Inter-layer and inter-subject variability of circadian gene expression in human skin

Marta del Olmo[1], Florian Spörl[2], Sandra Korge[2,3], Karsten Jürchott[3,4], Matthias Felten[5,6], Astrid Grudziecki[3], Jan de Zeeuw[7], Claudia Nowozin[7], Hendrik Reuter[2], Thomas Blatt[2], Hanspeter Herzel[1], Dieter Kunz[7], Achim Kramer[3], Bharath Ananthasubramaniam[1,8]

bharath.ananthasubramaniam@hu-berlin.de

Abstract

The skin is the largest human organ with a circadian clock that regulates its function. Although circadian rhythms in specific functions are known, rhythms in the proximal clock output, gene expression, in human skin have not been thoroughly explored. This work reports circadian gene expression in two skin layers, epidermis and dermis, in a cohort of young, healthy adults, who maintained natural, regular sleep schedules. 10% of the expressed genes showed rhythms at the population level, of which only a third differed between the two layers. Broadly, expression magnitudes of circadian genes were consistent across subjects in each layer. Amplitude and phases of circadian gene expression, however, varied more across subjects than layers, with amplitude being more variable than phases. Expression amplitudes in the epidermis were larger and more subject-variable, while they were smaller and more consistent in the dermis. Core clock gene expression was similar across layers at the population-level, but were heterogeneous in the their variability across subjects. We used this data to identify small sets of biomarkers for internal clock phase in each layer, which consisted of layer-specific non-core clock genes. This work provides a valuable resource to advance our understanding of human skin to realize the potential of circadian medicine as well as a novel methodology to quantify sources of variability in human circadian rhythms.

I. Introduction

The skin is the largest organ of the body and its main functions are protection against bacteria, radiation or temperature from the exterior, as well as against water loss from the interior [1]. It is morphologically complex and consists of many cell types that are organized into three main layers: epidermis, dermis and hypodermis [2]. The skin evolved a circadian clock [3] in response to the direct exposure to the rhythmic external environment to anticipate changes and to adjust its physiology accordingly.

The mammalian circadian clock is a hierarchical network with the central clock in the suprachiasmatic nucleus (SCN) and peripheral clocks in many tissues including the skin. The cell-autonomous molecular "core" clock [4, 5] consists of a number of interlocked transcriptional-translational negative feedback loops. Core clock genes *CLOCK* and *BMAL1* induce the expression of their own inhibitors, *PER* and *CRY* genes. Once translated, PER and CRY proteins form large complexes that travel back to the nucleus to repress CLOCK and BMAL1, thus repressing their own transcription and thereby creating self-sustained 24 h rhythms in gene and protein expression. In mammals, core clock components that are also transcription factors act at *cis*-regulatory sequences to drive rhythmic expression of a large number of output genes (about 10% of all genes) in a cell-autonomous and tissue-specific manner [6, 7].

The circadian gene expression profile of the skin remains nevertheless incompletely characterized. The presence of a skin circadian clock in humans was first inferred from circadian rhythms in biophysical skin parameters, such as sebum secretion [8], water loss [9] or response to allergens [10]. At the turn of the century, rhythmic expression of selected core clock genes in the skin was described in humans [11] and mice [12]. Over time, circadian expression of core clock genes was recorded in several skin cell types, including epidermal and hair follicle keratinocytes, dermal fibroblasts and melanocytes [13–17]. Spörl et al. [18] performed the first high-throughput analysis of circadian gene expression in the skin. That study identified $\sim$300 circadian genes from measurements at three time points in one layer (epidermis). More recently, Wu et al. [19, 20] identified $\sim$100-150 circadian genes each in the epidermis and dermis from samples collected every 6 h over one day. However, both these microarray studies provide only limited insight into circadian gene expression in human skin, since they lacked sufficient number of samples over one circadian cycle. To more thoroughly describe the impact of the human skin clock, we identified circadian genes in the two prominent skin layers, epidermis and dermis. To assess whether the complex and heterogeneous skin also results in a cell type-/layer-specific clock, we compared the circadian gene expression across layers.

As one of the few accessible tissue clocks, skin samples could be used for circadian phenotyping of humans. Evidence is gradually accumulating that therapeutic efficacy and the degree of side effects are dependent on the time of administration [21–23]. Such observations are likely to grow, since 50% of all drugs target the product of a circadian gene [6, 7]. One key challenge to implementing time-of-day-aware 'circadian medicine' strategies is the fact that internal clocks of humans are heterogeneous. Since rhythms in human physiology are determined by *internal* clock time and not on time according to the external environment, circadian studies in humans ought to record and present results relative to the internal phase of entrainment (termed chronotype) of subjects. This internal clock time in turn depends on genetic factors [16, 24], age [25], sex [25], level of light exposure [26, 27], the season [26, 28] and on the local time-zone [25]. Thus, circadian treatments need to be personalized to an subjects's clock. To evaluate the utility of skin samples for circadian phenotyping, we used our comprehensive gene expression profiles to identify biomarkers for circadian phase in each layer.

We measured gene expression using microarrays in a small cohort of young, healthy adults of both sexes, who maintained their natural yet regular sleep schedules. We first quantified the circadian expression expected in the general population in the epidermis and dermis. In contrast to previous studies on skin, we present our results with respect to internal time of subjects; our study design included chronotyping of subjects. We then analyzed the layer-specificity of *population* circadian rhythms. The population circadian rhythms are the most representative rhythms of an individual in the population. However, the inherent heterogeneity will result in individual rhythms diverging from the population rhythm. Therefore, we next quantified how much individual circadian rhythms deviate from the population rhythms in both layers. Finally, we identified a small set of biomarkers with circadian gene expression to predict internal clock phase in each layer.

II. RESULTS

Population circadian gene expression is layer-specific in healthy human skin

To explore molecular circadian rhythms in human dermis and epidermis, 11 healthy subjects (male and female) were biopsied in the upper back every 4 h across a 24 h duration (Figure 1A, Materials and Methods). In the two weeks leading up to the biopsies, subjects maintained their desired natural sleep-wake schedule. Samples were separated into dermis and epidermis and subsequently quantified using whole-genome microarrays. We adjusted sample collection times using their chronotypes to indicate *internal* time of subjects (Figure 1B). Chronotypes were estimated from sleep schedules (available in Supplementary Table 1) as the mid-sleep time on free days after correcting for sleep debt (MSF_{sc}) [29].

