UNIVERSITÀ DEGLI STUDI DI MILANO

SCUOLA DI DOTTORATO
TERRA, AMBIENTE E BIODIVERSITÀ
Dottorato di Ricerca in Biologia animale
Ciclo XXVIII

Genes on the move: candidate genes and long-distance migration in birds

GAIA BAZZI
R10085

Tutor:
Prof. DIEGO RUBOLINI

Anno accademico 2015/2016

Photo: Andrea Galimberti

CONTENTS

1. Introduction

1.1. Endogenous control of migration and other life-history events

The Earth undergoes circannual seasonal changes in ecological conditions due to the inclination of its rotation axis with respect to the ecliptic plane. Seasonality is expressed through oscillations in temperatures, precipitations, photoperiod and availability of food resources; such oscillations are more obvious at high latitudes, becoming gradually less noticeable towards the tropics. Organisms inhabiting seasonal environments must face periods in which food resources are predictably scant or absent, and weather conditions are adverse. Migration is an adaptive strategy, shared by several animal taxa, evolved to exploit the seasonal changes of resources by moving to different areas at different times of the year, as these become suitable (Gauthreaux 1982; Rappole 1995; Berthold 1996).

Bird migration is impressively widespread, involving every year billions of individuals of thousands of species (Berthold 2001; Hahn *et al.* 2009), and is accomplished by means of a set of morphological, physiological and behavioural adaptations that allow birds to successfully move from breeding to wintering grounds and back at the correct time of the year (Newton 2008). Migratory birds, indeed, reach the breeding grounds when the food supplies enable them to breed (Lack 1954), and leave towards more suitable areas before the ecological conditions deteriorate (Newton 2008). Migratory species, especially those travelling over long distances, must anticipate the changes in ecological conditions by weeks or even months to prepare for migration, for example by reaching the critical fuel load required for successfully completing their long journey across ecological barriers, and usually have very limited (if any) cues about the ecological conditions at their goal areas (Gwinner 1972; Hagan *et al.* 1991; Newton 2008; Saino & Ambrosini 2008). Hence, the annual cycle is timed to ensure that the birds' major life-history events, i.e. reproduction, migration and moult, occur in the proper sequence and at the right time of the year (Newton 2008).

This is allowed by the existence of an endogenous control of the annual schedule, that has been well documented in several migratory bird species by means of experimental studies on wild birds kept under controlled laboratory conditions (Gwinner 1968, 1971, 1972, 1981, 1986; Berthold *et al.* 1974; Berthold & Terrill 1991).

Relying entirely on internal rhythms, however, implies that any slight shift of the annual program would uncouple the life-history events from the prevailing ecological conditions (Newton 2008). Hence, the endogenous cycle is kept in phase with seasonal variation and interannual differences in ecological conditions by means of external cues, that act as time-keepers (*Zeitgeber*) (Merrow *et al.* 2005; Newton 2008). Due to its high reliability, variation in photoperiod has been suggested among the primary of these cues (e.g. Gwinner 1986; Gwinner & Helm 2003; Sharp 2005). Importantly, since both day length and daily differences in photoperiod broadly vary along a south-north axis, photoperiod also provides cues about the latitude to migratory birds that travel between distant areas, experiencing different photoperiodic regimes. Hence, photoperiod seems to be involved not only in the mechanism that control the timing of circannual events, but also in the setting of the whole spatio-temporal program of migration (Newton 2008).

1.2. The genetics of migratory traits

The occurrence of a tight endogenous control of migration hint that migratory traits are under a strong genetic influence (reviewed in Liedvogel & Lundberg 2014). Findings from heritability studies, crossbreeding and displacement experiments (e.g. Berthold 1984; Møller 2001; Pulido *et al.* 2001; Pulido & Bethold 2003; Thorup & Rabøl 2007), and as well as from studies of the response of wild populations to changing selective pressures (e.g. Brown & Brown 2000) corroborate this idea.

Migratory species and populations could broadly differ in the extent of migratory activity and in the expression of migratory traits (e.g. Berthold 1988; Pulido 2007). At increasing latitudes the

breeding season becomes shorter and delayed, and the migration lengthens in term of distance travelled and time taken to be accomplished; as a consequence, birds breeding at different latitudes must adaptively adjust their phenology and their response to day length to the latitude at which they nest (Newton 2008). Even within populations individuals could show large differences in traits linked to migration, and such variability may be maintained by different selective pressures that occur in different years (e.g. partial migration, Lundberg 1987). It has been shown that such differences in migratory traits among species, populations and individuals are at least partly triggered by genetic variation (Pulido & Berthold 2003; Pulido 2007; van Noordwijk *et al.* 2006).

Identifying which genes or gene groups underline the phenotypic variation observed in natural populations, and to which extent, is of broad evolutionary and conservation interest (e.g. Ellegren & Sheldon 2008; Møller *et al.* 2010), though for decades the topic has received less attention (Liedvogel & Lundberg 2014). The few studies trying to link phenotypic variability of migratory traits with the overall genetic differentiation in bird, fish and insect migratory populations have disclosed mosly weak associations (Liedvogel et al. 2011). Hence, it has been suggested that variation in migratory traits among individuals could be the outcome of polymorphism at relatively few regions or loci with additive effects on downstream signalling cascades of a variety other genes (Liedvogel et al. 2011). In recent years, several studies have focused on the polymorphism of phenological candidate genes (i.e. genes involved in the control of the circadian clock), that may explain the variability of timing of circannual events and other behavioural traits in birds and other vertebrates (e.g. birds: Liedvogel *et al.* 2009; Caprioli *et al.* 2012; Mueller *et al.* 2011; Chakarov *et al.* 2013; Peterson *et al.* 2013; Saino *et al.* 2013; Bourret & Garant 2015. Salmonid species: O'Malley & Banks 2008; O'Malley *et al.* 2010). Candidate genes, indeed, show short tandem repeats in functional or neutral regions that could affect the gene functions and ultimately may influence the photoperiodic response (Kashi *et al.* 1997; Comings 1998; Li *et al.* 2004; Fondon *et al.* 2008; Visser *et al.* 2010). The candidate gene approach relies on *a priori* knowledge of

functional genes which contribute to shape the phenotype of a trait of interest in model organisms, or that were identified in previous biochemical studies (Tabor *et al.* 2002; Fitzpatrick *et al.* 2005; Hoffmann & Willi 2008; Pardo-Diaz *et al.* 2015). Since several gene structures and functions are conserved among divergent taxa, such an approach could be useful to investigate the association between gene polymorphism and phenotypic variation in natural populations of non-model organisms (Fitzpatrick *et al.* 2005). Other approaches implemented to identify the genomic regions underlying adaptive phenotypic variability include the 'forward genetics', which searches for gene regions or genes underpinning known adaptive traits by surveying a large number of molecular markers in individuals segregating for a trait of interest (reviewed in Pardo-Diaz *et al.* 2015), and the 'reverse genetics', that refers to the genome-wide detection of loci with evidence of selection without *a priori* knowledge of their associated phenotype (Pardo-Diaz *et al.* 2015).

1.3. *Clock* gene polymorphism and phenology of circannual events

Most of the phenological candidate genes studied so far are involved in the 'core circadian oscillator' (CCO), a fully characterized, auto-regulated negative feedback loop that modulates the photoperiodic response and that is responsible for the onset and setting of circadian and circannual rhythms (Panda *et al.* 2002; Lincoln *et al.* 2003; Bell-Pedersen *et al.* 2005; Ko & Takahashi 2006). The *Clock* (*Circadian Locomotor Output Cycles Kaput*) gene products play a central role within the CCO: CLOCK protein heterodimerizes with a second protein, BMAL1, forming the transcriptional-activating complex of the CCO itself (Gekakis *et al.* 1998; Young & Kay 2001; Panda *et al.* 2002; Ko & Takahashi 2006); in addition, the CLOCK-BMAL1 complex acts as an output of the CCO by activating the transcription of downstream clock-controlled genes (Reppert & Weaver 2002; Iuvone *et al.* 2005). The avian *Clock* exonic region shows a C-terminal polyglutamine (Poly-Q) repeat sequence that could vary in length depending on the number of CAG repeats it bears, and that affects the binding affinity of CLOCK protein with its transmission factor (Darlington *et al.* 1998).

It has been suggested that polymorphism at Poly-Q, reported both among populations and among individuals within populations (e.g. Fidler & Gwinner 2003; Johnsen *et al.* 2007; Liedvogel *et al.* 2009; Caprioli *et al.* 2012; Peterson *et al.* 2013), may contribute to phenotypical variability in the timing of life-history events (among which migration) by differentially affecting circadian rhythms (Darlington *et al.* 1998; Gekakis *et al.* 1998; Hayasaka *et al.* 2002).

A first evidence of the possible role of *Clock* gene in the control of phenology come from an among-populations comparison of *Clock* allele size in the blue tit (*Cyanistes caeruleus*). The study revealed a latitudinal cline in the frequency of 'longer' and 'shorter' alleles (i.e. alleles carrying more and less CAG repeats, respectively), with Poly-Q length increasing at northern latitudes (Johnsen *et al.* 2007). It has been proposed that such a latitudinal cline may arise from an adaptation to local regimes of annual photoperiodic oscillations that birds experience at their breeding areas (Johnsen *et al.* 2007; O'Malley & Banks 2008; O'Malley *et al.* 2010). Moreover, as the breeding season becomes delayed and shorter moving polewards, the occurrence of Poly-Q of different lengths at different latitudes could be related to the timing of birds' life-history events, longer alleles being associated with delayed phenology. The latter hypothesis seems to be supported by the findings of several studies investigating the association between *Clock* allele size and phenology of life-history events within populations. Indeed, intrapopulation variation in *Clock* allele size predicted the timing of reproduction and moult in some avian species (Liedvogel *et al.* 2009; Caprioli *et al.* 2012; Saino *et al.* 2013; Bourret & Garant 2015). The direction of the *Clock* allele size-phenology association was coherent in all the studies that detected such relationship: individuals bearing longer *Clock* alleles showed a delayed timing, whereas those carrying shorter alleles had an advanced phenology. Specifically, a study of the blue tit reported that females with shorter *Clock* alleles started laying earlier and had shorter incubation periods than their counterparts bearing longer alleles (Liedvogel *et al.* 2009). Similar findings come from studies of two swallow species: in the barn swallow (*Hirundo rustica*) a rare, long allele was associated with delayed

breeding phenology (Caprioli *et al.* 2012); a study of a breeding tree swallow (*Tachycineta bicolor*) population showed a similar pattern of association, although the relationship was statistically significant only in combination with breeding density (Bourret & Garant 2015). Furthermore, the single study investigating the association between *Clock* polymorphism and moult showed that the timing of moult had a tendency to increase with *Clock* allele size in wintering barn swallows (Saino *et al.* 2013). Finally, findings from further research on candidate genes polymorphism suggested an involvement of *Clock* also in the control of behavioural traits linked to migration. Indeed, *Clock* allele size increased with migration distance within two subspecies groups belonging to the genus *Junco*, although not among populations (Peterson *et al.* 2013).

Conversely, studies of other bird species have failed to disclose any association between *Clock* genotype and breeding latitude or phenotype (Liedvogel & Sheldon 2010; Dor *et al.* 2011; Dor *et al.* 2012; Chakarov *et al.* 2013; Kuhn et al. 2013) and the role of *Clock* polymorphism in shaping avian phenotype variability still needs to be clarified.

1.4. *Adcyap1* allele size and variability of migratory behaviour

Besides *Clock*, other phenological candidate genes have been studied in wild birds to investigate their possible role in the control of circannual activities (e.g. Steinmeyer *et al.* 2009; Mueller *et al.* 2011; Chakarov *et al.* 2013; Peterson *et al.* 2013; Bourret & Garant 2015). *Adcyap1* (*Adenylate Cyclase-Activating Polypeptide 1*) is a second candidate gene whose polymorphism has been suggested to underline variability in behavioural traits (Mueller *et al.* 2011; Peterson *et al.* 2013; Chakarov *et al.* 2013; Bourret & Garant 2015). *Adcyap1* encodes PACAP (pituitary adenylate cyclase-activating polypeptide), a neuropeptide known to exert a plethora of biological functions (Nowak & Zawilska 2003; Vaudry *et al.* 2009; Olano-Marin *et al.* 2011). There are several indications of a possible role of PACAP in the control of some of the physiological and behavioural shifts associated with migration in avian species (Mueller *et al.* 2011). First, it has been suggested

that a determinant of migration onset in night-migratory birds (Fusani & Gwinner 2005), the phase-shift of the endogenous oscillators from day to night activity through melatonin modulation (Mueller *et al.* 2011), could be driven by PACAP expression, since PACAP stimulates the melatonin synthesis in the pineal gland (Simonneaux *et al.* 1993; Schwartz & Andrews 2013). Furthermore, PACAP seems to be involved in the photoperiodic control of the annual schedule in several ways (Schwartz & Andrews 2013): it modulates the expression of the CCO, by directly activating *Clock* and other circadian genes in the pineal gland (Nagy & Csernus 2007; Racz *et al.* 2008) and it conveys light information from the retina to the suprachiasmatic nucleus of hypothalamus (Hannibal *et al.* 1997), a key element of the circadian timing (Schwartz & Andrews 2013). Finally, findings from experimental manipulation of PACAP levels in avian brain are suggestive for a role of the neuropeptide in driving migration-related physiological changes, such as fuel shift to lipid metabolism, premigratory hyperphagia and migratory fasting (Schwartz & Andrews 2013). Indeed, administration of PACAP in the ventricular system of chicken brain influenced energy metabolism by increasing metabolic rate, body temperature and lipid utilization (Tachibana et al. 2007), whereas increased PACAP concentration in the brain inhibited the feeding behaviour (Tachibana et al. 2003).

The 3' UTR of *Adcyap1* gene, that consist of post-transcriptional regulatory elements, shows length polymorphism due to simple sequence repeat insertion (Steinmeyer *et al.* 2009). It has been proposed that microsatellite polymorphism at the 3' UTR may influence the post-transcriptional processes by modifying the structure of the region (Hirokawa & Takemura 2005; Bartel 2009; Chatterjee & Pal 2009). At the among-population level, *Adcyap1* polymorphism predicted the migration propensity and the migration distance in the blackcap (*Sylvia atricapilla*), with *Adcyap1* allele size increasing with migratory activity (Mueller *et al.* 2011). Nevertheless, no association between *Adcyap1* allele size and migration distance has been found among populations of the genus *Junco* (Peterson *et al.* 2013). In addition, it has been shown that *Adcyap1* allele size affected

migratory restlessness in blackcap (*Sylvia atricapilla*) (Mueller *et al.* 2011) and Oregon junco (*Junco hyemalis thurberi*) (Peterson *et al.* 2013) individuals kept in cage, longer *Adcyap1* alleles being associated with more persistent migratory restlessness (Mueller *et al.* 2011; Peterson *et al.* 2013). Since that migratory restlessness reflects migration propensity in wild populations and ultimately affects the distances travelled by migratory birds (Gwinner 1990; Berthold 1996; Maggini & Bairlein 2010), on the whole these findings foster the idea that longer *Adcyap1* alleles should be associated with longer migratory journeys.

A further cue of an involvement of *Adcyap1* polymorphism in shaping avian behavior come from three studies of the timing of life-history events in association with candidate genes polymorphism. Juvenile buzzards (*Buteo buteo*) bearing longer *Adcyap1* alleles dispersed earlier compared to those carrying shorter alleles (Chakarov *et al.* 2013). Furthermore, in female tree swallows *Adcyap1* allele size affected the timing of laying, although the association varied with latitude. Finally, *Adcyap1* allele size predicted timing of spring migration in female blackcaps migrating across Europe, although the genotype-phenology association emerged in combination with wing morphology only (Mettler *et al.* 2015). Such relationships could be partly mediated by an effect of *Adcyap1* polymorphism on migratory traits. Chakarov *et al.* (2013), indeed, suggested that *Adcyap1* genotype could affect dispersal in buzzards through an effect on migratory restlessness, whereas the latitudinal variation of the *Adcyap1* allele size-breeding phenology detected in the tree swallow could reflect spatial variation in migratory traits driven by *Adcyap1* variability (Bourret & Garant 2015).

