






OPEN

Population-specific association of *Clock* gene polymorphism with annual cycle timing in stonechats

Hannah Justen^{1,2}, Timo Hasselmann^{1,3}, Juan Carlos Illera⁴, Kira E. Delmore^{1,2}, David Serrano⁵, Heiner Flinks⁶, Masayuki Senzaki⁷, Kazuhiro Kawamura^{8,9}, Barbara Helm^{10,11} & Miriam Liedvogel^{1,12}

Timing is essential for survival and reproduction of organisms across the tree of life. The core circadian clock gene *Clk* is involved in the regulation of annual timing events and shows highly conserved sequence homology across vertebrates except for one variable region of poly-glutamine repeats. *Clk* genotype varies in some species with latitude, seasonal timing and migration. However, findings are inconsistent, difficult to disentangle from environmental responses, and biased towards high latitudes. Here we combine field data with a common-garden experiment to study associations of *Clk* polymorphism with latitude, migration and annual-cycle timing within the stonechat species complex across its trans-equatorial distribution range. Our dataset includes 950 records from 717 individuals from nine populations with diverse migratory strategies. Gene diversity was lowest in resident African and Canary Island populations and increased with latitude, independently of migration distance. Repeat length and annual-cycle timing was linked in a population-specific way. Specifically, equatorial African stonechats showed delayed timing with longer repeat length for all annual-cycle stages. Our data suggest that at low latitudes with nearly constant photoperiod, *Clk* genotype might orchestrate a range of consistent, individual chronotypes. In contrast, the influence of *Clk* on annual-cycle timing at higher latitudes might be mediated by its interactions with genes involved in (circadian) photoperiodic pathways.

Time is an essential dimension of the ecological niche of a species^{1–3}. In consequence, organisms have evolved internal (i.e., endogenous) time-keeping mechanisms to anticipate changes in the environment that recur rhythmically with high precision. One such endogenous mechanism is the circadian clock that enables organisms to anticipate diel (i.e. within a 24 h period) changes^{4,5}. Circadian clocks continue to run under experimental conditions of constant light or darkness, and in natural environments they are entrained by diel changes in light intensity and duration. The circadian core clock in many organisms is well characterised on the molecular level as a set of genes that regulate biochemical oscillations with a period of 24 h (here called "circadian genes"^{5–7}). For many species, for example fish, mammals and birds, we observe consistent variation in rhythmic behaviour on the level of the individual e.g., by being consistently early or late relative to the population⁸. Such so-called "chronotypes" have been linked to allelic variation in circadian genes, mostly, but not exclusively in humans⁹. The

¹MPRG Behavioural Genomics, Max Planck Institute for Evolutionary Biology, 24306 Plön, Germany. ²Department of Biology, Texas A&M University, 3258 TAMU, College Station, TX 77843, USA. ³Department of Animal Physiology, University of Cologne, Cologne, Germany. ⁴Biodiversity Research Institute (CSIC-Oviedo University-Principality of Asturias), Oviedo University, Oviedo, Spain. ⁵Department of Conservation Biology, Estación Biológica de Doñana (EBD-CSIC), Avda América Vespucio 26, 41092 Seville, Spain. ⁶Wagnerstr. 19, 46354 Suedlohn, Germany. ⁷Environmental Earth Science, Hokkaido University, Sapporo, Hokkaido, Japan. ⁸Graduate School of Agriculture, Hokkaido University, Kitaku Kita 9 Nishi 9, Sapporo, Hokkaido, Japan. ⁹Department of Wildlife Biology, Forestry and Forest Products Research Institute, Matsunosato 1, Tsukuba, Ibaraki 305-8687, Japan. ¹⁰Department of Bird Migration, Swiss Ornithological Institute, Seerose 1, 6204 Sempach, Switzerland. ¹¹Institute of Biodiversity, Animal Health and Comparative Medicine, University of Glasgow, Glasgow, UK. ¹²Institute of Avian Research "Vogelwarte Helgoland", 26386 Wilhelmshaven, Germany. email: h.justen@evolbio.mpg.de; liedvogel@evolbio.mpg.de

demonstrated links between behaviour and circadian gene regulators have made variation in rhythms particularly accessible to ecologists and evolutionary biologists¹⁰. Evolutionary interest so far has mostly focused on the gene *Clk*, a single copy gene that is constrained from evolving fast¹¹. The translated protein CLOCK is a key element in the positive arm of the circadian transcription-translation feedback loop¹². *Clk* is highly conserved among vertebrates throughout most of its sequence, except for a C-terminal exonic region that contains an often-variable poly-glutamine (poly-Q; CAG/CAA) trio-repeat motif. Length variation in this region varies between species, but in many species also shows substantial variation between and also within populations¹³. This variable poly-Q region has been shown to influence the transcription activating potential of the CLOCK protein in *Drosophila*¹⁴ and mouse¹⁵, where truncated versions of the *Clk*

Circadian clocks, and the molecular mechanisms that control them, are also involved in annual time-keeping, in conjunction with circannual clocks. Their main role is measurement of photoperiod (daylength)^{16,17}, whose predictable change over the year provides an external calendar of increasing amplitude pole-wards, while being absent at equatorial regions. Annual changes in photoperiod can time key life cycle events, such as reproduction or migration, either directly or by entrainment of the circannual rhythm, for which photoperiod is the main synchronising cue¹⁸. Thus, molecular changes in clock genes may affect annual timing either directly via pleiotropic effects on circadian and circannual rhythms, or indirectly via influencing photoperiodic responses^{19,20}. For example, in many organisms the circadian clock defines a photo-sensitive time window during which reproduction can be triggered or terminated^{16,17}. Hence, individuals with genetic disposition for early *versus* late-entraining circadian clocks are expected to respond differently to a given daylength²¹. Recent evidence supports corresponding diel and annual chronotypes, e.g. for migration-linked behaviour in a songbird²². However, photoperiod-dependent links between diel and annual time-keeping can be complex, since daylength increases in spring, but decreases in autumn.

So far, some studies have indeed demonstrated a correlation between *Clk* genotype and annual-cycle traits. For example, poly-Q repeat length correlated with timing of migration in fish^{23,24} and birds^{25,26}, timing of moult²⁷, as well as latitudinal clines and breeding phenology in some species^{28–30}. However, these results varied between study systems and a consistent pattern is lacking, e.g. no such correlation was detected between poly-Q repeat length and latitudinal clines in other bird and fish species^{31–33}. Furthermore, no correlation was found for timing of breeding³² and migration³⁴ in other bird species, and some species were altogether lacking detectable genetic variation at the poly-Q locus³⁵.

This lack of consistency has raised critical responses towards candidate gene approaches^{2,10,31,36}, especially when evaluating genotype–phenotype association of complex behavioural traits that are generally assumed to be polygenically controlled. Furthermore, circadian genes have broad, highly pleiotropic effects, similar to “house-keeping” genes¹⁰, and there is scepticism whether selection for a particular trait (e.g., timing of breeding) would modify genes with such broad functions, rather than acting on more specific physiological pathways^{20,37,38}. In addition, caveats have been raised against investigating single candidate genes that are part of wider pathways, like the *Clk* gene^{36,39}, and genotype–phenotype associations based on polymorphism of these genes. The reasons for the observed inconsistency between studies are so far not clear.

Inconsistent findings on the role of the *Clk* gene in different species emphasize that results from these candidate gene studies cannot easily be generalized across taxa²⁵. For different organisms, the role of *Clk* likely differs, for example due to differences in circadian organization, in genetic architecture, in photoperiodic environments, or in demands on time-keeping^{2,8,40,41}. Thus, in some organisms diel and annual timing might correlate, whereas in others separate and independent regulation may be beneficial^{21,22,42}. Furthermore, most studies are correlational, so that genetic contributions to timing are entangled with a host of responses to environmental conditions⁴³.

Here, we address uncertainty over *Clk* genotype–phenotype associations using a widely discussed subject, avian annual-cycle timing, by a combination of experimental approaches and comparison of closely related taxa. In birds, variation in allelic diversity and repeat length has been focally studied in the context of environmental and behavioural seasonality, specifically breeding latitude, annual-cycle timing and migratory distance. With increasing latitude and thus intra-annual amplitude in day length, reproductive seasons are delayed and shortened, moult timing is shifted, and migratory distance generally increases⁴⁴. Convincing support for a relationship between *Clk* genotype and annual-cycle timing, however is scarce^{45,46}.

Aiming to shed light on so far inconsistent results, we developed an experimental design that paired biogeographic associations with controlled experiments under common-garden conditions in captivity³⁸. We compared closely related taxa within a species complex across wide latitudinal breeding ranges and used a common-garden experiment to disentangle inherited timing from environmental responses⁴³.