A large majority of circadian genes have similar rhythms of gene expression in both layers. We identified and compared genes with circadian population rhythms in both human skin layers using differential rhythmicity analysis [30]. *Population* rhythms are circadian patterns of gene expression averaged across the entire cohort. We identified 1053 circadian genes in dermis and 1352 in epidermis (FDR < 0.05 and amplitude > 1.5 fold peak-to-trough in at least one layer, see Materials and Methods for details). 966 of these circadian genes were common to both skin layers (Figure 1C, inset, Supplementary Figure 1A). The number of circadian genes remained stable across a range of choices of FDR cutoff (Figure 1C). 386 genes were rhythmic only in the epidermis and 87 only in the dermis, as well as a further 49 genes that were rhythmic in both but with significantly different amplitude and/or phase. Thus, the expression of 917 circadian genes was indistinguishable between the two layers.

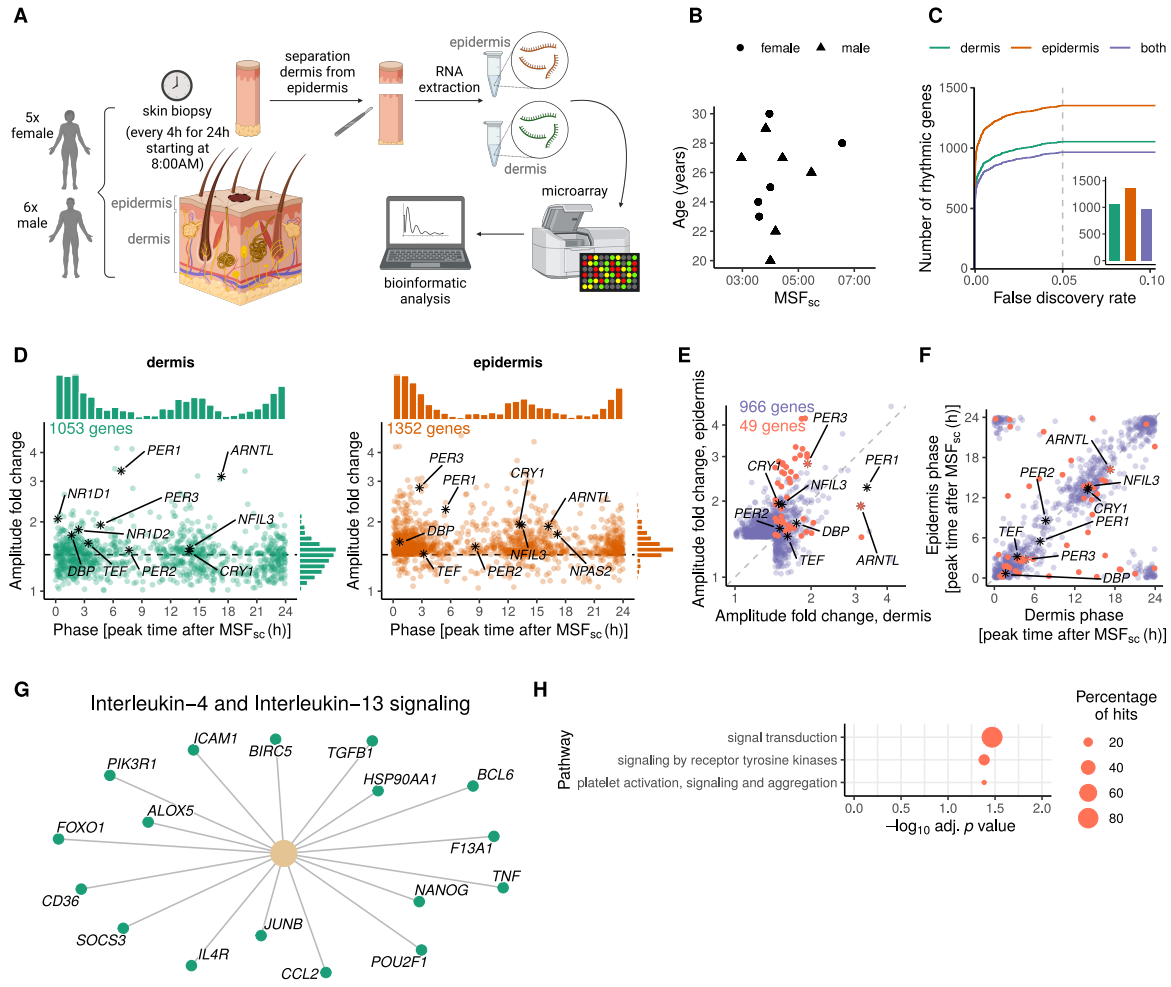

Figure 1: Functional and layer-specific clocks in human dermis and epidermis. A. Experimental setup: 11 healthy subjects were biopsied in the back every 4 h for 24 h starting at 8 AM. Dermis and epidermis were separated and gene expression was analyzed using whole-genome microarrays. **B.** Composition of the study cohort by sex, age and mid-sleep time on free days after correcting for sleep debt (MSF$_{sc}$). **C.** Number of circadian genes with respect to *internal* time in dermis, epidermis or in both layers were determined using differential rhythmicity analysis (minimum requirement of peak-to-trough fold change amplitude > 1.5 in at least one layer, see Materials and Methods for details). The number of genes for FDR $= 0.05$ is shown in the inset. **D.** Phase (as peak time after MSF$_{sc}$) and amplitude distributions of the circadian genes in human dermis (in green, left panel) and epidermis (orange, right panel) (FDR < 0.05, peak-to-trough fold change amplitude >1.5). Each gene is represented by a dot; clock genes are highlighted in black. **E.** Amplitude correlation of the 966 circadian genes in both layers, from which 49 show significant different rhythms (highlighted in coral). **F.** Phase correlation of circadian genes in dermis *and* epidermis, with differentially rhythmic genes highlighted in coral. Clock genes are shown in black asterisks (or coral if differentially rhythmic). **G.** Pathway enrichment analysis of circadian genes in dermis was performed with REACTOME pathway database. Rhythmic genes belonging to the pathway are shown. **H.** Reactome pathway enrichment analysis of differentially-rhythmic genes in dermis and epidermis tested against the background of all 1439 rhythmic genes in dermis or epidermis. Only pathways with a p value < 0.05 are shown.