2. Outline of the thesis

The findings described above are suggestive for a role of phenological candidate genes polymorphism in driving the variability in migratory behaviour both among populations and among individuals within populations, although the effect of genotype was somewhat inconsistent across the different studies (e.g. Liedvogel & Sheldon 2010, Dor *et al.* 2011). However, to the best of my knowledge, no study has investigated the effect of *Clock* and *Adcyap1* genotype, as well as that of other candidate genes, on migration phenology. Moreover, further studies focusing on the *Clock*- and *Adcyap1*-migratory behaviour associations are needed to shed light on the role of candidate genes polymorphism in shaping the inter-individual variation of other migratory traits.

The aim of this thesis is to investigate whether polymorphism at candidate genes affects the timing of migration and the migration distance in avian species, with particular reference to *Clock* and *Adcyap1* genes. I will focus on *Clock* and *Adcyap1* because information on variation at these genes has recently become available for several bird populations and species. Moreover, as outlined above, polymorphism at *Clock* and *Adcyap1* genes have been previously linked to phenotypic variability in avian species. However, I will also assess the genotype-phenology association for two other candidate genes, *Creb1* and *Npas2*, whose polymorphism has been recently linked to differences in the timing of breeding and of juvenile dispersal (Chakarov *et al.* 2013; Bourret & Garant 2015), and for a newly identified polymorphic *Clock* region. Since long-distance migratory species exhibit a tighter endogenous control of migratory traits (Debat & David 2001; Pulido & Widmer 2005), they are good candidates to study genotype-phenotype association. I will hence focus on long-distance migration.

In the first chapters of the thesis (**Chapter 1**, **2** and **3**) I investigated the association between *Clock* and *Adcyap1* polymorphism and migratory traits' variability in individuals from different

population migrating through stopover sites, with particular focus on spring migration. Specifically, **Chapter 1** focuses on whether *Clock* and *Adcyap1* allele size predicted the arrival date at a stopover site during spring migration in four trans-Saharan migratory species (nightingale, *Luscinia megarhynchos*; pied flycatcher, *Ficedula hypoleuca*; tree pipit *Anthus trivialis* and whinchat *Saxicola rubetra*). In addition, I tested for the allelic dominance of the long *Clock* allele, since this pattern was highlighted in previous studies for *Clock* and other genes characterized by Poly-Q polymorphism (Ross 2002; Fondon *et al.* 2008; Liedvogel *et al.* 2009). *Clock* allele size significantly and positively predicted migration timing in two out of four species; moreover, I found evidence that such relationship was driven by the size of the long rather than the short *Clock* allele. Conversely, *Adcyap1* allele size did not covary with spring migration timing in the study species.

Chapter 2 is focused on the association between *Clock* and *Adcyap1* allele size and two aspects of migratory phenotype: the timing of spring migration and the migration distance. By capitalizing on feather deuterium ratio (δ^2H) of individual Wilson's warblers (*Cardellina pusilla*) migrating through a stopover site halfway the wintering and the breeding grounds, I provided evidence that *Adcyap1* allele size was associated with migration distance in males, but not in females. Indeed, the frequency of longer *Adcyap1* alleles was higher in (male) Wilson's warblers populations breeding at northern latitudes and migrating over longer distances than in those breeding at southern latitudes. Furthermore, I reported that the genotype-breeding latitude association was stronger among the long-distance migrants from the northern breeding populations than among southern birds migrating over shorter distances. On the other hand, *Clock* gene showed very low variability, and nor *Adcyap1* neither *Clock* allele size significantly predicted the timing of spring migration in the Wilson's warbler.

In **Chapter 3** I considered variation at five candidate genes (*Clock*, *Adcyap1*, *Npas2*, *Creb1* and a newly identified *Clock* region that shows a tandem repeat motif, *Clock3*) in association with the timing of two major life-history events in a trans-Saharan migratory passerine bird, the willow

warbler (*Phylloscopus trochilus*). First, I tested whether candidate genes allele size predicted the timing of spring migration across the central Mediterranean sea, recorded at a stopover site. Secondly, I investigated the association between candidate genes polymorphism and moult speed, another important determinant of the timing of circannual events, as gauged by ptilochronological study of feather growth rate. *Npas2* allele size affected migration date in male willow warblers, and long *Clock* allele size predicted migration phenology in females (but not in males), although the relationship was opposite to our predictions in both cases, longer alleles being associated with early migration. However, wing length decreased with migration date in females (rather than increasing, as expected), hinting that such associations might emerge because of a geographic differentiation at *Npas2* and *Clock* combined with an unusual pattern of migration of different populations. Furthermore, *Creb1* polymorphism was related to variation in moult speed; indeed, males, but not females, bearing longer *Creb1* alleles grown their feather faster than those carrying shorter alleles. Conversely, *Clock3* and *Adcyap1* allele size did not predict timing of migration nor moult speed.

In the second part of thesis (**Chapter 4**) I considered the within-population variability in migration phenology in relation to *Clock* allele size. By taking advantage of year-round tracking data of individual barn swallows obtained by means of light-level geolocators, I tested whether *Clock* polymorphism predicted the timing of non-breeding period (from departure for fall migration to arrival at the breeding grounds in spring) within three different populations breeding in the Po plain (Italy) and in southern Switzerland. Despite the very low genetic variation at this locus, I showed *Clock* allele size predicted the timing of non-breeding season in the Swiss population. Moreover, I found that the effect of *Clock* allele size on phenology varied depending on the phase considered, being more pronounced for phenological variables related to the timing of spring migration. Conversely, no association between genotype and migration timing emerged for the two Italian populations.

Finally, the third part of the thesis (**Chapter 5**) consists of a comparative study of 23 trans-Saharan migratory species; I tested whether genetic variation (allele size and gene diversity) at *Clock* and *Adcyap1* loci covaried with the species' life-history traits, including distribution, phenology of spring migration across the central Mediterranean sea and distance between the wintering and the breeding ranges. I used phylogenetic confirmatory path analysis to disentangle the possible causal relationships behind the associations between migratory traits and gene diversity. Using such a comparative approach I highlighted that species breeding at northern latitudes had a higher frequency of longer *Clock* alleles. Moreover, I found that *Clock*, but not *Adcyap1*, gene diversity was lower among species migrating over longer distances and having a delayed and more concentrated spring migration, the most important variables directly affecting *Clock* gene diversity being the migration distance and the migration date.

3. Conclusions and future perspectives

On the whole, these findings contribute to broaden our understanding of the role of phenological candidate genes in the control of avian behaviour, from the inter-individual to the inter-species level. Capitalizing on several and in some cases novel methods used to infer migratory traits at different scales, I provided entirely novel evidence of the association between *Clock* and *Adcyap1* polymorphism and traits linked to migration.

In this thesis I showed for the first time that *Clock* allele size predicts migration timing in bird species (**Chapter 1**, **3** and **4**). The proximate mechanisms underpinning the *Clock* allele size-phenology association in migratory birds are still unknown and open to speculations, but could involve different photoperiodic response during migrations, for example at the time of departure

from the wintering grounds or during the spring migratory journey. This hypothesis seems to be corroborated by the findings of **Chapter 4**, along with previous evidence for a role of *Clock* allele size in the control of winter moult phenology in birds overwintering around the Equator (Saino *et al.* 2013); indeed, in barn swallows tracked from fall migration to arrival at the breeding grounds, the magnitude of the effect of *Clock* genotype on phenology increased along the non-breeding season, being less clear for the departure from the breeding grounds and becoming more marked during wintering and at the onset of spring migration. Despite the small circannual variation in photoperiod experienced by birds in the equatorial region, it has been demonstrated that they can detect even minor daily changes in photoperiod and in the time of sunrise and sunset and that they could set the timing of seasonal events accordingly (Hau *et al.* 1998; Goymann *et al.* 2012). If this is true, the association between *Clock* allele size and timing of breeding detected among migratory species (e.g. Caprioli *et al.* 2012) could originate from carry-over effects of events that occur during the non-breeding season, at low latitudes. Alternatively, it has been proposed that the photoperiodic response mediated by *Clock* gene may be restricted to the life-history events occurring at high latitudes, where the seasonal differences in photoperiod are marked (Johnsen *et al.* 2007; Liedvogel *et al.* 2009; Caprioli *et al.* 2012; Bourret & Garant 2015). By highlighting that the frequency of longer *Clock* alleles increased with the breeding latitude of a species (**Chapter 5**), I provided further support for a possible role of *Clock* length polymorphism in the adaptation to local photoperiodic regimes at the breeding grounds.

Although contentious and not in line with the expected, the evidence of an association between timing of spring migration and genetic variation at *Npas2* gene, another candidate gene involved in the photoperiodic response (**Chapter 3**), hint that polymorphism at other candidate genes, besides the well-studied ones, may contribute to phenotypic variation in birds, especially for those species that show low variability at *Clock* locus.

Furthermore, I found an association between *Adcyap1* allele size and migration distance among a Nearctic-Neotropical migratory passerine bird, the Wilson's warbler (**Chapter 2**). This is the first study to detect a direct association between *Adcyap1* genotype and migratory phenotype in individual wild birds and supported the expectations of previous research on migration propensity and migratory restlessness (Mueller *et al.* 2011; Peterson *et al.* 2013). Importantly, the relationship was remarkably stronger among the long-distance, northern breeding populations and weaker and statistically non-significant for the southern breeding ones. Thus, we may hypothesize that the tighter endogenous control of migration (and possibly of other seasonal activities) that characterizes the long-distance migrants (Debat & David 2001; Pulido & Widmer 2005) could arise from a stronger genotype-phenotype association at phenological candidate genes.

Finally, using a comparative approach (**Chapter 5**), I showed that longer, delayed and concentrated migratory journeys are associated with lower genetic diversity at *Clock* locus. Such result suggests that the reduced phenotypic variance observed among long-distance migratory species (Pulido & Widmer 2005; Rubolini *et al.* 2010; Knudsen *et al.* 2011), that has been previoulsy attributed to a tighter endogenous control of migration and hence to an environmental canalization of migratory traits (Pulido & Widmer 2005), may also be determined by depleted genetic variation at phenological candidate genes.

Taken together, these findings hint that candidate genes controlling migratory behaviour could be of pivotal evolutionary importance for long-distance migratory birds. For instance, the adaptive phenological shifts expected under climate change could be achieved by genetic variation at phenological candidate genes, since in these species the plastic phenotypic response appears to be reduced (Pulido & Widmer 2005; Rubolini *et al.* 2010; Knudsen *et al.* 2011). Selection towards shorter *Clock* alleles has been detected in two bird species that showed a *Clock* genotype-phenology association. Indeed, blue tits and barn swallows bearing shorter *Clock* alleles showed a higher fecundity, possibily mediated by the benefits to breed early, compared to those carrying longer

alleles (Liedvogel *et al.* 2009; Caprioli *et al.* 2012). However, the loss of genetic variation at genes controlling migratory traits observed in bird species migrating over longer distances suggests that the marked decline of populations of long-distance migratory birds driven by changing environmental conditions (Sanderson *et al.* 2006; Møller *et al.* 2008; Saino *et al.* 2011) may at least partly explained by the impossibility to accomplish micro-evolutionary changes in genetic composition. In this context, a deeper insight of phenological candidate gene variability and of the genotype-phenology association could help to predict the consequences of climate changes and to identify which species deserve more attention from a conservational perspective.

Studying candidate gene polymorphism in relation to migratory phenotype could also improve our understanding of evolutionary mechanisms. It has been suggested that the occurrence of among-population differences in the timing of seasonal events could drive reproductive isolation, possibly leading to speciation (e.g. Quinn et al. 2000; Adams & Thibauld 2006; Friesen et al. 2007; Devaux & Lande 2008; Ketterson *et al.* 2015). If genetic structure at phenological candidate genes actually underlies among-populations variation in timing of migration and breeding, at least in some bird species, the occurrence of latitudinal clines in the frequency of longer and shorter alleles could be involved in the mechanism of population divergence through allochrony. Other behavioural traits linked to migration could promote reproductive isolation in sympatric migratory populations; for instance, in blackcaps breeding in central Europe that have recently established a migratory divide, differences in migration distance and migratory orientation promoted reproductive isolation and genetic divergence in a few generations (Rolshausen *et al.* 2009). A study comparing genome-wide patterns of divergence between populations of Swainson's thrush (*Catharus ustulatus*) migrating over different distances and along distinct routes further suggests that that migratory-linked genes and migratory behaviour could be important drivers of population differentiation. In these populations, indeed, several genes linked with migratory behaviour, among which *Adcyap1*, were remarkably more differentiated than other autosomal genes (Ruegg et al. 2014).

Findings from this thesis (**Chapter 1**, **2**, **3** and **4**) and previous research (e.g. Peterson *et al.* 2013; Bourret & Garant 2015) highlighted that the effect of candidate genes on phenotype could vary according to the species or the population considered, to the environment experienced by individuals or to the sex. It has been suggested that the among-population differences in candidate genes-phenology associations could arise from the occurrence of different environmental effects or from variation in the genetic background of the populations (Peterson *et al.* 2013; Bourret & Garant 2015). Hence, future research should consider sex- or population-specific selective pressures or a broader set of candidate genes that could affect phenotype in wild populations. A further, intriguing, possibility could be to identify new polymorphic loci possibly determining variation in migratory traits. For instance, the restriction site-associated DNA sequencing (RADSeq) allows the detection of thousands of single nucleotide polymorphisms (SNPs) and other genetic markers simultaneously, without the need to rely on existing sequencing data and with rather low costs (Davey & Blaxter 2010). For such reasons, this technique is particularly useful for genome-wide studies of natural populations of non-model organisms (Davey & Blaxter 2010) and, along with novel methods that could efficiently analyze large SNP dataset (Gutenkunst *et al.* 2009; Excoffier *et al.* 2013) may represent the next frontier in the study of the genetic bases of migratory behaviour.

4. References

Adams RA, Thibault KM (2006) Temporal resource partitioning by bats at water holes. Journal of Zoology **270**, 466-472.

Bartel DP (2009) MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215-233.

Bell-Pedersen D, Cassone VM, Earnest DJ*, et al.* (2005) Circadian rhythms from multiple oscillators: lessons from diverse organisms. *Nature Reviews Genetics* **6**, 544-556.

Berthold P (1974). Circannual rhythms in birds with different migratory habits. pp. 44–94 in *Circannual clocks* (ed. E. T. Pengelly). New York, Academic Press.

Berthold P (1984) The endogenous control of bird migration: a survey of experimental evidence. *Bird Study* **31**, 19-27.

Berthold P (1988) The control of migration in European warblers. *Acta XIX Congressus Internationalis Ornithologici* **19**, 215-249.

Berthold P (1996) *Control of bird migration* Springer Science & Business Media.

Berthold P (2001) *Bird migration: a general survey* (2nd edn) Oxford University Press, Oxford.

Berthold, P. & Terrill, S. B. (1991) Recent advances in studies of bird migration. *Annual Review of Ecology, Evolution and Systematics* **22**, 357-378.

Bourret A, Garant D (2015) Candidate gene-environment interactions and their relationships with timing of breeding in a wild bird population. *Ecology and Evolution* **5**, 3628-3641.

Brown CR, Brown MB (2000) Weather-mediated natural selection on arrival time in Cliff Swallows (*Petrochelidon pyrrhonata*). *Behavioral Ecology and Sociobiology* **47**, 339–345.

Caprioli M, Ambrosini R, Boncoraglio G*, et al.* (2012) *Clock* gene variation is associated with breeding phenology and maybe under directional selection in the migratory barn swallow. *PLoS ONE* **7**, e35140.

Chakarov N, Jonker RM, Boerner M, Hoffman JI, Kruger O (2013) Variation at phenological candidate genes correlates with timing of dispersal and plumage morph in a sedentary bird of prey. *Molecular Ecology* **22**, 5430-5440.

Chatterjee S, Pal JK (2009) Role of 5'- and 3'-untranslated regions of mRNAs in human diseases. *Biology of the Cell* **101**, 251-262.

Comings DE (1998) Polygenic inheritance and micro/minisatellites. *Molecular Psychiatry* **3**, 21-31.

Darlington TK, Wager-Smith K, Ceriani MF*, et al.* (1998) Closing the Circadian Loop: CLOCK-Induced Transcription of Its Own Inhibitors *per* and *tim*. *Science* **280**, 1599-1603.

Davey JW, Blaxter M.L (2010) RADSeq: next-generation population genetics. *Briefings in functional genomics* **9**, 416-423.

Debat V, David P (2001) Mapping phenotypes: canalization, plasticity and developmental stability. *Trends in Ecology & Evolution* **16**, 555-561.

