How have our clocks evolved? Adaptive and demographic history of the out-of-African dispersal told by polymorphic loci in circadian genes

Arcady A. Putilov, Vladimir B. Dorokhov & Michael G. Poluektov

To cite this article: Arcady A. Putilov, Vladimir B. Dorokhov & Michael G. Poluektov (2017): How have our clocks evolved? Adaptive and demographic history of the out-of-African dispersal told by polymorphic loci in circadian genes, Chronobiology International

To link to this article: https://doi.org/10.1080/07420528.2017.1417314

Published online: 20 Dec 2017.
How have our clocks evolved? Adaptive and demographic history of the out-of-African dispersal told by polymorphic loci in circadian genes

Arcady A. Putilov⁎, Vladimir B. Dorokhov⁎, and Michael G. Poluektov†

⁎Research Group for Math-Modeling of Biomedical Systems, the Research Institute for Molecular Biology and Biophysics, Novosibirsk, Russia; †Laboratory of Sleep/Wake Neurobiology, The Institute of Higher Nervous Activity and Neurophysiology of the Russian Academy of Sciences, Moscow, Russia; ‡Department of Nervous Diseases, Institute of Professional Education, I.M. Sechenov 1-st Moscow State Medical University, Moscow, Russia

ABSTRACT

The mechanism of the molecular circadian clocks is currently understood as a transcription/translation feedback loop involving more than ten genes. Genetic variation at some of loci in these genes has been shaped by adaptation to environmental factors. In particular, latitudinal clines in allele frequency were documented in several animal species, but the contradictory conclusions were drawn from the results of rare human studies. Here we tested whether the out-of-African dispersal of human populations to higher latitudes of the Eurasian continent was associated with latitude-dependent shifts in allele frequency at polymorphic loci in genes of three (reference, circadian and skin pigmentation) groups. In order to detect the genetics-based signatures left by latitude-driven adaptation and to distinguish them from the confounding effects of population demographic history, we analyzed allele frequencies in 1594 individuals from 5 African and 11 Eurasian populations of the 1000 Genomes Project Phase 3. Up to 80 polymorphisms with global minor allele frequency > 0.2 were sampled from each of 36 genes (1665 polymorphisms in total). As expected, percentage of polymorphisms demonstrating both significantly enlarged differentiation of Eurasian populations on allele frequency and significant correlation between latitude and allele frequency was significantly higher in pigmentation genes compared to circadian genes and in circadian genes compared to reference genes. We also showed that the latitude-driven adaptation can be separated from genetic consequences of demographic perturbations by comparison of results obtained for the whole set of 16 African and Eurasian populations with results for only Eurasian populations that share the common demographic history. The revealed latitudinal clines in allele frequency seemed to be shaped by polygenic selection occurring by small allele frequency shifts spread across many loci in circadian and non-circadian genes. The present results provided a rationale for necessity to facilitate candidate gene studies by prioritizing genetic markers of chronotype.

Introduction

The intrinsic molecular clocks coordinate our physiology and behavior into circadian rhythms entrained to the 24-hour solar day. In human and other mammal species the mechanism of these clocks is currently understood as a transcription/translation feedback loop involving more than ten genes (Partch et al. 2014; Takahashi 2015). The genetic variation at some of particular loci in these genes was shaped by environmental factors. The most evident examples of such adaptation are latitudinal clines in allele frequency (Hut et al. 2011; Hut et al. 2013; Kyriacou et al. 2008). They were documented for the genes of circadian family in several animal species including birds (Johnsen et al. 2007; Liedvogel et al. 2009), fishes (Lemay and Russello 2014; O’Malley and Banks 2008), and flies (Costa et al. 2009; Sawyer et al. 2006). However, a set of rather contradictory conclusions was drawn from the results of rare human studies. Their authors generalized that the latitude-driven changes in allele frequency were either absent (e.g., Ciarleglio et al., 2008) or rare (e.g., Dall’Ara et al., 2016) or common (e.g., Forni et al. 2014).

For such phenotypic trait as chronotype a latitude-dependent variation was recognized in the analyses of questionnaire data collected in both Northern (Randler C 2017) and Southern Hemispheres (Leocadio-Miguez et al. 2017). However, the simplest explanation for the origin of the revealed shift toward eveningness at higher latitudes might be the latitude-
dependent reduction of the exposure to light (Leocadio-Miguel et al. 2017). On the other hand, significant differences in chronotype between people tracing their ancestry to different continents were also found in, at least, two multi-ethnic communities. In the USA, non-Hispanic European Americans differed from African Americans in reporting a more pronounced evening preference (Eastman et al. 2016; Malone et al. 2017), having a longer circadian period (Eastman et al. 2012; Eastman et al. 2016; Eastman et al. 2017) and a smaller impact of extreme circadian misalignment on sleep duration (Paech et al. 2017). In Brazil, a shift toward morningness was related to Amerindian but not African or European ancestry (Egan et al. 2017).

It is known that anatomically modern humans had evolved in near equatorial regions of Africa for more than 200,000 years (Richter et al. 2017; Skoglund et al. 2017). Only rather recently, 50,000–100,000 years ago, a relatively small group of the ancestors of present-day Eurasians had migrated out of Africa to pass through the genetic bottleneck along the way of this migration followed by a rapid population expansion (Carto et al. 2009; Lippold et al. 2014; Mellars et al. 2013). For the first time in their evolutionary history, these people have been exposed to such new environmental factor as seasonal variation in day length and other environmental conditions. Therefore, it is reasonable to expect that the dispersal from the near-equatorial African regions to various regions at higher latitudes of Eurasia had imposed selection pressure on human “chronophenotype”. The adjustment of the sleep-wake times to seasonal changes in times and duration of night and day seemed to require certain modification of basic characteristics of the endogenous circadian rhythms (e.g., of such parameters of these rhythms as intrinsic period and phase of entrainment). It is also reasonable to expect that this adaptation of the mechanism of circadian clocks shaped some of polymorphic loci in the genes of circadian family.

When alleles under selection increase in prevalence in a population, they leave distinctive genetics-based signatures (patterns of genetic variation) in DNA sequence. However, a great challenge for a search for such patterns is determining whether a given signature is due to selection or to the confounding effects of population demographic history that include bottlenecks (periods of reduced population size) and expansions (Kelley et al. 2006; Sabeti et al. 2006). Both the genetic bottleneck and the following rather rapid increase of effective population size had occurred on the way of out of Africa dispersal of the ancestors of the current Eurasians (Lippold et al. 2014). Consequently, the major purpose of the present analysis was to examine whether the expansion of human populations to higher latitudes of Eurasia led to latitude-driven shifts in allele frequency at some of polymorphic loci in circadian genes and whether such adaptive shifts might be distinguished from the shifts caused by the demographic changes in the Eurasian populations.