Specifically, we investigated the relationship between *Clk* poly-Q polymorphism and breeding latitude, migration and annual-cycle timing in a well-studied songbird complex with a trans-equatorial distribution range, the stonechat (*Saxicola* spp.; Fig. 1). Taxonomy of the former species *Saxicola torquata* is undergoing rapid reassessment. Currently, *Saxicola* is considered a species complex, of which we are studying populations or subspecies assigned to the species *Saxicola rubicola* (Austria, Ireland, Spain, Germany), *Saxicola torquatus* (Kenya, Tanzania), *Saxicola dacotiae* (Canary Islands), *Saxicola stejnegeri* (Japan), and *Saxicola maurus* (Kazakhstan). Henceforth, we here use the term “population” to refer to any group of individuals belonging to these taxa, except where otherwise stated. We compared *Clk* genotypes of nine populations with breeding ranges spanning 55° of latitude, including resident, short- and long-distance migratory populations from Europe, Asia and equatorial Africa (Table 1;^{47,48}). In addition to their latitudinal variation these populations differ in their annual timing of breeding, moult and migration. Furthermore, stonechats breeding at the equator experience a constant 12:12 h light/dark cycle throughout the year (Fig. 1), and thus present a natural scenario that allows to contrast populations coping with annually fixed and seasonally changing daylengths in the same species complex, minimising cross-taxa comparative noise. However, previous experiments have shown that African populations still adjust

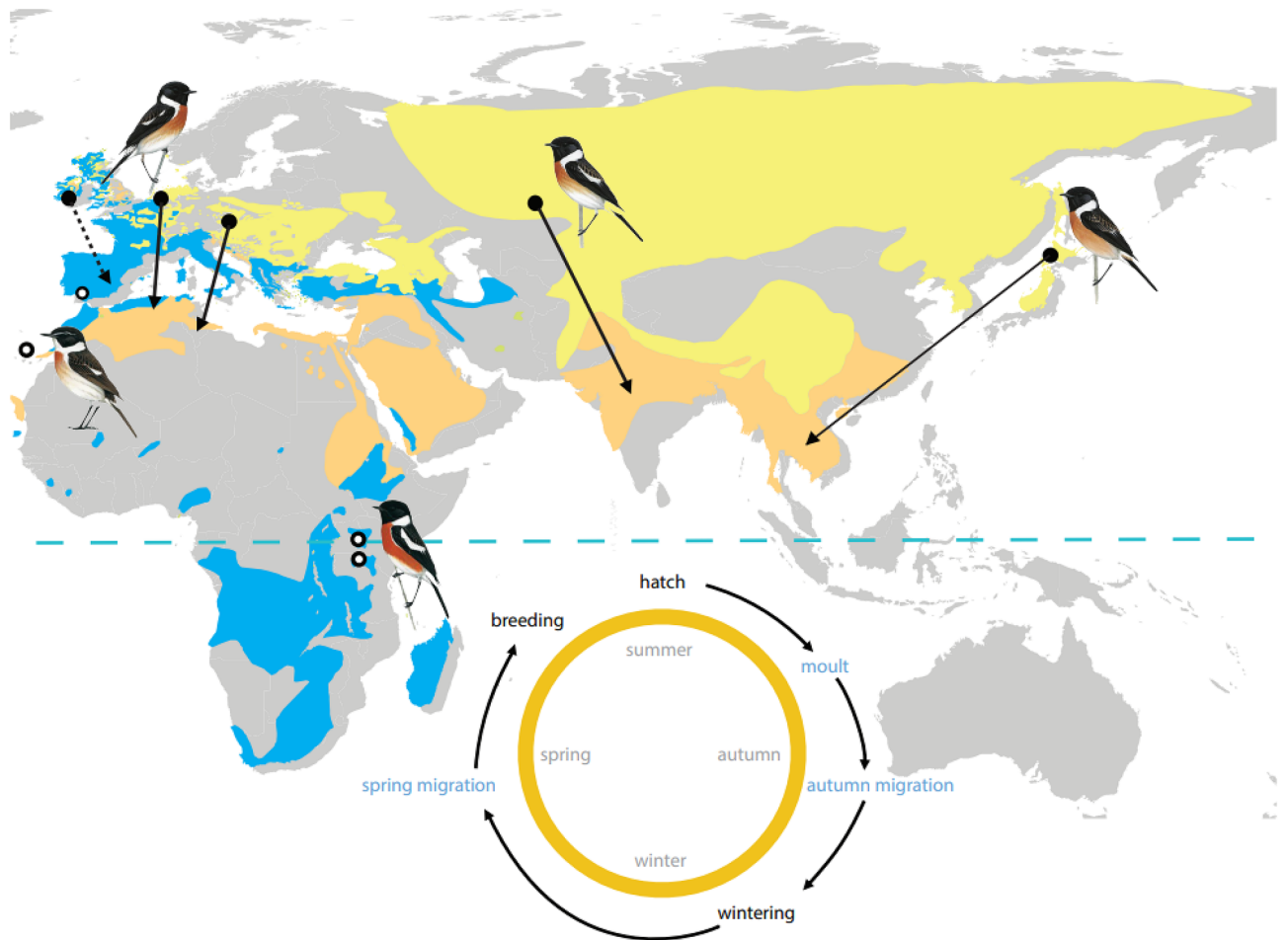


Figure 1. Overview of the geographical distribution and migratory behaviour within the *Saxicola* complex. Breeding range (yellow) and wintering range (orange) is shown for migrants, year-round range (blue) is shown for residents; the equator is indicated by a dashed blue line. Breeding location is indicated by filled circles for migratory populations and open circles for residents; arrows depict migratory direction and distance (dotted line indicates partial migrants). The circular inset schematic illustrates key life history events during the first annual cycle of a stonechat's life, starting with hatching, followed by moult, autumn migration (in case of migrants), wintering period, return migration in spring, and breeding, before the cycle starts all over again. Focal timing events investigated in this study are highlighted in blue. Bird illustrations show the European *S. rubicola*, the Fuerteventura *S. dacotiae*, the African *S. torquatus*, the Siberian *S. maurus* and the Japanese *S. m. stejnegeri*⁸⁶ taxa.

Populations	n	Latitude	Distance	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	p-value ^a	HO	GD
Kenya	125	0°14' S	0	0	0.004	0	0.116	0.024	0.844	0.012	0	0.102	0.280	0.275
Tanzania	47	3°5' S	0	0	0	0.021	0.043	0.021	0.883	0.032	0	0.162	0.213	0.219
Canary	61	28°46' N	0	0	0	0	0.008	0	0.008	0.984	0	0.008	0.016	0.033
Spain	93	37°39' N	0	0.022	0	0	0.258	0	0.21	0.511	0	0.911	0.645	0.631
Japan	11	43.6°N	4900	0	0	0	0.273	0	0.546	0.182	0	0.085	0.909	0.609
Austria	130	48°14' N	1700	0.008	0	0	0.127	0.004	0.25	0.612	0	0.109	0.608	0.549
Germany	47	51°47' N	1700	0.011	0.021	0	0.160	0.021	0.213	0.564	0.011	0.831	0.617	0.617
Kazakhstan	90	51.5°N	2600	0	0	0.028	0.128	0	0.656	0.033	0.156	0.233	0.644	0.530
Ireland	112	52°N	1200	0	0.005	0	0.098	0.058	0.455	0.384	0	0.709	0.661	0.635

Table 1. Breeding latitude, migration distance [km], *Clk* poly-Q repeat length frequencies (Q₈₋₁₅), Hardy-Weinberg exact test p-values, observed heterozygosity (HO) and gene diversity (GD) of nine stonechat (*Saxicola* spp.) populations listed by increasing latitude. ^aEstimated with Markov Chain parameters: dememorization, 10 000; batches, 10 000; iterations per batch, 10,000.

annual timing to photoperiodic information in similar ways to northern congeners, a finding that was interpreted to reflect a possibly retained, ancestral pattern⁴⁹.

A subset of four of these populations was studied under a common garden setting; specifically: residents (Kenya), partial migrants (Ireland), short-distance (Austria) and long-distance migrants (Kazakhstan). This approach required to keep all main known modifiers of timing consistent between populations. Thus, photoperiod, the main synchroniser of annual cycles, was set to simulate European daylength, under which all populations perform appropriate, population-specific annual activities^{47,49,50}. Our annual-cycle perspective ranged from late summer (postjuvenile moult), through autumn migration to spring migration. We investigated *Clk*-related differences in timing with the following key objectives:

- (1) Investigating geographic patterns of allelic diversity. Assuming that *Clk* genotype affects annual-cycle timing, we expect allelic diversity to vary with latitude and migration behaviour (here quantified as mean migration distance). We tested this hypothesis using all nine populations. If individuals within the population differ in timing or in their photoperiodic exposure, we would expect to find high allelic diversity in *Clk*. Evolutionarily, such differences could result from fluctuating selection on timing. For increasing breeding latitude and migration distance, we might expect increased allelic diversity due to fluctuating mortality linked to arrival timing and wintering latitude^{51,52}. In equatorial populations, constant photoperiodic conditions, but inter-annually variable breeding opportunities could either lead to canalisation and favour fixation of an optimised circadian *Clk* genotype, or, conversely, diversified time-keeping and hence, a broad range of genetically determined chronotypes.
- (2) Investigating associations between repeat length and population-level timing. If *Clk* allele lengths affect annual timing, we expect to detect differences in allele lengths with breeding latitude and migration distance in the nine populations of wild birds. Equatorial populations may again deviate from patterns at higher latitudes because of the absence of photoperiodic change.
- (3) Investigating associations between *Clk* genotype and individual-level timing of annual cycle traits. We test directly for genotype–phenotype associations using the four captive populations kept under identical photoperiodic conditions. Because *Clk* allele lengths might affect annual timing via photoperiodic time measurement, we examine whether relationships between repeat length and timing reverses between autumn (decreasing daylength), and spring (increasing daylength). Finally, because the role of *Clk* in annual timing might differ between populations, e.g., due to changes in photoperiodism, we test for population-specific relationships.