Circadian gene amplitudes are larger in the epidermis, but core clock genes are remarkably similar in the two layers. We observed a bimodal distribution of phases of all circadian dermal and epidermal genes, with peaks clustering at 1 h and 13 h after mid-sleep time on free days (MSF_{sc}) (Figure 1D and Supplementary Figure 1A). Despite the similarity of the distributions, amplitudes of individual genes varied between layers. Epidermal circadian genes oscillated with a higher amplitude than those in the dermis among the differentially-rhythmic genes and with a trend in this direction among the circadian genes with similar rhythms (Figure 1E). There was no systematic trend in the phase difference between circadian genes common to the two layers (Figure 1F). The core clock genes were remarkably consistent (statistically indistinguishable) in amplitude and phase between the two layers (Supplementary Figure 1B) with the exception of *ARNTL* and *PER3*, which had higher amplitudes in the dermis and epidermis, respectively. Note, *NR1D1* and *NR1D2* were rhythmic in both layers but with amplitudes just outside the amplitude cutoff in the epidermis.

The similar and dissimilar rhythms in the two layers are primarily involved in cellular signaling. Circadian genes in the dermis were enriched for immune response-related pathways; in particular, multiple analyses (Reactome, KEGG, MSigDB) revealed enrichment of IL-4 and IL-13 signaling pathways (Figure 1G). However, circadian genes in the epidermis did not show any statistically significant enrichment. The differentially-rhythmic genes (circadian genes with dissimilar rhythms in the two layers) were primarily involved in the three pathways: platelet activation & signaling, signal transduction and signaling by receptor tyrosine kinases (Figure 1H).

Population circadian rhythms are not affected by choice of time reference in this dataset. We repeated the above analysis without correcting the sample collection times for subject chronotypes using MSF_{sc}. Interestingly, the circadian genes determined with respect to external time (Supplementary Figure 2) did not deviate appreciably in number, amplitude or phase from the circadian genes (compare Figure 1C,D,E,F and Supplementary Figure 2) with respect to internal time, i.e, controlling for MSF_{sc} did not affect rhythmicity analysis significantly in our dataset.

In summary, we observed significant similarity in the circadian gene expression across these two adjacent skin layers complemented by some layer-specificity of circadian rhythms.

Amplitudes and phases of circadian genes are subject-specific, while magnitudes are layer-specific

The population (mean) circadian gene expression describes rhythmic gene expression at the level of the cohort. We quantified next how much circadian parameters of the rhythmic genes varied among individuals in the cohort. We present this variability within the cohort in relation to the variability across layers. We fit linear-mixed-effect models [31] followed by error propagation to obtain both the average circadian gene expression (fixed effects) and the variation across subjects and layer (random effects). This analysis was performed only on the 1439 genes that showed population rhythms in at least one layer. We computed the variability in circadian parameters (magnitude (MESOR), amplitude and phase) of individual rhythmic genes across both layers and subjects (available as an additional file in Supplementary Table 2).

Magnitudes vary more across layers, and amplitudes and phases more across subjects. The error propagation analysis produced estimates of the absolute variability (as standard deviations) of the circadian parameters across subjects and layers. When the variability is viewed in absolute terms (Figure 2A), magnitude of circadian genes varied more across layers than across subjects, while amplitudes and phases were more variable across subjects than across layers. In order to better quantify the relative contribution of layer and subject to the circadian rhythm variability, we defined the fraction of variance explained by subject as $\sigma^2_{\text{subj}}/(\sigma^2_{\text{subj}} + \sigma^2_{\text{layer}})$. More genes were variable in amplitude and phase across subjects (have fraction of variance explained closer to one in Figure 2B) consistent with Figure 2A. However, specific genes were more or less subject- (layer-) variable in all three circadian parameters. *PER1,2* and *NR1D2* showed highly variable magnitudes and amplitudes across subjects, while positive regulators *NPAS2*, *ARNTL* were very consistent

