially expressed and methylated genes. Points are scaled by the negative logarithmized FDR of differential expression and colored by confidence of interaction. Edges are scaled by log₂ fold-changes. (f) Significantly correlated differential expression and enhancer methylation of CARD14, expression and TSS200 methylation of IRF8, expression and TSS200 methylation of CSNK2A2, and expression and TSS200 methylation of KRT17. Plots were generated using R v3.6.1¹⁶ software.

vation or repression of gene transcription and analysis of simultaneously regulated genes on methylation and transcription level have been shown to yield higher prognostic values in several pathophysiological states^{23,24}. We thus mapped the most stably differentially expressed genes and CpGs by genomic position and observed a high proportion of genes with equally pronounced methylation changes. Analysis revealed that 29.3% of all upregulated genes (FDR < 0.05 and $\log_2FC_{\text{abs}} > 0.5$) harbored at least three differentially methylated CpGs (FDR < 0.05 and $\log_2FC_{\text{abs}} > 0.2$), whereas for downregulated genes this number increased significantly further to 34.3% (Fig. 1c,d), with a high number of genes exhibiting inverse correlations to their methylation state. Stratifying the significant CpGs within these genes by regulatory regions revealed they were most frequently located in enhancers and more seldom in exon regions (Supplementary Fig. S1d). Analyzing known protein–protein-interactions between these differentially regulated and methylated genes (DEMGs) using STRING²⁵ revealed a network of highly interconnected proteins surrounding ESR1, the estrogen receptor α , which was found downregulated following repetitive irradiation (Fig. 1e). Estrogen receptor α expression has been shown to be reduced following UV irradiation in vitro before²⁶, and its activity has been linked to photoimmune suppression in animal studies. In mice, estrogen receptor antagonists were found to exacerbate immune suppressive action in a dose-dependent manner with estradiol treatment exerting protective effects respectively²⁷, and the estrogenic compound equol protecting against irradiation-induced carcinogenesis²⁸. Notably, the core network also involved the similarly downregulated retinoic \times receptor α , previously linked to a functional vitamin A deficiency in the skin following UV irradiation²⁹ and thereby contributing to photoaging.

We next performed genome wide correlation analyses of all annotated genes and interrogated CpGs to identify significant linear correlations between gene expression and methylation in annotated functional gene regions. Modeling gene expression as a function of mean methylation for all CpGs in potentially regulatory gene regions (enhancers, 1,500 bp and 200 bp upstream of the TSS as well as exon regions) revealed 2,267 significant associations after multiple testing correction. Again, most of these associations were found with alterations in methylation patterns in enhancer regions. Among these highly correlated differentially expressed and methylated genes we identified several known and previously described actors in the UV response, such as CYP24A1, BRCA2, NOTCH2, FOXO3 and GATA3. Examples also included the observed hypomethylation and upregulation of CSNK2A2, a catalytic subunit of Casein kinase II, a ubiquitous serine/threonine protein kinase involved with a manifold of cellular processes, such as cell cycle control and apoptosis and the immune-modulatory keratin KRT17 (Fig. 1f), both of which have previously been associated with UV response and tumorigenesis. Remarkably some of the identified genes, e.g. CARD14 or IRF8 (Fig. 1f), have thus far not been associated with UV irradiation, possibly reflecting the variance between in vivo and in vitro generated data. Interestingly however, CARD14 mutations have been observed previously in psoriasis patients. Gain-of-function CARD14 mutations in mice lead to spontaneous psoriasis-like skin inflammation by inducing activation of the IL-23-IL-17 axis in keratinocytes and thereby immune cell infiltration³⁰. In contrast CARD14^{-/-} mice displayed attenuated skin inflammation in murine psoriasis models³¹. Demethylation-driven CARD14 activation in irradiated cells of the human epidermis might thus present a hitherto undiscovered mechanism of epidermal UV response. The transcription factor IRF8 was found concomitantly significantly hypermethylated and downregulated, which is significant given its function as a tumor suppressor and its frequent downregulation in various cancer types through epigenetic silencing^{32–34}. Recently IRF8 has further been implicated in cutaneous wound healing³⁵, the methylation-driven downregulation of IRF8 might therefore constitute a novel mechanism contributing to the observed impairment of wound healing following irradiation. Notably, IRF8 is located within one of the genomic regions differentially methylated in photoaged skin¹⁷, it would therefore be interesting to investigate its functional role in photoaging, even more so considering the age-associated impairment of wound healing in the skin and the increased risk of developing skin cancer that is associated with both chronic sun-exposure and higher age.

Pathway analysis shows distinct functional enrichments for methylation-associated transcriptional alterations.

Since the dissection of DEMGs revealed several genes which had previously not been connected to epidermal UV responses, we performed pathway analyses by means of gene set enrichment. Multiple pathways were strongly enriched with DEMGs, including DNA repair, immune signaling and stress response, strengthening the notion that DEMGs are at the heart of known and key response mechanisms to UV irradiation (Fig. 2a,b). In addition, a high number of enriched pathways were involved in metabolic processes, including some that had previously not been assigned to the canonical UV response pathways. Prominently these were linked to lipid biosynthetic and cofactor metabolic processes (Fig. 2c). Lipid synthesis in the epidermis is vital to skin permeability and barrier function, one of the skin's most crucial functions. Outer epidermal keratinocytes secrete lamellar bodies, which are unique to the epidermis³⁶ and contain phospholipids, glycosyl-ceramides, sphingomyelin, as well as cholesterol and numerous enzymes, including lipid hydrolases, such as β -glucocerebrosidase, acidic sphingomyelinase, secretory phospholipase A2 (sPLA2), and acidic/neutral lipases^{37,38}. When the permeability barrier is perturbed, both the secretion and synthesis of lamellar bodies is stimulated, which allows for the rapid repair and normalization of permeability barrier function³⁹. So far only few studies have evaluated the effect of UV on the stratum corneum. They provide evidence for an increased epidermal lipid synthesis in response to UV radiation and alterations of lipid profiles^{40–42}, however these studies gave no functional correlation to genes or mechanisms involved. In addition, atopic dermatitis and psoriasis patients display modified lipid profiles and both groups are known to benefit from UV therapy. Induced DEMGs related to lipid biosynthetic processes might therefore provide evidence of an understudied UV response mechanism and potentially aid in identifying novel targets to help the regeneration of diseased skin. Differentially upregulated genes involved with other notably positively enriched pathways, such as TYMS and DHFR (Fig. 2c), are mainly involved in nucleotide synthesis and alterations to their increased expression may be part of important cellular responses that ensure proper DNA repair through the replenishment of DNA precursor molecules.