Results

Clk gene diversity, breeding latitude and migration distance (objective 1). We characterised *Clk* poly-Q allelic lengths variation in nine breeding populations and identified eight length variants ranging from 8 to 15 poly-Q repeats (*Clk* poly-Q_{8–15}; subscript indicates the number of poly-Q repeats) found at medium to high frequencies (Table 1). The most common allele (MCA) was Q₁₃ for Kenya, Tanzania, Ireland, Japan, and stonechats from Kazakhstan; and Q₁₄ for Spain, Austria, Germany and the Canary Islands (Table 1). In both African as well as the Canary Islands populations MCAs accounted for the majority of observed alleles, resulting in reduced population specific allelic diversity: Canary Islands (MCA 98.4%), Kenya (MCA 84.4%) and Tanzania (MCA 88%). In contrast, the contribution of the MCA to overall allelic diversity was considerably lower in stonechats from the European and Asian continent: Kazakhstan (65.6%), Austria (61.2%), Germany (56.4%), Japan (54.5%), Spain (51.1%), Ireland (45.5%) (Table 1). Poly-Q allele frequency within stonechat taxa did not deviate significantly from Hardy–Weinberg equilibrium ($p > 0.5$; Table 1), except for the Canary Islands population ($p < 0.01$; Table 1).

Clk gene diversity was characterised as observed heterozygosity (defined as frequency of observed number of heterozygotes) as well as within population gene diversity (defined as unbiased gene diversity per sample and locus by Goudet⁵³). We analysed *Clk* gene diversity across the geographical range of stonechat populations in the context of breeding latitude and migration distance, two factors that were correlated, albeit not significantly (Pearson's correlation: $R = 0.56$; $p = 0.119$; $df = 7$). Equatorial populations from Africa and the Canary Islands showed lower levels of diversity for both measures compared to all other populations; highest diversity levels were found in the Irish population (Table 1). Gene diversity was predicted significantly by breeding latitude, but not migration distance in a linear model (linear regression: latitude: $F_{2,6} = 5.417$; $R^2 = 0.64$; $p = 0.017$; migration distance: $p = 0.694$). When we excluded the Canary Island population and restricted our analysis to continental populations, lowest levels of gene diversity were found in the African populations breeding close to the equator (Fig. 2). However, this model did not account for potential differences in neutral genome-wide nucleotide diversity.

We thus additionally analysed the five populations for which genome-wide nucleotide diversity (defined as mean autosomal pi) was available (Kenya, Canary Islands, Austria, Ireland and Kazakhstan; estimates from Van Doren et al.⁵⁴). This analysis tentatively evaluated if the pattern we observed in *Clk* gene diversity is due to selection in that particular region of the genome or the result of whole genome elevation of nucleotide diversity. When adding autosomal pi as a covariate to our models for this subset of populations, we neither found a correlation of *Clk* gene diversity with breeding latitude or migration distance, nor with genome-wide nucleotide diversity (linear regression: $F_{3,1} = 1.717$; $R^2 = 0.84$; latitude: $p = 0.305$; distance: $p = 0.522$; autosomal pi: $p = 0.589$).

Clk repeat length, breeding latitude and migration distance (objective 2). We detected no significant relationships of breeding latitude or migration distance of our nine stonechat populations with *Clk* poly-Q repeat length of a total of 716 individuals (mixed effects linear model with origin as a random factor; latitude:

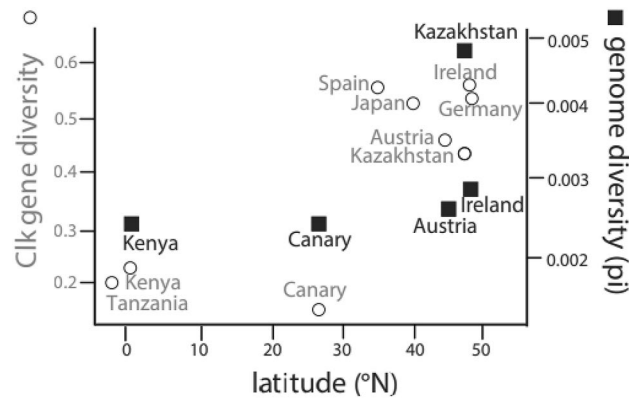


Figure 2. *Clk* locus specific and genome wide diversity of stonechat populations relative to breeding latitude. Gene diversity at the polymorphic *Clk* locus for the nine populations of this study is indicated by open circles (left axis); genome wide nucleotide diversity (autosomal π) is plotted by black squares (right axis) for five populations with available genome wide statistics⁵⁴. *Clk* gene diversity is lower in the African populations breeding at the equator than in the populations that breed at higher latitudes in Europe and Asia.

slope estimate \pm SE: 0.12 ± 0.18 ; df: 6.63; $p = 0.515$; migratory distance: slope estimate \pm SE: -0.083 ± 0.16 ; df: 7.61; $p = 0.625$); see Supplementary Material, Fig. S1).

Genotype–phenotype association: annual-cycle timing (objective 3). To test for associations between *Clk* genotype, characterised as the mean number of poly-Q repeats at the variable locus, and timing of different focal traits we used the four captive populations and included all available data from a common-garden experiment⁴⁷ (for sample sizes per model, see Fig. 3). We ran linear mixed effects models including *Clk* repeat length, origin, sex, and hatch date, as well as selected two-way interactions, for onset, peak and end of moult and spring and autumn migratory restlessness.

Timing of postjuvenile moult. During this first annual-cycle stage in young birds, timing correlated positively with *Clk* allele length for onset, peak and end in the Kenyan population (Fig. 3a–c, Table 2). These three time points were delayed by 9, 11 and 12 days per additional poly-Q repeat, respectively. In contrast, the relationship between *Clk* mean allele length and timing of the Austrian, Irish and Kazakh populations differed significantly from those of Kenyans and was slightly negative (Fig. 3, Table 2). Hatch date had significant effects on moult timing, and this association was population-specific (Table 2). In contrast, moult timing was not affected by sex.

Timing of autumn migratory restlessness. For the subsequent annual-cycle stage, autumn migration, we found a positive relationship between *Clk* allele length and timing (Fig. 3d–f, Table 2), which did not differ significantly between populations. Slopes were steepest in Kenyan stonechats, whose onset, peak and end were delayed by 33, 25 and 11 days per poly-Q repeat, respectively. Slopes in the remaining populations were far less steep, but overall positive, except for end of autumn restlessness in Irish stonechats (Table 2). Sex showed no association with autumn timing.

Timing of spring migratory restlessness. Population-specific patterns for spring migratory restlessness were similar (Fig. 3g–i, Table 2) to those observed for moult timing, and were clearest for the onset of restlessness. Kenyan stonechats delayed onset, peak and end per poly-Q repeat by 17, 12 and 13 days, respectively, whereas Austrian, Irish and Kazakh populations showed no association or slightly advanced timing with increasing poly-Q repeat. Additionally, we confirmed protandry (i.e. males started migratory restlessness slightly earlier than females) during spring migration (Table 2).

Discussion

We characterised *Clk* gene poly-Q variation in 950 records from 717 individuals from nine closely related populations of stonechats, including residents, short-distance and long-distance migrants. The latitudinal range covered by these populations also included equatorial populations, which to our knowledge, were not investigated in previous studies. Our study system thus newly allowed us within the same species-complex to contrast *Clk* gene variation across a latitudinal gradient including the equator.

All stonechat populations were in Hardy–Weinberg equilibrium, suggesting random mating within populations of the same subspecies ($p > 0.5$, Table 1), except for Canary Island stonechats, which diverged from the mainland stonechats 1.6 mya⁵⁵ and are endemic to the island of Fuerteventura⁵⁶. A bottleneck event in the Canary Island population after colonization may explain their low genetic diversity⁵⁴, and their resident behaviour⁵⁷, along with non-overlapping ranges with other stonechat taxa, may explain the deep genetic differentiation. This