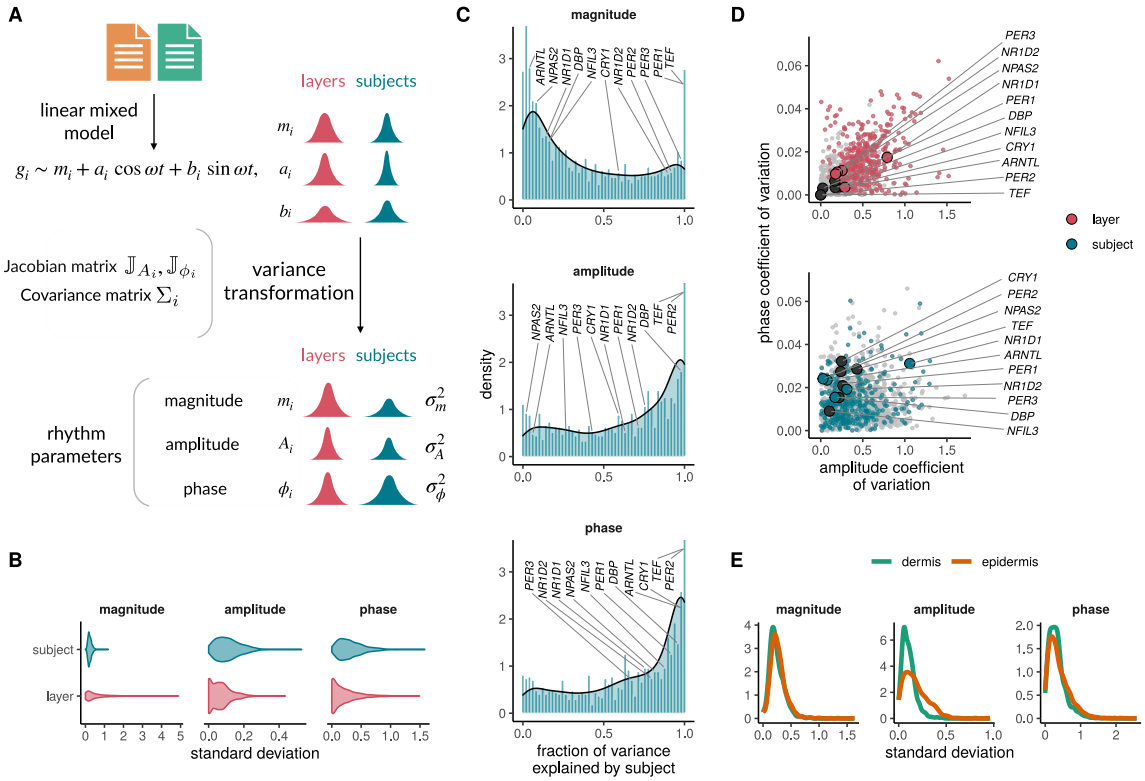

Figure 2: Amplitudes and phases of circadian genes differ across subjects, while magnitudes differ across skin layers. A. Methodology to compute variability in circadian parameters across layers and subjects. **B.** Quantification of the variability in magnitude (MESOR), amplitude and phase across subjects (blue) and layers (red) in genes (1439) rhythmic in at least one layer. **C.** Relative contribution of layer and subject to variation in magnitudes, amplitudes and phases. The fraction of variance explained by subject was defined as $\sigma^2_{\text{sub}}/(\sigma^2_{\text{sub}} + \sigma^2_{\text{lay}})$. Circadian genes falling in bins closer to 0 represent cycling genes with low fraction of variance attributed to subject differences (and high fraction of variance explained by layer differences). **D.** Correlations of amplitude and phase relative variability (using the coefficient of variation) across layers (left panel) or subjects (right panel). Clock genes are shown with larger dots. Differentially-rhythmic genes (from the previous analysis) across layers are shown in color (red, blue), while genes with non-significant rhythm differences in dermis and epidermis are shown in grey. **E.** Quantification of the variability in circadian parameters across subjects in each skin layer separately. Linear mixed models were fit in all 1439 circadian genes in at least one layer to plot panels B-D; in all 1053 rhythmic genes in dermis and 1352 in epidermis to plot panel E.

across subjects. The phases of core clock genes were also much more variable across subjects than across layers.

Amplitudes of circadian genes are more variable than phases. The relative variation of amplitudes exceeded relative variation of phases both across layers and across subjects (Figure 2C). Genes with different population rhythms between layers (Figure 1) showed high amplitude and phase variability across layers as expected. Core clock genes had remarkably low variability in amplitude and phases across layers consistent with indistinguishable population rhythms we observed across layers. *NR1D1* was the only exception that showed larger variation in amplitude across layers; it was rhythmic in both layers but with an amplitude just below our cutoff in epidermis. However, the core clock genes had higher variability in phase across subjects, but with comparable amplitude variability across subjects and layers.

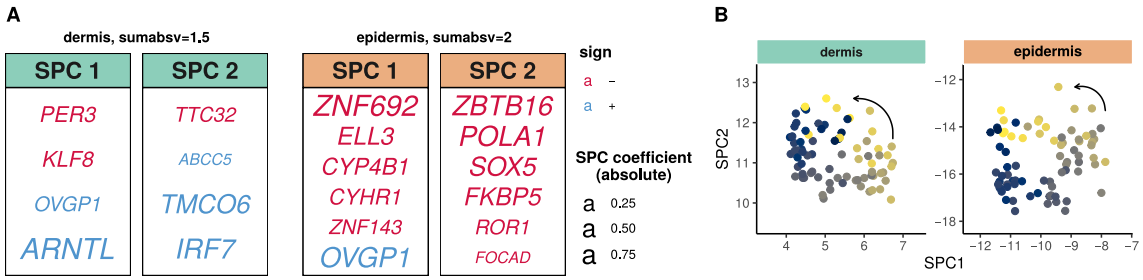

Figure 3: Identification of internal time-telling genes in human dermis and epidermis. A. Predictive biomarkers from human dermis (left panels) and epidermis (right panels) for optimal parameter choice (Supplementary Figure 3A). Genes assigned to SPC1 or SPC2 as well as their coefficients are shown. **B.** Expression profiles of the predictor genes in our cohort in dermis (left) and epidermis (bottom) represented in SPC space. Colors indicate the internal time of the subjects. `ZeitZeiger` was used to identify biomarkers of internal phase and was run with all $\sim$ 11000 expressed genes, separately for dermis and epidermis.

Amplitudes of circadian genes in the epidermis differ more than in dermis. We observed previously that epidermal population rhythms had greater amplitudes than dermal rhythms. In addition, when the variability was quantified in the two layers separately, amplitude variability in the epidermis across subjects exceeded that in the dermis (Figure 2D). Thus, amplitudes in the dermis were smaller but more consistent than amplitudes in the epidermis, which were larger and more variable. Nevertheless, magnitude and phase variabilities were similar in the two layers.

Predictive biomarkers of internal time in human dermis and epidermis

Finally, we assessed the viability of skin samples to be used for circadian phenotyping. Expression of biomarker genes in the skin can serve as predictors of internal phase of entrainment or chronotype, if a fixed phase relationship between the skin clock and the central SCN clock can be assumed. To predict *internal* time from a single sample, we identified biomarkers among genes expressed in either layer individually (as suggested by the layer-specificity) using `ZeitZeiger` [32].

A small set of population rhythmic genes accurately predict internal time. For an optimal parameter choice (Supplementary Figure 3A), 8-12 rhythmic genes were sufficient to predict internal time with a median absolute error (MAE) of $\sim$ 1.2 h (Figure 3A). The ability to infer internal time from a single sample can be seen in the counter-clockwise arrangement of samples projected on the two sparse principal components (Figure 3B). The biomarkers found in the two layers all exhibited robust population circadian rhythms (Supplementary Figure 3B,C) with the exception of one gene *FOCAD*, which was also rhythmic but with an amplitude below our cutoff. The genes chosen as biomarkers all showed particularly low variability in amplitude, magnitude and phase across subjects according to our analysis in the previous section (Supplementary Figure 3D).

The biomarkers sets contain layer-specific genes, but are depleted of core clock genes. Curiously, the biomarkers for internal time included only two canonical clock genes (*PER3*, *ARNTL*) and that too only in the dermal set. Moreover, the smaller biomarker set for the dermis shared only one circadian gene (*OVGP1*) with the larger set for the epidermis. The biomarkers in the dermis consisted of genes that were circadian (at the population level) in both layers. However, the epidermal biomarker set included several genes that were rhythmic only in the epidermis (*POLA1*, *ROR1*, *SOX5*, *ZNF143*). These sets also overlapped poorly with previously published biomarkers for epidermis (*ZBTB16*, *FKBP5*, [19]) and dermis (*PER3*, *ARNTL*, [20]).