Chronobiological traits are not the only traits that supposed to be shaped by the latitude-related environmental factors. The well-known examples are the genetic networks underlying the differences between populations in skin pigmentation and vitamin D synthesis (e.g., Tiosano et al. 2016). The first observations of the association between skin color and latitude have been published as early as in eighteenth century (e.g., Smith 1787) and then refined in more recent scientific reposts in the second part of twentieth century (e.g., Roberts and Kahlon 1976; Walter 1971). After the transition of modern biology from genetics to the genomic era, polymorphisms in skin pigmentation genes have been always listed among the forefront genetic variants showing distinctive signatures of natural selection (Coop et al. 2009; Huerta-Sánchez et al. 2014; Lao et al. 2007; Pickrell et al. 2009; Soejima and Koda 2007; Sulem et al. 2007; Voight et al. 2006; Williamson et al. 2007). Such findings come as no surprise because, despite being a polygenic trait, genetic architecture of skin pigmentation was shown to be much simpler (i.e., fewer genes with stronger effect) as compared to the genetic architecture of vast majority of other such traits (Crawford et al. 2017). Therefore, signatures left by natural selection in these genes are expected to be stronger than the signatures in genes of circadian and most other genetic networks.

Demographic history of populations confounds inferences of selection history due to similar effects of both processes on the distribution of genetic variation. Population demographic history is a genome-
wide force that affects patterns of variation at all loci in a genome in a similar manner, whereas selection acts upon specific loci (Cavalli-Sforza 1966; Lewontin and Krakauer 1973; Przeworski et al. 2000). Therefore, the effects of demography and selection might be disentangled by comparing the groups of genes representing different genetic networks. Thus, an intuitively appealing approach for detecting genes that have been targets of natural selection is the identification of genes that seem exceptionally unusual (referred to as outliers) compared with most other loci (Biswas and Jm 2006; Sabeti et al. 2006).

Here we compared a set of polymorphisms from 12 circadian genes with polymorphisms in 12 pigmentation and 12 reference genes (i.e., the groups of genes for which clear evidence for latitude-driven selection was and was not previously provided, respectively). We predicted that, in agreement with the hypothesis of natural selection, some of polymorphic loci in circadian genes are characterized by a high level of population differentiation and a clear latitude-dependent shift of major allele frequency (MaAF). Since the loci associated with chronotype and sleep times in the previous candidate gene studies might also demonstrate such patterns of genetic variation, we predicted that MaAF of these loci more likely yield a pattern consistent with the hypothesis of natural selection than MaAF of loci that failed to show association with chronotype and sleep times in the results of previous candidate gene studies. Moreover, we predicted that MaAF changes that can be attributed to population demographic history are roughly similar for polymorphisms sampled from reference, circadian and skin pigmentation genes, but the amount of adaptive MaAF changes is significantly different in the three groups of polymorphisms. Namely, we predicted to detect the latter changes more often in circadian rather than reference genes and in pigmentation rather than circadian genes.

Methods

Samples

Genotype frequencies of 2504 individuals were taken from the dataset of the 1000 Genomes Project Phase 3 (Sudmant et al. 2015). Of 26 samples from the populations sampled for this project, 10 were not included in the analysis of genetic signatures of adaptation (910 individuals) due to a rather recent rapid change in a place of residence and/or a rather modern origin of population through admixture of people from different continents. These were people of African, Mexican and European ancestry in the USA, Gujarati Indian in the USA, Puerto Rican in Puerto Rico, African Caribbean in Barbados, Peruvian in Peru, Colombian in Colombia, Sri Lankan Tamil and Indian Telugu in the UK. The near equatorial regions of Africa always remained the places of residence of five out of 16 included populations (Yoruba and Esan in Nigeria, Luhya in Kenya, Mende in Sierra Leone and Gambian in Gambia). Other 11 populations (Figures 1 and 6) had evolved within the continent of Eurasia after the out-of-African exodus of the common for all Eurasians ancestral population (Japanese in Tokyo, Han Chinese in Beijing, Southern Han Chinese, Chinese Dai in Xishuangbanna, Kinh in Ho Chi Minh City, Punjabi in Lahore, Bengali in Bangladesh, British in the UK, Toscani in Italy, Iberian in Spain and Finnish in Finland).

Genes

The analyzed set of 1665 polymorphisms (Table 1, upper part) included single nucleotide polymorphisms (SNPs) and short indels (deletions and insertions) mapped in 12 reference genes (DBH, SLC6A3, DRD3, NPSR1, BDNF, CACNA1C, ACE, ACTN3, PPARA, GRIK3, TMEM132D and BRAF), 12 circadian genes (PER1, PER2, PER3, CLOCK, TIM, RORC, RORA, ARNTL, NPAS2, NFIL3, NR1D1 and CSNK1E) and 12 skin pigmentation genes (TMEM138, DDB1, TYR, MCIR, SLC24A5, SLC45A2, MFSD12, KITLG, TYR, OCA2, GRM5 and HERC2). Previously, the polymorphisms in the first three on the reference genes (DBH, SLC6A3 and DRD3) were associated with individual variation and alteration in dopaminergic neurotransmission (e.g., Corominas et al. 2009; Gray and MacKillop 2014; Mandelli and Serretti 2013; Paclt et al. 2004). Reliable evidence from several candidate gene studies (e.g., Bhat et al. 2012; Howe et al. 2016; Muglia et al. 2003; Schumacher et al. 2005) supported the associations.
of the polymorphic loci in the next three genes (NPSR1, BDNF and CACNA1C) with mental disorders. The polymorphic loci in ACE, ACTN3 and PPARA can be regarded as the three most replicable genetic markers of achievements in sport (e.g., reviewed by Ahmetov et al., 2015; Ahmetov et al. 2016). The last three reference genes (GRIK3, TMEM132D and BRAF) were shown to be involved in regulation of neurobehavioral functioning and, in particular, in the response to the processes of domestication of different animal species (e.g., dog and cattle) and to the process of self-domestication of anatomically modern humans (Erhardt et al. 2012; Hodgson et al. 2016; Minelli et al. 2009; Theofanopoulou et al. 2017). As for the structural genetic markers of variation in skin color, they were reliably identified in several genome-wide association studies in, at least, 12 genes (Beleza et al. 2013; Crawford et al. 2017; Jablonski and Chaplin 2013; Sturm 2009).

**Polymorphisms**

For each polymorphism included in the present analysis (Table 1), Global Minor Allele Frequency (GMAF) of 0.2 or higher was required. Since hundreds of such polymorphisms were found in some of these genes, not more than 80 polymorphic loci per gene were taken for the analysis (Table 1). Given that polymorphisms are sorted by default order at each webpage (e.g., https://www.ncbi.nlm.nih.gov/snp/?term=TMEM132D), the selection process always started from the last of the listed polymorphisms, continued toward the 1st listed polymorphism, and stopped anyway after selection of the 80th polymorphism with GMAF > 0.2. A preliminary analysis was performed by splitting each set of 80 polymorphisms into two halves. We applied χ²-test to examine whether two 40-polymorphism halves provide significantly different results of division of polymorphisms into two subgroups of polymorphisms that either met or did not meet criteria reported in Tables 2 (lower part) and 3.
The results of such statistical tests did not yield significant differences between two halves of a set of 80-polymorphisms thus pointing at the within-gene invariability of the studied features of the polymorphic loci.