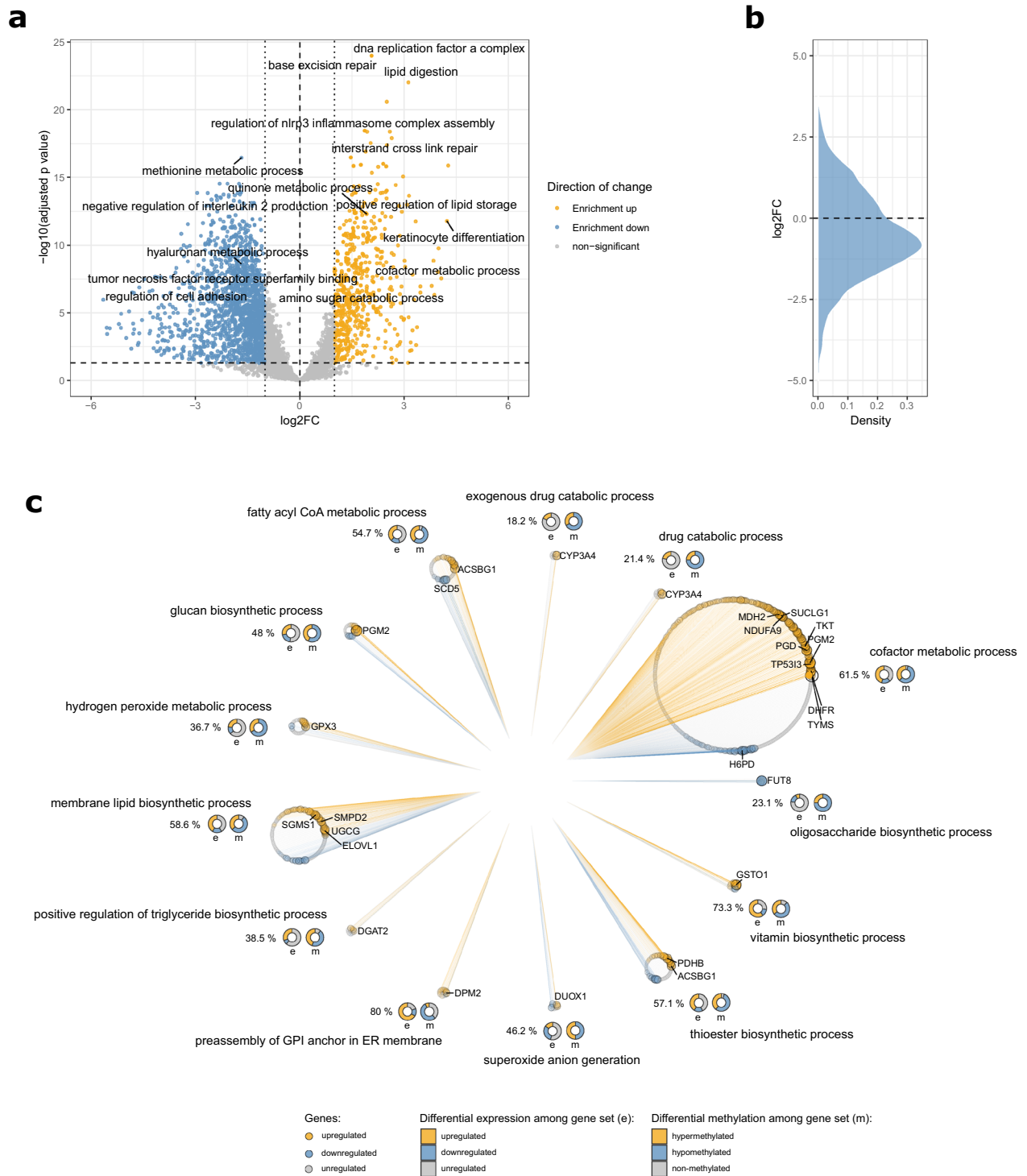


Figure 2. Biological pathways affected by simultaneous changes in both methylation and expression patterns in response to UV irradiation: **(a)** Volcano plot of enriched GO terms based on the analysis of differentially expressed genes with concomitant changes in methylation patterns (≥ 3 CpGs) with positively enriched pathways colored in yellow and negatively enriched pathways in blue **(b)** Distribution of \log_2 fold-changes in pathway enrichment after irradiation. **(c)** Enriched pathways involved with lipid biosynthesis and cofactor metabolic processes. Pathways are shown as circles with points corresponding to genes annotated to each respective pathway. Genes are colored by up- (yellow) or downregulation (blue) with size scaled to the negative \log_{10} of the FDR derived from differential expression analysis and ordered by \log_2 fold-changes. Circles underneath pathway names represent proportions of differentially regulated (e) and methylated (m) genes within each gene set. Numbers to the left of the circles summarize the overall percentage of differentially expressed genes per pathway (FDR < 0.05). Plots were generated using R v3.6.1⁷⁶ software.

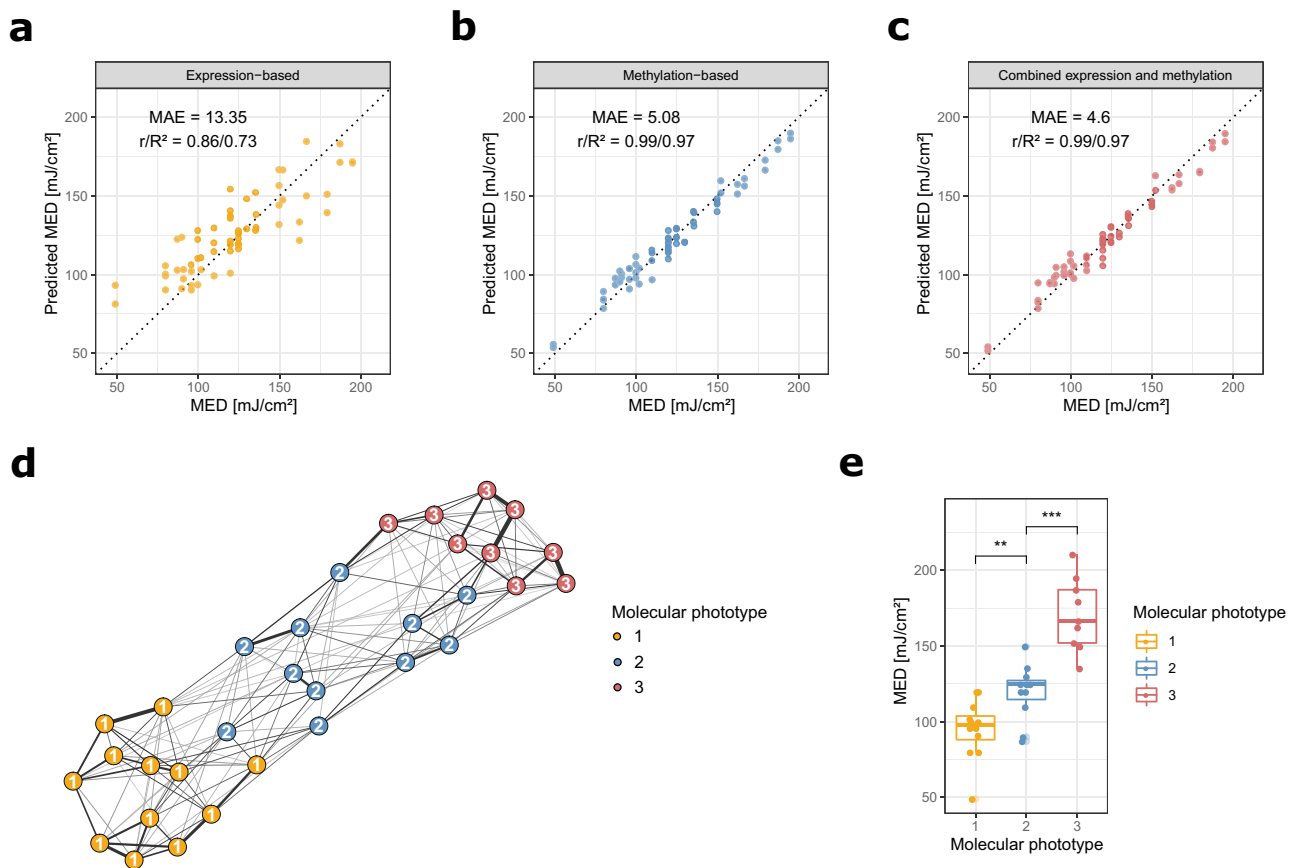


Figure 3. Prediction of inter-individual UV sensitivity from molecular data and identification of MED-correlated molecular phototypes: (a) Cross-validated predictions of MED from gene expression data using lasso regression models. (b) Cross-validated predictions of MED from DNA methylation data. (c) Cross-validated predictions of MED from combination of gene expression and DNA methylation data. (d) Fused similarity network generated from gene expression and DNA methylation data of irradiated samples, with nodes colored by molecular phototypes identified through spectral clustering. (e) Distribution of MED stratified by the molecular phototypes identified through spectral clustering on the fused similarity network. Statistical comparison was performed using unpaired two-sided t-tests. Plots were generated using R v3.6.1⁷⁶ software.