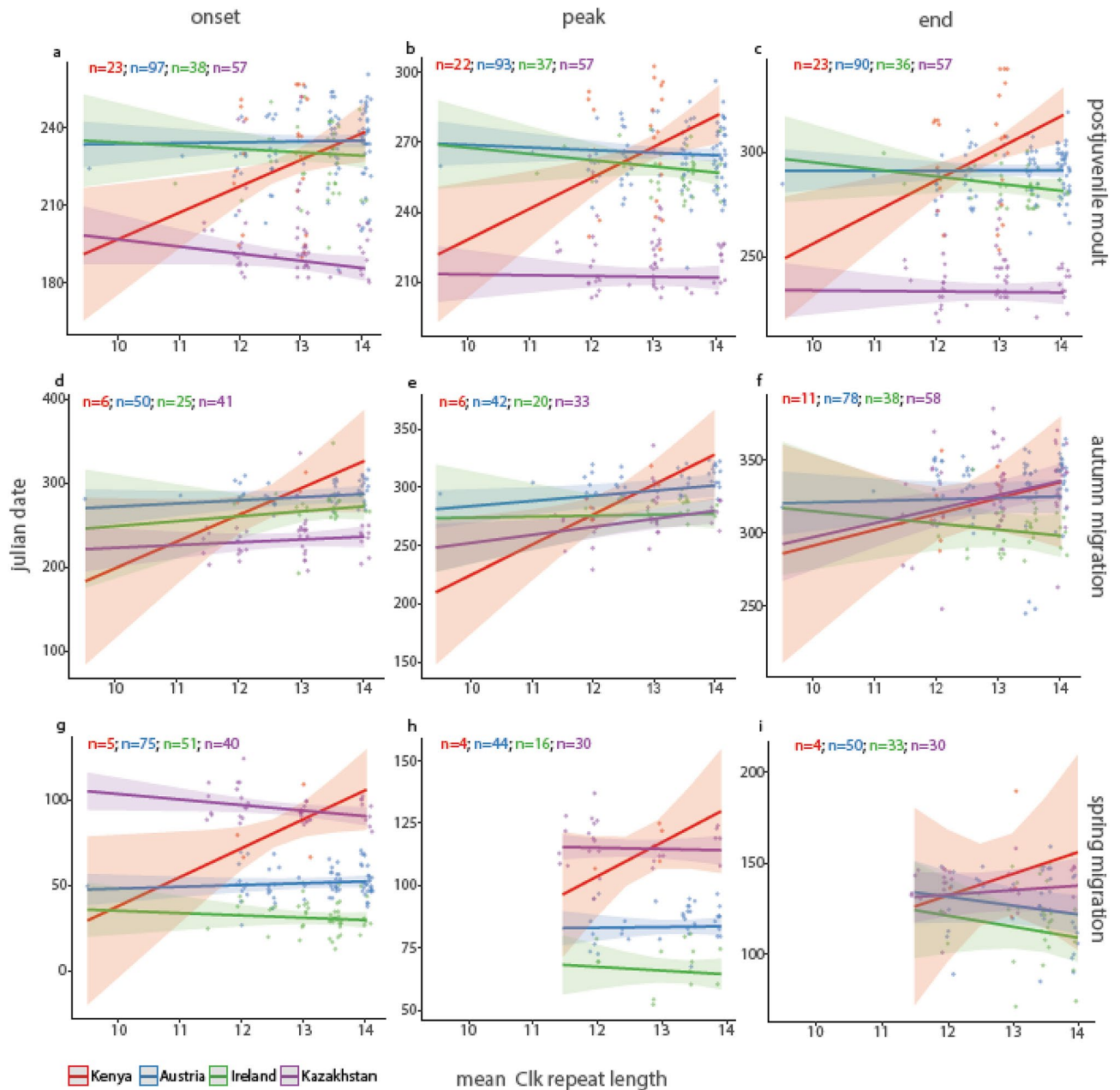


Figure 3. Timing of moult and migration in relation to population-specific *Clk* gene mean repeat length in stonechats. Onset, peak and end of postjuvenile moult (a–c) and of migratory restlessness exhibited during autumn migration (d–f) and spring migration (g–i) are shown in relation to mean repeat length for four populations: Kenyan stonechat populations plotted in red, Austrian in blue, Irish in green, and Kazakh in purple.

history is consistent with significant deviation from Hardy–Weinberg equilibrium ($p = 0.008$; Table 1) and advises caution when interpreting population genetics results from the Canary Island population.

Clk gene diversity in different stonechat populations across a latitudinal gradient revealed substantial variation in different *Clk* poly-Q allelic variants ranging from 8 to 15 repeats at the variable Q-locus (Table 1). Variation across study systems varies considerably and in *Saxicola spp.* levels of diversity at the *Clk* locus (here as observed heterozygosity [$HO = 0.016–0.909$]) and frequency of eight different allelic variants is high compared to other passerines. For example, a couple of earlier studies on a variety of study species only showed four different poly-Q variants including barn swallows (*Hirundo rustica*, Q_{5-8} , $HO = 0.066$) and pied flycatcher (*Ficedula hypoleuca*, Q_{10-13} , $HO = 0.478$)²⁶ while five different poly-Q variants were reported in other species including tree pipit (*Anthus trivialis*, Q_{6-10}) and nightingales (*Luscinia megarhynchos*, Q_{9-13} , $HO = 0.55$)²⁶. With exclusion of rare outliers, six was the highest number of poly-Q variants previously reported in several species, including whinchats (*Saxicola rubetra*, Q_{11-16} , $HO = 0.125$)²⁶, blue tits (*Cyanistes caeruleus*, Q_{9-14} , $HO = 0–0.637$)²⁸, blue-throats (*Luscinia svecica*, Q_{10-15} , $HO = 0–0.476$)²⁸ and great tits (*Parus major*, Q_{10-15} , $HO = 0.077$)³⁵. One possible

	Onset			Peak			End		
	Estimate ± SE	DF	p	Estimate ± SE	DF	p	Estimate ± SE	DF	p
Postjuvenile moult									
Intercept	87.09 ± 54.03	202	0.109	-19.70 ± 62.07	196	0.751	101.93 ± 60.35	193	0.093
<i>Clk</i> length	10.98 ± 4.26	202	0.011	13.501 ± 4.60	196	0.004	15.48 ± 4.76	193	0.001
Austria	145.34 ± 56.30	202	0.011	235.58 ± 64.24	196	<0.001	188.90 ± 62.98	193	0.003
Ireland	163.24 ± 63.28	202	0.011	246.16 ± 72.03	196	<0.001	227.92 ± 70.76	193	0.001
Kazakhstan	140.01 ± 58.21	202	0.017	61.48 ± 71.06	196	0.388	133.73 ± 65.01	193	0.041
Hatch date	16.73 ± 1.58	202	<2e-16	0.73 ± 0.08	196	<2e-16	14.75 ± 1.78	193	1.85e-14
Sex	-0.44 ± 1.35	202	0.744	0.55 ± 1.41	196	0.695	0.42 ± 1.53	193	0.786
<i>Clk</i> : Austria	-10.63 ± 4.42	202	0.017	-14.64 ± 4.77	196	0.002	-15.42 ± 4.95	193	0.002
<i>Clk</i> : Ireland	-12.35 ± 4.94	202	0.013	-16.21 ± 5.26	196	0.002	-18.93 ± 5.52	193	<0.001
<i>Clk</i> : Kazakhstan	-13.94 ± 4.58	202	0.003	-13.84 ± 4.91	196	0.005	-15.73 ± 5.11	193	0.002
Austria : hatch date	-6.35 ± 1.80	202	<0.001	-0.31 ± 0.09	196	<0.001	-7.92 ± 2.04	193	<0.001
Ireland : hatch date	-2.10 ± 2.46	202	0.394	-0.29 ± 0.13	196	0.023	-12.51 ± 2.82	193	1.57e-05
Kazakhstan : hatch date	7.58 ± 3.40	202	0.027	0.40 ± 0.17	196	0.021	7.31 ± 3.84	193	0.058
Autumn migratory restlessness									
Intercept	1.06 ± 0.56	110.80	0.058	0.78 ± 0.58	79.42	0.183	0.05 ± 0.53	138.76	0.925
<i>Clk</i> length	0.83 ± 0.46	112.08	0.074	1.09 ± 0.45	89.19	0.018	0.35 ± 0.42	173.25	0.406
Austria	-0.50 ± 0.54	111.25	0.358	-0.41 ± 0.54	88.99	0.442	-0.06 ± 0.49	172.69	0.907
Ireland	-1.03 ± 0.55	111.49	0.064	-1.48 ± 0.55	88.92	0.008	-0.96 ± 0.50	173.26	0.058
Kazakhstan	-2.13 ± 0.55	111.96	<0.001	-1.64 ± 0.54	89.12	0.003	0.06 ± 0.50	173.68	0.899
Sex	0.08 ± 0.13	112.75	0.540	-0.13 ± 0.14	91.10	0.336	-0.11 ± 0.13	174.34	0.365
<i>Clk</i> : Austria	-0.73 ± 0.47	111.89	0.119	-0.90 ± 0.46	89.14	0.054	-0.32 ± 0.43	173.23	0.460
<i>Clk</i> : Ireland	-0.68 ± 0.52	112.60	0.189	-1.05 ± 0.52	89.65	0.044	-0.48 ± 0.46	173.40	0.301
<i>Clk</i> : Kazakhstan	-0.75 ± 0.47	112.00	0.113	-0.81 ± 0.47	89.45	0.085	-0.06 ± 0.44	173.06	0.895
	Variance	SD		Variance	SD		Variance	SD	
Random effects									
Bird (intercept)	0	0		0	0		0	0	
Cal.year (intercept)	0.02	0.15		0.16	0.40		0.13	0.36	
Residual	0.42	0.65		0.40	0.64		0.69	0.83	
Spring migratory restlessness									
Intercept	1.36 ± 0.24	155.57	4.98e-08	1.24 ± 0.28	84.11	2.78e-05	0.90 ± 0.67	66.40	0.181
<i>Clk</i> length	0.54 ± 0.26	160.91	0.035	0.53 ± 0.38	84.32	0.165	0.50 ± 0.86	94.04	0.56
Austria	-1.58 ± 0.23	160.20	2.77e-10	-1.72 ± 0.28	83.84	3.66e-08	-1.02 ± 0.64	93.82	0.115
Ireland	-2.39 ± 0.24	161.36	<2e-16	-2.59 ± 0.29	83.39	1.52e-13	-1.63 ± 0.65	93.22	0.014
Kazakhstan	0.08 ± 0.24	160.62	0.748	-0.18 ± 0.29	83.74	0.528	-0.54 ± 0.65	93.15	0.411
Sex	0.13 ± 0.06	161.60	0.021	0.19 ± 0.09	71.06	0.049	-0.08 ± 0.19	78.20	0.653
<i>Clk</i> : Austria	-0.51 ± 0.26	161.03	0.050	-0.52 ± 0.39	83.93	0.179	-0.70 ± 0.87	93.43	0.423
<i>Clk</i> : Ireland	-0.59 ± 0.26	160.57	0.027	-0.59 ± 0.40	83.49	0.142	-0.76 ± 0.90	92.36	0.401
<i>Clk</i> : Kazakhstan	-0.64 ± 0.26	160.41	0.014	-0.55 ± 0.39	84.05	0.158	-0.40 ± 0.87	93.41	0.65
	Variance	SD		Variance	SD		Variance	SD	
Bird (intercept)	0	0		0.01	0.11		0.36	0.60	
Cal.year (intercept)	0.01	0.06		0	0		0.14	0.37	
Residual	0.12	0.34		0.15	0.39		0.48	0.69	

Table 2. Relationship between *Clk* repeat length and annual cycle timing. Estimates from nine linear mixed effects models are shown for the analysis of postjuvenile moult, and autumn and spring migratory restlessness. Reference population for the estimates is Kenyan, reference sex is male. Significant values are in bold.

explanation for the high variation observed in the stonechats could be that our study comprised populations that varied not only in breeding latitude, but also in their migratory behaviour.