In the present analysis, when frequency of an allele was found to become higher than 0.500 after averaging frequencies obtained for five African samples, this allele was defined as the major allele. Frequency of this allele (MaAF) was subjected to further analysis of its geographic variation.

Detecting signatures of adaptation

Numerous approaches have been previously developed to exploit signatures left by natural selection for identification of regions in the human genome harboring adaptations (Akey 2009; Sabeti et al. 2006), including the methods combining several such approaches (Grossman et al. 2010). Here, we combined two traditional approaches relying on results of analysis of differences between populations (e.g., Sabeti et al. 2006). The first approach is aimed on detection of a heightened level of population differentiation and another is aimed on detection of a strong and reliable correlation between geographic and genetic variables.

The development of the approach examining levels of population differentiation has been started by Sewall Wright in the 1920s (the so-called F-statistics). It was thereafter applied in his seminal work (Wright, 1950), publications of Lewontin (Lewontin and Krakauer 1973), Harpending (2002) as well as in many other more recent publications (Akery 2009; Sabeti et al. 2006). Given that African
populations always evolved in near equatorial regions thus contrasting to Eurasian populations exposed to various novel environments at distinct latitudes, we tested whether standard deviation (SD) of MaAF obtained for Eurasian samples is significantly heightened compared to that obtained for African samples. SD was calculated separately for 11 Eurasian (Figure 3) and 5 African samples and significance of difference between them was examined with Levene’s test for equality of variances.

The second approach was also previously widely applied in various studies on genetics of human adaptation (Coop et al. 2009; Fraser 2013; Hancock et al. 2011; Pritchard et al. 2010; Thompson et al. 2004). Here, we measured spatial relationship between allele frequency and the distance of the recent place of populations’ residence from the equator (latitude, degree North). Spearman rank coefficient of correlation (p) was obtained to relate latitude to MaAF (Figures 2–5).

<table>
<thead>
<tr>
<th>Gene</th>
<th>n</th>
<th>%</th>
<th>Gene</th>
<th>n</th>
<th>%</th>
<th>Gene</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBH</td>
<td>2</td>
<td>5.1</td>
<td>PER1</td>
<td>11</td>
<td>45.8</td>
<td>TMEM138</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>SLC6A3</td>
<td>0</td>
<td>0.0</td>
<td>PER2</td>
<td>0</td>
<td>0.0</td>
<td>DDB1</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>DRD3</td>
<td>0</td>
<td>0.0</td>
<td>PER3</td>
<td>11</td>
<td>14.3</td>
<td>TLRP1</td>
<td>21</td>
<td>91.3</td>
</tr>
<tr>
<td>NPSR1</td>
<td>6</td>
<td>7.5</td>
<td>CLOCK</td>
<td>0</td>
<td>0.0</td>
<td>MC1R</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>BDNF</td>
<td>2</td>
<td>6.5</td>
<td>TIM</td>
<td>6</td>
<td>9.4</td>
<td>SLC24A5</td>
<td>6</td>
<td>75.0</td>
</tr>
<tr>
<td>CACNA1C</td>
<td>3</td>
<td>7.7</td>
<td>RORC</td>
<td>0</td>
<td>0.0</td>
<td>SLC45A2</td>
<td>24</td>
<td>54.5</td>
</tr>
<tr>
<td>ACE</td>
<td>0</td>
<td>0.0</td>
<td>RORA</td>
<td>5</td>
<td>6.3</td>
<td>MFSD12</td>
<td>22</td>
<td>59.5</td>
</tr>
<tr>
<td>ACTN3</td>
<td>0</td>
<td>0.0</td>
<td>ARNTL</td>
<td>28</td>
<td>35.0</td>
<td>KITLG</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>PPARA</td>
<td>0</td>
<td>0.0</td>
<td>NPSA2</td>
<td>11</td>
<td>13.8</td>
<td>TYR</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>GRIK3</td>
<td>9</td>
<td>30.0</td>
<td>NFIL3</td>
<td>0</td>
<td>0.0</td>
<td>OCA2</td>
<td>31</td>
<td>38.8</td>
</tr>
<tr>
<td>TMEM132D</td>
<td>6</td>
<td>7.5</td>
<td>NR1D1</td>
<td>2</td>
<td>15.4</td>
<td>GRM5</td>
<td>1</td>
<td>1.3</td>
</tr>
<tr>
<td>BRAF</td>
<td>0</td>
<td>0.0</td>
<td>CSNK1E</td>
<td>0</td>
<td>0.0</td>
<td>HERC2</td>
<td>2</td>
<td>2.5</td>
</tr>
</tbody>
</table>

The same with p < 0.05

<table>
<thead>
<tr>
<th>Gene</th>
<th>n</th>
<th>%</th>
<th>Gene</th>
<th>n</th>
<th>%</th>
<th>Gene</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBH</td>
<td>24</td>
<td>61.5</td>
<td>PER1</td>
<td>19</td>
<td>79.2</td>
<td>TMEM138</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>SLC6A3</td>
<td>32</td>
<td>82.1</td>
<td>PER2</td>
<td>5</td>
<td>23.8</td>
<td>DDB1</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>DRD3</td>
<td>4</td>
<td>10.3</td>
<td>PER3</td>
<td>48</td>
<td>62.3</td>
<td>TLRP1</td>
<td>21</td>
<td>91.3</td>
</tr>
<tr>
<td>NPSR1</td>
<td>18</td>
<td>22.5</td>
<td>CLOCK</td>
<td>64</td>
<td>80.0</td>
<td>MC1R</td>
<td>9</td>
<td>56.3</td>
</tr>
<tr>
<td>BDNF</td>
<td>20</td>
<td>64.5</td>
<td>TIM</td>
<td>62</td>
<td>96.9</td>
<td>SLC24A5</td>
<td>8</td>
<td>100.0</td>
</tr>
<tr>
<td>CACNA1C</td>
<td>14</td>
<td>35.9</td>
<td>RORC</td>
<td>26</td>
<td>76.5</td>
<td>SLC45A2</td>
<td>34</td>
<td>77.3</td>
</tr>
<tr>
<td>ACE</td>
<td>25</td>
<td>64.1</td>
<td>RORA</td>
<td>41</td>
<td>51.3</td>
<td>MFSD12</td>
<td>36</td>
<td>97.3</td>
</tr>
<tr>
<td>ACTN3</td>
<td>9</td>
<td>32.1</td>
<td>ARNTL</td>
<td>61</td>
<td>76.3</td>
<td>KITLG</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>PPARA</td>
<td>21</td>
<td>36.2</td>
<td>NPSA2</td>
<td>37</td>
<td>46.3</td>
<td>TYR</td>
<td>21</td>
<td>26.3</td>
</tr>
<tr>
<td>GRIK3</td>
<td>20</td>
<td>66.7</td>
<td>NFIL3</td>
<td>13</td>
<td>68.4</td>
<td>OCA2</td>
<td>61</td>
<td>76.3</td>
</tr>
<tr>
<td>TMEM132D</td>
<td>39</td>
<td>48.8</td>
<td>NR1D1</td>
<td>12</td>
<td>92.3</td>
<td>GRM5</td>
<td>69</td>
<td>86.3</td>
</tr>
<tr>
<td>BRAF</td>
<td>2</td>
<td>5.7</td>
<td>CSNK1E</td>
<td>3</td>
<td>20.0</td>
<td>HERC2</td>
<td>39</td>
<td>48.8</td>
</tr>
</tbody>
</table>