The extent and magnitude of differential regulation in these pathways indicates high cellular priorities of these processes. These findings might warrant further investigation, as these pathways may be vital to maintaining genomic stability after UV irradiation.

Molecular data allow precise inter-individual prediction of UV tolerance without experimental irradiation.

Prediction of UV response is an important tool for risk assessment and prognostication of sun tolerance, photoaging, skin cancer and phototherapy. As a proxy for UV sensitivity, the Fitzpatrick scale is often used⁷. The Fitzpatrick scale or Fitzpatrick phototypes are a subjective, semi-quantitative scale made up of six phototypes that describe skin color by basal complexion, melanin level, and subjective assessment of inflammatory response to UV⁴³. A more accurate way to measure UV tolerance is the experimental determination of the MED, which includes the acute irradiation of a test area with different UV dosages and a subsequent assessment of the minimal dose leading to erythema manifestation^{13,44}. This method produces accurate and objective results, but is potentially harmful as it exposes the test subject to UV irradiation during the assessment. In the present study, subjects ranging from Fitzpatrick phototype 1 to 4 were analyzed and their MED assessed. As expected from previously published data^{8,9,11}, stratification of donors using the Fitzpatrick classification was a relatively poor predictor of MED. For instance, the measured MED values for subjects of Fitzpatrick phototype 4 varied from 99.7 up to 210.4 mJ/cm². We thus set out to explore if the assessment of individual UV sensitivity could be improved using molecular markers, forgoing the necessity to expose test subjects to harmful UV irradiation in the first place. We employed lasso regression models to attempt the prediction of individual UV sensitivity, as measured by MED, based on gene expression and DNA methylation data and a dataset combining expression and methylation features. The data included both irradiated and control samples, in order for the models to select features that would allow reliable estimation of UV tolerance irrespective of prior sun exposure of the tissue. The tenfold cross-validated predictions showed a high accuracy achieved by both expression- and methylation-based models (Fig. 3a,b), far outperforming the Fitzpatrick classifications with median absolute errors of 13.35 mJ/cm² (expression-based) and 5.08 mJ/cm² (methylation-based). Models built using DNA methylation

features in particular were able to predict individual UV sensitivity to a remarkable degree, indicating a strong epigenetic component associated with UV tolerance. The combination of both expression and methylation data yielded the most accurate prediction model with a median absolute error of 4.6 mJ/cm², suggesting further complementarity in the two data levels (Fig. 3c). Model performance was similar on irradiated and control samples, demonstrating the utility of the method irrespective of exposure status (Supplementary Fig. S3a–c). To our knowledge this is the first attempt to derive an accurate estimation of UV sensitivity purely from molecular data, which provides a reliable tool to assess individual UV tolerance, which importantly does not necessitate putting patients at risk of prior irradiation of the skin, as is the case with regular MED assessment.

Multi-omics integration allows the identification of latent subgroups among irradiated samples. Considering the predictivity of the multi-omics data with regard to UV sensitivity, we decided to use an integrative approach to search for heterogeneity in the biological UV response. For this we integrated gene expression and methylation data from irradiated samples using similarity network fusion⁴⁵. Similarity network fusion is a flexible network-based method for integrating different levels biological data, otherwise mostly employed in cancer research: First, a separate similarity network is created from each data level, with samples represented as nodes and similarities in profiles as edges. In a second step, the separate networks are then integrated using an iterative algorithm that strengthens edges between individual samples present in several levels of data, and finally converges into a fused similarity network that incorporates information from all the different data levels. In our case, this led to a fused network incorporating information from both gene expression and DNA methylation data of the irradiated samples (Fig. 3d). Spectral clustering on the fused network then identified three latent subgroups in the multi-omics data, indicating differences in the biological responses to UV by different test subjects, and allowing their classification into distinct subtypes. The identified subtypes showed very high association to the MED (Fig. 3e) and allowed a better stratification of subjects based on UV sensitivity than Fitzpatrick phototypes, especially in the higher MED range (Supplementary Fig. S4a,b). Molecular subtyping of the skin with regards to UV response using these molecular phototypes (MPs) could prove helpful in developing preventive interventions, stratifying patients for risk factors (e.g. skin cancer and disease) and yielding deeper insights into molecular response mechanisms to irradiation.

Molecular phototypes reveal divergent biological responses to UV irradiation connected to cytokine response, programmed cell death and DNA damage sensing and repair. To characterize the biological processes underpinning the variability of UV responses exhibited by the identified MPs, we assessed the importance of all pathways in the GO term collection with regard to UV response. For this we employed pathway-based machine-learning classifiers that were based on support vector machines using radial basis function kernels, capable of learning non-linear patterns from high-dimensional data. These classifiers were trained to predict irradiation status of a sample using gene expression data from a given pathway, and each “pathway model” was subsequently scored for how well it enabled discrimination between the groups in a repeated cross-validation scheme. This yielded a predictivity score for every gene set, ranking all pathways on a common scale whilst also capturing non-linear gene regulation patterns. Predictivity was assessed for all pathways stratified by the three identified MPs, allowing the identification of biological processes predictive for the UV response for a given subtype and thus also revealing pathways whose regulation diverges between the three subgroups. This resulted in a mapping of the whole pathway landscape with regard to UV response relevance within the three subtypes (Supplementary Fig. S5).

The on average most predictive pathways were involved with DNA damage response mechanisms such as cell cycle transition, DNA replication and chromosome condensation in concordance with the top pathways obtained using gene set enrichment earlier. Further analysis of the involved pathways however revealed divergent patterns between the three molecular subtypes (Fig. 4a). MP 1 and 2 for instance exhibited stronger signals in pathways associated with inflammatory and immune signaling in comparison to MP 3. In case of MP 1, the subgroup with the lowest average MED, these related strongly to inflammasome activation and cytokine response, both generally well-described responses in regards to UV irradiation in human skin^{46–49}. In comparison, MP 2 exhibited decreased inflammasome predictivity scores but on the other hand a stronger type I interferon response than either MP 1 or MP 3. MP 2 was further singled out by stronger signals detected in apoptotic and autophagy pathways compared to the other subgroups. This might be connected to a stronger regulation in p53 related signaling pathways, as signaling by p53 class mediators showed increased predictivity in this subtype accordingly. Taken together this could indicate a higher efficiency in clearing cells with unrepairable DNA damage from the tissue. Both MP 2 and MP 3 further showed higher activities in pigment metabolic processes, which is in concordance with the stronger tanning responses observed in more UV tolerant skin⁵⁰. MP 3 on the other hand, incorporating subjects with the highest recorded UV resilience in our cohort, was defined by the strongest pathway signals detected in cell cycle checkpoint and DNA synthesis pathways, as well as genes involved with chromosome condensation. These findings are indicative of a higher sensitivity of the DNA damage sensing machinery in MP 3 subjects in response to irradiation, which would provide a more tightly regulated cessation of DNA replication and thus more time for the repair of UV-induced DNA damage. This hypothesis led us to investigate the extent of DNA photodamage in the samples of study subjects from the different molecular phototypes. We profiled the most common and important form of UV-induced damage to the DNA, the formation of cyclobutane pyrimidine dimers (CPDs), a frequent cause of mutation in the skin after UV irradiation, that directly links UV damage to carcinogenesis⁵¹. Analysis of the extent of CPDs detectable in the samples revealed lower abundances of CPD-alterations in the DNA of MP 3 subjects compared to the other molecular phototypes (Fig. 4b). This supports not only the model predictions but also the hypothesis of a pigmentation-independent UV protective mechanism in highly UV tolerant skin after repetitive irradiation. The identification of the direct