Devaux C, Lande R (2008) Incipient allochronic speciation due to non-selective assortative mating by flowering time, mutation and genetic drift. *Proceedings of the Royal Society B: Biological Sciences* **275**, 2723-2732.

Dor R, Lovette IJ, Safran RJ*, et al.* (2011) Low variation in the polymorphic *Clock* gene poly-Q region despite population genetic structure across barn swallow (*Hirundo rustica*) populations. *PLoS ONE* **6**, e28843.

Dor R, Cooper CB, Lovette IJ *et al.* (2012) *Clock* gene variation in *Tachycineta* swallows. *Ecology and Evolution* **2**, 95-105.

Ellegren H, Sheldon BC (2008) Genetic basis for fitness differences in natural populations. Nature, **452**, 169-175.

Excoffier L, Dupanloup I, Huerta-Sanchez E, Sousa VC, Foll M. (2013) Robust demographic inference from genomic and SNP data. *PLOS Genetics* **9**, e1003905.

Fidler AE, Gwinner E (2003) Comparative analysis of Avian BMAL1 and CLOCK protein sequences: a search for features associated with owl nocturnal behaviour. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* **136**, 861-874.

Fitzpatrick MJ, Ben-Shahar Y, Smid HM, *et al.* (2005) Candidate genes or behavioural ecology. *Trends in Ecology and Evolution* **20**, 96-104.
Fondon JW, 3rd, Hammock EA, Hannan AJ, King DG (2008) Simple sequence repeats: genetic modulators of brain function and behavior. *Trends in Neurosciences* **31**, 328-334.
Friesen VL, Smith AL, Gómez-Díaz E, *et al.* (2007) Sympatric speciation by allochrony in a seabird. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 18589-18594.
Fusani L, Gwinner E (2005) Melatonin and nocturnal migration. Annals of the New York Academy of Sciences **1046**,264-270.
Gauthreaux SA (1982) The ecology and evolution of avian migration systems. pp. 93–167 in *Avian Biology*, Vol. 6 (eds DS Farner & JR King). Academic Press, New York.
Gekakis N, Staknis D, Nguyen HB*, et al.* (1998) Role of the CLOCK Protein in the Mammalian Circadian Mechanism. *Science* **280**, 1564-1569.
Goymann W, Helm B, Jensen W, Schwabl I, Moore IT (2012) A tropical bird can use the equatorial change in sunrise and sunset times to synchronize its circannual clock. *Proceedings of the Royal Society B: Biological Sciences* **279**, 3527-3534.
Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD (2009) Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genetics* **5**, e1000695.
Gwinner, E. (1968). Circannuale Periodik als Grundlage des jahreszeitlichen Funktionswandels bei Zugvögeln. Untersuchungen am Fitis (*Phylloscopus trochilus*) und am Waldlaubsänger (*P. sibilatrix*). *Journal of Ornithology* **109**, 70-95.
Gwinner, E. (1971). A comparative study of circannual rhythms in warblers. pp. 405–427 in *Biochronometry* (ed. M Menaker.). National Academy of Sciences, Washington, D.C.
Gwinner E (1972) Adaptive functions of circannual rhythms in warblers. In: *XVth International Ornithological Congress*, pp. 218-236.
Gwinner, E. (1981). Circannual rhythms: Their dependence on the circadian system. pp. 153–169 in *Biological clocks in seasonal reproductive cycles* (eds BK Follett & DE Follett). Wright Publishers, Bristol.
Gwinner E (1986) *Circannual rhythms* Springer-Verlag, Berlin.
Gwinner E (1990) *Bird migration: physiology and ecophysiology* Springer Verlag, Berlin.
Gwinner E, Helm B (2003) Circannual and circadian contributions to the timing of avian migration. In: *Avian migration* (eds. Berthold P, Gwinner E, Sonnenschein E), pp. 81-95. Springer-Verlag, Berlin.
Hagan JM, Lloyd-Evans L, Alwood JL (1991) The relationship between latitude and the timing of spring migration of North American landbirds. *Ornis Scandinavica* **22**, 129-136.
Hahn S, Bauer S, Liechti F (2009) The natural link between Europe and Africa–2.1 billion birds on migration. *Oikos* 1**18**, 624-626.
Hannibal J, Ding JM, Chen D*, et al.* (1997) Pituitary adenylate cyclase-activating peptide (PACAP) in the retinohypothalamic tract: a potential daytime regulator of the biological clock. *Journal of Neuroscience* **17**, 2637-2644.
Hau M, Wikelski M, Wingfield JC (1998) A neotropical forest bird can measure the slight changes in tropical photoperiod. *Proceedings of the Royal Society B: Biological Sciences* **265**, 89-95.
Hayasaka N, LaRue SI, Green CB (2002) In Vivo Disruption of Xenopus CLOCK in the Retinal Photoreceptor Cells Abolishes Circadian Melatonin Rhythmicity without Affecting Its Production Levels. *Journal of Neuroscience* **22**, 1600-1607.
Hirokawa N, Takemura R (2005) Molecular motors and mechanisms of directional transport in neurons. *Nature Reviews: Neuroscience* **6**, 201-214.
Hoffmann AA, Willi Y (2008) Detecting genetic responses to environmental change. *Nature Reviews Genetics* **9**, 421-432.
Iuvone PM, Tosini G, Pozdeyev N, *et al.* (2005) Circadian clocks, clock networks, arylalkylamine N-acetyltransferase, and melatonin in the retina. *Progress in retinal and eye research* **24**, 433-456.
Johnsen A, Fidler AE, Kuhn S*, et al.* (2007) Avian Clock gene polymorphism: evidence for a latitudinal cline in allele frequencies. *Molecular Ecology* **16**, 4867-4880.
Kashi Y, King D, Soller M (1997) Simple sequence repeats as a source of quantitative genetic variation. *Trends in Genetics* **13**, 74-78.

Ketterson ED, Fudickar AM, Atwell JW, Greives TJ (2015) Seasonal timing and population divergence: when to breed, when to migrate. *Current Opinion in Behavioral Sciences*, **6**, 50-58.

Knudsen E, Linden A, Both C*, et al.* (2011) Challenging claims in the study of migratory birds and climate change. *Biological Reviews* **86**, 928-946.

Ko CH, Takahashi JS (2006) Molecular components of the mammalian circadian clock. *Human Molecular Genetics* **15 Spec No 2**, R271-277.

Kuhn K, Schwenk K, Both C, et al. (2013) Differentiation in neutral genes and a candidate gene in the pied flycatcher: using biological archives to track global climate change. *Ecology and Evolution* **3**, 4799-4814.

Lack, D. (1954). *The natural regulation of animal numbers* Oxford University Press, Oxford.

Li YC, Korol AB, Fahima T, Nevo E (2004) Microsatellites within genes: structure, function, and evolution. *Molecular Biology and Evolution* **21**, 991-1007.

Liedvogel M, Sheldon BC (2010) Low variability and absence of phenotypic correlates of *Clock* gene variation in a great tit *Parus major* population. *Journal of Avian Biology* **41**, 543-550.

Liedvogel M, Lundberg M (2014) The genetics of migration. In: *Animal movements across scales* (eds. Hansson LA, Åkesson S). Oxford University Press, Oxford.

Liedvogel, M., Åkesson, S., & Bensch, S. (2011). The genetics of migration on the move. *Trends in Ecology & Evolution*, **26**, 561-569.

Liedvogel M, Szulkin M, Knowles SC, Wood MJ, Sheldon BC (2009) Phenotypic correlates of *Clock* gene variation in a wild blue tit population: evidence for a role in seasonal timing of reproduction. *Molecular Ecology* **18**, 2444-2456.

Lincoln GA, Andersson H, Loudon A (2003) Clock genes in calendar cells as the basis of annual timekeeping in mammals - a unifying hypothesis. *Journal of Endocrinology* **179**, 1-13.

Lundberg, P. (1987). Partial bird migration and evolutionarily stable strategies. *Journal of Theoretical Biology*, **125**, 351-360.

Maggini I, Bairlein F (2010) Endogenous rhythms of seasonal migratory body mass changes and nocturnal restlessness in different populations of Northern Wheatear *Oenanthe oenanthe*. *Journal of Biological Rhythms* **25**, 268-276.

Merrow M, Spoelstra K, Roenneberg T (2005) The circadian cycle: daily rhythms from behaviour to genes. *EMBO Reports* **6**, 930-935.

Mettler R, Segelbacher G, Schaefer HM (2015) Interactions between a candidate gene for migration (ADCYAP1), morphology and sex predict spring arrival in blackcap populations. *PLoS ONE* **10**, e0144587.

Møller AP (2001) Heritability of arrival date in a migratory bird. *Proceedings of the Royal Society B: Biological Sciences* **268**, 203-206.

Møller AP, Rubolini D, Lehikoinen E (2008) Populations of migratory bird species that did not show a phenological response to climate change are declining. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 16195-16200.

Møller AP, Fiedler W, Berthold P (2010) *Effects of Climate Change on Birds*. Oxford University Press, Oxford.

Mueller JC, Pulido F, Kempenaers B (2011) Identification of a gene associated with avian migratory behaviour. *Proceedings of the Royal Society B: Biological Sciences* **278**, 2848-2856.

Nagy AD, Csernus VJ (2007) The role of PACAP in the control of circadian expression of clock genes in the chicken pineal gland. *Peptides* **28**, 1767-1774.

Newton I (2008) *The migration ecology of birds.* Academic Press, London.

Nowak JZ, Zawilska JB (2003) PACAP in avians: origin, occurrence, and receptors-pharmacological and functional considerations. *Current Pharmaceutical Design*.

Olano-Marin J, Mueller JC, Kempenaers B (2011) Heterozygosity and survival in blue tits (*Cyanistes caeruleus*): contrasting effects of presumably functional and neutral loci. *Molecular Ecology* **20**, 4028-4041.

O'Malley KG, Banks MA (2008) A latitudinal cline in the Chinook salmon (*Oncorhynchus tshawytscha*) *Clock* gene: evidence for selection on PolyQ length variants. *Proceedings of the Royal Society B: Biological Sciences* **275**, 2813-2821.

O'Malley KG, Ford MJ, Hard JJ (2010) Clock polymorphism in Pacific salmon: evidence for variable selection along a latitudinal gradient. *Proceedings of the Royal Society B: Biological Sciences* **227**, 3703-3714.

Panda S, Hogenesch JB, Kay SA (2002) Circadian rhythms from flies to human. *Nature* **417**, 329-335.

Pardo-Diaz C, Salazar C, Jiggins CD (2015) Towards the identification of the loci of adaptive evolution. *Methods in Ecology and Evolution* **6**, 445-464.

Peterson MP, Abolins-Abols M, Atwell JW*, et al.* (2013) Variation in candidate genes CLOCK and ADCYAP1 does not consistently predict differences in migratory behavior in the songbird genus Junco. *F1000Research* **2**, 115.

Pulido F (2007) The Genetics and Evolution of Avian Migration. *Bioscience* **57**, 165.

Pulido F, Berthold P (2003) Quantitative genetic analysis of migratory behaviour. In: *Avian migration* (eds. Berthold P, Gwinner E, Sonnenschein E), pp. 53-77. Springer Science & Business Media, Berlin.

Pulido F, Widmer M (2005) Are long-distance migrants constrained in their evolutionary response to environmental change? Causes of variation in the timing of autumn migration in a blackcap (*S. atricapilla*) and two garden warbler (*Sylvia borin*) populations. *Annals of the New York Academy of Sciences* **1046**, 228-241.

Pulido F, Berthold P, Mohr G, Querner U (2001) Heritability of the timing of autumn migration in a natural bird population. *Proceedings of the Royal Society B: Biological Sciences* **268**, 953-959.

Quinn TP, Unwin MJ, Kinnison MT (2000) Evolution of temporal isolation in the wild: genetic divergence in timing of migration and breeding by introduced chinook salmon populations. *Evolution* **54**,1372-1385.

Racz B, Horvath G, Faluhelyi N*, et al.* (2008) Effects of PACAP on the circadian changes of signaling pathways in chicken pinealocytes. *Journal of Neuroscience* **36**, 220-226.

Rappole J (1995) *The ecology of migrant birds. A neotropical perspective* Smithsonian Institution Press, Washington, D.C.

Reppert SM, Weaver DR (2002) Coordination of circadian timing in mammals. *Nature* **418**, 935-941.

Rolshausen G, Segelbacher G, Hobson K A, Schaefer HM (2009) Contemporary evolution of reproductive isolation and phenotypic divergence in sympatry along a migratory divide. *Current Biology* **19**, 2097-2101.

Ross CA (2002) Polyglutamine Pathogenesis. *Neuron* **35**, 819-822.

Rubolini D, Saino N, Møller AP (2010) Migratory behaviour constrains the phenological response of birds to climate change. *Climate Research* **42**, 45-55.

Rubolini D, Spina F, Saino N (2005) Correlates of timing of spring migration in birds: a comparative study of trans-Saharan migrants. *Biological Journal of the Linnean Society* **85**, 199-210.

Ruegg K, Anderson EC, Boone J, Pouls J, Smith TB (2014) A role for migration-linked genes and genomic islands in divergence of a songbird. Molecular Ecology **23**, 4757-4769.

Saino N, Ambrosini R (2008) Climatic connectivity between Africa and Europe may serve as a basis for phenotypic adjustment of migration schedules of trans-Saharan migratory birds. *Global Change Biology* **14**, 250-263.

Saino N, Ambrosini R, Rubolini D*, et al.* (2011) Climate warming, ecological mismatch at arrival and population decline in migratory birds. *Proceedings of the Royal Society B: Biological Sciences* **278**, 835-842.

Saino N, Romano M, Caprioli M*, et al.* (2013) Timing of molt of barn swallows is delayed in a rare *Clock* genotype. *PeerJ* **1**, e17.

Sanderson FJ, Donald PF, Pain DJ, Burfield IJ, van Bommel FPJ (2006) Long-term population declines in Afro-Palearctic migrant birds. *Biological Conservation* **131**, 93-105.

Schwartz C, Andrews MT (2013) Circannual transitions in gene expression: lessons from seasonal adaptations. In: *Developmental timing* (eds. Rougvie AE, O'Connor MB) pp. 247–273. Elsevier, Philadelphia.

Sharp PJ (2005) Photoperiodic regulation of seasonal breeding in birds. *Annals of the New York Academy of Sciences* **1040**, 189-199.

Simonneaux V, Ouichou A, Pévet P (1993) Pituitary adenylate cyclase-activating polypeptide (PACAP) stimulates melatonin synthesis from rat pineal gland. *Brain Research* **603**, 148-152.

Steinmeyer C, Mueller JC, Kempenaers B (2009) Search for informative polymorphisms in candidate genes: clock genes and circadian behaviour in blue tits. *Genetica* **136**, 109-117.

Tabor HK, Risch NJ, Myers RM (2002) Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nature Reviews Genetics* **3**, 391-397.

Tachibana T, Saito S, Tomonaga S, *et al.* (2003) Intracerebroventricular injection of vasoactive intestinal peptide and pituitary adenylate cyclaseactivating polypeptide inhibits feeding in chicks. Neuroscience Letters **339**, 203-206.

Tachibana T, Oikawa D, Adachi N, Boswell T, Furuse M (2007) Central administration of vasoactive intestinal peptide and pituitary adenylate cyclase-activating polypeptide differentially regulates energy metabolism in chicks. Comparative Biochemistry and Physiology A **147**, 156-164.

Thorup K, Rabøl J (2007). Compensatory behaviour after displacement in migratory birds. *Behavioral Ecology and Sociobiology*, ***61***, 825-841.

van Noordwijk AJ, Pulido F, Helm B, *et al.* (2006) A framework for the study of genetic variation in migratory behaviour. *Journal of Ornithology* **147**, 221-233.

Vaudry D, Falluel-Morel A, Bourgault S*, et al.* (2009) Pituitary adenylate cyclase-activating polypeptide and its receptors: 20 years after the discovery. *Pharmacological Reviews* **61**, 283-357.

Visser ME, Caro SP, van Oers K, Schaper SV, Helm B (2010) Phenology, seasonal timing and circannual rhythms: towards a unified framework. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**, 3113-3127.

Young MW, Kay SA (2001) Time zones: a comparative genetics of circadian clocks. *Nature Reviews: Genetics* **2**, 702-715.