Across stonechats, we found a latitudinal pattern in *Clk* gene diversity, whereby genetic variation was reduced in populations breeding at the equator. This is particularly interesting in a chronobiological context as the scope for light entrainment at the equator dramatically differs from seasonally varying light-dark cycles at higher latitudes. The two equatorial breeding populations from Kenya and Tanzania, which experience constant 12-h days throughout the year, differed from all other populations included here by significant reduction in *Clk* gene diversity. Previous research on African stonechats has revealed that life-cycle timing shows robust annual cycles,

both in free-living birds and under restrictive experimental conditions. Captive African stonechats held under constant 12-h light–dark cycles retained clear circannual rhythms of breeding condition and moult for several years^{58,59}. Thus, although African stonechats have retained the ability to respond to changing photoperiod⁴⁹, they appear to rely on an endogenous timing mechanism. Hence, the lower variability in *Clk* repeat length could be signature of selection facilitating adaptation to a constant photoperiod in the birds' environment resulting in a putatively optimised and less variable genotype^{58,59}.

Circannual rhythms of European stonechat populations that live permanently in the northern hemisphere⁶⁰ are less rigid, and the substantial changes in photoperiod they experience may play a dominant role for their annual time-keeping^{50,61}. Hence, greater variability in *Clk* repeat length might have resulted from fluctuating inter-annual selection for photoperiodic timing, for example, due to rough winters or sudden weather changes^{51,52}. Thus, we speculate that breeding populations in Germany, Austria, Spain and Ireland require integration of photoperiodism²⁰, and thereby a higher degree of genetic diversity in *Clk* genotype.

We consider it unlikely that the observed genotype–phenotype patterns for the focal *Clk* locus are instead caused by random drift. While similar *Clk* gene diversities in Austrian, German and Irish stonechat populations could be due to ongoing gene flow between these populations (e.g. resulting from geographic proximity and breeding dispersal^{54,62}), phylogeny cannot explain the similarly high *Clk* gene diversity we find in Kazakh stonechats, which diverged over 2.5 mya⁵⁵. Furthermore, the pattern in genetic diversity we observed for the focal *Clk* locus differs from a genome-wide characterisation across populations, which showed similar levels across all populations except for Kazakhstan, where levels were elevated (see Fig. 2; ⁵⁴). However, when we included genome-wide diversity of five of the nine populations as a covariate, we no longer observed a significant relationship between *Clk* gene diversity and latitude. We thus cannot exclude that the higher *Clk* gene diversity at higher latitudes could also be a result of higher genome-wide diversity. However, power for these results from only five populations was low, and results might be driven by the high genome-wide diversity in Kazakhstan⁵⁴. The other populations showed similarly low diversity values, although the Austrian and Irish populations breed at similar latitude as the Kazakhstan population. The Kazakhstan stonechat has a vast breeding range and high effective population size⁶³, which may have caused its outlying, elevated genome-wide diversity⁵⁴. Conversely, despite its low genome diversity, the Irish population shows high *Clk* gene diversity. It is tempting to speculate that the particular diversity in the *Clk* gene is associated with the partial migratory nature of the Irish population⁶². Among partial migrants, the resident fraction of a population experiences different photoperiods, and typically breeds earlier in the season, compared to the migrant fraction⁴⁷.

Our results on stonechats provide little support for effects of migration on *Clk* gene diversity, independently of breeding latitude. The only indicative effects of migratory phenotype are the differences in level of heterozygosity between migrants and residents (see HO in Table 1), where all migrant populations show at least twofold higher levels of heterozygosity compared to most residents (Table 1, Fig. 2). However, reduced levels of *Clk* gene diversity in the population from the Canary Islands are like explained by demographic history and although the Spanish population consists of residents, *Clk* gene diversity is comparably high as in populations of migrants. In addition, *Clk* gene diversity was significantly predicted only by breeding latitude, but not by migration distance.

In a cross-species comparative approach across different trans-Sahara migrants Bazzi et al.⁴⁶ hypothesise that selection mechanisms for longer repeats in species with small northern breeding ranges could restrict their postglacially acquired *Clk* diversity, while selection forces in species with larger breeding ranges should be weaker resulting in higher *Clk* diversity. They suggest that analyses of *Clk* gene variation in nonmigratory African relatives of Afro-Palaearctic migrants, such as included in our study, should provide additional insights for the hypothesised evolutionary scenario. Results in stonechats, however, are contrary to what Bazzi et al.⁴⁶ hypothesized: genetic diversity in African stonechat is significantly lower compared to migratory stonechat populations, and thus does not provide support for this hypothesis within the stonechat complex.

A latitudinal cline in *Clk* gene variation with longer repeats at higher latitudes has been demonstrated in various bird systems²⁸, suggesting a functional link between changes in daylength and *Clk* gene variation. However, our comparison of mean *Clk* length between populations showed no significant relationship with latitude, in contrast to earlier studies from different species of birds and fish^{23,28,46}. In our comparison, specifically long-distance migratory stonechats from the Kazakh population that breed at similar latitudes as European short-distance migrants, showed high frequencies of the longest *Clk* alleles we observed (15% of alleles are Q₁₅, Table 1). This observation points towards contributions of additional factors, such as environmental or climatic variation at the breeding or wintering area, or other characteristics of migratory routes, in correlative studies.

To separate genotype–phenotype associations between *Clk* allele length and annual timing from environmental influences, we compared captive individuals from four populations in a common garden, mimicking European daylength changes. These analyses revealed clear population-specific patterns that depended to lesser degree on time of year. *Clk* allele length in the equatorial stonechats from Kenya correlated positively with timing. Individuals delayed onset, peak and end of the three investigated life-cycle stages, moult and migratory restlessness in autumn and spring, by 9–33 days per additional poly-Q repeat. For migratory restlessness, timing of onset showed the closest relationship to *Clk* allele length. This fits with earlier findings from stonechats and other species showing that especially onset of this behaviour is under particularly strong, presumably genetic, regulation^{47,64}.

Overall, genotype–phenotype associations were retained across a broad range of photoperiods and in traits as different as seasonal nocturnality and plumage renewal. Although sample sizes varied between populations (range: 4–23 individuals), the consistent results suggest that *Clk* allele length is robustly associated with individual chronotypes in the equatorial population. In contrast, the three populations breeding at higher latitudes showed weaker and seasonally variable genotype–phenotype associations. Similar to Kenyan stonechats, high-latitude individuals delayed autumn migratory restlessness with increasing poly-Q repeat numbers, but patterns were inverse for moult, and varied for spring migratory restlessness.

Overall, our results fit well with evidence that circadian clock gene variants can also be relevant for coordination of annual timing, but their roles and effects depend on species, populations and time of year⁶⁵. More broadly, genetic contributions to annual cycle timing have been shown in several common garden studies, in quantitative genetic analyses and in breeding experiments^{43,61}. Heritability (h^2) estimates from captive migratory songbirds were medium to high for onset of migratory restlessness in blackcaps *Sylvia atricapilla* (0.34–0.45) and garden warblers *Sylvia borin* (0.67 onset for spring and autumn migration); heritability estimates for termination of migratory activity however were lower (0.16–0.44 in blackcaps^{66,67}; for a summary also see⁶⁸). Heritability estimates of migration timing traits from the wild are scarce, and estimates of repeatability and heritability are generally moderate, but sometimes lower, for example in Collared flycatchers (*Ficedula albicollis*)⁶⁵. It is thus clear that to a varying degree, flexibility and non-genetic factors, such as learning, state and ontogenetic factors, also contribute to individual variation⁶⁹. Inherited timing programs need to integrate information from multiple physiological pathways, such as metabolism and photic input⁷⁰. Hence, evolutionary change in annual cycle timing can involve several pathways including, but not limited to, genes with known circadian roles such as *Clk*⁷¹.

Previous findings from the captive stonechat populations used in this study indicate high individual consistency and high heritability estimates of annual timing traits in the *Saxicola* complex, which also differed between populations^{47,50,72}. Our new data suggest the possibility that the underlying genetic basis for individual timing might differ between populations from different latitudes. In equatorial populations, variation in annual chronotype might be partly regulated through a limited range of variants in the gene *Clk* that exert major effects, whereas at higher latitudes, *Clk* variants may mainly act in conjunction with photoperiodic pathways, and hence show weaker genotype–phenotype associations. Despite the lack of a mechanistic framework and reservations regarding the candidate genes approach, our results provide important and novel insight into understanding the possible genetic basis of annual-cycle timing.