Notes. Single criterion: Significant increase in standard deviation (SD) of MaAF in 11 Eurasian samples compared to five African samples, p < 0.001 and p < 0.05 (upper and lower part, respectively); n and %: Amount and percentage of the polymorphisms meeting this criterion.

Either 16 or 11 samples were included in correlation analysis (either all 16—African and Eurasian—or 11—only Eurasian—populations, p16 or p11, respectively).

Our expectations were that the complementary results will be provided by these criteria relying on analysis of differences between populations, either by the population differentiation criterion (SD) or by the correlation criterion (ether p16 or p11), and, therefore, a stronger multiple (dual or triple criterion) might be obtained by combining such single criteria (Tables 4 and 5).

Detecting confounding effects of demographic history

Because both applied approaches might be complicated by the confounding effects of population demographic history, we also compared correlation results obtained for all 16
(Eurasian and African) populations with results obtained for only 11 (Eurasian) populations to separate these effects from the effects of selection. All people of Eurasian ancestry share the common demographic history including the genetic bottleneck occurred soon after the out of African migration and the following rapid increase of effective population size (Lippold et al. 2014). Therefore, those features that differentiate the results obtained for all samples from the results for only Eurasian samples might be regarded as the representatives of demographic rather than adaptive genetic history.

**SNPs-correlates of chronotype**

In several candidate gene studies, the strength of association between such well-known phenotypic trait as chronotype and polymorphisms in, at least, 10 of 12 analyzed circadian genes was previously examined. These published reports provided possibility to compare two sets of, at least, 13 SNPs on their MaAF and ρ. Each of 13 SNPs included in the first set was previously found to be significantly associated with chronotype or sleep times (Table 6, left) whereas such an association was found to be non-significant for each of 13 other SNPs in the same genes included in the second set (Table 6, right part).
Statistical analyses

The SPSS statistical software package was used for all statistical analyses (IBM, Armonk, NY, USA, version 22.0). For the vast majority of analyses, $\chi^2$-test was applied to compare the percentages of polymorphisms meeting the applied criteria of differences between populations (e.g., Table 5). Additionally, principal component analysis was performed to reduce a set of MaAFs to a more manageable number of scores on the largest principal components (Figures 1 and 6).

Results

Principal component structure of MaAFs representing three groups of genes

Figure 1 illustrates similarity between results of principal components scoring of MaAFs of polymorphic loci sampled from references, circadian and pigmentation genes (Figure 1A, 1B and 1C, respectively). This analysis pointed at the major division between African and Eurasian populations (score on the 1st principal...
component) and at the following secondary division between East and West Eurasian populations (score on the 2nd principal component). None of the axes of this plot clearly corresponds to the South–North direction, and, therefore, this analysis did not provide evidence that latitudinal cline in MaAFs is a very common feature of polymorphisms in these genes. Instead, this analysis pointed at profound effects of demographic history that, in particular, led to differentiation of African populations from Eurasian populations.

**Reduction of MaAF in Eurasians**

The signatures of this history are clarified in Figures 2–4. In Figure 2, correlation coefficients between latitude and MaAF are plotted against

---

**Figure 4.** SD of MaAF for Eurasian samples and correlation with latitude for 16 samples. Data on polymorphisms from reference, circadian and pigmentation genes (A, B and C, respectively) with standard deviation (SD) showing significant increase in 11 Eurasian samples relative to standard deviation in 5 African samples (Levene’s test for equality of variances, $F > 4.59$, df = 14, $p < 0.05$). Each such polymorphism is represented by two measures of difference between populations: this SD and $\rho_{16}$ (Spearman coefficients of correlation between latitude and MaAF calculated for all 16 - African and Eurasian – samples). The vertical lines show the borderlines between non-significant ($-0.5 < \rho_{16} < 0.5$) and significant coefficients ($\rho_{16} < -0.5$ and $> 0.5$, $p < 0.05$, $n = 16$). Line illustrates a quadratic trend for relationship between SD and $\rho_{16}$.
changes in MaAF in Eurasia relative to Africa. Results on MaAF suggested that the out-of-African dispersal was associated with a general tendency of reduction of MaAF (Figure 2). Mean reduction of MaAF obtained by averaged over reference, circadian and pigmentation polymorphisms ($N = 537, 587$ and $541$) attained the values of $17.4\%, 15.0\%$ and $29.6\%$ (standard error $= 1.1\%, 0.7\%$ and $1.1\%$, respectively).

### Association of latitude with MaAF

As a result, the distribution of correlation coefficients between latitude and MaAF calculated for all 16 populations was found to be prominently skewed toward the negative pole and these coefficients were positively associated with reduction of MaAF (Figure 2). Such a relationship, however, was not yielded by results of a similar analysis of correlation coefficients obtained for 11 Eurasian populations.

---

**Figure 5.** SD for MaAF in Eurasian samples and correlation with latitude for 11 samples. The same data on SD as in the previous Figure 4. Again, each polymorphism is represented by two measures: SD and $p$, but $p_{16}$ was replaced by $p_{11}$ (Spearman coefficients of correlation between latitude and MaAF calculated for 11 - only Eurasian - samples). The vertical lines show the borderlines between non-significant ($-0.6 < p_{11} < 0.6$) and significant coefficients ($p_{11} < -0.6$ and $> 0.6$, $p < 0.05$, $n = 11$). Line illustrates a quadratic trend for relationship between SD and $p_{11}$. 
populations (Figure 3). As can be seen in Figure 3, the skew of coefficients toward the negative pole and their association with reduction of MaAF disappeared after replacing the correlations based on 16 samples (Figure 2) by the correlations based on 11 samples (Figure 3). Figure 3 also demonstrates that many polymorphisms were characterized by a rather strong positive correlation between latitude and MaAF in Eurasia despite the opposing shift (reduction) of MaAF after the out-of-African exodus. Therefore, these two patterns of variation observed at most of loci in any group of genes, the skew of correlation coefficients calculated for 16 samples toward the negative pole and the tendency of reduction of MaAF in Eurasia relative to Africa, might be attributed to the confounding effects of the population demographic history rather than to the adaptive shifts in allele frequency in response to the northward migration of ancestors of the current Eurasians. Northward movement of the migrants from Africa might lead to the latitude-dependent increase of MaAF of one set of loci and to the latitude-dependent decrease of MaAF of another set of loci (Figure 3).