5. Acknowledgements

I would like to express my gratitude to my supervisor Prof. Diego Rubolini for the opportunity to study migratory birds, for the precious support during the past three years and for helping me to publish my researches in high-level international scientific journals. I warmly thank Prof. Nicola Saino for his helpful supervision and suggestions. I am sincerely grateful to Manuela Caprioli, Luca Gianfranceschi and Emanuele Gatti for the precious assistance with the lab work. I thank Marco Parolini, Andrea Romano, Maria Romano and my friends and colleagues Margherita Corti, Alessandra Costanzo and Stefano Podofillini with whom I shared the field and lab work in the past years and who gave me very much appreciated support and suggestions. I wish also to say special thanks to Cristina Possenti; without her the field work at Ventotene would have been much more harder and less funny. I would also like to thank the thesis students who helped me to collect the data on the field, especially Sara Fusetti, Maria Patrascanu, Arianna Bazzocchi, Giuseppe Chialà and Stefano dell'Oro. I am grateful to Elisa Mancuso, Lisa Carrera, Giulia Capobianco Dondona, Violetta Longoni, Sara Riello, Pierfrancesco Micheloni, Dario Piacentini, Luca Ilahiane, Valerio Orioli, Annarita Matrone, Davide De Rosa, Ilaria Fozzi and all the ringers and field assistants I met at Ventotene for the funny days on the island. I thank Fernando Spina and the CNI, as well as the Riserva Naturale Isole di Ventotene e Santo Stefano, for the logistic support and for allowing me to collect the data on Ventotene, making this research possible. I wish to thank Andrea Galimberti for the really appreciated suggestions (and for the precious support and friendship) and Ilaria Bruni, who performed a part of the lab work. I have also benefitted from suggestions by Roberto Ambrosini, who helped me with statistical analyses and with the interpretation of results. I am grateful to Jacopo G. Cecere, Ivan Maggini, Christopher Guglielmo, Yolanda E. Morbey, Keith H. Hobson and Quentin R. Hays, which allowed me to analyze a very interesting dataset and to enlarge my PhD work to Nearctic species. I am also grateful to Piotr Matyjasiak, who hosted me in Poland. Finally, I thank my family and Giacomo, which supported me during these years.

Photo: Andrea Capobianco Dondona

PART 1

CHAPTER 1

Polymorphism at the *Clock* gene predicts phenology of long-distance migration in birds

Nicola Saino, Gaia Bazzi, Emanuele Gatti, Manuela Caprioli, Jacopo G. Cecere, Cristina D. Possenti, Andrea Galimberti , Valerio Orioli, Luciano Bani, Diego Rubolini, Luca Gianfranceschi and Fernando Spina

Molecular Ecology **24** (2015): 1758-1773

Photo: Alberto Erba, Andrea Galimberti and Francesco Renzi

MOLECULAR ECOLOGY

Molecular Ecology (2015) **24**, 1758–1773 doi: 10.1111/mec.13159

Polymorphism at the *Clock* gene predicts phenology of long-distance migration in birds

NICOLA SAINO,* GAIA BAZZI,* EMANUELE GATTI,* MANUELA CAPRIOLI,* JACOPO G. CECERE,† CRISTINA D. POSSENTI,* ANDREA GALIMBERTI,‡ VALERIO ORIOLI,§ LUCIANO BANI,§ DIEGO RUBOLINI,* LUCA GIANFRANCESCHI* and FERNANDO SPINA†

*Department of Biosciences, University of Milan, via Celoria 26, Milan I-20133, Italy, †ISPRA—Italian Institute for Environmental Protection and Research, Via Cà Fornacetta 9, Ozzano dell'Emilia I-40064, Italy, ‡Department of Biotechnology and Biosciences, University of Milano-Bicocca, Piazza della Scienza 2, Milan I-20126, Italy, §Department of Earth and Environmental Sciences, University of Milano-Bicocca, Piazza della Scienza 1, Milan I-20126, Italy

Abstract

Dissecting phenotypic variance in life history traits into its genetic and environmental components is at the focus of evolutionary studies and of pivotal importance to identify the mechanisms and predict the consequences of human-driven environmental change. The timing of recurrent life history events (phenology) is under strong selection, but the study of the genes that control potential environmental canalization in phenological traits is at its infancy. Candidate genes for circadian behaviour entrained by photoperiod have been screened as potential controllers of phenological variation of breeding and moult in birds, with inconsistent results. Despite photoperiodic control of migration is well established, no study has reported on migration phenology in relation to polymorphism at candidate genes in birds. We analysed variation in spring migration dates within four trans-Saharan migratory species (*Luscinia megarhynchos*; *Ficedula hypoleuca*; *Anthus trivialis*; *Saxicola rubetra*) at a Mediterranean island in relation to *Clock* and *Adcyap1* polymorphism. Individuals with larger number of glutamine residues in the poly-Q region of *Clock* gene migrated significantly later in one or, respectively, two species depending on sex and whether the within-individual mean length or the length of the longer *Clock* allele was considered. The results hinted at dominance of the longer *Clock* allele. No significant evidence for migration date to covary with *Adcyap1* polymorphism emerged. This is the first evidence that migration phenology is associated with *Clock* in birds. This finding is important for evolutionary studies of migration and sheds light on the mechanisms that drive bird phenological changes and population trends in response to climate change.

Keywords: *Adcyap1*, *Anthus trivialis*, birds, climate change, *Clock*, *Ficedula hypoleuca*, *Luscinia megarhynchos*, migration, phenology, *Saxicola rubetra*

Received 31 March 2014; revision received 13 March 2015; accepted 13 March 2015

Introduction

Information on phenotypic variation in life-history traits is plentiful and detailed, but our knowledge of the genetic mechanisms and diversity that underpin such variation is patchy, at best. For example, most organisms living in seasonal environments need to match the timing of their main activities, such as leafing, emerging from diapause, breeding or migrating, with temporal variation in ecological conditions, as even small deviations from appropriate timing may have serious fitness consequences (Clutton-Brock 1988; Nussey *et al.* 2005; Both *et al.* 2006). Individuals from a same or different populations often show large variation in the timing of such activities (i.e. in their phenology), but the extent to which phenological differences are hard-wired in genetic differences or rest on phenotypically plastic

Correspondence: Nicola Saino, Fax: +39 0250314802; E-mail: nicola.saino@unimi.it

responses to environmental heterogeneity is still largely unknown (see Møller *et al.* 2010). Yet, dissecting the sources of phenological variation is relevant to the generality of evolutionary studies and is of pivotal importance for our understanding of the impact that changes in ecological conditions (Ellegren & Sheldon 2008), including anthropogenic climate change at global scale (see Møller *et al.* 2010), will have on populations and communities of organisms.

Adaptive timing of switching between major activities in order to track seasonal variation often requires that organisms anticipate ecological conditions days to months ahead and undergo physiological and behavioural changes that may take long to be accomplished (Bradshaw & Holzapfel 2007, 2010; Foster & Kreitzman 2009). Environmental seasonality is predictably associated with changes in photoperiod because both are ultimately driven by the revolution of the earth around the sun. Diverse organisms have evolved the ability to sense photoperiodic changes as a cue to prepare for the forthcoming new conditions (Vaz Nunes & Saunders 1999; Bradshaw & Holzapfel 2007; Helm *et al.* 2013). This is the case, for example, in salmonid fish where gonadal maturation depends on photoperiod (Beacham & Murray 1988). All main annual activities in birds, including breeding, migration and plumage moult, are at least partly under photoperiodic control (Gwinner 1986, 2003; Dawson *et al.* 2001; Berthold *et al.* 2003; Dawson 2004; Sharp 2005; Vágási *et al.* 2010). The endogenous biochemical, physiological and behavioural circadian cycles which exist in most organisms are controlled by well-characterized biological clocks (Bell-Pedersen *et al.* 2005), and genetic polymorphism in circadian clock genes has been associated with behavioural variation in many organisms (Tauber & Kyriacou 2005; Johnsen *et al.* 2007; O'Malley & Banks 2008; Liedvogel *et al.* 2009).

Recent search for candidate genes has hinted at a core role of the *Clock* (*Circadian Locomotor Output Cycles Kaput)* gene in controlling the phenology of photoperiodic traits (Steinmeyer *et al.* 2009; see Mueller *et al.* 2011; O'Brien *et al.* 2013), although the evidence is inconsistent and contentious (see below). The circadian *Clock* gene encodes a protein (CLOCK) that heterodimerizes with the protein encoded by the *Bmal1* gene to produce a transcription factor which serves as a positive driver and as an 'output' signal in the core circadian oscillators (Panda *et al.* 2002; Fidler & Gwinner 2003; Ko & Takahashi 2006). *Clock* is highly conserved in vertebrates, but length polymorphism has been frequently documented particularly at the exonic carboxyl-terminal polyglutamine ($(CAG)_n$) repeat region (poly-Q) (Saleem *et al.* 2001; Fidler & Gwinner 2003; Johnsen *et al.* 2007). Poly-Q polymorphism has been found to be associated with variation in circadian rhythms in mammals (Gekakis *et al.* 1998) and amphibians (Hayasaka *et al.* 2002), by influencing the transcription-activating potential of the heterodimer complex. In birds, for which information on *Clock* variation has recently become available from several populations and species, length polymorphism at the poly-Q region is a largely prevailing condition among passerines (65/69 populations from 12 species; Chakarov *et al.* 2013; e.g. Johnsen *et al.* 2007; Liedvogel *et al.* 2009; Dor *et al.* 2011; Caprioli *et al.* 2012; Peterson *et al.* 2013). The extent of polymorphism, however, can vary among conspecific populations as well as between syntopic populations of closely related species (e.g. Johnsen *et al.* 2007; Liedvogel & Sheldon 2010; Dor *et al.* 2011; Kuhn *et al.* 2013).

A first line of evidence for an involvement of *Clock* poly-Q in the control of phenology of vertebrates stems from the observation of latitudinal variation in *Clock* allele frequencies: in two anadromous Pacific salmon species and in a mostly resident passerine bird, the blue tit (*Cyanistes caeruleus*), poly-Q allele length increases with latitude, implying that an adaptive delay in homologous phenophases with increasing latitude across populations is associated with longer poly-Q repeats (Johnsen *et al.* 2007; O'Malley & Banks 2008; O'Malley *et al.* 2010). However, studies of other Pacific salmon species and birds have failed to disclose consistent evidence for latitudinal variation at *Clock* poly-Q (Johnsen *et al.* 2007; O'Malley & Banks 2008; Dor *et al.* 2011, 2012; Kuhn *et al.* 2013; O'Brien *et al.* 2013). A second line of evidence is the covariation between *Clock* polymorphism and bird breeding phenology within populations: blue tit and barn swallow (*Hirundo rustica*) females with longer alleles show delayed breeding (Liedvogel *et al.* 2009; Caprioli *et al.* 2012). In both species, shorter alleles are at a fecundity advantage, and barn swallow individuals with shorter alleles also appear to be more viable (Liedvogel *et al.* 2009; Caprioli *et al.* 2012). Again, however, counterevidence exists from other passerines (Liedvogel & Sheldon 2010; Dor *et al.* 2011). Hence, the role of *Clock* poly-Q polymorphism in driving phenological variation in birds and other vertebrates is far from clear. In addition, the information on the association of individual variation in *Clock* and the phenology of migration and moult published to date is very limited. This is surprising because moult and migration are also known to be under photoperiodic control (see above). Barn swallows with rare long poly-Q alleles have been found to postpone moult compared to individuals with shorter alleles (Saino *et al.* 2013). Studies comparing migration strategies of different conspecific populations in relation to variation at *Clock* have failed to identify consistent relationships (Mueller *et al.* 2011; Peterson *et al.* 2013). In two subspecies groups identified within two species of the passerine

genus *Junco*, which shows considerable geographical variation in migratoriness, Peterson *et al.* (2013) found a positive relationship between migration distance and *Clock* allele length. Although this association was not consistent across all subspecies, these findings suggest an involvement of *Clock* also in the control of bird migration behaviour. To the best of our knowledge, however, no study has reported on variation in migration phenology according to *Clock* poly-Q polymorphism at the individual level.

Besides *Clock*, other regulatory genes are candidates as part of the mechanisms that control migration in birds (see e.g. Mueller *et al.* 2011). *Adcyap1* (*Adenylate Cyclase-Activating Polypeptide 1*) (Racz *et al.* 2008; Vaudry *et al.* 2009), for example, encodes pituitary adenylate cyclase-activating polypeptide (PACAP), a neuropeptide which is involved in diverse biological functions, such as the regulation of gastrointestinal functions (Mirabella *et al.* 2002) and neuroprotective potential (Dejda *et al.* 2005). PACAP stimulates release of melatonin and directly activates *Clock* in chickens (*Gallus gallus*) (Nagy & Csernus 2007), thereby affecting the signalling pathway involved in the circadian rhythms of the molecular clock (Racz *et al.* 2008). *Adcyap1* shows microsatellite polymorphism at the 3′-UTR region of the gene (Steinmeyer *et al.* 2009), which affects several processes at post-transcriptional level (Li *et al.* 2004; Riley & Krieger 2009). In the blackcap (*Sylvia atricapilla*), microsatellite polymorphism at *Adcyap1* has been shown to predict variation in migratory tendency among 14 geographical populations. In addition, polymorphism at *Adcyap1* predicts variation in migratory restlessness (*Zugunruhe*) in two blackcap populations, with longer alleles being positively associated with migratory activity (Mueller *et al.* 2011). Consistently with the results from some blackcap populations (Mueller *et al.* 2011), a positive relationship has been documented between migratory restlessness and *Adcyap1* allele length among individuals from one of two Oregon junco (*Junco hyemalis thurberi*) populations that were tested in captivity. However, no consistent variation in migratoriness was found among geographical populations belonging to different *Junco* subspecies (Peterson *et al.* 2013).

Hence, several pieces of evidence hint at an important role of *Clock* and *Adcyap1* in the control of bird phenology and migratoriness, but the information on the association between timing of migration and polymorphism at these candidate genes is lacking.

We therefore choose to investigate the association between timing of migration and polymorphism at *Clock* poly-Q and *Adcyap1* using four species of obligate, long-distance (trans-Saharan) migrants (nightingale, *Luscinia megarhynchos*; pied flycatcher, *Ficedula hypoleuca*; tree pipit, *Anthus trivialis*; whinchat, *Saxicola rubetra*) as models. We focused on these genes because of their role in controlling avian phenology and because of their potential reciprocal interactions. Timing of migration at the individual level was gauged from the date of capture during spring migration at a small Mediterranean island (Ventotene), off the western coast of southern Italy. These species were chosen because they are the only ones that can be sampled in large numbers at the study site, and second-year individuals (i.e. approximately 1-year-old individuals hatched during the spring/summer preceding that of capture) can be distinguished from older individuals. This allowed us to control for age effects on migration phenology (see also Methods). Individuals were sexed morphologically or by typing of the *CHD* gene (see Saino *et al.* 2010a), to account for sex differences in timing of migration. We predicted that individuals with a larger number of *Clock* poly-Q repeats averaged between the two alleles (i.e. individuals with larger mean allele length; *Clock*-MAL hereafter) would exhibit delayed migration. Dominance of the longer allele has been suggested for other genes that show poly-Q polymorphism (Ross 2002). We therefore also tested for allelic dominance. In addition, we tested whether any positive relationship between migration date and poly-Q polymorphism was stronger when we considered the length of the longer rather than of the shorter allele within each individual genotype. Previous evidence pointed at larger migratory distance and restlessness in individuals/populations with longer *Adcyap1* alleles (see above). Larger migration distance may translate, other things being equal, into later spring migration, due to longer migration journey (Rubolini *et al.* 2005; Liechti *et al.* 2015). We thus expected longer *Adcyap1* alleles to be associated with relatively delayed migration. However, we had no unequivocal predictions on any differential relationship between migration date and *Clock* or *Adcyap1* genotype in either sex or age classes.

By modulating circadian rhythms in physiology and behaviour (e.g. foraging activity), *Clock* and *Adcyap1* genotypes may influence general body condition. We therefore also tested whether *Clock* and *Adcyap1* genotype predicted total body mass and two indexes of condition, that is the extent of subcutaneous fat stores and the size of the pectoral muscle (Kaiser 1993; Jenni *et al.* 2000).