In conclusion, we found a latitudinal cline in *Clk* gene diversity in a large dataset of 950 records from 717 individuals distributed over a wide geographical range, confirming findings in other species. Making use of a common garden setting our study highlights that the relationship between *Clk* polymorphism and annual-cycle timing in captive stonechats depended on population and on time of year. Our findings also allow us to speculate that in populations that live under unchanging photoperiods of the equator, *Clk* genotype may be less variable, but exert strong association with annual chronotype. Conversely, at higher latitude, other evolutionary forces may favour *Clk* gene diversity to be higher, but its association with phenotype may be obscured by additional molecular inputs into annual timing that furthermore depend on time of year.

Material and methods

Study populations. The breeding distribution of stonechats covers a wide geographical range (35° S–75° N; 20° W–180° E). Different populations exhibit a variety of migratory behaviours (Fig. 1) and differ geographically in the timing of annual processes such as breeding, moult and migratory activity^{47,50}. In captivity, differences in timing between free-living populations largely persist when the birds are kept under common-garden captive conditions⁵⁰, and even resident populations from equatorial Kenya display migratory restlessness, albeit at a low level^{47,50}. Combined, this makes stonechats an ideal system to study associations of *Clk* gene polymorphisms with different processes. Here, we capitalise on the following phenotypically and geographically distinct populations to study *Clk* gene polymorphism within one species complex (Table 1)^{54,55,73}:

We studied two African mainland populations [*Saxicola torquata axillaris* Shelley 1884: 135 individuals originating from Kenya (0°14' S, 36°0' E) and 52 individuals originating from Tanzania (3°5' S, 36°5' E)], which are year-round residents. We further included a resident island population (n = 61) endemic on Fuerteventura, Canary Islands, Spain (28°46' N, 14°31' W; *Saxicola dacotiae dacotiae* Meade-Waldo 1889; referred to as 'Canary' in the main text⁵⁶).

We further included four populations of European stonechat *Saxicola rubicola* ssp: (1) one population (n = 217) sampled in Austria (Neusiedel; 48°14' N, 16°22' E; *Saxicola rubicola rubicola* L. 1776), (2) one population (n = 47) from north-west Germany (Borken; 51°47' N, 6°01' E *Saxicola rubicola rubicola* L. 1776;⁷⁴), (3) one population (n = 184) from Ireland (Killarney; 52°N, 10° W; *Saxicola rubicola hibernans* Hartert 1910), and (4) one population (n = 93) from mainland Spain (Seville; 37°39' N, 5°34' W *Saxicola rubicola rubicola* L. 1776). Birds from Austria and Germany are obligatory medium-distance migrants that spend the winter in Mediterranean regions including North Africa. Irish stonechats are partial migrants that either stay on the breeding grounds year round or migrate over short distances^{62,75}. Spanish birds originate from a population of non-migratory stonechats (Serrano, D. unpublished data).

Lastly, we included two populations of long-distance migrants, Eurasian stonechat *Saxicola maurus* ssp: one population (n = 150) from Kazakhstan (Kustanaj; 51.5° N, 63° E; *Saxicola maurus maurus* Pallas 1773) and a second population (n = 11) from Japan (Hokkaido; 43.6° N, 141.23° E; *Saxicola maurus stejnegeri* L. 1766). Kazakh birds migrate to India, southern continental China and North East Africa^{54,76}, while Japanese birds migrate from Hokkaido over a continental path to winter in mainland southeast Asia⁷⁷.

Sample origin. We successfully genotyped 717 individuals from the nine study populations for variable *Clk* poly-Q repeat length. Most samples (n = 504), including birds from populations in Austria (n = 130), Ireland (n = 112), Kazakhstan (n = 90) and East Africa (Kenya and Tanzania; n = 172), originated from a common-garden experiment that Eberhard Gwinner initiated in 1981 at the Max Planck Institute for Ornithology in Andechs, Germany, and continued until 2005^{47,50,78}. These samples were thus collected from birds born over a long time span, including some individuals that were born in captivity, and others were brought in from the wild. African and European Stonechats hatched between 1982 and 2004, with similar spread over time, whereas Irish and Siberian stonechats hatched between 1997 and 2005. Additional genotype samples (n = 213) were collected

from birds sampled in the wild at their breeding grounds. Specifically, these included populations from the Canary Islands⁵⁶, Spain (D. Serrano, unpublished data), Germany⁷⁹ and Japan⁷⁷. These samples were collected between 1998 and 2017.

All methods were carried out in accordance with relevant guidelines and regulations, all experimental procedures were approved and conformed to the relevant regulatory standards under permit (Number: 55.2-1-54-2531-119-05) by the Ethics committee of the state of Upper Bavaria. The study was carried out in compliance with the ARRIVE guidelines (<https://arriveguidelines.org>).

Genetic analyses. Genomic DNA was isolated from blood (n=706) and feathers (n=11) using a salt extraction protocol and diluted to a working concentration of 25 ng/ μ l. Genomic DNA samples were genotyped for length polymorphism in the variable poly-Q repeat region of the *Clk* gene using a polymerase chain reaction (PCR) amplification protocol and lengths characterisation of the variable region. PCR amplification was carried out in a 10 μ l total volume using a previously published primer set Johnsen et al.²⁸ (forward primer: 5'-labelled with the 'blue' fluorescent dye 6-FAM 5'-6-FAM-TGGAGCGGTAATGGTACCAAGTA-3'; reverse primer: 5'-TCAGCTGTGACTGAGCTGGCT-3'). PCR conditions were optimised for stonechats following conditions published in Liedvogel et al.²⁹. Amplification reaction conditions for Taq DNA polymerase catalysed PCR for the stonechats: Mg²⁺ = 2 mM; 95 °C/2 min; 95 °C/30 s, 56.8 °C/30 s, 72 °C/30 s, 40 cycles; 72 °C/5 min, 4 °C hold. PCR products were prepared for capillary electrophoresis by adding HIDI-formamide-LIZ 500 mixture. Length polymorphism of amplified products were analysed using Geneious version 10.1.3. Repeat counts of lengths characterised samples was confirmed by Sanger sequencing of 50 samples across populations using an adapted version of a previous published protocol^{28,29} with following primers: forward: 5'-TTTTCTCAAGGTCAGCAGCTTGT-3'; reverse: 5'-CTGTAGGAAGTGTTCGGGTGCTG-3'). Amplification conditions using Taq DNA polymerase for Sanger sequencing were optimised for the stonechat material: Mg²⁺ = 2 mM; 95 °C/2 min; 95 °C/30 s, 61.1 °C/30 s, 74 °C/30 s, 40 cycles; 74 °C/5 min, 4 °C hold. Nucleotide sequences of with Exo/SAP purified PCR fragment were determined with BigDye Terminator ready reaction mix, version 3.1 (Applied Biosystems) under standard sequencing conditions according to the manufacturer's protocol²⁹.

Genotype characterisation and population genetic analyses. We characterised *Clk* genotype as mean allele length ($p + q/2$; as previously used in comparative studies^{26,28}) and did additional sensitivity tests using minimum and maximum allele length in our models. Results from these sensitivity tests were qualitatively the same, therefore we only present results from models with mean allele length in the main text. Genotype frequency data for all populations were tested statistically for deviation from Hardy–Weinberg equilibrium using GENEPOP 4.2 (web version; <http://genepop.curtin.edu.au/>; accessed 1 November 2020). We adapted to use the same Markov chain parameters that were used in previous studies^{28,29}: dememorization (10,000), batches (10,000) and iterations per batch (10 000). For each population we calculated observed heterozygosity as a measure of *Clk* allelic diversity as the proportion of observed number of heterozygotes of the total number of individuals. To account for the differences in sample size we used the program FSTAT version 2.9.3.2⁵³ to calculate F_{st}, an index of unbiased gene diversity. Results of gene diversity remained similar when we excluded captive-bred individuals from our dataset (see Supplementary Material Table S1).

To examine whether the *Clk* locus is under selection we compared genetic differentiation at the *Clk* locus with presumably neutral, genome wide nucleotide diversity. Levels of genome wide nucleotide diversity were calculated as autosomal nucleotide diversity (π) from publicly available whole genome re-sequencing data for five of the included stonechat populations (Kenya, Canary Islands, Austria, Ireland and Kazakhstan⁵⁴). To identify if differences in gene diversity were due to differences in breeding latitude or migratory phenotype, we ran linear models predicting gene diversity by covariables latitude, migratory distance and π .

Within individuals, we tested for an effect of latitude or migratory distance on *Clk* gene length by running linear mixed effects models with *Clk* mean repeat length as response variable, and either latitude or migratory distance as predictor variable as well as origin as a random effect.

Phenotypes of captive stonechats. Data on migratory phenotypes were previously presented for our four captive stonechat populations⁴⁷. Briefly, stonechats were hand-raised in the Max Planck Institute for Ornithology, as offspring of captive or wild stonechats belonging to the Kenyan, Austria, Irish or Kazakh populations (for detail see⁴⁷). The birds were kept indoors under simulated natural daylength changes in south Germany (47.5° N during the breeding season, and 40° N during winter⁵⁰). Here we include all birds with available phenotype and genotype data (migration: n = 142 for spring and n = 188 for autumn; moult: n = 215 for onset, n = 209 for peak and n = 206 for end).

Postjuvenile moult in stonechats is the first moult after hatching and equates to the first prebasic moult⁸⁰. A precise description of how moult was recorded was published previously⁷². In brief, immature stonechats were regularly checked for body moult in 19 plumage areas. The number of plumage areas moulting provided a moult score for each date to determine the timing of onset and completion of moult⁷².