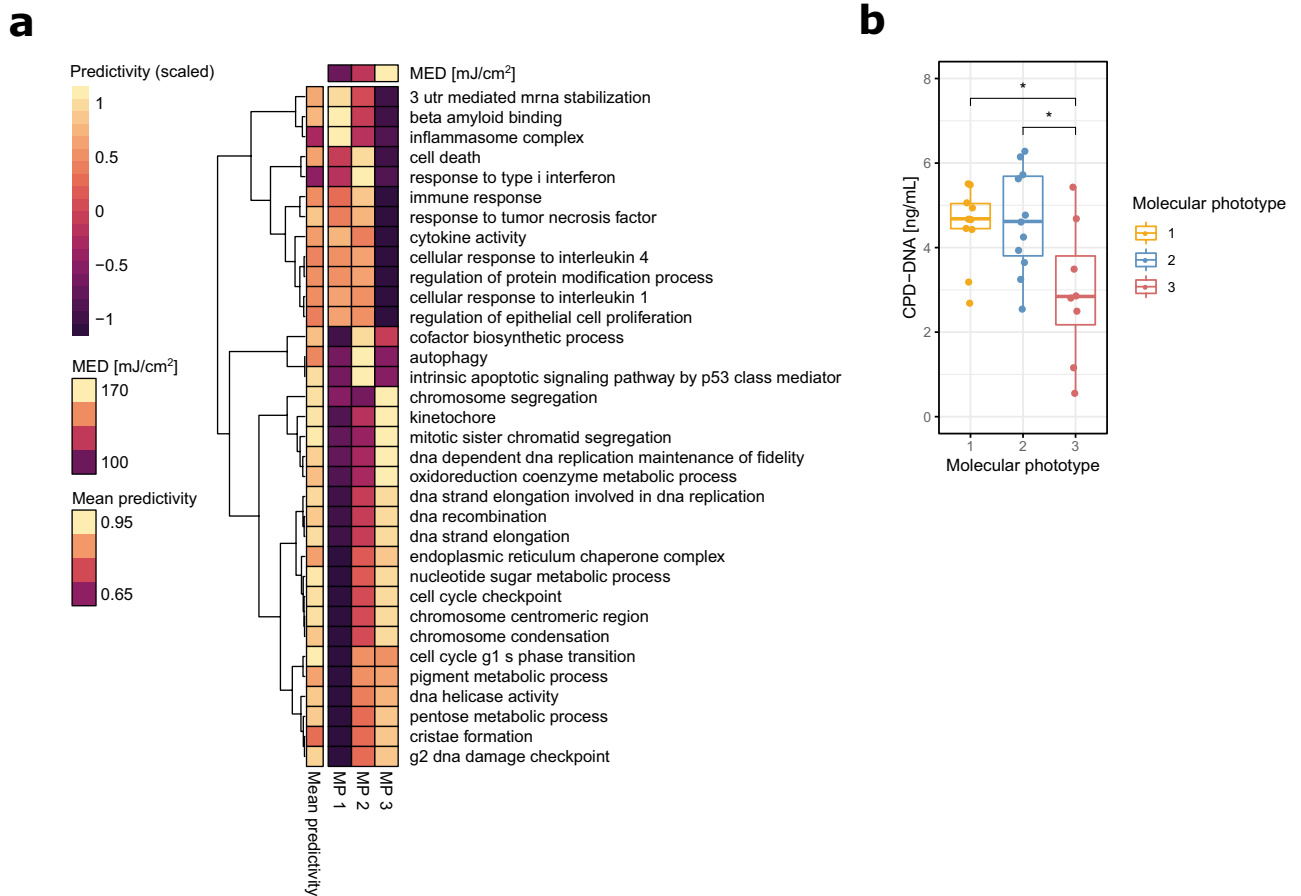


Figure 4. Molecular subtyping identifies heterogeneous biological responses to irradiation that correlate with innate UV sensitivity: **(a)** Heatmap showing the predictivity of the most defining pathways for each of the molecular phototypes to UV irradiation. The heatmap is scaled by pathway to enhance readability, average predictivity of a given pathway over all three molecular phototypes is shown to the left of the heatmap in original scale. **(b)** Extent of DNA damage in the form of cyclobutane pyrimidine dimers (CPDs) measured in the molecular phototypes 24 h after the last irradiation. Statistical comparison was performed using unpaired two-sided t-tests. Plots were generated using R v3.6.1⁷⁶ software.

mechanics and the elucidation of key players involved with this response will be important directions for future studies, as they may have potential implications for skin cancer prevention.

Discussion

Epigenetic changes are considered to play a fundamental role in establishing gene expression patterns and providing a genomic response mechanism towards extrinsic influences. However, the experimental evidence describing the extent of this response still remains somewhat limited in many biological processes. We have generated comprehensive methylation and expression profiling data to enable a more comprehensive examination of the intricacies of epidermal UV responses. Our results provide first hints towards an immediate acquisition of aging and cancer related epigenetic patterns in response to UV irradiation. In accordance with these findings, epidemiological studies have previously established a causal role for short term UV exposure (e.g. blistering sunburns) during childhood and adolescence in the late epidermal cancer pathogenesis^{52,53}. The spectrum of driver mutations related to skin cancer provides unequivocal genomic evidence for a direct mutagenic role of UV light in carcinogenesis^{54–56}. Meanwhile, genomic sites of mutation in skin cancer frequently coincide with CpG-islands^{57,58}, regions of high DNA methylation density, which has been attributed to the higher vulnerability of 5-methylcytosine bases to CPD-formation^{59,60}. Apart from potential mutagenic effects, recent publications also revealed that actinic keratosis samples already bear the classical methylation features of cutaneous squamous cell carcinomas⁶¹. These reports are consistent with the notion that epigenetic imprinting might present another common mechanism of both photoaging and carcinogenesis.

In general, substantial inter-individual variation in UV tolerability and cancer risk can be observed among Caucasian subjects. Genetic factors like polymorphisms of the melanocortin 1 receptor (MC1R) gene correlate with fairness of skin, UV sensitivity and enhanced cancer risk, however do not fully explain the diversity of UV

responses, suggesting the possibility of epigenetic involvement in UV sensitivity and pathogenesis⁴⁴. In support of this, the molecular data and methylation-based features in particular allowed the highly precise prediction of individual UV sensitivity without any prior irradiation, delivering new and strong evidence of an epigenetic component to individual UV tolerance. Further molecular evidence of the heterogeneity in response to irradiation was delivered by the molecular subtyping analysis, where clustering on the integrated multi-omics data revealed three different molecular phototypes (MPs) among irradiated samples with distinct biological signatures. The MPs differed most prominently in their association with pathways regarding cellular stress response, apoptosis/autophagy and DNA damage sensing and repair (MP 1–3, respectively). In an attempt to validate the predicted improved damage sensing and DNA repair of the most UV resilient MP 3, we analyzed the extent of cyclobutane pyrimidine dimers (CPDs) at 24 h after the last irradiation, as readout of UV-induced DNA damage and mutagenic potential. Significantly, this data indeed revealed a decreased amount of CPDs after UV irradiation in MP 3 compared to lower MPs. This is in line with some previous studies which showed lower CPD counts in irradiated samples derived from higher phototypes^{62,63}, although no molecular mechanisms could so far be elucidated. Interestingly, studies of this type often suffered from substantial biological variation within each phototype as well, once more highlighting the need for better stratification and potentially explaining why the mechanisms leading to these observations could so far not be elucidated in *ex vivo* tissue. Notably, our study setup differed slightly from most of the previous by making use of a repetitive irradiation scheme, potentially widening the window for detecting inter-individual differences in response mechanisms.