Methods

Study species

The four species we studied are small (ca. 10–28 g) passerines that breed in Eurasia and overwinter in sub-Saharan Africa (Cramp 1998). The pied flycatcher and

the whinchat are sexually dichromatic, whereas the tree pipit and the nightingale show no obvious sexual dichromatism and were therefore sexed molecularly (see Saino *et al.* 2008). Individuals hatched in the spring before that of capture (i.e. second-year individuals) can be distinguished from older individuals based on plumage (Svensson 1992; Jenni & Winkler 1994).

Field methods

We captured migrants at the ringing station run by ISPRA on Ventotene island (40°48′N–13°25′E), a very small (1.75 km^2) isle located ca. 50 km off the western coast of southern Italy, during spring 2013. Birds were trapped using mist nets in the period 22 March–27 May, which covers the entire migration period at Ventotene of the species that we studied (Spina *et al.* 1993; Saino *et al.* 2010a). Nets were operated according to standard protocols and constant capture effort, except under adverse meteorological conditions. The species we considered are exclusively migratory in coastal southern Italy (see Meschini & Frugis 1993), with the exception of the nightingale, which, however, does not breed on Ventotene (Fernando Spina, personal observation). Trans-Saharan migrants that pass southern Italy during spring migration mostly breed in Northern, Central and Eastern Europe (Jonzén *et al.* 2006; Spina & Volponi 2008). However, the relative timing of passage of the geographical populations is unknown. Migratory birds stopover at Ventotene during spring migration through the Mediterranean. Stopover of individual migrants on Ventotene, however, is rather short (mostly <2 days), as suggested by recapture data of individually marked birds and by radio-tracking data (Goymann *et al.* 2010; Tenan & Spina 2010). Given the spread of capture dates within species (see Supporting Information), we are therefore confident that any error in estimating migration date over Ventotene was negligible compared to among-individual variation in 'true' migration date.

Trapped birds were ringed and subjected to measurements of body mass and 3rd primary length ('wing length') (Spina *et al.* 1993). In addition, the extent of subcutaneous fat depots and the size of the pectoral muscle were recorded to serve as proxies for individual general state (see Kaiser 1993; Bairlein 1995; Jenni *et al.* 2000). Because wing length increases with breeding latitude in several European passerines (Cramp 1998; Peiró 2003; Evans *et al.* 2009; Tarka *et al.* 2010), wing length was entered as a covariate in the analyses as a proxy for latitude of the population of origin/destination. A small blood sample was taken using capillary tubes after puncturing the ulnar vein. In cases of large numbers of birds due to be processed simultaneously, we rather sampled 3–4 undertail coverts as a source of DNA. The samples were kept in a cool bag while in the field and then stored at −20 °C until use for analyses. When feathers, rather than blood, were collected, they were stored in plastic tubes containing 1 mL of 99% ethanol.

We aimed at sampling a maximum of 250 individuals per species. Based on the species-specific total number of birds trapped during the previous years (2006–2011), we therefore decided to sample every captured nightingale and tree pipit, every third pied flycatcher and two of every three whinchat, according to capture sequence (sample 'A'). In addition, for the pied flycatcher and whinchat, we also sampled all individuals that did not happen to be included in sample 'A', until the 50th that was captured since the start of the season (sample 'B'). This was decided to better resolve the genotypic frequencies among early migrants. The analyses of capture date in relation to *Clock* and *Adcyap1* genotype were conducted on samples A and B pooled (but see *Analysis of* Clock *and* Adcyap1 *genes* section for constraints on the sample for *Adcyap1*). Importantly, this approach did not bias the results because sampling was independent of sex, age and genotype. Thus, there was no reason to speculate that our sampling design inflated or, conversely, decreased the strength of any relationship between migration date and genotype.

In the analyses, date was expressed as Julian day. The temporal frequency distribution of the individuals that we considered (see Fig. S1, Supporting information) suggests that the entire main migration period at Ventotene of all the species we studied was effectively sampled. Indeed, the modes of the distributions were far from either extreme of the sampling periods, and capture frequencies at the temporal extremes of the sampling periods were small or equal to 0 for all species. Thus, the chances of missing rare alleles were not larger at either extreme compared to the core of the species-specific migration period.

Analysis of Clock *and* Adcyap1 *genes*

Total genomic DNA was extracted by alkaline lysis using 6 μL of blood in 100 μL of a 50 mM NaOH at 100 °C for 20 min. Extracted DNA was quantified using a spectrophotometer and diluted to a final concentration of 50–100 ng/μL.

Genomic DNA from feathers was extracted using the commercial 5 PRIME, ArchivePure DNA purification kit (5 PRIME, Hilden, Deutschland). The samples were initially rehydrated using a 0.9% NaCl solution for 10 min, with constant agitation. Only the basal portion of the calamus was taken and chopped into small pieces. The calamus fragments were placed into 300 μL

of the kit Cell Lysis Solution. Subsequently, the 'Protocol 14: DNA purification from 50 to 100 mg fresh or frozen solid tissue' of the ArchivePure DNA Purification Manual was followed adjusting quantities to 1/10 of the suggested volumes. Genomic DNA was finally resuspended in 30 μL of DNA hydration solution and stored at −20 °C. DNA extracted from feathers was in all cases very little; thus, the maximum quantity allowed in each PCR (5 μL) was used.

Polymorphism at the *Clock* gene and ultimately the number of glutamine residues present in the poly-Q domain of the CLOCK protein was determined by fragment analysis after PCR amplification as described in Caprioli *et al.* (2012). The genetic analysis of the microsatellite sequence present in the 3′-UTR of the *Adcyap1* gene was conducted using PCR primers reported in Steinmeyer *et al.* (2009), with the forward primer labelled either with 6-FAM or with TAMRA dyes, allowing mixing the amplification products from two individuals into one sample. PCRs were conducted in a final volume of 15 μL containing 100–200 ng of genomic DNA and 1 U Taq DNA polymerase (Genespin, Milan, Italy), a final concentration of 200 μM dNTPs, 0.3 μM of each of the forward and reverse primers, 1.5 mM $MgCl_2$ and 1× of supplied Taq buffer. PCR cycling was as follows: an initial denaturation at 95 °C for 3 min, then 35 cycles of 95 °C for 40 s, annealing at 57 °C for 40 s and 72 °C for 1 min, followed by a final extension of 72 °C for 7 min.

As no prior sequence information about the *Clock* and *Adcyap1* genes was available for the investigated species, the PCR-amplified fragment of few homozygous individuals per species was sequenced (BMR genomics, Padua, Italy). The sequencing allowed determining the number of glutamine residues present in the most frequent allele of the *Clock* gene. The identified *Clock* alleles were named according to the number of glutamine residues in the poly-Q region of the predicted mature protein, as Q_n, with *n* being the number of poly-Q repeats, whereas the length of *Adcyap1* alleles was expressed in base pairs. It should be noted that within the *Adcyap1* genomic sequence, two microsatellites are present: one is a dinucleotide repeat, whereas the other, although shorter, is a three base pair repeat. This resulted in the identification of alleles that could differ by one base pair, as previously reported by Steinmeyer *et al.* (2009). We also point out that we identified a 5-bp difference in size of the *Adcyap1* alleles between the DNA sequence and the estimated PCR fragment size. In particular, the sequenced fragment is larger than the genescan-estimated fragment size used in the analyses.

Because of previous evidence from diverse vertebrate species showing associations between phenology and *Clock* genotype (see Introduction), the study was originally designed to investigate the association between migration phenology and *Clock* gene. Recent evidence (Peterson *et al.* 2013) for an involvement of *Adcyap1* in the control of migration behaviour (though not phenology), besides that originally presented by Mueller *et al.* (2011), prompted us to also extend the analyses to *Adcyap1*, based on the genetic material originally extracted for *Clock* analyses. However, most of the PCRs from DNA extracted from feathers did not produce PCR fragments detectable by the automatic sequencer. This was probably due to the limiting available DNA quantity in the material extracted from feathers. In addition, PCR amplification efficiency was suboptimal because the *Adcyap1* primers we used were those designed for other species (see Steinmeyer *et al.* 2009). This was the cause of the smaller sample available for *Adcyap1* than for *Clock* analyses (see Results for species-specific sample sizes for either gene). However, the samples of individuals available for the analyses of either gene were similarly representative of the temporal spread of migrants within-individual species. Indeed, the frequency distributions of the samples used in the analyses of either gene did not differ (Kolmogorov–Smirnov test; $Z < 0.73$, $P > 0.66$ for all species).

Molecular sexing

Sexing from genomic DNA was performed using P2/P8 primers (Saino *et al.* 2008). As the use of P2/P8 primers on DNA extracted from feathers did not produce reliable results (due to a limited PCR amplification efficiency), a different set of primers, also from the *CHD* gene, was used. The following pair of primers, CHD1F 5′-TATCGTCAGTTTCCHTTTCAGGT-3′ and CHD1Rs 5′-CCTTTTATTGATCCATCAAGTC-3′, modified from Lee *et al.* (2010), was used. PCR amplification was performed in a 15-μL reaction mixture, with 1–5 ng of genomic DNA, 1× PCR buffer (5 PRIME), 1.5 mM of Mg^{2+}, 0.3 μL of each primer (stock 10 mM), 1.5 μL of dNTPs (stock 2 mM) and 1 U Taq DNA polymerase (5 PRIME). Amplification was conducted under the following conditions: 94 °C for 5 min, 35 cycles at 94 °C for 45 s, 55 °C for 45 s and 72 °C for 1 min and further extension at 72 °C for 7 min. The amplification products were separated on 2.5% agarose gel and visualized after ethidium bromide staining.

Statistical analyses

We mostly relied on linear models of migration date where age (second year or older) and sex were included as two-level factors, and wing length and mean allele length (MAL) of either candidate gene (hereafter *Clock*-MAL or, respectively, *Adcyap1*-MAL) as

covariates. In addition, we ran linear models including the length of the longer or, respectively, the length of the shorter *Clock* or *Adcyap1* allele as covariates. The two-way interactions among sex, age and allele length variables were initially included in the models. Models were subsequently simplified by excluding all the non-significant ($P > 0.05$) interaction terms in a single step, in order to reduce the risk of incurring type I statistical errors due to inflation of the number of tests. The effect size of individual terms is given as partial η^2. Fisher's exact test was used to test for within-species differences in *Clock* genotype frequencies between age classes or sexes. Cochran–Mantel–Haenszel test was used to investigate sex-dependent differences in *Clock* genotype frequencies while controlling for age effects and, reciprocally, for age-related differences while controlling for sex effects. Tests on sex- and age-dependent variation in *Adcyap1* frequencies were not practical because large polymorphism (see Results) reduced the frequency of individuals within any particular genotype group.

Following Liedvogel *et al.* (2009), we tested for allelic dominance by comparing migration date of birds that were heterozygous or homozygous for one of the two most common *Clock* alleles in an analysis of variance. Given the frequencies of homozygotes, this test was practical only among the Q_{11}/Q_{11}, Q_{11}/Q_{12} and Q_{12}/Q_{12} nightingale genotype groups. The frequency of the second most common homozygote genotype was too low to permit such test in the other species (see Results). Owing to large polymorphism (see Results), tests for allelic dominance with this approach were not practical for *Adcyap1*. However, in the absence of established a priori information on allelic dominance at *Clock* or *Adcyap1* depending on allele length in birds (see Liedvogel *et al.* 2009 for the single test on *Clock* in the blue tit), we repeated the analyses using three different indicators of genotype at these loci: length of the shorter allele, length of the longer allele and mean length (MAL; see also above) of the two alleles.

The covariation between morphological traits and allele length indicators was analysed in linear models where we controlled for the effects of sex and age. The analyses were then repeated while also controlling for the effect of migration date.

Sampling effort was stronger during early migration, implying that mean migration date estimates may not accurately reflect the 'true' (i.e. the absolute) migration parameters for each sex and age class of individuals. Because of this sampling design, we refrained from making inferences on sex- and age-dependent variation in migration dates. We entered sex and age as factors in the analyses merely to control for their potentially confounding effects on the association between migration date and genotype.

Throughout the text, parameter estimates are given with their standard error in parentheses. The α-level of statistical tests was set at $P < 0.05$. Statistical analyses were performed using SPSS 13.0, 15.0 or SAS 9.2 statistical packages.

Results

Migration date and Clock *polymorphism*

The sample included 729 individuals from the four species (see Table 1 for sex- and age-specific sample sizes). *Clock* polymorphism, in terms of both number of alleles and heterozygosity, varied among species. The gaps in the distributions of the alleles for the pied flycatcher and the whinchat, whereby allele Q_{14} and Q_{10}, respectively, were not recorded (see Table 1), suggest that at least one additional allele could exist in each of these species (see Kuhn *et al.* 2013). This is unlikely to have affected the results because these 'missing' alleles were likely to be rare, particularly in the whinchat, as suggested by the frequency of the two alleles adjacent to the missing one. The difference in the number of alleles between species could be due to sampling effects, because the two species with the largest number of alleles (including gaps) were those with the largest sample (Table 1).

Table 1 Number of individuals (*n*), number of *Clock* poly-Q alleles (*K*), poly-Q allele frequencies (Q_n), mean allele size (SE in parentheses) and observed heterozygosity (H_o). The information for all the individuals from samples A and B (see Methods) that could be successfully genotyped is reported

	n	*K*	Q_6	Q_7	Q_8	Q_9	Q_{10}	Q_{11}	Q_{12}	Q_{13}	Q_{14}	Q_{15}	Q_{16}	Mean allele size	H_o
Nightingale	151	5				0.043	0.030	0.440	0.480	0.007				11.38 (0.12)	0.550
Pied flycatcher	226	5					0.013	0.104	0.704	0.177		0.002		12.05 (0.03)	0.478
Tree pipit	144	5	0.007	0.063	0.035	0.854	0.042							8.86 (0.04)	0.250
Whinchat	208	7				0.007		0.005	0.026	0.012	0.938	0.005	0.007	13.90 (0.04)	0.125

Sample sizes by sex and age classes: nightingale: males-second year: 40, males-older: 34, females-second year: 39, females-older: 38; pied flycatcher: 57, 50, 77, 42; tree pipit: 27, 35, 38, 44; whinchat: 58, 26, 106, 18.

Genotype frequencies did not deviate from the Hardy–Weinberg equilibrium (Fisher's tests; $P > 0.11$ for all species). The frequency of the genotypes did not vary between age classes nor between males and females (Table 2). In addition, no significant variation in genotype frequencies between age classes existed after controlling for sex differences (Table 2). Reciprocally, no differences existed between males and females after controlling for age effects (Table 2).

Table 2 Fisher's exact tests (*P*-value) for age and sex effects and *P*-values (Cochran–Mantel–Haenszel statistic, degrees of freedom) for sex or age differences in genotypic frequencies (see Statistical analyses)

	Nightingale	Pied flycatcher	Tree pipit	Whinchat
Age	0.901	0.966	0.128	0.990
Sex	0.998	0.923	0.084	0.629
Age correcting for sex	0.779 (6.42, 10)	0.946 (1.68, 6)	0.258 (10.11, 8)	0.708 (3.77, 6)
Sex correcting for age	0.910 (4.71, 10)	0.873 (2.28, 6)	0.180 (11.40, 8)	0.408 (6.17, 6)

In linear models, migration date was not predicted by two-way interaction terms between *Clock*-MAL and sex or age ($P > 0.19$ in all cases). Simplified models where we controlled for the significant main (and interaction, if significant) effects of sex and age showed that in the tree pipit migration date increased with *Clock*-MAL (Table 3; Fig. 1). No statistically significant effect of *Clock*-MAL on migration date was observed for the other species (Table 3; Fig. 1). Moreover, these models showed that nightingales and pied flycatchers with longer wings migrated earlier, whereas no significant wing length effect on migration date emerged for the tree pipit and the whinchat (Table 3).