Taken as a whole, our analysis indicates that either dermis or epidermis can be used for phenotyping circadian phase using a small set of mostly skin-specific circadian genes.

III. Discussion

This study aimed to characterize circadian gene expression in human skin. Human studies have to cope with the heterogeneity of individuals and their clocks [25], in contrast to circadian gene expression studies in mice [33]. We addressed this challenge in several ways: First, our study only included young, healthy subjects with intermediate chronotypes and stable natural sleep-wake rhythms; we nevertheless included male and female subjects to meaningfully describe circadian rhythms in the population. Second, we corrected sampling times for chronotype differences between subjects (using mid-sleep time on free days) to present results with respect to internal time; this was only possible due to the chronobiological profiles included in our study design. Third, we structured our characterization to describe (a) the circadian gene expression on average in a random healthy member of the population and (b) extent of the deviation of the circadian gene expression of that random member from the average circadian expression.

We found ~ 1200 genes with population circadian rhythms in either layer, thus, significantly expanding on the list of known clock output genes in the skin [19, 20]. This represents $\sim 10\%$ of the expressed genes like in most circadian mammalian tissues [6, 7] suggesting that our results capture most of the rhythmic genes. This improvement resulted from the higher frequency of sampling (every 4h) in our study in relation to past studies [18–20]. However, two-thirds of these genes had indistinguishable rhythms between the layers. On the one hand, this is unsurprising given the physical proximity between the layers. On the other, it is unexpected given the well-documented heterogeneity of the skin [34] and tissue-specificity of circadian programs in physiology [7]. Core clock gene expression too was consistent between the layers with rare exceptions. Despite this general similarity, a third of the circadian genes did display expression in only one or the other layer. Epidermal circadian genes were more in number and tended to have higher amplitudes than the dermal circadian genes (as suggested previously [20]). This might be due to the greater cellular heterogeneity of the dermis or the direct exposure of the epidermis to the external environment.

To quantify the deviations of subjects from the average circadian expression, we developed a novel method based on linear-mixed models and error propagation. Unexpectedly, the magnitude of circadian genes overall remained consistent across subjects in each layer, even though it varied across layers. This result supports the claim that circadian rhythms in humans can be constructed from population sampling, i.e., one sample per individual and thus side-step cumbersome and potentially unethical longitudinal sampling [7, 19, 35]. However, amplitudes and phases of circadian genes varied more across subjects than layers (in absolute terms and as a fraction of variance). But, the expression of specific circadian genes were relatively more or less subject-variable. For instance, negative core clock members (*PER1*, *PER2*) had highly subject-variable magnitudes and amplitudes, while positive core clock members (*ARNTL*, *NPAS2*) were the opposite. One consistent feature of the core clock genes was the high subject-variability of their phases in both layers. This might reflect the fact that the $\mathrm{MSF_{sc}}$ does not fully account for the chronotype differences between the subjects. However, amplitudes generally varied more than phases measured relative to population means across subjects in each layer. Amplitudes of circadian rhythms are expected to decrease with age [36], but were not previously known to vary more than phases in similarly-aged young subjects. Finally, dermal rhythm amplitudes were smaller but less variable across subjects, while epidermal rhythms possessed larger amplitudes and were more subject-variable. This represents an interesting dichotomy: the dermis might be a better source of circadian biomarkers, but the epidermis might be more indicative of amplitude differences between individuals with a larger dynamic range.

In recent years, a number of novel approaches have been introduced to assess circadian parameters and, in particular, circadian phase in humans (see [37] for a nice review) based on machine-learning on high-dimensional -omics data. We therefore explored the suitability of these two skin layers to provide

biomarkers to predict internal clock phase from single samples. Similar to our previous work on blood-based circadian phase determination [38], the expression of a small set of 8-12 circadian genes at a single time point was sufficient in either skin layer to predict internal clock phase with a median accuracy of $\sim$1 h. This accuracy is probably optimistic as it is based on internal cross-validation and a separate validation is necessary to estimate its true accuracy. Even with some loss of accuracy these biomarkers might be expected to perform as well as biomarker sets previously proposed for the epidermis and dermis [19, 20]. Genes in biomarker sets must possess consistent magnitudes and amplitudes across subjects in addition to robust rhythmicity in order for the inference from a single sample to be feasible. Our biomarker set for each layer showed low magnitude and amplitude variability across subjects as is desirable. Unlike other identified circadian biomarkers for phase [19, 20, 38, 39], the sets discovered in this study were almost devoid of core clock genes. Moreover, most biomarker genes were either rhythmic only in the layer in question or the genes differed significantly in amplitude and/phase from the other layer. This raises the unexpected possibility that biomarker sets involving tissue/layer-specific circadian genes might be better suited for internal phase prediction than core clock genes.

Our results leave open questions that need to be addressed in future studies. Our cohort was small, healthy, young and Caucasian and it is unclear how much our results can be extrapolated to a diverse population. Our inability to find differences between an analysis based on internal and external time was surprising, but is probably due to both the lack of extreme chronotypes in our cohort and insufficiency of 4 h sampling resolution to accurately reflect the $\sim$6h range of 95% of human chronotypes. The estimates of the variability in gene expression across subjects are affected by the small cohort size. In fact, the inter-subject variability subsumes inter-sex variability and cannot be reliably separated in this small cohort. Variance estimates across just two layers are well defined, but are likely less accurate compared to variance estimates across subjects. The error propagation analysis to quantify variation of circadian parameters is based on linearization and hence, assumes estimated mixed effects are "small". We identified circadian biomarkers for healthy young individuals maintaining natural yet regular sleep schedules. Whether these are also good markers in elderly and sick individuals and those with disrupted sleep schedules, such as shift workers, remains unanswered. It is unknown whether the human skin clock has a fixed phase relationship too or is independent of the central clock [40]. If the former, then the identified biomarkers can be used to predict central clock phase, else they would only be able to predict peripheral skin clock phase.

IV. Conclusions

We presented the most complete description to date of the transcriptional output of the circadian clock in two human skin layers. We reported both how average gene expression rhythmically varies in the population and the inherent variability in these average rhythms due to population heterogeneity. Our consideration of internal time in the analysis makes our results applicable to humans regardless of chronotype. Not only does our work provide a comprehensive resource of circadian gene expression and circadian biomarkers for phase in human skin, but also provides a methodology to describe human circadian rhythms in a population.