**Differentiation of Eurasians on MaAF**

The confounding effects of demographic history on MaAF were also confirmed by the patterns of association between results of application of two approaches to examination of differences between populations. As in the case of pattern of association between reduction of MaAF and correlation with latitude shown in Figures 2 and 3, they are evident from the pattern of association between SD in Eurasia and correlation with latitude illustrated in Figures 4 and 5. Figure 4 illustrates the absence of the expected “butterfly-like” form of relationship between the two measures of population differentiation, SD of MaAF in Eurasia and correlation of allele frequency with geographical distribution of 16 populations, \( \rho_{16} \). Despite clear statistical evidence for the increased SD of MaAF in Eurasia relative to Africa (Table 2), a heightened SD did not lead to a stronger \( \rho_{16} \), either negative or positive. However, again, the expected pattern of relationship was revealed by replacement of the correlations obtained for 16 populations (Figure 4) by the correlations calculated for 11 (only Eurasian) populations, \( \rho_{11} \) (Figure 5). As indicated by the quadratic trends in this Figure 5, when an SD of MaAF in Eurasia for a given polymorphism was larger than for most other polymorphisms, a correlation coefficient, either negative or positive, tended to be stronger. In other words, this relationship between the results of two tests emerged when the analysis was limited to the samples from 11 Eurasian populations with similar demographic history and disappeared when the samples from populations with different demographic history (African vs. Eurasians) were merged in the analysis.

**Difference between the groups of genes in differentiation of Eurasians on MaAF**

We subdivided polymorphisms selected from each gene into two subgroups (Table 1, lower part, and 2–5). Polymorphisms of one subgroup met a particular criterion of adaptation while polymorphisms of another subgroup did not meet this criterion. For instance, Table 2 gives results on such single criterion as enlarged population differentiation in Eurasia. Upper and lower parts of Table 2 provide the results for two levels of significance of difference between SD of MaAF in the Eurasian and African samples \((p < 0.001 \text{ and } <0.05, \text{ respectively})\). The fraction of polymorphisms obtained by applying each criterion was subjected to statistical analysis aimed on comparison of the three groups of polymorphisms (see two upper lines of Table 5). The results suggested that reference, circadian and pigmentation polymorphisms were significantly different in percentage of MaAFs meeting this single criterion \((p < 0.001)\). As indicated by \( \chi^2 \)-test, when the level of significance was fixed at \( p < 0.001 \), this percentage was higher for circadian genes compared to reference genes \((\chi^2 = 18.57, p < 0.001)\) as well as for pigmentation genes compared to circadian genes \((\chi^2 = 10.75, p < 0.001)\). Such a result for reference genes was not challenged by lowering \( p \) to \(< 0.05 \). Percentage of cases of significantly higher differentiation in non-Africans was low in this group of genes as compared to the circadian genes \((\chi^2 = 6.611 \text{ and } 17.19, p < 0.001 \text{ for both})\). However, this percentage was the highest in circadian rather than pigmentation genes \((\chi^2 = 15.74, p < 0.001)\). The percentage obtained for circadian genes (66.6) suggested that two out of every three polymorphic loci in these genes demonstrated an enlarged
population differentiation between Europeans (Table 2, bottom line).

**Difference between the groups of genes in association of latitude with MaAF**

Moreover, circadian polymorphisms met a single criterion of correlation with latitude in the Eurasian samples (ρ₁₁) more often than polymorphisms in other genes when the level of significance was fixed at p < 0.001 (Table 3, lower part, and Table 5, middle). In this respect, the circadian polymorphisms significantly differed from polymorphisms in both reference and pigmentation genes (χ² = 4.16, p < 0.05, and χ² = 7.49, p < 0.01, respectively). As for a single criterion of correlation with latitude applied to the whole set of 16 samples (ρ₁₆), the pigmentation polymorphisms over-numbered other polymorphisms in the subgroups of significant correlates of latitude irrespective of the level of significance set for ρ₁₆ (Tables 3 and 5).

**Difference between the groups of genes revealed by multiple criteria**

Tables 1 and 4 (lower parts) contain the results of applying triple and dual criteria for dividing the polymorphisms into subgroups. The vast majority of such results allowed the distinguishing between each pair of the gene’s groups. As indicated by χ²-test, percentage of polymorphisms meeting these criteria was significantly higher for circadian genes compared to reference genes and, in turn, it was significantly higher for pigmentation genes compared to circadian genes. When the strictest criteria of significance were applied, none of the reference polymorphisms met these criteria (Table 5). For instance, as indicated by results of applying the dual criterion of significant correlation with latitude in 16 samples and significant correlation with latitude in 11 samples (ρ₁₁), the pigmentation polymorphisms over-numbered other polymorphisms in the subgroups of significant correlates of latitude.
increase of SD in Eurasia, the percentage of polymorphisms meeting this criterion was higher for circadian genes than for reference genes ($\chi^2 = 7.37$, $p < 0.01$) and, in turn, it was higher for pigmentation genes than for circadian genes ($\chi^2 = 4.11$, $p < 0.05$) when the level of significance was fixed at $p < 0.001$. Very similar differences between the three groups of genes were found after fixation of the level of significance at $p < 0.05$ ($\chi^2 = 15.74$, $p < 0.001$, and 7.37, $p < 0.01$, respectively).

### Principal component scoring of MaAFs-correlates of latitude

As many as 112 correlation coefficients obtained for 16 samples remained significant after Bonferroni’s adjustment for the whole number of 1665 tests (Table 5). Principal component analysis of three sets of MaAFs representing these polymorphisms in reference, circadian and pigmentation genes (15, 25 and 72, respectively) yielded the first principal component accounting for 85.61%, 89.19% and 90.21% of the total variation in MaAF, respectively. In order to illustrate the strength of relationship between genetic and geographic variations, scores on this 1st principal component were calculated and plotted against latitude in Figure 6A, 6B and 6C, respectively. As expected, the 1st principal component scores displayed strong latitudinal clines. Spearman’s coefficients of correlation in 16 populations attained the values of $-0.938$, $-0.925$ and $-0.906$, respectively ($p < 0.001$ for all). Moreover, the coefficients for scores of reference and pigmentation genes strongly correlated one with another (0.943) whereas such correlations with the score for circadian genes were somewhat weaker (0.804 and 0.848, respectively, $p < 0.001$ for all).