The most intensely explored UV-protection mechanism of the human skin is melanin pigmentation. Melanin serves as a physical barrier that scatters UVR and as an absorbent filter that reduces the penetration of UV through the epidermis⁶⁴. The efficacy of melanin as a sunscreen in darker skin is two- to four-fold higher compared to Caucasians⁶⁵. However individuals with highly pigmented skin have been found 16–500 times less likely to present with skin cancer compared to individuals with fair skin^{53,66–68}. The type of melanin produced also plays an important role in skin cancer risk determination. The photoprotective effects of melanin are mainly attributed to eumelanin. Pheomelanin on the other hand has only weak photoprotective properties and has even been found to contribute to carcinogenesis by a mechanism of oxidative damage⁶⁹. Still, even less deeply pigmented ethnicities such as Asians present far lower skin cancer rates compared to Caucasians⁵³, hinting towards the existence of additional cancer protective mechanisms apart from melanin. One possible explanation involves MC1R variants, which have been shown to confer an increased risk of melanoma and non-melanoma skin cancers, independently of skin pigment (including red hair phenotype)⁷⁰. The increased expression of transcripts which are associated with nucleotide metabolism and DNA repair in our dataset might present another previously uncharacterized mechanism leading to higher cancer protection afforded by skin with high UV tolerance. The detailed characterization of these biological pathways and the analysis of their clinical significance will be important aspects for future studies.

Taken together, our analyses demonstrate the benefit of using multi-omics integration for elucidating complex and diverse responses by disentangling inter-individual variation caused by insufficiently precise subject groupings, such as Fitzpatrick phototypes. The presented data illuminates the diverse and interconnected impacts of repetitive UV irradiation on both transcriptomic and epigenetic patterning in the human skin and provides new insights on protective mechanisms of subjects with high innate UV resilience, that might have further-reaching implications for UV-induced carcinogenesis.

Material and methods

Recruiting. 32 healthy female Caucasian subjects belonging to Fitzpatrick phototypes 1–4 were recruited, with twelve subjects belonging to phototype 1+2, ten to phototype 3 and ten to phototype 4. Subjects were aged between 30 and 65 years, with homogenous age distributions in each phototype group. Similar to previous studies⁷¹, exclusion criteria included tattoos or scars in the test area, pigmentation disorders, pregnancy and medication such as anti-histamines or anti-inflammatory drugs within two weeks prior to study start. A detailed listing of exclusion criteria can be found in the Supplementary information.

Minimal erythema dose determination. Minimal erythema dose (MED) estimation is a quantitative method to report the amount of UV (particularly UVB) needed to induce sunburn in the skin 24–48 h after exposure, by determining erythema (redness) and edema (swelling) as endpoints. Individual MED was determined for every subject on the first day of the study following the protocols described in DIN EN ISO 24444¹³.

Repetitive irradiation of test sites and sampling. The study sites were located in a sun-protected area on the subjects' lower backs and were randomly split into control and test areas. On the second day of the study, the first irradiation of the test sites was performed using a SOL 500 full spectrum solar simulator (Hönlle UV Technology). Intensities were chosen individually to reach 0.9 MED for all subjects or in other words 90% of the required minimal dose causing erythema in a given test subject. Irradiation to 0.9 MED was repeated in the same manner on the third day and once again on the fourth day of the study, leading to a cumulative irradiation of all test sites three times. On the fifth day of the study and 24 h after the last irradiation session of each subject, epidermal samples were taken using the suction blister method, as previously described⁷². For each subject, two suction blister roofs of 7 mm diameter were extracted from both control and test sites, one of each to be used to extract RNA for sequencing, the other to extract DNA for the DNA methylation analysis. This amounted to four suction blister roofs extracted per subject and a total of 128 samples.

Nucleic acid extraction. Nucleic acid extraction was performed as previously described⁷¹. Tissue samples were suspended in the respective lysis buffers for DNA or RNA extraction and homogenized using an MM 301 bead mill (Retsch). DNA was then extracted using the QIAamp DNA Investigator Kit (Qiagen) according to manufacturer's instructions. RNA was extracted using the RNeasy Fibrous Tissue Mini Kit (Qiagen) according to manufacturer's instructions.

Transcriptome sequencing. Transcriptome libraries were prepared using TruSeq Library Prep Kit (Illumina) and sequencing was performed at 1 × 50 bp on Illumina's HiSeq system to a final sequencing depth of approximately 100 million reads per sample. Sequencing data was processed using a pipeline including Fastqc v0.11.7⁷³ for quality control, Trimmomatic v0.36⁷⁴ for quality based read trimming and Salmon v0.8.1⁷⁵ for read mapping and quantification of transcript expression in the form of read counts and transcripts per million (TPM).

Differential expression analysis. Differential gene expression analysis was performed based on the quantified read counts in R v3.6.1⁷⁶ using DESeq2⁷⁷. Linear models were fitted using a paired design matrix to account for inter-individual variation unrelated to the irradiation treatment. Genes were considered significantly differentially regulated with FDR < 0.05 after multiple testing adjustment by the Benjamini–Hochberg procedure.

Array based methylation profiling. Methylation profiling was performed using Infinium MethylationEPIC arrays (Illumina)⁷⁶. Methylation data was processed using the minfi package⁷⁸ in R. Normalization was carried out using the functional normalization method⁷⁹, which makes use of internal control probes present on the array to infer and correct for technical variation between arrays. Subsequent analyses used M values to describe CpG methylation levels, as their approximate homoscedasticity renders them superior for statistical testing compared to Beta values⁸⁰.

Differential methylation analysis. Differential CpG methylation analysis was performed in R using limma⁸¹. Linear models were fitted using a paired design matrix to account for inter-individual variation unrelated to the irradiation treatment. CpGs were considered significantly differentially methylated with FDR < 0.05 after multiple testing adjustment by the Benjamini–Hochberg procedure. To compare DNA methylation patterns with those previously described in chronically sun-exposed skin and cancer, we used lists of the respective genomic regions and their methylation status in photoaged skin¹⁷ and different types of cancer²², which were available from the Supplementary information. The originally reported methylation changes within these regions were then compared to the average difference in methylation of all significantly differentially methylated CpGs (FDR < 0.05) annotated to the respective genomic regions in our data, allowing for a region-wise comparison of differential methylation.

Gene expression and methylation overlap and correlation analysis. For the calculation of overlap between genes and CpGs, only differentially expressed genes with absolute log₂ fold-changes above 0.5 with at least three differentially methylated CpGs with absolute log₂ fold-changes above 0.2 were considered, in order to uncover the most reliably differentially expressed and methylated genes. Pearson's correlation coefficients of gene-CpG pairs were calculated as the sum of all gene transcripts for a given gene correlated with the mean of all CpGs belonging to a functionally annotated group (i.e. all enhancer CpGs) annotated to a given gene. Annotations such as enhancer status, location in transcription start sites or within exons were extracted from the official manifest files for the Infinium MethylationEPIC array provided by Illumina via their website. Significance was assessed using linear models in R, with p-values being adjusted for multiple testing using the Benjamini–Hochberg procedure.