In linear models of migration date of individual species with the same design as for *Clock*-MAL (see above) where length of the longer *Clock* allele, rather than *Clock*-MAL, was entered as a predictor, length of the longer allele differentially predicted migration date of male and female tree pipits (sex by allele length interaction effect: $F_{1,138} = 12.30$, $P = 0.001$): the relationship was significantly positive in females ($t = 4.07$, $P < 0.001$) and nonsignificant in males (Fig. 2; Table S1,

Table 3 Linear models of migration date in relation to sex, age, mean *Clock* poly-Q allele length and wing length. η^2 is the effect size of the individual terms. Two-way interactions between mean allele length and sex or age were excluded when nonsignificant. Sample size for the nightingale is 150 rather than 151 (see Table 1) due to one missing datum for wing length. F: females; M: males; SY: second-year individuals

	F	d.f.	*P*	η^2	Estimated marginal means or coefficients (SE)	
Nightingale						
Sex	25.00	1,144	<0.001	0.148		
Age	16.15	1,144	<0.001	0.101		
Sex × Age*	4.36	1,144	0.039	0.029		
Mean allele length	0.13	1,144	0.718	0.001	0.414 (1.146)	
3rd primary length	5.17	1,144	0.024	0.035	−0.064 (0.028)	
Pied flycatcher						
Sex	56.76	1,221	<0.001	0.204	F: 118.7 (0.86)	M: 108.8 (0.91)
Age	47.82	1,221	<0.001	0.178	SY: 118.0 (0.77)	Older: 109.4 (0.94)
Mean allele length	0.00	1,221	0.976	0.000	−0.043 (1.456)	
3rd primary length	4.16	1,221	0.043	0.018	−0.084 (0.041)	
Tree pipit						
Sex	23.90	1,139	<0.001	0.147	F: 112.3 (1.02)	M: 104.1 (1.19)
Age	5.18	1,139	0.024	0.036	SY: 109.8 (1.06)	Older: 106.5 (0.97)
Mean allele length	6.07	1,139	0.015	0.042	4.057 (1.646)	
3rd primary length	0.14	1,139	0.708	0.001	−0.014 (0.036)	
Whinchat						
Sex	9.43	1,203	0.002	0.044	F: 121.2 (0.99)	M: 116.9 (1.11)
Age	3.48	1,203	0.063	0.017	SY: 120.5 (0.72)	Older: 117.6 (1.38)
Mean allele length	2.62	1,203	0.107	0.013	2.371 (1.465)	
3rd primary length	0.46	1,203	0.497	0.002	−0.030 (0.044)	

*Estimated marginal means: Second-year females: 109.8 (1.28), Second-year males: 105.7 (1.20), Older females: 107.3 (1.24), Older males: 98.1 (1.36).

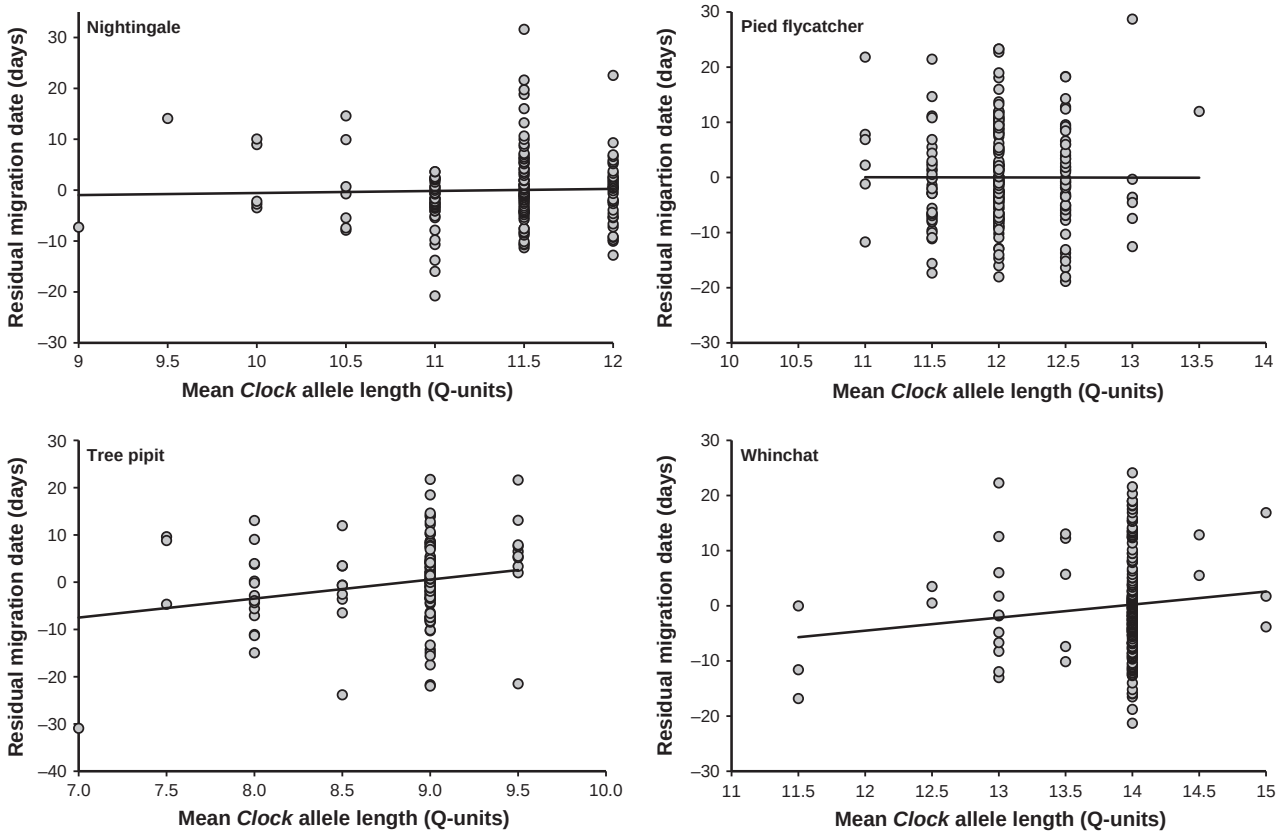


Fig. 1 Migration date in relation to mean *Clock* allele length (*Clock*-MAL). Migration date is expressed as residuals from a linear model where we controlled for sex, age (and also age by sex interaction effect in the nightingale) and wing length. The relationship is significantly positive for the tree pipit. Regression lines fitted to the residuals are shown merely to represent overall trends. For appropriate statistics, see Table 3.

Supporting information). In addition, the relationship was statistically significant and positive among nightingales ($F_{1,144} = 5.17$, $P = 0.024$, coefficient: 2.640 (1.161), $\eta^2 = 0.035$; Table S1, Supporting information) (Fig. 2). The length of the longer allele in the other species did not significantly predict migration date ($P > 0.13$ in both cases; Table S1, Supporting information). In models with the length of the shorter allele, no significant effect on migration date was observed in of the four species ($P > 0.07$ in all cases; Table S2, Supporting information).

The test for allelic dominance (see Statistical analyses) in nightingales that were homozygous or heterozygous for Q_{11} or Q_{12} partly supported dominance of the longer allele. In fact, migration date significantly varied among the three genotypic groups in a model where we controlled for the main effects of sex, age and wing length ($F_{2,123} = 4.36$, $P = 0.015$; estimated marginal mean (SE, *n*) for Q_{11}/Q_{11}: 101.9 (1.34, 31); Q_{11}/Q_{12}: 106.8 (0.94, 63); Q_{12}/Q_{12}: 104.8 (1.26, 36)). Thus, homozygotes for the shorter (Q_{11}) allele had the earliest migration date and migrated significantly earlier than heterozygote individuals (LSD test: $P = 0.004$). In addition, no significant difference in migration date existed between the heterozygote and the homozygote individuals for the longer allele (Q_{12}) ($P = 0.19$). However, no significant difference existed between migration date of either homozygote groups ($P = 0.14$). This approach (see also Liedvogel *et al.* 2009) for testing allelic dominance could only be taken for the two most frequent alleles in the nightingale because in no other species more than one common homozygote was recorded. In fact, the maximum frequency of the second most common homozygote recorded in the other species was 3.1% for Q_{13}/Q_{13} in the pied flycatcher.

Migration date and Adcyap1 *polymorphism*

The sample included 460 individuals from the four species (see Table 4 for sex- and age-specific sample sizes). As observed for *Clock*, polymorphism at *Adcyap1* varied among species, with the number of alleles ranging from 7 to 13. The observed heterozygosity was relatively low

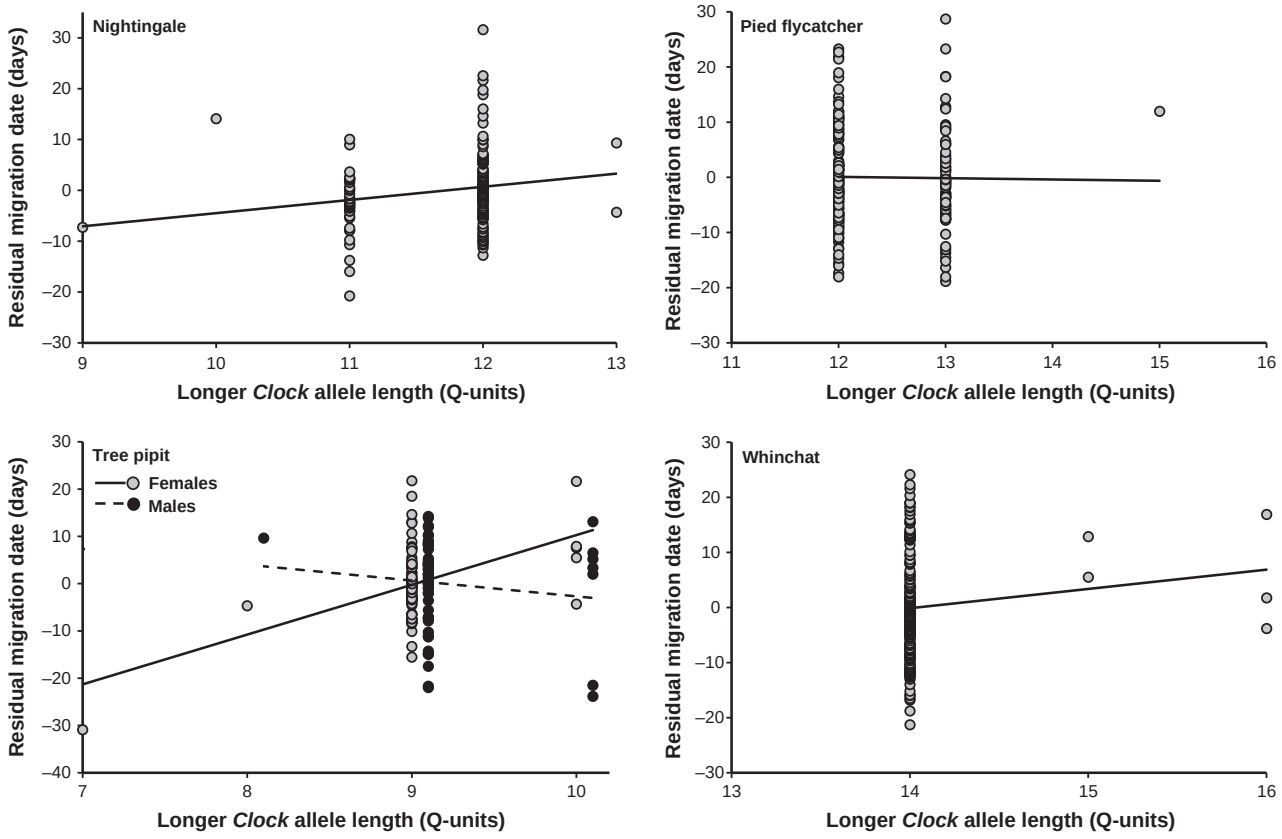


Fig. 2 Migration date in relation to length of the longer *Clock* allele. Migration date is expressed as residuals from a linear model where we controlled for sex, age (and also age by sex interaction effect in the nightingale) and wing length effects. The relationship is significantly positive for the nightingale. For the tree pipit, the slopes of the relationships were significantly different for either sex, and the relationship for females was significantly positive. Regression lines fitted to the residuals are shown merely to represent the overall trends. For appropriate statistics, see Table 3.

in the nightingale, the species with the smallest number of alleles, as compared to the other species (Table 4).

Genotype frequencies did not deviate from the Hardy–Weinberg equilibrium (Fisher's tests; $P > 0.57$ for all species). Large polymorphism and heterozygosity, and relatively small sample size per each sex by age group, prevented an analysis of sex- and age-dependent variation of genotype frequencies (see Statistical analyses).

Linear models showed that migration date was not predicted by *Adcyap1*-MAL while controlling for the effects of sex, age and wing length ($P > 0.27$ for all species; see Table 3 for sample sizes; see also Table S3 and Fig. S2, Supporting information). The largest effect size ($\eta^2 = 0.013$) for *Adcyap1*-MAL, which was observed in the pied flycatcher, was associated with a negative effect, that is opposite to the expectation. The two-way interaction terms among sex, age and *Adcyap1*-MAL were removed from the models as their effect never attained significance ($P > 0.15$ in all cases). Models of migration date including the length of the longer allele disclosed a significant sex by allele length interaction in the pied flycatcher ($F_{1,89} = 4.90$, $P = 0.029$; Table S4, Supporting information). The relationship within either sex was nonsignificant. In addition, the stronger relationship, which was observed among males, was negative and thus opposite to the expectation. Hence, this effect will not be discussed further. No other main or interaction effects were observed in models of migration date in relation to longer or, respectively, shorter *Adcyap1* allele length in any species (Tables S4 and S5).

Covariation of morphology with Clock *and* Adcyap1 *genotype*

We found a significant negative association between wing length and *Clock*-MAL in the pied flycatcher ($F_{1,222} = 3.97$, $P = 0.048$, coefficient: −4.668 (2.343)) (Fig. 3), but not in the other species ($P > 0.05$ in all cases) in models where we controlled for the effect of sex and age. The statistically nonsignificant sex by age interaction ($P > 0.33$ in all cases) was excluded from

Table 4 Number of individuals (*n*), number of *Adcyap1* alleles (*K*), frequencies of alleles of different length, mean allele length (SE) and observed heterozygosity (H_o). The information for all the individuals from samples A and B (see Methods) that could be successfully genotyped is reported

	n	*K*	157	160	161	163	164	165	166	167	168	169	170	171	172	173	174	175	176	178	180	182	183	184	186	187	188	Mean allele size	H_o
Nightingale	99	7	0.010	0.010	0.046	0.707		0.126		0.086		0.015																163.5 (0.12)	0.455
Pied flycatcher	95	11					0.005				0.005						0.026		0.337	0.163	0.390	0.042	0.005	0.017	0.005		0.005	178.3 (0.19)	0.695
Tree pipit	97	12					0.031		0.088	0.031	0.253	0.016	0.330	0.067	0.113	0.046	0.016	0.005	0.005									169.4 (0.16)	0.763
Whinchat	169	13			0.012	0.044	0.030	0.036	0.018	0.228	0.012	0.376	0.012	0.183	0.015	0.033											0.003	168.4 (0.14)	0.787

Sample sizes by sex and age classes: nightingale: males-second year: 25, males-older: 23, females-second year: 25, females-older: 26; pied flycatcher: 26, 23, 30, 16; tree pipit: 17, 23, 26, 31; whinchat: 50, 21, 82, 16.

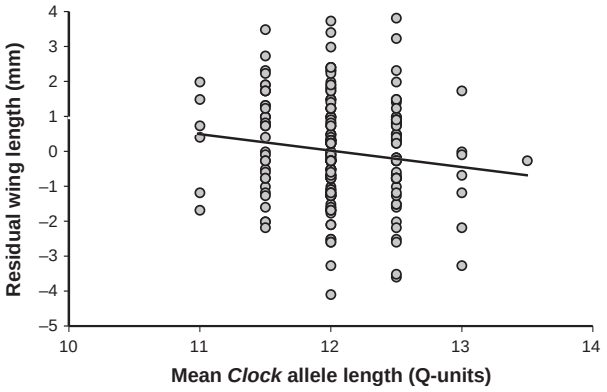


Fig. 3 Wing length in relation to mean *Clock* allele length in the pied flycatcher. Wing length is expressed as residuals from a model controlling for sex and age effects. The relationship is negative and statistically significant.

models on all species. Analyses where *Clock* longer allele length, rather than *Clock*-MAL, was considered did not confirm the significant relationship between wing length and *Clock* in the pied flycatcher ($F_{1,222} = 2.60$, $P = 0.109$), but disclosed a significant negative association between wing length and longer allele length in the whinchat [$F_{1,204} = 8.36$, $P = 0.004$, coefficient: −10.997 (3.803)]. Analyses where *Clock* shorter allele length was considered did not show any significant effect on wing length in any species ($P > 0.14$ in all cases). Fat stores and body mass (also after controlling for wing length as a proxy for body size) were not predicted by *Clock*-MAL nor by *Clock* shorter or, respectively, longer allele length in any species ($P > 0.11$). Pectoral muscle score was significantly and positively predicted by *Clock*-MAL [$F_{1,204} = 4.78$, $P = 0.030$, coefficient: 0.195 (0.089)] and *Clock* shorter allele length [$F_{1,204} = 3.97$, $P = 0.048$, coefficient: 0.095 (0.048)] in the whinchat, whereas no significant association was observed for all the *Clock* allele length indicators in the other species ($P > 0.05$). *Clock*-MAL or *Clock* shorter or longer allele length did not significantly predict body mass, and pectoral muscle and fat stores ($P > 0.06$ in all cases) while controlling for capture date (besides sex and age) in any species.