We used nocturnal migratory restlessness behaviour as proxy for migratory timing. Migratory restlessness, the nocturnal activity of captive birds during migration periods, generally mirrors the migratory timing of free-living conspecifics⁸¹, although in stonechats and other songbird species, resident populations also show some migratory restlessness⁴⁷. A detailed description of how migratory restlessness of the stonechats was recorded and characterised is published elsewhere^{47,50}.

Analysis of genotype–phenotype association. To investigate correlation between *Clk* genotype and annual-cycle phenotype we ran linear mixed effects models including mean *Clk* repeat length, origin, sex and

a two-way interaction of *Clk* repeat length with origin as fix effects and year and individual as random effects in our comparisons with migratory timing. For our analysis of postjuvenile moult we added hatch date and a two-way interaction of origin and hatch date as we expected a potential effect of hatch date on moult timing as it had been demonstrated in a previous study⁷². Random effects were not used in this analysis, since we only had values from 1 year.

We tested the genotype effect on timing of postjuvenile moult, spring and autumn migration, as response variables, using three different time-points, i.e. onset, peak and end (given as Julian date). Analysis was performed in RStudio interface of R version 3.3.3^{82,83} using packages 'lme4'⁸⁴ and 'sjPlot'⁸⁵.

All trends remained unchanged in alternative models where one individual with a rare short poly-Q repeat length of Q_8 was initially excluded. However, Q_8 genotype for this individual (as classified by size fragment determination) was confirmed by Sanger sequencing, giving us no reason to exclude it from the analysis.

Received: 30 October 2021; Accepted: 19 April 2022

Published online: 13 May 2022

References

- Kronfeld-Schor, N. & Dayan, T. Partitioning of time as an ecological resource. *Annu. Rev. Ecol. Evol. Syst.* **34**, 153–181 (2003).
- Tauber, E. & Kyriacou, C. P. Review: Genomic approaches for studying biological clocks. *Funct. Ecol.* **22**, 19–29 (2008).
- White, E. R. & Hastings, A. Seasonality in ecology: Progress and prospects in theory. *Ecol. Complex.* **44**, 100867 (2020).
- Ko, C. H. & Takahashi, J. S. Molecular components of the mammalian circadian clock. *Hum. Mol. Genet.* **15**, R271–R277 (2006).
- Cassone, V. M. Avian circadian organization: A chorus of clocks. *Front. Neuroendocrinol.* **35**, 76–88 (2014).
- Kyriacou, C. P., Peixoto, A. A., Sandrelli, F., Costa, R. & Tauber, E. Clines in clock genes: Fine-tuning circadian rhythms to the environment. *Trends Genet.* **24**, 124–132 (2008).
- Partch, C. L., Green, C. B. & Takahashi, J. S. Molecular architecture of the mammalian circadian clock. *Trends Cell Biol.* **24**, 90–99 (2014).
- Helm, B. *et al.* Two sides of a coin: ecological and chronobiological perspectives of timing in the wild. *Philos. Trans. R. Soc. B Biol. Sci.* **372**, 20160246 (2017).
- Kalmbach, D. A. *et al.* Genetic basis of chronotype in humans: Insights from three landmark GWAS. *Sleep* <https://doi.org/10.1093/sleep/zsw048> (2017).
- Takahashi, J. S., Shimomura, K. & Kumar, V. Searching for genes underlying behavior: Lessons from circadian rhythms. *Science* **322**, 909–912 (2008).
- Yoshimura, T. *et al.* Molecular analysis of avian circadian clock genes. Published on the World Wide Web on 23 May 2000. *Mol. Brain Res.* **78**, 207–215 (2000).
- Gekakis, N. *et al.* Role of the CLOCK Protein in the Mammalian circadian mechanism. *Science* **280**, 1564–1569 (1998).
- Saleem, Q., Anand, A., Jain, S. & Brahmachari, S. K. The polyglutamine motif is highly conserved at the Clock locus in various organisms and is not polymorphic in humans. *Hum. Genet.* **109**, 136–142 (2001).
- Darlington, T. K. *et al.* Closing the circadian loop: CLOCK-induced transcription of its own inhibitors *per* and *tim*. *Science* **280**, 1599–1603 (1998).
- King, D. P. *et al.* Positional cloning of the mouse circadian clock gene. *Cell* **89**, 641–653 (1997).
- Follett, B. Rhythms and photoperiodism in birds. *Biological rhythms and photoperiodism in plants* (1998).
- Hazlerigg, D. G. & Wagner, G. C. Seasonal photoperiodism in vertebrates: from coincidence to amplitude. *Trends Endocrinol. Metab.* **17**, 83–91 (2006).
- Gwinner, E. Circadian and circannual programmes in avian migration. *J. Exp. Biol.* **199**, 39–48 (1996).
- Stirland, J. A., Mohammad, Y. N. & Loudon, A. S. I. A mutation of the circadian timing system (*tau* gene) in the seasonally breeding Syrian hamster alters the reproductive response to photoperiod change. *Proc. R. Soc. London Ser. B Biol. Sci.* **263**, 345–350 (1996).
- Bradshaw, W. E. & Holzapfel, C. M. Evolution of animal photoperiodism. *Annu. Rev. Ecol. Evol. Syst.* **38**, 1–25 (2007).
- Graham, J. L., Cook, N. J., Needham, K. B., Hau, M. & Greives, T. J. Early to rise, early to breed: A role for daily rhythms in seasonal reproduction. *Behav. Ecol.* **28**, 1266–1271 (2017).
- Rittenhouse, J. L., Robart, A. R. & Watts, H. E. Variation in chronotype is associated with migratory timing in a songbird. *Biol. Lett.* **15**, 20190453 (2019).
- O'Malley, K. G., Ford, M. J. & Hard, J. J. Clock polymorphism in Pacific salmon: Evidence for variable selection along a latitudinal gradient. *Proc. R. Soc. B Biol. Sci.* **277**, 3703–3714 (2010).
- O'Malley, K. G. & Banks, M. A. A latitudinal cline in the Chinook salmon (*Oncorhynchus tshawytscha*) Clock gene: Evidence for selection on PolyQ length variants. *Proc. R. Soc. B Biol. Sci.* **275**, 2813–2821 (2008).
- Peterson, M. P. *et al.* Variation in candidate genes CLOCK and ADCYAP1 does not consistently predict differences in migratory behavior in the songbird genus *Junco*. *PLoS ONE* **8**, 115 (2013).
- Saino, N. *et al.* Polymorphism at the Clock gene predicts phenology of long-distance migration in birds. *Mol. Ecol.* **24**, 1758–1773 (2015).
- Saino, N. *et al.* Timing of molt of barn swallows is delayed in a rare Clock genotype. *PeerJ* **1**, e17 (2013).
- Johnsen, A. *et al.* Avian Clock gene polymorphism: Evidence for a latitudinal cline in allele frequencies. *Mol. Ecol.* **16**, 4867–4880 (2007).
- Liedvogel, M., Szulkin, M., Knowles, S. C. L., Wood, M. & Sheldon, B. C. Phenotypic correlates of Clock gene variation in a wild blue tit population: Evidence for a role in seasonal timing of reproduction. *Mol. Ecol.* **18**, 2444–2456 (2009).
- Caprioli, M. *et al.* Clock gene variation is associated with breeding phenology and maybe under directional selection in the migratory barn swallow. *PLoS ONE* **7**, e35140 (2012).
- Dor, R. *et al.* Clock gene variation in *Tachycineta swallows*. *Ecol. Evol.* **2**, 95–105 (2012).
- Dor, R. *et al.* Low variation in the polymorphic Clock gene poly-Q region despite population genetic structure across barn swallow (*Hirundo rustica*) populations. *PLoS ONE* **6**, e28843 (2011).
- O'Brien, C. *et al.* Geography of the circadian gene clock and photoperiodic response in western North American populations of the three-spined stickleback *Gasterosteus aculeatus*. *J. Fish Biol.* **82**, 827–839 (2013).
- Mueller, J. C., Pulido, F. & Kempenaers, B. Identification of a gene associated with avian migratory behaviour. *Proc. R. Soc. B Biol. Sci.* **278**, 2848–2856 (2011).
- Liedvogel, M. & Sheldon, B. C. Low variability and absence of phenotypic correlates of Clock gene variation in a great tit *Parus major* population. *J. Avian Biol.* **41**, 543–550 (2010).
- Lugo-Ramos, J. S., Delmore, K. E. & Liedvogel, M. Candidate genes for migration do not distinguish migratory and non-migratory birds. *J. Comp. Physiol. A* **203**, 383–397 (2017).