### Table 4. Polymorphisms meeting dual criteria of correlation and differentiation.

<table>
<thead>
<tr>
<th>Gene</th>
<th>n</th>
<th>%</th>
<th>Gene</th>
<th>n</th>
<th>%</th>
<th>Gene</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBH</td>
<td>2</td>
<td>5.1</td>
<td>PER1</td>
<td>1</td>
<td>4.2</td>
<td>TMEM138</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>SLC6A3</td>
<td>21</td>
<td>53.8</td>
<td>PER2</td>
<td>4</td>
<td>19.0</td>
<td>DDB1</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>DRD3</td>
<td>0</td>
<td>0.0</td>
<td>PER3</td>
<td>15</td>
<td>19.5</td>
<td>TNP1</td>
<td>3</td>
<td>13.0</td>
</tr>
<tr>
<td>NPSR1</td>
<td>7</td>
<td>8.8</td>
<td>CLOK</td>
<td>15</td>
<td>18.8</td>
<td>MCIR</td>
<td>5</td>
<td>31.3</td>
</tr>
<tr>
<td>BDNF</td>
<td>5</td>
<td>16.1</td>
<td>TIM</td>
<td>8</td>
<td>12.5</td>
<td>SLCA</td>
<td>8</td>
<td>100.0</td>
</tr>
<tr>
<td>CACNA1C</td>
<td>6</td>
<td>15.4</td>
<td>RORC</td>
<td>15</td>
<td>44.1</td>
<td>SLCA</td>
<td>8</td>
<td>18.2</td>
</tr>
<tr>
<td>ACE</td>
<td>2</td>
<td>5.1</td>
<td>RORA</td>
<td>16</td>
<td>20.0</td>
<td>MFS1D2</td>
<td>13</td>
<td>35.1</td>
</tr>
<tr>
<td>ACTN3</td>
<td>6</td>
<td>21.4</td>
<td>ARNTL</td>
<td>17</td>
<td>21.3</td>
<td>KITLG</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>PPARA</td>
<td>1</td>
<td>1.7</td>
<td>NPS2</td>
<td>20</td>
<td>25.0</td>
<td>TYR</td>
<td>13</td>
<td>16.3</td>
</tr>
<tr>
<td>GRIK3</td>
<td>7</td>
<td>23.3</td>
<td>NF1L3</td>
<td>13</td>
<td>68.4</td>
<td>OCA2</td>
<td>20</td>
<td>25.0</td>
</tr>
<tr>
<td>TMEM132D</td>
<td>12</td>
<td>15.0</td>
<td>NR1D1</td>
<td>5</td>
<td>38.5</td>
<td>GRM5</td>
<td>69</td>
<td>86.3</td>
</tr>
<tr>
<td>BRAF</td>
<td>2</td>
<td>5.7</td>
<td>CSNK1E</td>
<td>2</td>
<td>13.3</td>
<td>HERC2</td>
<td>20</td>
<td>25.0</td>
</tr>
<tr>
<td>Sum of 12</td>
<td>71</td>
<td>13.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes. Dual criterion: Significant increase in standard deviation (SD) of MaAF in 11 Eurasian samples compared 5 African samples and correlation between latitude and MaAF in either 16 or 11 samples ($\rho_{16}$ or $\rho_{11}$, upper or lower part, respectively, $p < 0.05$ for all); n and %: Amount and percentage of the polymorphisms meeting this criterion.
Table 5. Significant differences between three groups of genes in fractions of polymorphisms.

<table>
<thead>
<tr>
<th>Group Criterion</th>
<th>Reference genes</th>
<th>Circadian genes</th>
<th>Pigmentation genes</th>
<th>χ²-test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>Single</td>
<td>28</td>
<td>5.2</td>
<td>74</td>
<td>12.6</td>
</tr>
<tr>
<td></td>
<td>228</td>
<td>42.5</td>
<td>391</td>
<td>66.6</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>2.8</td>
<td>25</td>
<td>4.3</td>
</tr>
<tr>
<td></td>
<td>61</td>
<td>11.4</td>
<td>75</td>
<td>12.8</td>
</tr>
<tr>
<td></td>
<td>255</td>
<td>47.5</td>
<td>269</td>
<td>45.8</td>
</tr>
<tr>
<td>Duel</td>
<td>10</td>
<td>1.9</td>
<td>23</td>
<td>3.9</td>
</tr>
<tr>
<td></td>
<td>71</td>
<td>13.2</td>
<td>131</td>
<td>21.2</td>
</tr>
<tr>
<td></td>
<td>116</td>
<td>21.6</td>
<td>195</td>
<td>33.2</td>
</tr>
<tr>
<td></td>
<td>49</td>
<td>9.1</td>
<td>75</td>
<td>12.8</td>
</tr>
<tr>
<td>Triple</td>
<td>20</td>
<td>3.7</td>
<td>50</td>
<td>8.5</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Notes. The polymorphisms from each of three 12-gene groups meeting a criterion; n and %: Amount of the polymorphisms that met a given criterion and its percentage relative to the whole amount of the polymorphisms in this 12-gene group (N = 537, 587, and 541, respectively). Significant differences between three groups of genes in fractions of polymorphisms. χ²-test: Comparison of distributions of reference, circadian, and pigmentation polymorphisms into subgroups.

Table 6. SNPs in circadian genes examined for their link to chronotype in previous reports.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Significant link was reported</th>
<th>Significant link was not reported</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>rs2735611 (Carpen et al. 2006)</td>
<td>rs2855405 (Carpen et al. 2016)</td>
</tr>
<tr>
<td></td>
<td>rs934945 (Lee et al. 2011; Ojeda et al. 2013)</td>
<td>rs880140 (Kripke et al. 2014)</td>
</tr>
<tr>
<td></td>
<td>rs228697 (Kripke et al. 2014; Hida et al. 2014)</td>
<td>rs228727 (Dmitrzak-Weglarz et al. 2016)</td>
</tr>
<tr>
<td></td>
<td>rs2640909 (Ojeda et al. 2013)</td>
<td>rs228729 (Mansour et al. 2017)</td>
</tr>
<tr>
<td></td>
<td>rs1801260 (Katzenberg et al. 1998)</td>
<td>rs346544 (Kripke et al. 2014)</td>
</tr>
<tr>
<td></td>
<td>rs2291738 (Jankowski and Dmitrzak-Weglarz 2017)</td>
<td>rs10876890 (Jankowski and Dmitrzak-Weglarz 2017)</td>
</tr>
<tr>
<td></td>
<td>rs774045 (Etain et al. 2014)</td>
<td>rs774049 (Etain et al. 2014)</td>
</tr>
<tr>
<td></td>
<td>rs3828057 (Kripke et al. 2014)</td>
<td>rs9826 (Etain et al. 2014)</td>
</tr>
<tr>
<td></td>
<td>rs1159814 (Dorokhov et al. 2018)</td>
<td>rs782931 (Etain et al. 2014)</td>
</tr>
<tr>
<td></td>
<td>rs1481892 (Dmitrzak-Weglarz et al. 2016)</td>
<td>rs4146388 (Jankowski and Dmitrzak-Weglarz 2017)</td>
</tr>
<tr>
<td></td>
<td>rs3816358 (Evans et al. 2013)</td>
<td>rs11022780 (Jankowski and Dmitrzak-Weglarz 2017)</td>
</tr>
<tr>
<td></td>
<td>rs3768984 (Evans et al. 2013)</td>
<td>rs17024874 (Kripke et al. 2014)</td>
</tr>
<tr>
<td></td>
<td>rs2482705 (Kripke et al. 2014)</td>
<td>rs13297268 (Kripke et al. 2014)</td>
</tr>
</tbody>
</table>