V. Materials and Methods

Experimental design and collection of human skin samples

The study to obtain human skin biopsies was approved by the local Ethical Review Board at Charité Universitätsmedizin Berlin (EA4/019/11). Tissue samples were collected according to the recommendations of the Declaration of Helsinki and to applicable laws for a non-drug study. All donors provided written and informed consent. The study was performed at the Clinic for Sleep & Chronomedicine, St. Hedwig-Krankenhaus Berlin. Eleven healthy volunteers (six males, five females, aged 20-30 years in 2011) participated in the study. Individual chronotypes were assessed using the Munich Chronotype Questionnaire (MCTQ) by calculating the mid sleep time on free days adjusted for the sleep-debt accumulated during the workweek MSF_{sc} [29]. All information about the subjects is provided in Supplementary Table 1.

Three millimeter biopsies were obtained from the upper back for seven time points over a period of 24 h (8:00, 12:00, 16:00, 20:00, 00:00, 04:00 and 8:00 the following day). Skin biopsies were subsequently incubated in PBS at 55 ^{o}C for 3 min to separate epidermis and dermis. Tissue samples were then frozen in liquid nitrogen and stored at -80 ^{o}C. RNA extraction and quality control from skin biopsies was performed by Miltenyi Biotec using the TRIzol method. Linear amplification and labeling of RNA and hybridization of Agilent Whole Genome Oligo Microarrays 4x44k (Agilent Technologies) using 1.2–1.65 μg of Cy3-labeled cRNA was performed by Miltenyi Biotec, essentially as reported in [41].

Gene expression analysis

The microarray gene expression analysis was conducted in R. The RMA (Robust Multichip Average) algorithm was used to pre-process and extract expression profiles from the raw CEL files. Genes were annotated with ENSEMBL and ENTREZ IDs using Agilent "Human Genome, Whole" annotation data (hgug4112a.db, v3.2). The raw gene expression data has been deposited in the Gene Expression Omnibus (GEO) under the accession number GSExxxx.

Raw data of the hybridized microarrays were normalized and processed using the Bioconductor R-Project package Linear Models for Microarray Data (`limma`). Background correction, normalization between the different arrays, removal of non-annotated probes, lowly-expressed genes as well as control probes was performed as suggested by the `limma` user-guide. The filtered and annotated dataset contains 11578 expressed genes.

Principal Component Analysis (PCA) was performed in order to remove outliers. The expressed genes nicely organized in two clusters in PCA space, separated by skin layer (data not shown). Nevertheless, the epidermis sample from subject 109 taken at 8:00 the following day did not cluster with the rest of epidermal samples and for this reason was removed from further analyses.

Rhythmicity analysis and functional annotation of circadian gene lists

To detect genes exhibiting rhythmic behavior with a 24 h period in their expression we used cosinor analysis and differential rhythmicity analysis as decribed in [30]. In short, we first tested the null hypothesis that the sine and cosine terms from cosinor analysis are equal to 0. Acrophases and amplitudes were estimated from the analysis and used to identify significantly oscillating genes. If the null hypothesis could be rejected under a false discovery rate threshold < 0.05 and a minimum amplitude requirement (i.e., that either the amplitude of the oscillating gene in dermis *or* in epidermis is above a peak-to-trough fold change amplitude > 1.5), we classified that gene as rhythmic in at least one of the layers. Next, we tested the differential rhythmicity null hypothesis, namely that, among the genes that were rhythmic in at least one layer, sine and cosine terms are equal across skin layers. If this hypothesis could be rejected under a FDR< 0.05, the gene was considered to have significantly *different* rhythms in dermis compared to epidermis. If, on the contrary, the null hypothesis

could not be rejected, we defined the gene to have *indistinguishable rhythms* across layers. Importantly, genes with amplitudes below the threshold in one of the layers but above the threshold in the other layer (that passed the minimum amplitude requirement) and with statistically indistinguishable rhythms were considered rhythmic and were included in our analyses despite the amplitude value being below the threshold in the first step of the analysis.

In analyses where *internal* time was used, sampling (wall) time was corrected to internal time in each subject by subtracting the mid-sleep time on free days after correcting for sleep debt during week days (MSF_{sc} [29]) to wall time.

Assessment of variability in circadian parameters across subjects and skin layers

In order to analyze how magnitudes, amplitudes and phases of individual circadian genes vary across subjects and layers, we analyzed each gene separately using the linear mixed models [31,42,43]. The expression of gene i, g_i, is modeled as

$$g_i = (m_i + \Delta m_{i,subj} + \Delta m_{i,layer}) + (a_i + \Delta a_{i,subj} + \Delta a_{i,layer}) \cos \omega t + (b_i + \Delta b_{i,subj} + \Delta b_{i,layer}) \sin \omega t,$$

where m_i, a_i and b_i represent the coefficients of the fixed effects for gene i; and Δm_i, Δa_i and Δb_i represent the random effects attributed to differences across layers or subjects, which are drawn from a normal distribution, whose variance is estimated.

While m_i is a direct readout of the gene's magnitude (and its respective uncertainty attributed to layers/subjects), amplitude A_i and phase ϕ_i of a gene i were calculated from the coefficients a_i and b_i for each gene as $A_i = \sqrt{a_i^2 + b_i^2}$ and $\phi_i = \arctan \frac{b_i}{a_i}$. To determine the variability in amplitude and phase across subjects and layers, we used error propagation. The variance of amplitude and phase across subjects and layers ($x = layer, subj$) was computed from the Jacobian matrices $\mathbb{J}_{i,x}$ and the covariance matrices $\Sigma_{i,x}$ of the rhythm parameters (obtained from the linear mixed-model fits) as

$$\sigma^2_{A,i,x} = \mathbb{J}_{A,i,x} \Sigma_{A,i,x} \mathbb{J}^{\mathrm{T}}_{A,i,x} \text{ and } \sigma^2_{\phi,i,x} = \mathbb{J}_{\phi,i,x} \Sigma_{\phi,i,x} \mathbb{J}^{\mathrm{T}}_{\phi,i},$$

where $\mathbb{J}_{A,i,x} = \left(\frac{\partial A_i}{\partial m_{i,x}} \; \frac{\partial A_i}{\partial a_{i,x}} \; \frac{\partial A_i}{\partial b_{i,x}} \right)$ (and $\mathbb{J}_{\phi,i,x} = \left(\frac{\partial \phi_i}{\partial m_{i,x}} \; \frac{\partial \phi_i}{\partial a_{i,x}} \; \frac{\partial \phi_i}{\partial b_{i,x}} \right)$).