The analyses of wing length, fat stores, pectoral muscle score and body mass, where we controlled for sex and age effects, did not reveal significant covariation with the indicators of *Adcyap1* allele length in any species, with the only exception for a negative association of pectoral muscle score with *Adcyap1* shorter allele length in the whinchat [$F_{1,165} = 4.99$, $P = 0.027$, coefficient: −0.039 (0.017); $P > 0.06$ in the other cases]. However, no significant covariation between morphological variables and *Adcyap1* allele length indicators persisted in the models where we

controlled for migration date, besides sex and age effects ($P > 0.05$ in all cases).

Discussion

We studied the phenology of spring migration of four passerine species and found evidence that timing of migration covaries with the number of poly-Q repeats at the *Clock* gene. This association was due to the number of poly-Q repeats of the 'longer' rather than of the 'shorter' allele. Moreover, polymorphism at another candidate gene for the control of migratory behaviour, *Adcyap1*, did not predict migration date.

Polymorphism at Clock *and migration date*

Polymorphism at *Clock* poly-Q has been documented in most bird populations/species studied so far (Chakarov *et al.* 2013). The four species we studied mirror the large diversity in *Clock* polymorphism among passerines (Johnsen *et al.* 2007; Liedvogel *et al.* 2009; Dor *et al.* 2011; Caprioli *et al.* 2012). The only species for which data on *Clock* are specifically available for comparison is the pied flycatcher: 3–5 *Clock* alleles have been reported in 6 European populations (Spain to Norway), with a total of 5 alleles (Q_{10}–Q_{14}). The present data thus expand allelic diversity to a rare longer allele (Q_{15}) and suggest that Q_{14} was not reported here due to sampling effects. The heterozygosity observed in the present study (0.478) is intermediate to those reported previously (0.412–0.563; Kuhn *et al.* 2013), suggesting that at least the pied flycatcher population(s) we sampled is representative of genetic variation at *Clock* at continental scale.

Timing of migration was significantly later for tree pipits with relatively large *Clock*-MAL. Models of migration date in relation to *Clock* longer allele length disclosed a significant difference in the relationship between the two sexes. Migration date positively covaried with longer *Clock* allele length in females but not in males in the tree pipit. In addition, these models disclosed a significant positive relationship between migration date and longer *Clock* allele length in the nightingale. We have no unequivocal explanation for sex dependence of migration phenology in relation to longer *Clock* allele length in the tree pipit, as sex dependence in *Clock* photoperiodic responses has not been previously documented. Conversely, the statistical effect of shorter allele length was always nonsignificant. A test for allelic dominance (see also Liedvogel *et al.* 2009) where we compared migration date of two homozygote (Q_{11}/Q_{11} and Q_{12}/Q_{12}) groups and of the heterozygote group Q_{11}/Q_{12} in the nightingale (the only species for which this approach was feasible) partly supported dominance of the longer allele: migration date of Q_{11}/Q_{12} individuals was significantly later than for Q_{11}/Q_{11} individuals and not significantly different from that of Q_{12}/Q_{12} individuals. These findings are consistent with genetic dominance of the longer allele in most of human disorders which are caused by the expansion of $(CAG)_n$ repeats coding for polyglutamine protein segments, such as Spinal and Bulbar Muscular Atrophy, Huntington's disease, DentatoRubral and PallidoLuysian Atrophy and Spino-Cerebellar Ataxia (Ross 2002). Poly-Q tracts function as protein interaction domains and regulate various signalling processes at the transcriptional level. The presence of more glutamine residues affects transcription, either by interacting directly with DNA or by competing with polyglutamine-containing transcription factors (e.g. Riley & Orr 2006; Hands *et al.* 2008), as demonstrated in many polyglutamine-containing proteins. We have no evidence suggesting that the *Clock* alleles we identified are anything but fully functional. Although the number of glutamine residues is far from that detected in the disease-causing proteins in human poly-Q disorders, it can be argued that as the poly-Q tract determines protein interactions, longer alleles are dominant over shorter ones.

Sources of variation in the Clock*–migration date relationship*

The size of the effects (η^2) of *Clock*-MAL and also of *Clock* longer allele length on migration date was generally small (see e.g. Mueller *et al.* 2011). Migration data obtained at bird ringing stations are inherently affected by large random 'noise' due to interference of local and 'upstream' meteorological conditions on actual migration flux as well as on stopover/capture probability (Saino *et al.* 2010b). While such nuisance effects could partly confound existing associations between migration date and *Clock* genotype, we deem it unlikely that they produced spurious evidence for the associations we observed. In fact, this effect would require that individuals with longer alleles were differentially more likely to stopover at Ventotene late as compared to early in the migration period. While any such association between *Clock* and stopover probability would be interesting on its own, this interpretation is not supported by the present data. This is the case because there were generally no relationships between *Clock*-MAL and proxies of body condition, which could affect stopover decisions. Yet, the causes of the considerable, though statistically nonsignificant variation among species in the slope of the association between migration date and *Clock*-MAL (Table 3), deserve consideration. One possibility is that such variation partly arises if 'noise' in capture data differs among species because, for example,

some species are more sensitive to weather *en route* (Saino *et al.* 2010b). If this is the case, interspecific differences in the interference of weather on landing decisions may differentially affect the relationship we observed between migration date and *Clock*-MAL, producing larger confounds for some species than others. Alternatively, interspecific variation in the relationship between migration date and *Clock*-MAL could be the genuine reflection of environmental 'masking' effects on endogenous rhythms that differ among species. Species are believed to differ in the strength of circannual clock and, thus, in the extent to which environmental inputs affect the expression of such rhythms (Dawson *et al.* 2001; Bradshaw & Holzapfel 2007; Helm 2009). Species which are less susceptible to external influences may display stronger association between migration date and *Clock*, producing interspecific variation in slopes and effect sizes. Unfortunately, no information is available on the strength of the circannual clock for any of the species we considered.

Proximate mechanisms of Clock*–migration date association*

A proximate explanation for the covariation between migration date and *Clock* may consist in differences in timing of migration among geographical populations that differ in genetic composition at *Clock*. If *Clock*-mediated response to photoperiod orchestrates migration phenology, latitudinal variation in *Clock* frequencies might be expected (Johnsen *et al.* 2007). Trans-Saharan migrants that pass southern Italy during spring migration mostly breed in Northern, Central and Eastern Europe (Jonzén *et al.* 2006; Spina & Volponi 2008). However, marked confounding effects of latitudinal variation in *Clock* are very unlikely because most bird species studied so far do not show geographical population structuring at *Clock* (e.g. Dor *et al.* 2011; but see Johnsen *et al.* 2007). Because latitudinal clines exist in wing morphology across Europe, whereby wing pointedness and length increase with latitude (Cramp 1998; Peiró 2003; Evans *et al.* 2009; Tarka *et al.* 2010), in the analyses, we included wing length as a covariate. This partly accounted for variation in geographical origin/destination. Moreover, we found no significant positive covariation between *Clock* genotype and wing length, which would be expected if both wing length and *Clock* allele length increased with latitude. The negative sign of this association in the pied flycatcher (*Clock*-MAL) and in the whinchat (*Clock* longer allele) was opposite to that expected. In combination with the lack of latitudinal population structure at *Clock* (Dor *et al.* 2011; Kuhn *et al.* 2013), this suggests no interference of latitudinal variation on the results. To sum up, poor geographical structuring of bird populations at *Clock*, the evidence that *Clock* allele length is not positively associated with wing length and the fact that genotype–phenotype associations persisted when controlling for wing length suggest that the present findings are not the spurious results of the confounding effect of latitude of origin/destination.

Mechanistic links between *Clock* variation and migration phenology may operate via timing of departure from the sub-Saharan wintering grounds or during migration. It has been argued that photoperiodic responses in equatorial regions are hindered by small circannual variation in photoperiod, implying no role of *Clock* in departure decisions from equatorial Africa. On the other hand, even minor changes in photoperiod at low latitudes elicit physiological responses in birds and other vertebrates (Hau *et al.* 1998; see also Burns 1985; Wayne & Rissman 1991). In addition, *Clock* genotype was found to be associated with the phenology of moult at the tropics (Saino *et al.* 2013): in the barn swallow, winter moult close to the Equator was delayed in individuals with a rare long *Clock* allele. Hence, *Clock* may indeed control photoperiodic responses also in Equatorial areas. Alternatively, *Clock* may affect the duration of staging at stopover sites in North Africa (around 35°N), before crossing of the Mediterranean in early spring. Finally, an additional alternative interpretation is that apparent *Clock* effects at tropical latitudes represent carry-over effects of the influence that *Clock* had on previous phenological events, including breeding and/or autumn migration.

Implications of the Clock*–migration date association*

Long-distance migrants are believed to have evolved hardwired phenological control of migration. These are expected because they need to set their circannual life cycle without having any cue to the current and future conditions at either opposite extreme of their migratory journey (Pulido & Widmer 2005; Pulido 2007; but see Saino & Ambrosini 2008). Some evidence suggests that lower phenotypic variation in migratory traits observed among long-distance migrants is due to environmental canalization functioning to adaptively buffer environmental variation (Pulido & Widmer 2005). Evidence for variation in timing of migration being associated with genetic variation at candidate genes in long-distance migrants is scant (see Gwinner 1996; Knudsen *et al.* 2011; Mueller *et al.* 2011; Peterson *et al.* 2013). Our present findings are entirely novel in that they hint at an association between timing of migration and *Clock*. The present results therefore have a bearing for the interpretation of the mechanisms of response of birds to climate change (Møller *et al.* 2010). Populations can accommodate the

consequences of climate change by phenotypically plastic and/or micro-evolutionary responses (Dunn & Winkler 2010). Phenotypic plasticity should result in phenological shifts most readily. Conversely, micro-evolutionary change is mainly expected to occur whenever selection exceeds the range of phenological change which is allowed by phenotypic plasticity and to proceed at lower pace in promoting adaptation to changing conditions (Visser *et al.* 2010). Environmental canalization of migratory traits may hinder the ability of long-distance migrants to adjust their migration phenology to rapid advancement in the spring phenophases at their breeding grounds (Rubolini *et al.* 2007, 2010; Knudsen *et al.* 2011; but see Jonzén *et al.* 2006) and cause marked population decline in long-distance Afro-Palearctic migratory passerines (Sanderson *et al.* 2006; Møller *et al.* 2008; Both *et al.* 2009; Saino *et al.* 2011). This is ultimately the case because photoperiodic cues, which are not affected by climate change, lose their predictive value as to the optimal time to migrate towards the breeding grounds (Visser *et al.* 2010). In the present study, a unit decrease in the number of repeats at *Clock* poly-Q was associated with an advancement in migration date ranging from around 0 to ca. 4 days (Table 3). Over the last decades, warming of spring in continental Europe has led to an advancement of spring phenophases in the order of 1–2 days per decade (Schwartz *et al.* 2006). Thus, the difference in migration date putatively caused by a unit difference in the number of *Clock* poly-Q repeats roughly corresponds to the shift in spring phenophases that has occurred per each decade over the last half-century. We should therefore expect selection to have occurred for shorter *Clock* alleles (Dunn & Winkler 2010; Lehikoinen & Sparks 2010). Retrospective analysis of micro-evolution of *Clock* is obviously a difficult task in the absence of ad hoc sampling to estimate allele frequencies in past and extant populations. Recently, such an analysis has been attempted in the pied flycatcher over a century: no evidence of change in *Clock* frequencies was recorded, suggesting that selection for early arrival has not caused micro-evolutionary change at *Clock* (Kuhn *et al.* 2013). It should be noticed, however, that any inadvertent bias towards sampling of individuals with *Clock*-based early or, conversely, late arrival in the past and/or recent populations may have confounded these analyses. Hence, we feel that micro-evolutionary change at *Clock* in association with current climate change awaits more empirical tests.

Polymorphism at Adcyap1 *and migration date*

No statistically significant association between migration date and *Adcyap1* polymorphism could be detected. The analyses of phenology in relation to *Adcyap1* polymorphism were based on a smaller sample compared to those of phenological variation in relation to *Clock* polymorphism. Lack of significant relationships between phenology and *Adcyap1* polymorphism could therefore be caused by reduced power of the statistical tests. This interpretation is unlikely to apply, however, because the size of the effects *Adcyap1* allele length on migration date was very small. In addition, the strongest observed relationship within any species was associated with a negative, rather than a positive effect of *Adcyap1* allele length on migration date. Previous studies have provided evidence for an association between migratory distance and restlessness at the population or at the individual level and *Adcyap1* (Mueller *et al.* 2011; Peterson *et al.* 2013). Thus, our study suggests that *Adcyap1* may only be involved in the control of those migration traits, but not of timing of migration. Admittedly, the marginally significant negative association between pectoral muscle score and *Adcyap1* shorter allele length has no obvious interpretation. We may speculate that the PACAP protein coded by the *Adcyap1* gene might have a weak pleiotropic effect, not necessarily associated with migration.

Thus, this study adds a novel facet to the growing body of studies which are providing mixed evidence for a role of *Clock* in the control of phenological variation in vertebrates. We showed for the first time that spring migration date covaries with *Clock* genotype in migratory birds. These results disclose a new perspective on the ecological and evolutionary study of phenological response of birds to current climate change.

Acknowledgements

We warmly thank the Riserva Naturale Isole di Ventotene e Santo Stefano for logistic support. We also thank many volunteers ringers and field assistants that helped collecting the data on Ventotene, and especially M. Caffi, S. Fasano, S. Fusetti, L. Ilahiane, V. Longoni, E. Mancuso, P. Micheloni, D. Piacentini, and M. Romano. We are grateful to three anonymous reviewes for their most constructive comments. Results from the Progetto Piccole Isole (INFS-ISPRA): study No. 54.

References

Bairlein F (1995) *European-African Songbird Migration Network. Manual of field Methods*. IFV, Wilhelmshaven, Germany.

Beacham TD, Murray CB (1988) Influence of photoperiod and temperature on timing of sexual maturity of pink salmon (*Oncorhynchus gorbucha*). *Canadian Journal of Zoology*, **66**, 1729–1732.

Bell-Pedersen D, Cassone VM, Earnest DJ *et al.* (2005) Circadian rhythms from multiple oscillators: lessons from diverse organisms. *Nature Reviews Genetics*, **6**, 544–556.

Berthold P, Gwinner E, Sonnenschein E (2003) *Avian Migration*. Springer, Berlin.

Both C, Bouwhuis S, Lessells CM *et al.* (2006) Climate change and population declines in a long-distance migratory bird. *Nature*, **441**, 81–83.

Both C, van Asch M, Bijlsma LG *et al.* (2009) Climate change and unequal phenological changes across four trophic levels: constraints or adaptations? *Journal of Animal Ecology*, **78**, 73–83.

Bradshaw WE, Holzapfel CM (2007) Evolution of animal photoperiodism. *Annual Review of Ecology, Evolution, and Systematics*, **38**, 1–25.

Bradshaw WE, Holzapfel CM (2010) Light, time, and the physiology of biotic response to rapid climate change in animals. *Annual Review of Physiology*, **72**, 149–166.

Burns JR (1985) The effect of low-latitude photoperiods on the reproduction of female and male *Poeciliopsis gracilis* and *Poecilia sphenops*. *Copeia*, **4**, 961–965.

Caprioli M, Ambrosini R, Boncoraglio G *et al.* (2012) *Clock* gene variation is associated with breeding phenology and maybe under directional selection in the migratory barn swallow (*Hirundo rustica*). *PLoS One*, **7**, e35140.

Chakarov N, Jonker RM, Boerner M *et al.* (2013) Variation at candidate genes correlates with timing of dispersal and plumage morph in a sedentary bird of prey. *Molecular Ecology*, **22**, 5430–5440.

Clutton-Brock TH (1988) *Reproductive Success*. University of Chicago Press, Chicago.

Cramp S (1998) *The Complete Birds of the Western Palearctic on CD-ROM*. Oxford University Press, Oxford.