Notes. Left side: Findings on significant link to chronotype/sleep times were previously reported for 13 SNPs in 10 circadian genes. Right side: For 13 other SNPs in the same 10 genes the previously reported findings were negative; p₁₁ and p₁₂: Spearman coefficient of correlation between latitude and MAAP in 16 (African and Eurasian) and 11 (only Eurasian) samples (line above and line below, respectively).
SNPs-correlates of chronotype and latitude

A remarkable similarity was found between the set of 26 SNPs from the previously reported candidate gene studies (Table 6) and the set of 587 circadian polymorphisms (Tables 2 and 3) included in the above analyses (e.g., patterns of correlation of MaAF with latitude, reduction of MaAF in Eurasia, increase of its SD in Eurasia, etc.). No statistically significant difference in changes of MaAF was obtained in the comparison of MaAFs of two sets of SNPs for which evidence for significant association with chronotype/sleep time either was or was not provided in the previously reported candidate gene studies. However, MaAFs of 13 chronotype-associated SNPs seemed to be more strongly associated with latitude than MaAFs of 13 other SNPs. For instance, the left and right sides of Table 6 illustrate that Spearman coefficients of correlation with latitude ($\rho_{16}$) reached a statistically significant level for 10 of 13 chronotype-associated SNPs and for only 4 of 13 other SNPs (77% vs. 30%, respectively). Consequently, the comparison made with $\chi^2_{1}$-test suggested that the difference between two sets of 13 SNPs in the amounts of established correlations with latitude (at $p < 0.05$) was significant ($\chi^2_{1} = 5.571$, $p = 0.0183$). Moreover, principal component analyses of these two sets yielded two scores on the 1st (largest) principal component in 13 MaAFs one of which was a significant correlate of latitude (the first set) while another did not correlate so strongly with latitude (the second set). Spearman coefficients of correlation for $\rho_{16}$ attained the values of $-0.898$ and $-0.414$ ($p < 0.001$ and $p = 0.111$, respectively) and of $-0.834$ and 0.610 for $\rho_{11}$ ($p < 0.001$ and $p = 0.046$, respectively).

Discussion

The results of present analysis yielded the genetics-based signatures left by the latitude-driven adaptation and allowed their distinguishing from the confounding effects of population demographic history. We predicted that these signatures might or might not be found in DNA sequences within reference genes, that they have to be found in
circadian and pigmentation genes, and that they have to be found to be more distinctive in pigmentation than circadian genes due to a simpler genetic architecture of the skin color trait (i.e., fewer genes with stronger effect) as compared to most other quantitative traits including "chrono-phenotype". Our predictions were clearly supported by the results obtained by combining two traditional approaches to exploration of signatures of natural selection, that is, when the statistical tests suggested both a heightened level of population differentiation and a reliable correlation between geographic and genetic variables. The majority of comparisons of polymorphisms in genes of three groups indicated that percentage of those polymorphisms that showed both a statistically significant enlargement of differentiation of the Eurasian populations on MaAF and a statistically significant correlation between latitude and MaAF was higher in pigmentation genes as compared to circadian genes and in circadian genes as compared to reference genes. We also separated such genetics-based signatures of adaptation from changes associated with demographic perturbations by comparing the results obtained for the whole set of 16 samples (African and Eurasians) with the results on samples from populations sharing the common demographic history (only Eurasians). Finally, the present results of additional analysis indicated that an allele frequency more likely was a significant correlate of latitude when it was associated with chronotype or sleep times in the previously published reports.

Therefore, such results lent support for the suggestions that the latitudinal clines in allele frequency are not uncommon among the polymorphisms in the genes of circadian family and that, in response to the northward expansion of the ancestors of the current Eurasians, the genetic variation underlying the circadian clock mechanism has been modified relative to variation established in African populations. The most impressive finding suggested that 66.6% (two out of every three) loci in circadian genes responded to the out of African dispersal by significant increase in differentiation of Eurasian populations on MaAF (Table 2, last line). Usually, an inter-population variation in MaAF in circadian genes was not as high as that demonstrated by many MaAFs of pigmentation genes (Figure 5B and 5C), but only approximately a half of loci in pigmentation genes responded by significant increase in differentiation of Eurasian populations on MaAF (Table 2, last line). Possibly, the adaptive changes in skin color were still necessary in Africa after divergence of the studied populations and this adaptation led to further differentiation of these populations on polymorphisms in skin pigmentation genes. In contrast, the adaptive changes in the mechanism of molecular clocks were not required in Africa and these changes in the form of relatively small shifts in allele frequency at many loci in circadian genes occurred mostly in Eurasia. In general, such results supported the assumption that many polymorphisms in circadian genes might be shaped by polygenic adaptation to seasonal fluctuations in environmental factors (e.g., seasonality of timing and duration of night and day in the most of Eurasian regions). These fluctuations were absent during the long previous period of evolution in the near equatorial African regions and become a new selective factor at higher latitudes of Eurasia. The key feature of adaptation driven by polygenic selection is that it occurs by small allele frequency shifts spread across many loci (Pritchard and Di Rienzo 2010). Therefore, polygenic selection might be a cause of even those of these shifts that were too small to reach a statistically significant level in the present analysis.