Identification of predictive biomarkers of molecular skin phase

We used ZeitZeiger [32] to identify skin biomarkers of circadian phase. We tested two sets of predictors using the whole set of expressed genes in epidermis or dermis separately. The predicted variable was, in both cases, internal time. To evaluate the performance of the predictors, we followed a leave-one-subject-out cross-validation approach in the lines of [32,38]. To do this, predictors are trained with data from all subjects except one and internal time from the subject who is left-out is predicted. The process is iterated along all subjects and for different values of the two main parameters of ZeitZeiger, `sumabsv` and `nSPC`. The first parameter `sumabsv` controls how many genes form each sparse principal component (SPC) and the second parameter, `nSPC`, controls how many SPCs are used for prediction. Large values of either parameter result in more genes being needed for prediction. For each set of values of `sumabsv` and `nSPC` from the leave-one-subject-out cross-validation, we calculated the median absolute difference between the predicted and the observed internal time stamp across all subjects. We refer to this parameter as median absolute error (MAE), and it serves as a measure of accuracy of the prediction: the lower the error, the better the prediction.

Abbreviations

SCN: suprachiasmatic nucleus; MSF_{sc}: mid sleep time on free days after correcting for sleep debt during week days; FDR: false discovery rate; MAE: median absolute error; GEO: Gene Expression Omnibus; SPC: sparse principal component.

Conflict of interest

AK received consulting fees from Beiersdorf AG. The rest of the authors have on conflict of interest to declare.

Data accessibility

The gene expression microarray data of dermis and epidermis generated in this study are available at the Gene Expression Omnibus (GEO) database and can be accessed under the accession number GSExxxx. The source code is available through GitHub: xxxx.

Funding

This study was supported by Deutsche Forschungsgemeinschaft (DFG) grant AN 1553/2-1 to BA and by the DFG Project Number 278 001 972–TRR 186 to MdO. We acknowledge support from the German Research Foundation (DFG) and the Open Access Publication Fund of Charité Universitätsmedizin Berlin.

Author contributions

Study design and conceptualization: BA, MdO, AK, DK, HH, HR and TB. Experimental Methodology: SK, FS, AG, JdZ and CN. Bioinformatic methodology: MdO, BA and KJ. Investigation: BA, MdO, KJ, AK and HH. Writing (original draft and editing): MdO and BA. Writing (review): MdO, BA, MF, AK, HH, HR, TB, MF and JdZ. Funding acquisition: BA, AK, TB and HH.

Author affiliations

[1]Institute for Theoretical Biology – Laboratory of Theoretical Chronobiology, Humboldt Universität zu Berlin and Charité Universitätsmedizin Berlin, Philippstraße 13, House 4, 10115 Berlin, Germany. [2]Research and Development, Beiersdorf AG, 20245 Hamburg, Germany. [3]Institute for Medical Immunology – Laboratory of Chronobiology, Charité Universitätsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Germany. [4]Berlin Institute of Health – Center for Regenerative Therapies (BCRT), Charité Universitätsmedizin Berlin, Augustenburger Platz 1, 13353 Berlin, Germany. [5]Division of Pulmonary Inflammation, Charité Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt Universität zu Berlin, 10117 Berlin, Germany. [6]Department of Infectious Diseases and Respiratory Medicine, Charité Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt Universität zu Berlin, 10117 Berlin, Germany. [7]Institute of Physiology – Sleep Research & Clinical Chronobiology, Charité Universitätsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Germany. [8]Institute for Theoretical Biology – Laboratory of Systems Chronobiology and Chronomedicine, Humboldt Universität zu Berlin and Charité Universitätsmedizin Berlin, Philippstraße 13, House 20, 10115 Berlin, Germany.

REFERENCES

[1] Wong, R., Geyer, S., Weninger, W., Guimberteau, J.-C. & Wong, J. K. The dynamic anatomy and patterning of skin. *Experimental dermatology* **25**, 92–98 (2016).

[2] Plikus, M. V. *et al.* The circadian clock in skin: implications for adult stem cells, tissue regeneration, cancer, aging, and immunity. *Journal of biological rhythms* **30**, 163–182 (2015).

[3] Allada, R. & Bass, J. Circadian Mechanisms in Medicine. *New England Journal of Medicine* **384**, 550–561 (2021). URL http://www.nejm.org/doi/10.1056/NEJMra1802337.

[4] Takahashi, J. S. Transcriptional architecture of the mammalian circadian clock. *Nature Reviews Genetics* **18**, 164–179 (2017).

[5] Dibner, C., Schibler, U. & Albrecht, U. The mammalian circadian timing system: organization and coordination of central and peripheral clocks. *Annual review of physiology* **72**, 517–549 (2010).

[6] Zhang, R., Lahens, N. F., Ballance, H. I., Hughes, M. E. & Hogenesch, J. B. A circadian gene expression atlas in mammals: Implications for biology and medicine. *Proceedings of the National Academy of Sciences* **111**, 16219–16224 (2014). URL http://www.pnas.org/lookup/doi/10.1073/pnas.1408886111.

[7] Ruben, M. D. *et al.* A database of tissue-specific rhythmically expressed human genes has potential applications in circadian medicine. *Science Translational Medicine* **10**, eaat8806 (2018). URL http://stm.sciencemag.org/lookup/doi/10.1126/scitranslmed.aat8806.

[8] Burton, J. L., Cunliffe, W. J. & Shuster, S. Circadian Rhythm in Sebum Excretion. *British Journal of Dermatology* **82**, 497–501 (1970). URL https://onlinelibrary.wiley.com/doi/10.1111/j.1365-2133.1970.tb02211.x.

[9] Spruit, D. The Diurnal Variation of Water Vapour Loss from the Skin in Relation to Temperature. *British Journal of Dermatology* **84**, 66–70 (1971). URL https://onlinelibrary.wiley.com/doi/10.1111/j.1365-2133.1971.tb14198.x.

[10] Reinberg, A., Sidi, E. & Ghata, J. Circadian reactivity rhythms of human skin to histamine or allergen and the adrenal cycle. *Journal of Allergy* **36**, 273–283 (1965). URL https://linkinghub.elsevier.com/retrieve/pii/0021870765900869.