37. Majoy, S. B. & Heideman, P. D. Tau differences between short-day responsive and short-day nonresponsive white-footed mice (*Peromyscus leucopus*) do not affect reproductive photoreponsiveness. *J. Biol. Rhythms* **15**, 501–513 (2000).
38. O'Brien, C. *et al.* Geography of the circadian gene clock and photoperiodic response in western North American populations of the threespine stickleback *Gasterosteus aculeatus*. *J. Fish Biol.* **82**, 827–839 (2013).
39. Contina, A., Bridge, E. S., Ross, J. D., Shipley, J. R. & Kelly, J. F. Examination of clock and *Adcyap1* gene variation in a neotropical migratory passerine. *PLoS ONE* **13**, e0190859 (2018).
40. Herzog, E. D. Neurons and networks in daily rhythms. *Nat. Rev. Neurosci.* **8**, 790–802 (2007).
41. Chahad-Ehlers, S. *et al.* Expanding the view of clock and cycle gene evolution in Diptera. *Insect Mol. Biol.* **26**, 317–331 (2017).
42. Denlinger, D. L., Hahn, D. A., Merlin, C., Holzapfel, C. M. & Bradshaw, W. E. Keeping time without a spine: What can the insect clock teach us about seasonal adaptation?. *Philos. Trans. R. Soc. B Biol. Sci.* **372**, 20160257 (2017).
43. van Noordwijk, A. J. *et al.* A framework for the study of genetic variation in migratory behaviour. *J. Ornithol.* **147**, 221–233 (2006).
44. Newton, I. *The Migration Ecology of Birds* (Academic Press, 2008).
45. Gohli, J., Liffield, J. T. & Albrecht, T. Migration distance is positively associated with sex-linked genetic diversity in passerine birds. *Ethol. Ecol. Evol.* **28**, 42–52 (2016).
46. Bazzi, G. *et al.* Clock gene polymorphism, migratory behaviour and geographic distribution: A comparative study of trans-Saharan migratory birds. *Mol. Ecol.* **25**, 6077–6091 (2016).
47. Doren, B. M. V., Liedvogel, M. & Helm, B. Programmed and flexible: Long-term Zugunruhe data highlight the many axes of variation in avian migratory behaviour. *J. Avian Biol.* **48**, 155–172 (2017).
48. Helm, B., Gwinner, E. & Trost, L. Flexible seasonal timing and migratory behavior: Results from stonechat breeding programs. *Ann. N. Y. Acad. Sci.* **1046**, 216–227 (2005).
49. Helm, B. & Gwinner, E. Migratory restlessness in an equatorial nonmigratory bird. *PLoS Biol.* **4**, e110 (2006).
50. Helm, B. Geographically distinct reproductive schedules in a changing world: Costly implications in captive Stonechats. *Integr. Comp. Biol.* **49**, 563–579 (2009).
51. Dhondt, A. A. Variations in the number of overwintering stonechats possibly caused by natural selection. *Ring. Migr.* **4**, 155–158 (1983).
52. Brown, C. R. & Brown, M. B. Weather-mediated natural selection on arrival time in cliff swallows (*Petrochelidon pyrrhonota*). *Behav. Ecol. Sociobiol.* **47**, 339–345 (2000).
53. GOUDET, J. FSTAT, a program to estimate and test gene diversities and fixation indices, version 2.9.3. <http://www2.unil.ch/popgen/softwares/fstat.htm> (2001).
54. Van Doren, B. M. *et al.* Correlated patterns of genetic diversity and differentiation across an avian family. *Mol. Ecol.* **26**, 3982–3997 (2017).
55. Illera, J. C., Richardson, D. S., Helm, B., Atienza, J. C. & Emerson, B. C. Phylogenetic relationships, biogeography and speciation in the avian genus *Saxicola*. *Mol. Phylogenet. Evol.* **48**, 1145–1154 (2008).
56. Illera, J. C. & Díaz, M. Reproduction in an endemic bird of a semiarid island: A food-mediated process. *J. Avian Biol.* **37**, 447–456 (2006).
57. Illera, J. C. & Díaz, M. Site fidelity in the Canary Islands stonechat *Saxicola dacotiae* in relation to spatial and temporal patterns of habitat suitability. *Acta Oecol.* **34**, 1–8 (2008).
58. Gwinner, E. & Dittami, J. Endogenous reproductive rhythms in a tropical bird. *Science* **249**, 906–908 (1990).
59. Dittami, J. & Gwinner, E. Annual cycles in the African stonechat *Saxicola torquata* axillaris and their relationship to environmental factors. *J. Zool.* **207**, 357–370 (1985).
60. Gwinner, E. Circannual rhythms in tropical and temperate-zone stonechats: A comparison of properties under constant conditions. *Ökologie der Vögel* **13**, 5–14 (1991).
61. Gwinner, E. *Circannual Rhythms: Endogenous Annual Clocks in the Organization of Seasonal Processes* (Springer, 2012).
62. Helm, B., Fiedler, W. & Callion, J. Movements of European stonechats *Saxicola torquata* according to ringing recoveries. *ARDEA-WAGENINGEN* **94**, 33 (2006).
63. Opaev, A., Red'kin, Y., Kalinin, E. & Golovina, M. Species limits in Northern Eurasian taxa of the common stonechats, *Saxicola torquatus* complex (Aves: Passeriformes, Muscicapidae). *Vertebrate Zool.* **68**, 199 (2018).
64. Gwinner, E. & Czeschlik, D. On the significance of spring migratory restlessness in caged birds. *Oikos* **30**, 364–372 (1978).
65. Krist, M., Munclinger, P., Briedis, M. & Adamik, P. The genetic regulation of avian migration timing: combining candidate genes and quantitative genetic approaches in a long-distance migrant. *Oecologia* <https://doi.org/10.1007/s00442-021-04930-x> (2021).
66. Berthold, P. & Pulido, F. Heritability of migratory activity in a natural bird population. *Proc. R. Soc. London Ser. B Biol. Sci.* **257**, 311–315 (1994).
67. Pulido, F. & Berthold, P. Current selection for lower migratory activity will drive the evolution of residency in a migratory bird population. *Proc. Natl. Acad. Sci.* **107**, 7341–7346 (2010).
68. Liedvogel, M. & Lundberg, M. The Genetics of Migration. In *Animal Movement Across Scales* (eds Hansson, L.-A. & Åkesson, S.) 219–231 (Oxford University Press, 2014). <https://doi.org/10.1093/acprof:oso/9780199677184.003.0012>.
69. Åkesson, S. & Helm, B. Endogenous programs and flexibility in bird migration. *Front. Ecol. Evol.* **8**, 78 (2020).
70. Stevenson, T. J. & Kumar, V. Neural control of daily and seasonal timing of songbird migration. *J. Comp. Physiol. A* **203**, 399–409 (2017).
71. Verhagen, I. *et al.* Genetic and phenotypic responses to genomic selection for timing of breeding in a wild songbird. *Funct. Ecol.* **33**, 1708–1721 (2019).
72. Helm, B. & Gwinner, E. Timing of Postjuvenile molt in African (*Saxicola Torquata Axillaris*) and European (*Saxicola Torquata Rubicola*) stonechats: Effects of genetic and environmental factors. *Auk* **116**, 589–603 (1999).
73. Zink, R. M., Pavlova, A., Drovetski, S., Wink, M. & Rohwer, S. Taxonomic status and evolutionary history of the *Saxicola torquata* complex. *Mol. Phylogenet. Evol.* **52**, 769–773 (2009).
74. Flinks, H. & Pfeifer, F. Brutzeit, Gelegegröße und Bruterfolg beim Schwarzkehlchen (*Saxicola torquata*). *Charadrius* **23**, 128–140 (1987).
75. Urquhart, E. *Stonechats* (Christopher Helm, 2002).
76. Glutz von Blotzheim, U. *Bauer Handbuch der Vögel Mitteleuropas KM: Bd. 11. Aula, Wiesbaden* (1988).
77. Yamaura, Y. *et al.* Tracking the Stejneger's stonechat *Saxicola stejnegeri* along the East Asian-Australian Flyway from Japan via China to southeast Asia. *J. Avian Biol.* **48**, 197–202 (2017).
78. Gwinner, E., Neusser, V., Engl, D., Schmid, D. & Bals, L. Haltung, Zucht und Eiaufzucht afrikanischer und europäischer Schwarzkehlchen *Saxicola torquata*. *Gefiederte Welt* **111**, 118–120 (1987).
79. Flinks, H., Helm, B. & Rothery, P. Plasticity of moult and breeding schedules in migratory European Stonechats *Saxicola rubicola*. *Ibis* **150**, 687–697 (2008).
80. Humphrey, P. S. & Parkes, K. C. An approach to the study of molts and plumages. *Auk* **76**, 1–31 (1959).
81. Berthold, P. *Bird Migration: A General Survey* (Oxford University Press, 2001).
82. RStudio | Open source & professional software for data science teams. <https://rstudio.com/>.
83. R Core Team. *R: A language and environment for statistical computing*. (R Foundation for Statistical Computing, 2013).
84. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. <http://arxiv.org/abs/1406.5823> (2014).

85. Lüdecke, D. & Lüdecke, M. D. Package 'sjPlot'. (2015).

86. del Hoyo, J., Elliott, A., Sargatal, J., Christie, D. A. & de Juana, E. *Handbook of the Birds of the World Alive* (Lynx Edicions, 2018).

Acknowledgements

We thank the Max Planck Society (MPRG grant MFFALIMN0001 to ML), the DFG (Project Z02 within SFB 1372—Magnetoreception and Navigation in Vertebrates to ML), and Eberhard Gwinner, whose previous work on the stonechats and the avian clock forms the basis for this study. Sylvia Kuhn and Andrew Fidler have inspired this work by their piloting investigations. Additional thanks to Benjamin van Doren for helpful discussions concerning the incorporation of the phenotype data, Eva Stukenbrock for comments on an earlier version of this manuscript and all members of the Max Planck Research Group Behavioural Genomics for their comments and support.

Author contributions

B.H. and M.L. conceived and planned the study. Molecular work was carried out by H.J. and T.H., statistical analyses by H.J. with input from M.L., B.H., K.E.D. B.H., J.C.I., D.S., H.F., M.S., K.K. contributed sample material. H.J., M.L. and B.H. wrote the manuscript, all co-authors commented and agreed on the final draft.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-11158-z>.

Correspondence and requests for materials should be addressed to H.J. or M.L.

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

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



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

© The Author(s) 2022