Protein–protein-interaction analysis. Information on protein–protein-interactions (PPI) retrieved from the STRING²⁵ database, accessed through the STRINGdb package²⁵ in R. PPI information was retrieved for all differentially expressed genes (FDR < 0.05) with absolute log₂ fold-changes above 0.7 with at least three differentially methylated CpGs (FDR < 0.05) with absolute log₂ fold-changes above 0.5. The interaction query was performed using the standard combined interaction score threshold of 400. The resulting network was refined using in R using igraph⁸² by retaining only the top 20% of the most reliable edges based on the combined interaction score, with nodes disconnected from the core network being trimmed in the process. The resulting PPI-network was then visualized using igraph⁸².

GO term enrichment analysis. Enrichment analyses were performed using the z-score method⁸³ as implemented in the GSEA R package⁸⁴, applied to the log₂ transformed TPMs. GO term gene sets⁸⁵ were downloaded originated from the Molecular Signatures Database v6.2⁸⁶ and included all three sub-ontologies: biological processes (BP), molecular functions (MF) and cellular compartments (CC).

MED regression models. Lasso regression models⁸⁷ for the prediction of MED were built in R using the implementation provided in the glmnet package⁸⁸ and interfaced using the machine learning framework mlr⁸⁹. As lasso regression models perform automatic feature weighting by regularizing the absolute magnitude of coefficients, the models were trained on the full datasets, forgoing the necessity of prior feature selection. Further-

more, data from both control and irradiated samples was included in the training process, in order to allow accurate predictions irrespective of previous UV or sun exposure of a given sample. Model predictions and accuracy scores were extracted from tenfold cross-validation to avoid overfitting and derive unbiased predictions and estimates for the quality of model fit. Metrics used for judging model performance were median absolute error (MAE), as well as the Pearson correlation coefficient (r) and the coefficient of determination (R^2).

Similarity network fusion and clustering. After a filtering step, removing features which showed little correlation to MED and reducing feature matrices to 10% of their original size, gene expression (log₂ transformed TPMs) and CpG methylation data (M values) from irradiated samples were integrated via similarity network fusion as previously described⁴⁵ using parameter settings of $k = 10$ (number of neighbors), $t = 20$ (number of iterations) and $\alpha = 0.5$ (hyperparameter). Clustering on the fused network was then performed via spectral clustering as previously described⁴⁵. Measures used for the selection of cluster numbers were the eigen-gap statistic and rotation cost as proposed in the original method description⁴⁵.

Pathway predictivity analysis. Pathway predictivity analysis was performed using GO term gene sets⁸⁵ downloaded from the Molecular Signatures Database v6.2⁸⁶. The pathway models were based on the support vector machine (SVM) implementation from the e1071 R package⁹⁰, interfaced via the mlr⁸⁹ machine learning framework. The models were trained by restricting the expression data (log₂ transformed TPMs) to that of genes annotated within a given pathway and trained to predict sample irradiation status (control or irradiated to 0.9 MED) stratified by molecular phototype. The SVMs used the radial basis function kernel with hyperparameters set to $\gamma = \frac{1}{\text{sizeofgeneset}}$ and $C = 1$. Accuracy of prediction was derived from 5 × 5-fold repeated cross-validation for each pathway model, giving insight on how well genes within the gene set allow a discrimination between UV irradiated and control samples while controlling for overfitting, and used as a measure of predictivity of the respective pathway to irradiation status.

Profiling of cyclobutane pyrimidine dimers (CPDs). CPD concentrations were determined using the OxiSelect UV-Induced DNA Damage ELISA Kit (Cell Biolabs) according to the manufacturer's instructions.

General data analysis and visualization. Data analysis in R further included the usage of the package data.table⁹¹ and dplyr⁹² for diverse data handling tasks, as well as the packages ggplot2⁹³, ggpubr⁹⁴, circlize⁹⁵, and pheatmap⁹⁶ for visualization. Mapping and annotation of gene identifiers was performed using the biomaRt⁹⁷ and org.Hs.eg.db⁹⁸ packages, utilizing the GRCh37 (hg19) human genome build.

Ethics. The study was performed in agreement with the recommendations of the Declaration of Helsinki and all test subjects provided written, informed consent. Approval of the study protocol was granted by the Ethics Committee of the University of Freiburg (study code 016/1672).

Data availability

Data generated within this study has been deposited online at ArrayExpress, under the accessions E-MTAB-9251 and E-MTAB-9249.

Received: 8 April 2020; Accepted: 16 July 2020

Published online: 31 July 2020

References

- Mostafa, W. Z. & Hegazy, R. A. Vitamin D and the skin: Focus on a complex relationship: A review. *J. Adv. Res.* **6**, 793–804 (2015).
- Leccia, M.-T., Lebbe, C., Claudel, J.-P., Narda, M. & Basset-Seguin, N. New vision in photoprotection and photorepair. *Dermatol. Ther.* **9**, 103–115 (2019).
- Guerra, K. C. & Crane, J. S. Sunburn. (2019).
- Ikehata, H. & Ono, T. The mechanisms of UV mutagenesis. *J. Radiat. Res. (Tokyo)* **52**, 115–125 (2011).
- Mohania, D. *et al.* Ultraviolet radiations: Skin defense-damage mechanism. *Adv. Exp. Med. Biol.* **996**, 71–87 (2017).
- Costin, G.-E. & Hearing, V. J. Human skin pigmentation: Melanocytes modulate skin color in response to stress. *FASEB J.* **21**, 976–994 (2007).
- Fitzpatrick, T. B. Soleil et peau. *J. Méd. Esthét.* **2**, 33–34 (1975).
- Rampen, F. H. J., Fleuren, B. A. M., de Boo, T. M. & Lemmens, W. A. J. G. Unreliability of self-reported burning tendency and tanning ability. *Arch. Dermatol.* **124**, 885 (1988).
- Ravnbak, M. H. Objective determination of Fitzpatrick skin type. *Dan. Med. Bull.* **57**, B4153 (2010).
- Hemminki, K. & Snellman, E. How fast are UV-dimers repaired in human skin DNA in situ? *J. Investig. Dermatol.* **119**, 699 (2002).
- Wulf, H. C., Philipsen, P. A. & Ravnbak, M. H. Minimal erythema dose and minimal melanogenesis dose relate better to objectively measured skin type than to Fitzpatrick's skin type. *Photodermatol. Photoimmunol. Photomed.* **26**, 280–284 (2010).
- Harrison, G. I., Young, A. R. & McMahon, S. B. Ultraviolet radiation-induced inflammation as a model for cutaneous hyperalgesia. *J. Investig. Dermatol.* **122**, 183–189 (2004).
- International Organization for Standardization. *DIN EN ISO 24444—In vivo determination of the sun protection factor (SPF)*. (2010).
- Smith, Z. D. & Meissner, A. DNA methylation: Roles in mammalian development. *Nat. Rev. Genet.* **14**, 204–220 (2013).
- Baubec, T. & Schübeler, D. Genomic patterns and context specific interpretation of DNA methylation. *Curr. Opin. Genet. Dev.* **25**, 85–92 (2014).
- Jaenisch, R. & Bird, A. Epigenetic regulation of gene expression: How the genome integrates intrinsic and environmental signals. *Nat. Genet.* **33**(Suppl), 245–254 (2003).
- Vandiver, A. R. *et al.* Age and sun exposure-related widespread genomic blocks of hypomethylation in nonmalignant skin. *Genome Biol.* **16**, 80 (2015).