[11] Bjarnason, G. A. *et al.* Circadian expression of clock genes in human oral mucosa and skin: association with specific cell-cycle phases. *The American journal of pathology* **158**, 1793–1801 (2001).

[12] Oishi, K. *et al.* Differential expressions of mper1 and mper2 mrnas under a skeleton photoperiod and a complete light–dark cycle. *Molecular brain research* **109**, 11–17 (2002).

[13] Zanello, S. B., Jackson, D. M. & Holick, M. F. Expression of the circadian clock genes clock and period1 in human skin. *Journal of investigative dermatology* **115**, 757–760 (2000).

[14] Kawara, S. *et al.* Low-dose ultraviolet b rays alter the mrna expression of the circadian clock genes in cultured human keratinocytes. *Journal of investigative dermatology* **119**, 1220–1223 (2002).

[15] Brown, S. A. *et al.* The period length of fibroblast circadian gene expression varies widely among human individuals. *PLoS biology* **3**, e338 (2005).

[16] Brown, S. A. *et al.* Molecular insights into human daily behavior. *Proceedings of the National Academy of Sciences* **105**, 1602–1607 (2008).

[17] Spörl, F. *et al.* A circadian clock in hacat keratinocytes. *Journal of Investigative Dermatology* **131**, 338–348 (2011).

[18] Spörl, F. *et al.* Krüppel-like factor 9 is a circadian transcription factor in human epidermis that controls proliferation of keratinocytes. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 10903–10908 (2012).

[19] Wu, G. *et al.* Population-level rhythms in human skin with implications for circadian medicine. *Proceedings of the National Academy of Sciences* **115**, 12313–12318 (2018).

[20] Wu, G. *et al.* A population-based gene expression signature of molecular clock phase from a single epidermal sample. *Genome medicine* **12**, 1–12 (2020).

[21] Montaigne, D. *et al.* Daytime variation of perioperative myocardial injury in cardiac surgery and its prevention by rev-erbα antagonism: a single-centre propensity-matched cohort study and a randomised study. *The Lancet* **391**, 59–69 (2018).

[22] Dallmann, R., Okyar, A. & Lévi, F. Dosing-time makes the poison: circadian regulation and pharmacotherapy. *Trends in molecular medicine* **22**, 430–445 (2016).

[23] Long, J. E. *et al.* Morning vaccination enhances antibody response over afternoon vaccination: a cluster-randomised trial. *Vaccine* **34**, 2679–2685 (2016).

[24] Hsu, P.-K., Ptáček, L. J. & Fu, Y.-H. Genetics of human sleep behavioral phenotypes. *Methods in enzymology* **552**, 309–324 (2015).

[25] Roenneberg, T. *et al.* Epidemiology of the human circadian clock. *Sleep medicine reviews* **11**, 429–438 (2007).

[26] Stothard, E. R. *et al.* Circadian entrainment to the natural light-dark cycle across seasons and the weekend. *Current Biology* **27**, 508–513 (2017).

[27] Wright Jr, K. P. *et al.* Entrainment of the human circadian clock to the natural light-dark cycle. *Current Biology* **23**, 1554–1558 (2013).

[28] Allebrandt, K. V. *et al.* Chronotype and sleep duration: the influence of season of assessment. *Chronobiology international* **31**, 731–740 (2014).

[29] Vetter, C., Winnebeck, E. C. & Roenneberg, T. Asking the clock: How to use information from questionnaires for circadian phenotyping. In *Circadian Clocks*, 79–85 (2021).

[30] Pelikan, A., Herzel, H., Kramer, A. & Ananthasubramaniam, B. Venn diagram analysis overestimates the extent of circadian rhythm reprogramming. *The FEBS Journal* (2021).

[31] Hoffman, G. E. & Schadt, E. E. variancepartition: interpreting drivers of variation in complex gene expression studies. *BMC bioinformatics* **17**, 1–13 (2016).

[32] Hughey, J. J., Hastie, T. & Butte, A. J. ZeitZeiger: supervised learning for high-dimensional data from an oscillatory system. *Nucleic Acids Research* **44**, e80–e80 (2016).

[33] Geyfman, M. *et al.* Brain and muscle arnt-like protein-1 (bmal1) controls circadian cell proliferation and susceptibility to uvb-induced dna damage in the epidermis. *Proceedings of the National Academy of Sciences* **109**, 11758–11763 (2012).

[34] Duan, J., Greenberg, E. N., Karri, S. S. & Andersen, B. The circadian clock and diseases of the skin. *FEBS letters* **595**, 2413–2436 (2021).

[35] Anafi, R. C., Francey, L. J., Hogenesch, J. B. & Kim, J. Cyclops reveals human transcriptional rhythms in health and disease. *Proceedings of the National Academy of Sciences* **114**, 5312–5317 (2017).

[36] Blacher, E. *et al.* Aging disrupts circadian gene regulation and function in macrophages. *Nature Immunology* **23**, 229–236 (2022). URL `https://www.nature.com/articles/s41590-021-01083-0`.

[37] Dijk, D.-J. & Duffy, J. F. Novel approaches for assessing circadian rhythmicity in humans: a review. *Journal of Biological Rhythms* **35**, 421–438 (2020).

[38] Wittenbrink, N. *et al.* High-accuracy determination of internal circadian time from a single blood sample. *The Journal of clinical investigation* **128**, 3826–3839 (2018).

[39] Laing, E. E. *et al.* Blood transcriptome based biomarkers for human circadian phase. *Elife* **6**, e20214 (2017).

[40] Welz, P.-S. *et al.* BMAL1-Driven Tissue Clocks Respond Independently to Light to Maintain Homeostasis. *Cell* **177**, 1436–1447.e12 (2019). URL `https://linkinghub.elsevier.com/retrieve/pii/S0092867419305070`.

[41] Duggan, D. J., Bittner, M., Chen, Y., Meltzer, P. & Trent, J. M. Expression profiling using cdna microarrays. *Nature genetics* **21**, 10–14 (1999).

[42] Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* **67**, 1–48 (2015). URL `https://www.jstatsoft.org/index.php/jss/article/view/v067i01`.

[43] Laird, N. M. & Ware, J. H. Random-effects models for longitudinal data. *Biometrics* **38**, 963–974 (1982). URL `http://www.jstor.org/stable/2529876`.