Our results can be implicated into candidate gene studies in the field of chronobiology to facilitate a search for genetic markers of chronotype. Dozens or even hundreds of polymorphisms were identified within each circadian gene. Since most of these genetic variants seem to have only tiny effect on such complex trait as chronotype, false positive findings are not excluded in candidate gene association studies of this trait. So far, not so many such studies have been published (Goel 2017; Kalmbach et al. 2017; Zhang et al. 2013) and just a few replicated findings. The only well-studied polymorphism appears to be a variable number tandem repeat in the coding region of PER3 (rs57875989). Positive findings on association of this polymorphism with chronotype and sleep phase delay syndrome were reported in the earliest publications (e.g., Archer et al. 2003).
following publications, such results were either confirmed (Kunorozva et al. 2012) or only partially replicated (Jones et al. 2007; Pereira et al. 2005) or the association was found to be significant in the opposing direction (Lázár et al. 2012; Liberman et al. 2017). Besides, negative findings were reported in many other publications (An et al. 2014; Barclay et al. 2011; Kang et al. 2011; Kripke et al. 2014; Mansour et al. 2017; Osland et al. 2011; Perea et al. 2014; Shawa and Roden 2016; Turco et al. 2017; Voinescu and Coogan 2012; etc.). It is also unlikely that genome-wide association studies (GWAS) can help in identification of the most promising loci within any of the genes of circadian family. Most GWAS peaks map to non-protein-coding sequences, where their molecular consequences can be difficult to evaluate. In general, for peaks within protein-coding and non-protein-coding sequences, the ratio was found to be approximately 1:6 (e.g., Grossman et al. 2010). Therefore, it comes as no surprise that, although significant association with chronotype were reliably identified at DNA regions located in closed proximity to several circadian genes, such as PER2 and PER3 (e.g., Hu et al. 2016; Jones et al. 2016; Lane et al. 2016), their mapping near but not within these genes pointed at their regulatory rather than protein-coding function. It is reasonable to suggest that, in the case of the vast majority of causal variants within a circadian gene, such a variant explains just a small amount of variation, and, hence, its effect does not reach a stringent significance threshold of a GWAS study. Therefore, the problem of identification of most promising genetic candidate for chronotype among thousands such candidates in dozens of circadian genes can be, at least partly, solved by prioritization of genetic markers of this trait, and it seems that the present results on the signatures of natural selection in the structural genes of circadian family can be used for this purpose.

The rationale for using these results for such a prioritization might be the following. Let us consider the fate of minor alleles in circadian genes after the out-of-Africa exodus. The results of analyses of big samples of polymorphic loci (e.g., Gravel et al. 2011; Keinan et al. 2007; Tennessen et al. 2012) indicated that many very rare alleles had been mostly washed out during the following bottleneck and, thereafter, there was no enough evolutionary time to regain this part of allele spectrum via de novo mutations. In contrast, frequency of less rare minor alleles including common minor alleles (frequency ≥20%) had increased because an expanding population tends to increase the fraction of rare alleles (Coventry et al. 2010; Tennessen et al. 2012). In the present results, such an increase is evident for the studied minor alleles of the vast majority of polymorphisms that is in agreement with the previously reported differences between African and Eurasian populations in allele frequency spectrum (e.g., Gravel et al. 2011; Keinan et al. 2007). If during further population migration and expansion the latitude-driving adaptation imposed selective pressure on a common minor allele, its frequency is expected to decrease thus allowing the emergence of a positive correlation between latitude and MaAF in Eurasian populations. If, in contrast, this adaptation process favored a common minor allele, its frequency increased and, importantly, this increase led to an appearance of very strong negative correlation between latitude and MaAF due to cumulative effects of natural selection and population expansion on this common minor allele frequency. In fact, some of such alleles were minor in African populations to become major in Northern European populations, like rs1159814 in RORA (Table 6; Dorokhov et al. 2018). Therefore, compared to any other minor allele in the same genes, this rather small fraction of alleles is expected to contribute more to the genetic variation in chronobiological traits shaped by the latitude-driven adaptation. It seems that a lucky choice of such an allele might explain the present result suggesting that chronotype-associated loci significantly correlated with latitude more often than other loci. Consequently, the minor alleles that became major only in the North Eurasia, that show a strong negative correlation with latitude, and that demonstrate a heightened level of population differentiation in Eurasia seemed to be the most promising loci among the whole pool of thousands of candidates offered by circadian genes.

Several hypothetical reasons might be suggested for explaining why allele frequencies of
some of loci in reference genes demonstrated significant correlation with latitude. The pleotropic effects of the circadian genes are now well-documented. Particularly, the wide-ranging pleotropic effects were uncovered in studies exploring the associations of polymorphic variants of circadian genes with human metabolic diseases (Pappa et al. 2013), infertility (Hodžić et al. 2013), cancer (Karantanos et al. 2013), addictions (Blomeyer et al. 2013), psychiatric disorders (KarthykeyKarthykeyan et al. 2014), Parkinson’s disease (Gu et al. 2015), sleep disorders (Veatch et al. 2017) and so on. However, the pleotropic effects of the polymorphisms in other genetic networks on the circadian phenotypes remain unexplored. We previously published results (Dorokhov et al. 2018; Taranov et al. 2017) of examination of whether individual differences in such chronobiological characteristics as morning-evening preference and sleep times might be linked to polymorphic variants in either circadian (rs2640909 in PER3, rs12649507 in CLOCK, rs4851377 in NPAS2 and rs1159814 in RORA) or reference genes (rs1611125 in DBH, rs6347 in SLC6A3, rs6280 in DRD3, rs324981 in NPSR1 and rs6265 in BDNF). Although the significant associations with score on one of two morning-evening scales were confirmed only for variants in circadian genes (Dorokhov et al. 2018; Taranov et al. 2017), two of the variants in the reference genes (rs6347 in SLC6A3 and rs324981 in NPSR1) showed significant associations with state-like variation in sleep times (Taranov et al. 2017). The present results on correlation of some of polymorphisms in reference genes with latitude allow us to expect that future studies might reveal significant pleotropic effects on circadian phenotypes imposed by the polymorphisms in these and some other non-circadian genes have.

More general hypothetical reason for correlation between latitude and allele frequency showed by some of loci in reference genes might be sufficient interconnections of the gene regulatory networks that were yielded by the analysis of genetic background of complex traits. All genes expressed in a given cell are liable to affect the functions of core trait-related genes. Therefore, most heritability can be explained by effects on genes outside core pathways (Boyle et al. 2017). Some of polymorphisms in reference genes might be involved in such indirect way in regulation of either chronobiological traits or skin pigmentation or some other traits that were shaped to more or less extent by the latitude-driven polygenic selection. Moreover, it cannot be fully excluded that latitude-dependent environmental factors played, at least, secondary role in selection of complex traits associated with these genes.

In sum, the results of the present analysis indicated that the out-of-African dispersal of human populations had led to the enlargement of interpopulation difference in allele frequency at the vast majority of polymorphic loci in circadian genes and that such an enlargement was often accompanied by establishment of a strong link to latitude. The revealed pattern of geographic variation in allele frequency might be shaped by polygenic adaptation to seasonal variation in day length and other environmental factors.

Declaration of interest
The authors report no conflicts of interest. The authors alone are responsible for the content and writing of this article.

Funding
The studies of AAP were supported by grants from the Russian Foundation for Basic Research (grant numbers 07-06-00263-a, 10-06-00114-a, 13-06-00042-a and 16-06-00235-a) and the Russian Foundation for Humanities (grant numbers 06-06-00375-a, 12-06-18001-e and 15-06-10403-a). The studies of VBD were supported by grants from the Russian Foundation for Humanities (grant numbers 15-06-10909a, 15-06-10874a, and 14-06-00963a).

References
with objectively measured sleep traits in the elderly. Sleep 36:431–46.