18. Shen, Y., Stanislauskas, M., Li, G., Zheng, D. & Liu, L. Epigenetic and genetic dissections of UV-induced global gene dysregulation in skin cells through multi-omics analyses. *Sci. Rep.* **7**, 42646 (2017).
19. Storzaker, C., Taberlay, P. C., Statham, A. L. & Clark, S. J. Mining cancer methylomes: Prospects and challenges. *Trends Genet.* **30**, 75–84 (2014).
20. Costa-Pinheiro, P., Montezuma, D., Henrique, R. & Jerónimo, C. Diagnostic and prognostic epigenetic biomarkers in cancer. *Epigenomics* **7**, 1003–1015 (2015).
21. Gröniger, E. *et al.* Aging and chronic sun exposure cause distinct epigenetic changes in human skin. *PLoS Genet.* **6**, e1000971 (2010).
22. Hansen, K. D. *et al.* Increased methylation variation in epigenetic domains across cancer types. *Nat. Genet.* **43**, 768–775 (2011).
23. Liu, D. *et al.* Discovery and validation of methylated-differentially expressed genes in *Helicobacter pylori*-induced gastric cancer. *Cancer Gene Ther.* <https://doi.org/10.1038/s41417-019-0125-7> (2019).
24. Tong, Y., Song, Y. & Deng, S. Combined analysis and validation for DNA methylation and gene expression profiles associated with prostate cancer. *Cancer Cell Int.* **19**, 50 (2019).
25. Szklarczyk, D. *et al.* STRING v10: Protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **43**, D447–D452 (2015).
26. Toillon, R. A. *et al.* Interaction between estrogen receptor alpha, ionizing radiation and (anti-) estrogens in breast cancer cells. *Breast Cancer Res. Treat.* **93**, 207–215 (2005).
27. Widyarini, S., Domanski, D., Painter, N. & Reeve, V. E. Estrogen receptor signaling protects against immune suppression by UV radiation exposure. *Proc. Natl. Acad. Sci.* **103**, 12837–12842 (2006).
28. Widyarini, S., Husband, A. J. & Reeve, V. E. Protective effect of the isoflavonoid equol against hairless mouse skin carcinogenesis induced by UV radiation alone or with a chemical cocarcinogen. *Photochem. Photobiol.* **81**, 32–37 (2005).
29. Wang, Z., Boudjelal, M., Kang, S., Voorhees, J. J. & Fisher, G. J. Ultraviolet irradiation of human skin causes functional vitamin A deficiency, preventable by all-trans retinoic acid pre-treatment. *Nat. Med.* **5**, 418–422 (1999).
30. Mellett, M. *et al.* CARD14 gain-of-function mutation alone is sufficient to drive IL-23/IL-17-mediated psoriasiform skin inflammation in vivo. *J. Investig. Dermatol.* **138**, 2010–2023 (2018).
31. Wang, M. *et al.* Gain-of-function mutation of Card14 leads to spontaneous psoriasis-like skin inflammation through enhanced keratinocyte response to IL-17A. *Immunity* **49**, 66–79.e5 (2018).
32. Mattei, F. *et al.* IRF-8 controls melanoma progression by regulating the cross talk between cancer and immune cells within the tumor microenvironment. *Neoplasia N. Y. N* **14**, 1223–1235 (2012).
33. Tshuikina, M., Jernberg-Wiklund, H., Nilsson, K. & Oberg, F. Epigenetic silencing of the interferon regulatory factor ICSBP/IRF8 in human multiple myeloma. *Exp. Hematol.* **36**, 1673–1681 (2008).
34. Luo, X. *et al.* The tumor suppressor interferon regulatory factor 8 inhibits β -catenin signaling in breast cancers, but is frequently silenced by promoter methylation. *Oncotarget* **8**, 48875–48888 (2017).
35. Guo, Y. *et al.* Inhibition of IRF8 Negatively regulates macrophage function and impairs cutaneous wound healing. *Inflammation* **40**, 68–78 (2017).
36. van Smeden, J. & Bouwstra, J. A. Stratum corneum lipids: Their role for the skin barrier function in healthy subjects and atopic dermatitis patients. *Curr. Probl. Dermatol.* **49**, 8–26 (2016).
37. Feingold, K. R. Thematic review series: Skin lipids. The role of epidermal lipids in cutaneous permeability barrier homeostasis: Fig. 1. *J. Lipid Res.* **48**, 2531–2546 (2007).
38. Feingold, K. R. The outer frontier: The importance of lipid metabolism in the skin. *J. Lipid Res.* **50**(Suppl), S417–S422 (2009).
39. Lehmann, P., Hölzle, E., Melnik, B. & Plewig, G. Effects of ultraviolet A and B on the skin barrier: A functional, electron microscopic and lipid biochemical study. *Photodermatol. Photoimmunol. Photomed.* **8**, 129–134 (1991).
40. Biniek, K., Levi, K. & Dauskardt, R. H. Solar UV radiation reduces the barrier function of human skin. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 17111–17116 (2012).
41. Wefers, H. *et al.* Influence of UV irradiation on the composition of human stratum corneum lipids. *J. Investig. Dermatol.* **96**, 959–962 (1990).
42. Jungersted, J. M., Høgh, J. K., Hellgren, L. I., Jemec, G. B. E. & Agner, T. The impact of ultraviolet therapy on stratum corneum ceramides and barrier function. *Photodermatol. Photoimmunol. Photomed.* **27**, 331–333 (2011).
43. Scherer, D. & Kumar, R. Genetics of pigmentation in skin cancer: A review. *Mutat. Res.* **705**, 141–153 (2010).
44. D’Orazio, J., Jarrett, S., Amaro-Ortiz, A. & Scott, T. UV radiation and the skin. *Int. J. Mol. Sci.* **14**, 12222–12248 (2013).
45. Wang, B. *et al.* Similarity network fusion for aggregating data types on a genomic scale. *Nat. Methods* **11**, 333–337 (2014).
46. Sand, J. *et al.* Expression of inflammasome proteins and inflammasome activation occurs in human, but not in murine keratinocytes. *Cell Death Dis.* **9**, 24 (2018).
47. Hasegawa, T., Nakashima, M. & Suzuki, Y. Nuclear DNA damage-triggered NLRP3 inflammasome activation promotes UVB-induced inflammatory responses in human keratinocytes. *Biochem. Biophys. Res. Commun.* **477**, 329–335 (2016).
48. Faustin, B. & Reed, J. C. Sunburned skin activates inflammasomes. *Trends Cell Biol.* **18**, 4–8 (2008).
49. Sontheimer, C., Liggitt, D. & Elkon, K. B. Ultraviolet B irradiation causes stimulator of interferon genes-dependent production of protective type I interferon in mouse skin by recruited inflammatory monocytes. *Arthritis Rheumatol. Hoboken NJ* **69**, 826–836 (2017).
50. Miller, S. A. *et al.* Evidence for a new paradigm for ultraviolet exposure: A universal schedule that is skin phototype independent. *Photodermatol. Photoimmunol. Photomed.* **28**, 187–195 (2012).
51. Pfeifer, G. P. & Besaratinia, A. UV wavelength-dependent DNA damage and human non-melanoma and melanoma skin cancer. *Photochem. Photobiol. Sci. Off. J. Eur. Photochem. Assoc. Eur. Soc. Photobiol.* **11**, 90–97 (2012).
52. Henrikson, N. B. *et al.* Behavioral counseling for skin cancer prevention. *JAMA* **319**, 1143 (2018).
53. Armstrong, B. K. & Krickler, A. The epidemiology of UV induced skin cancer. *J. Photochem. Photobiol. B* **63**, 8–18 (2001).
54. Garibyan, L. & Fisher, D. E. How sunlight causes melanoma. *Curr. Oncol. Rep.* **12**, 319–326 (2010).
55. Wu, S., Han, J., Laden, F. & Qureshi, A. A. Long-term ultraviolet flux, other potential risk factors, and skin cancer risk: A cohort study. *Cancer Epidemiol. Biomark. Prev.* **23**, 1080–1089 (2014).
56. Hodis, E. *et al.* A landscape of driver mutations in melanoma. *Cell* **150**, 251–263 (2012).
57. Drouin, R. & Therrien, J.-P. UVB-induced cyclobutane pyrimidine dimer frequency correlates with skin cancer mutational hotspots in p53. *Photochem. Photobiol.* **66**, 719–726 (1997).
58. You, Y.-H. & Pfeifer, G. P. Similarities in sunlight-induced mutational spectra of CpG-methylated transgenes and the p53 gene in skin cancer point to an important role of 5-methylcytosine residues in solar UV mutagenesis 11 Edited by J. Miller. *J. Mol. Biol.* **305**, 389–399 (2001).
59. Tommasi, S., Denissenko, M. F. & Pfeifer, G. P. Sunlight induces pyrimidine dimers preferentially at 5-methylcytosine bases. *Cancer Res.* **57**, 4727–4730 (1997).
60. Martinez-Fernandez, L., Banyasz, A., Esposito, L., Markovits, D. & Improta, R. UV-induced damage to DNA: Effect of cytosine methylation on pyrimidine dimerization. *Signal Transduct. Target. Ther.* **2**, 17021 (2017).
61. Rodríguez-Paredes, M. *et al.* Methylation profiling identifies two subclasses of squamous cell carcinoma related to distinct cells of origin. *Nat. Commun.* **9**, 577 (2018).

62. Cragg, N., Chadwick, C. A., Potten, C. S., Sheehan, J. M. & Young, A. R. Repeated ultraviolet exposure affords the same protection against DNA photodamage and erythema in human skin types II and IV but is associated with faster DNA repair in skin type IV. *J. Investig. Dermatol.* **118**, 825–829 (2002).
63. Tadokoro, T. *et al.* UV-induced DNA damage and melanin content in human skin differing in racial/ethnic origin. *FASEB J.* **17**, 1177–1179 (2003).
64. Kollias, N., Sayre, R. M., Zeise, L. & Chedekel, M. R. Photoprotection by melanin. *J. Photochem. Photobiol. B* **9**, 135–160 (1991).
65. Kaidbey, K. H., Agin, P. P., Sayre, R. M. & Kligman, A. M. Photoprotection by melanin: A comparison of black and Caucasian skin. *J. Am. Acad. Dermatol.* **1**, 249–260 (1979).
66. Halder, R. M. & Bang, K. M. Skin cancer in blacks in the United States. *Dermatol. Clin.* **6**, 397–405 (1988).
67. Cress, R. D. & Holly, E. A. Incidence of cutaneous melanoma among non-Hispanic whites, Hispanics, Asians, and blacks: An analysis of California cancer registry data, 1988–93. *Cancer Causes Control CCC* **8**, 246–252 (1997).
68. National Cancer Institute (NCI). *SEER Cancer Statistics Review (CSR) 1975–2014*. (2018).
69. Mitra, D. *et al.* An ultraviolet-radiation-independent pathway to melanoma carcinogenesis in the red hair/fair skin background. *Nature* **491**, 449–453 (2012).
70. Demenais, F. *et al.* Association of MC1R variants and host phenotypes with melanoma risk in CDKN2A mutation carriers: A GenoMEL study. *JNCI J. Natl. Cancer Inst.* **102**, 1568–1583 (2010).
71. Holzschek, N. *et al.* Multi-omics network analysis reveals distinct stages in the human aging progression in epidermal tissue. *Aging* **12**, 12393–12409 (2020).
72. Südel, K. M. *et al.* Tight control of matrix metalloproteinase-1 activity in human skin. *Photochem. Photobiol.* **78**, 355–360 (2003).
73. Andrews, S. FastQC: A quality control tool for high throughput sequence data. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (2010).
74. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
75. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).
76. R Development Core Team. R: The R Project for Statistical Computing. <https://www.r-project.org/> (2008).
77. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
78. Aryee, M. J. *et al.* Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* **30**, 1363–1369 (2014).
79. Fortin, J.-P. *et al.* Functional normalization of 450k methylation array data improves replication in large cancer studies. *Genome Biol.* **15**, 503 (2014).
80. Du, P. *et al.* Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinform.* **11**, 587 (2010).
81. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47–e47 (2015).
82. Csardi, G. & Nepusz, T. The igraph software package for complex network research. *Interj. Complex Syst.* **1695**, 1–9 (2006).
83. Lee, E., Chuang, H.-Y., Kim, J.-W., Ideker, T. & Lee, D. Inferring pathway activity toward precise disease classification. *PLoS Comput. Biol.* **4**, e1000217 (2008).
84. Hänzelmann, S., Castelo, R., Guinney, J. & Castelo, R. GSEA: Gene set variation analysis for microarray and RNA-seq data. *BMC Bioinform.* **14**, 7 (2013).
85. Liberzon, A. *et al.* The molecular signatures database hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).
86. Liberzon, A. *et al.* Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011).
87. Tibshirani, R. & Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B* **58**, 267–288 (1994).
88. Simon, N., Friedman, J., Hastie, T. & Tibshirani, R. Regularization paths for Cox’s proportional hazards model via coordinate descent. *J. Stat. Softw.* **39**, 1–13 (2011).
89. Bischl, B. *et al.* mlr: Machine learning in R. *J. Mach. Learn. Res.* **17**, 1–5 (2016).
90. Dimitriadou, E., Hornik, K., Leisch, F., Meyer, D. & Weingessel, A. *e1071: Misc Functions of the Department of Statistics (e1071), TU Wien. R package version vol. 1* (2011).
91. Dowle, M. & Srinivasan, A. data.table: Extension of ‘data.frame’. (2018).
92. Wickham, H., François, R., Henry, L. & Müller, K. dplyr: A grammar of data manipulation. (2019).
93. Wickham, H. *ggplot2: elegant graphics for data analysis* (Springer-Verlag, New York, 2016).
94. Kassambara, A. ggpvr: ‘ggplot2’ Based Publication Ready Plots. (2018).
95. Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. Circlize implements and enhances circular visualization in R. *Bioinformatics* **30**, 2811–2812 (2014).
96. Kolde, R. Package ‘pheatmap’. *Bioconductor* 1–6 (2012).
97. Smedley, D. *et al.* BioMart—Biological queries made easy. *BMC Genom.* **10**, 22 (2009).
98. Carlson, M. org.Hs.eg.db: Genome wide annotation for Human. (2018).

Author contributions

T.S., L.T. and M.W. conceived the original idea for the study and planned the experiments. J.S. carried out the experiments. N.H. and K.G. conceived the analysis and performed the computations. N.H., K.G., L.K., M.W., T.S., C.F., E.G., L.K. and H.W. contributed to the interpretation of the results. N.H. and K.G. wrote the manuscript with contributions from all other authors. All authors discussed the results and commented on the manuscript.

Competing interests

All authors apart from LK are employees of Beiersdorf AG. LK received consultation fees from Beiersdorf AG.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-69683-8>.

Correspondence and requests for materials should be addressed to N.H. or K.G.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020