Dawson A (2004) The effects of delaying the start of moult on the duration of moult, primary feather growth rates and feather mass in Common Starlings *Sturnus vulgaris*. *Ibis*, **3**, 493–500.

Dawson A, King VM, Bentley GE *et al.* (2001) Photoperiodic control of seasonality in birds. *Journal of Biological Rhythms*, **16**, 365–380.

Dejda A, Sokolowska P, Nowak JZ (2005) Neuroprotective potential of three neuropeptides PACAP, VIP and PHI. *Pharmacological Reports*, **57**, 307–320.

Dor R, Lovette IJ, Safran RJ *et al.* (2011) Low variation in the polymorphic *Clock* Poly-Q region despite population genetic structure across Barn Swallow (*Hirundo rustica*) populations. *PLoS One*, **6**, e28843.

Dor R, Cooper CB, Lovette IJ *et al.* (2012) *Clock* gene variation in *Tachycineta* swallows. *Ecology and Evolution*, **2**, 95–105.

Dunn PO, Winkler DW (2010) Effects of climate change on timing of breeding and reproductive success in birds. In: *Effects of Climate Change on Birds* (eds Møller AP, Fiedler W, Berthold P), pp. 129–147. Oxford University Press, Oxford.

Ellegren H, Sheldon BC (2008) Genetic basis for fitness differences in natural populations. *Nature*, **452**, 169–175.

Evans K, Gaston KJ, Sharp SP *et al.* (2009) The effect of urbanisation on avian morphology and latitudinal gradients in body size. *Oikos*, **118**, 251–259.

Fidler AE, Gwinner E (2003) Comparative analysis of avian BMAL1 and CLOCK protein sequences: a search for features associated with owl nocturnal behaviour. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, **136**, 861–874.

Foster RG, Kreitzman L (2009) *Seasons of Life: the Biological Rhythms That Enable Living Things to Thrive and Survive*. Yale University Press, New Haven, Connecticut.

Gekakis N, Staknis D, Nguyen HB *et al.* (1998) Role of the CLOCK protein in the mammalian circadian mechanism. *Science*, **280**, 1564–1569.

Goymann W, Spina F, Ferri A, Fusani L (2010) Body fat influences departure from stopover sites in migratory birds: evidence from whole-island telemetry. *Biology Letters*, **6**, 478–481.

Gwinner E (1986) Circadian rhythms in the control of avian migration. *Advances in the Study of Behavior*, **16**, 191–228.

Gwinner E (1996) Circadian and circannual programmes in avian migration. *The Journal of Experimental Biology*, **199**, 39–48.

Gwinner E (2003) Circannual rhythms in birds. *Current Opinion in Neurobiology*, **13**, 770–778.

Hands S, Sinadinos C, Wyttenbach A (2008) Polyglutamine gene function and dysfunction in the ageing brain. *Biochimica et Biophysica Acta-Gene Regulatory Mechanisms*, **1779**, 507–521.

Hau M, Wikelski M, Wingfield JC (1998) A Neotropical forest bird can measure the slight changes in tropical photoperiod. *Proceedings of the Royal Society B*, **265**, 89–95.

Hayasaka N, LaRue SI, Green CB (2002) In vivo disruption of *Xenopus* CLOCK in the retinal photoreceptor cells abolishes circadian melatonin rhythmicity without affecting its production levels. *Journal of Neuroscience*, **22**, 1600–1607.

Helm B (2009) Temporal coordination of life cycle stages: an avian chronobiology perspective. *Integrative and Comparative Biology*, **49**, E73.

Helm B, Ben-Shlomo R, Sheriff MJ *et al.* (2013) Annual rhythms that underlie phenology: biological time-keeping meets environmental change. *Proceedings of the Royal Society B: Biological Sciences*, **280**, 1–10.

Jenni L, Winkler R (1994) *Moult and Ageing of European Passerines*. Academic Press, London.

Jenni L, Jenni-Eiermann S, Spina F *et al.* (2000) Regulation of protein breakdown and adrenocortical response to stress in birds during migratory flight. *American Journal of Physiology*, **278**, R1182–R1189.

Johnsen A, Fiedler AE, Kuhn S *et al.* (2007) Avian *Clock* gene polymorphism: evidence for a latitudinal cline in allele frequencies. *Molecular Ecology*, **16**, 4867–4880.

Jonzén N, Piacentini D, Andersson A *et al.* (2006) The timing of spring migration in trans-Saharan migrants: a comparison between Ottenby, Sweden and Capri, Italy. *Ornis Svecica*, **16**, 27–33.

Kaiser A (1993) A new multi-category classification of subcutaneous fat deposits of songbirds. *Journal of Field Ornithology*, **64**, 246–255.

Knudsen E, Linden A, Both C *et al.* (2011) Challenging claims in the study of migratory birds and climate change. *Biological Reviews*, **86**, 928–946.

Ko CH, Takahashi JS (2006) Molecular components of the mammalian circadian clock. *Human Molecular Genetics*, **15**, 271–277.

Kuhn K, Schwenk K, Both C *et al.* (2013) Differentiation in neutral genes and a candidate gene in the pied flycatcher: using biological archives to track global climate change. *Ecology and Evolution*, **3**, 4799–4814.

Lee JCI, Tsai L-C, Hwa P-Y *et al.* (2010) A novel strategy for avian species and gender identification using the CHD gene. *Molecular and Cellular Probes*, **24**, 27–31.

Lehikoinen E, Sparks TH (2010) Changes in migration. In: *Effects of Climate Change on Birds* (eds Møller AP, Fiedler W, Berthold P), pp. 89–112. Oxford University Press, Oxford.

Li YC, Korol AB, Fahima T, Nevo E (2004) Microsatellites within genes: structure, function, and evolution. *Molecular Biology and Evolution*, **21**, 991–1007.

Liechti F, Scandolara C, Rubolini D *et al.* (2015) Timing of migration and residence areas during the non-breeding period of barn swallows *Hirundo rustica* in relation to sex and population. *Journal of Avian Biology*, **45**, 1–12.

Liedvogel M, Sheldon BC (2010) Low variability and absence of phenotypic correlates of *Clock* gene variation in a great tit *Parus major* population. *Journal of Avian Biology*, **41**, 543–550.

Liedvogel M, Szulkin M, Knowles SCL *et al.* (2009) Phenotypic correlates of Clock gene variation in a wild blue tit population: evidence for a role in seasonal timing of reproduction. *Molecular Ecology*, **18**, 2444–2456.

Meschini E, Frugis S (1993) *Atlante degli uccelli nidificanti in Italia*. Istituto Nazionale di Biologia della Selvaggina, Ozzano Emilia.

Mirabella N, Squillacioti C, Colitti M *et al.* (2002) Pituitary adenylate cyclase activating peptide (PACAP) immunoreactivity and mRNA expression in the duck gastrointestinal tract. *Cell and Tissue Research*, **308**, 347–359.

Møller AP, Rubolini D, Lehikoinen E (2008) Populations of migratory bird species that did not show a phenological response to climate change are declining. *Proceedings of the National Academy of Sciences*, **105**, 16195–16200.

Møller AP, Fiedler W, Berthold P (eds) (2010) *Effects of Climate Change on Birds*. Oxford University Press, Oxford.

Mueller J, Pulido F, Kempenaers B (2011) Identification of a gene associated with avian migratory behaviour. *Proceedings of the Royal Society B*, **278**, 2848–2856.

Nagy AD, Csernus VJ (2007) The role of PACAP in the control of circadian expression of clock genes in the chicken pineal gland. *Peptides*, **28**, 1767–1774.

Nussey DH, Postma E, Gienapp P *et al.* (2005) Selection on heritable phenotypic plasticity in a wild bird population. *Science*, **310**, 304–306.

O'Brien C, Unruh L, Zimmerman C *et al.* (2013) Geography of the circadian gene clock and photoperiodic response in western North American populations of the three-spined stickleback *Gasterosteus aculeatus*. *Journal of Fish Biology*, **82**, 827–839.

O'Malley KG, Banks MA (2008) A latitudinal cline in the Chinook salmon (*Oncorhynchus tshawytscha*) *Clock* gene: evidence for selection on PolyQ length variants. *Proceedings of the Royal Society B*, **275**, 2813–2821.

O'Malley KG, Ford MJ, Hard JJ (2010) *Clock* polymorphism in Pacific salmon: evidence for variable selection along a latitudinal gradient. *Proceedings of the Royal Society B*, **227**, 3703–3714.

Panda S, Hogenesch JB, Kay SA (2002) Circadian rhythms from flies to human. *Nature*, **417**, 329–335.

Peiró G (2003) Intraspecific variation in the wing shape of the long-distance migrant reed warbler *Acrocephalus scirpaceus*: effects of age and distance of migration. *Ardeola*, **50**, 31–37.

Peterson MP, Abolins-Abols M, Atwell JW *et al.* (2013) Variation in candidate genes CLOCK and ADCYAP1 does not consistently predict differences in migratory behavior in the songbird genus *Junco*. *F1000Research*, **2**, 115.

Pulido F (2007) The genetics and evolution of avian migration. *BioScience*, **57**, 165–174.

Pulido F, Widmer M (2005) Are long-distance migrants constrained in their evolutionary response to environmental change? Causes of variation in the timing of autumn migration in a blackcap (*S. atricapilla*) and two garden warbler (*Sylvia borin*) populations. *Annals of the New York Academy of Sciences*, **1046**, 228–241.

Racz B, Horvath G, Faluhelyi N *et al.* (2008) Effects of PACAP on the circadian changes of signaling pathways in chicken pinealocytes. *Journal of Molecular Neuroscience*, **36**, 220–226.

Riley DE, Krieger JN (2009) UTR dinucleotide simple sequence repeat evolution exhibits recurring patterns including regulatory sequence motif replacements. *Gene*, **429**, 80–86.

Riley BE, Orr HT (2006) Polyglutamine neurodegenerative diseases and regulation of transcription: assembling the puzzle. *Genes & Development*, **20**, 2183–2192.

Ross CA (2002) Polyglutamine pathogenesis: emergence of unifying mechanisms for Huntington's disease and related disorders. *Neuron*, **35**, 819–822.

Rubolini D, Spina F, Saino N (2005) Correlates of timing of spring migration in birds: a comparative study of trans-Saharan migrants. *Biological Journal of the Linnean Society*, **85**, 199–210.

Rubolini D, Møller AP, Rainio K *et al.* (2007) Intraspecific consistency and geographic variability in temporal trends of spring migration phenology among European bird species. *Climate Research*, **35**, 135–146.

Rubolini D, Møller AP, Saino N (2010) Migratory behaviour constrains the phenological response of birds to climate change. *Climate Research*, **42**, 45–55.

Saino N, Ambrosini R (2008) Climatic connectivity between Africa and Europe may serve as a basis for phenotypic adjustment of migration schedules of trans-Saharan migratory birds. *Global Change Biology*, **14**, 250–263.

Saino N, Martinelli R, Romano M (2008) Ecological and phenological covariates of offspring sex ratio in barn swallows. *Evolutionary Ecology*, **22**, 659–674.

Saino N, Rubolini D, Serra L *et al.* (2010a) Sex-related variation in migration phenology in relation to sexual dimorphism: a test of competing hypotheses for the evolution of protandry. *Journal of Evolutionary Biology*, **23**, 2054–2065.

Saino N, Rubolini D, von Hardenberg J *et al.* (2010b) Spring migration decisions in relation to weather are predicted by wing morphology among trans-Mediterranean migratory birds. *Functional Ecology*, **24**, 658–669.

Saino N, Ambrosini R, Rubolini D *et al.* (2011) Climate warming, ecological mismatch at arrival and population decline in migratory birds. *Proceedings of the Royal Society B*, **278**, 835–842.

Saino N, Romano M, Caprioli M *et al.* (2013) Timing of molt of barn swallows is delayed in a rare *Clock* genotype. *PeerJ*, **1**, e17.

Saleem Q, Anand A, Jain S *et al.* (2001) The polyglutamine motif is highly conserved at the *Clock* locus in various organisms and is not polymorphic in humans. *Human Genetics*, **109**, 136–142.

Sanderson FJ, Donald PF, Pain DJ *et al.* (2006) Long-term population declines in Afro-Palearctic migrant birds. *Biological Conservation*, **131**, 93–105.

Schwartz MD, Ahas R, Aasa A (2006) Onset of spring starting earlier across the Northern Hemisphere. *Global Change Biology*, **12**, 343–351.

Sharp PJ (2005) Photoperiodic regulation of seasonal breeding in birds. *Annals of the New York Academy of Sciences*, **1040**, 189–199.

Spina F, Volponi S (2008) *Atlante della Migrazione degli Uccelli in Italia. Vol. 2: Passeriformi*, pp. 629. ISPRA – MATTM, Roma.

Spina F, Massi A, Montemaggiori A *et al.* (1993) Spring migration across Central Mediterranean: general results from the 'Progetto Piccole Isole'. *Vogelwarte*, **37**, 1–94.

Steinmeyer C, Mueller JC, Kempenaers B (2009) Search for informative polymorphism in candidate genes: clock genes and circadian behaviour in blue tits. *Genetica*, **136**, 109–117.

Svensson L (1992) *Identification Guide to European Passerines*. Privately published, Stockholm.

Tarka M, Åkesson M, Beraldi D *et al.* (2010) A strong quantitative trait locus for wing length on chromosome 2 in a wild population of great reed warblers. *Proceedings of the Royal Society B*, **277**, 2361–2369.

Tauber E, Kyriacou CP (2005) Molecular evolution and population genetics of circadian clock genes. *Methods in Enzymology*, **393**, 797–817.

Tenan S, Spina F (2010) Timing and condition-related effects on recapture probability, mass change and stopover length of spring migrating songbirds on a small Mediterranean island. *Ardeola*, **57**, 121–132.

Vágási CI, Pap PL, Barta Z (2010) Haste makes waste: accelerated molt adversely affects the expression of melanin-based and depigmented plumage ornaments in house sparrows. *PLoS One*, **5**, e14215.

Vaudry D, Falluel-Morel A, Bourgault S *et al.* (2009) Pituitary adenylate cyclase-activating polypeptide and its receptors: 20 years after the discovery. *Pharmacological Reviews*, **6**, 283–357.

Vaz Nunes M, Saunders D (1999) Photoperiodic time measurement in insects: a review of clock models. *Journal of Biological Rhythms*, **14**, 84–104.

Visser ME, Caro SP, van Oers K *et al.* (2010) Phenology, seasonal timing and circannual rhythms: towards a unified frame work. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **365**, 3113–3127.

Wayne NL, Rissman EF (1991) Tropical photoperiods affect reproductive development in the musk shrew, *Suncus murinus*. *Physiology & Behavior*, **50**, 549–553.

N.S., D.R., L.G., F.S., G.B. and L.B. designed the study. G.B., V.O., J.G.C., C.D.P., D.R. and N.S. did the fieldwork. M.C., L.G., E.G., C.D.P., G.B. and A.G. did the laboratory work. N.S., D.R., L.G. and G.B. did the statistical analyses. N.S., D.R., L.G. and C.D.P. wrote the paper.

Data accessibility

Partial *Clock* and *Adcyap1* genomic sequences: Dryad doi:10.5061/dryad.j22j1.

Timing of migration, morphology, sex, age, *Clock* and *Adcyap1* genotypes: Dryad doi:10.5061/dryad.j22j1.

Supporting information

Additional supporting information may be found in the online version of this article.

Fig. S1 Frequency distribution of captures dates of males and females of the four focal species entered in the analyses of migration date in relation to *Clock* genotype.

Fig. S2 Relationships between migration date (1 = January 1st) and mean *Adcyap1* allele length in the four species.

Table S1 Linear models of migration date in relation to sex, age, longer *Clock* allele length and wing length.

Table S2 Linear models of migration date in relation to sex, age, shorter *Clock* allele length and wing length.

Table S3 Linear models of migration date in relation to sex, age, mean *Adcyap1* allele length and wing length.

Table S4 Linear models of migration date in relation to sex, age, longer *Adcyap1* allele length and wing length.

Table S5 Linear models of migration date in relation to sex, age, shorter *Adcyap1* allele length and wing length.

SI_1

Supporting online information for

"Polymorphism at the *Clock* gene predicts phenology of long-distance migration in birds"

Saino et al.

Figure SI_1. Frequency distribution of captures dates of males and females of the four focal species entered in the analyses of migration date in relation to *Clock* genotype.

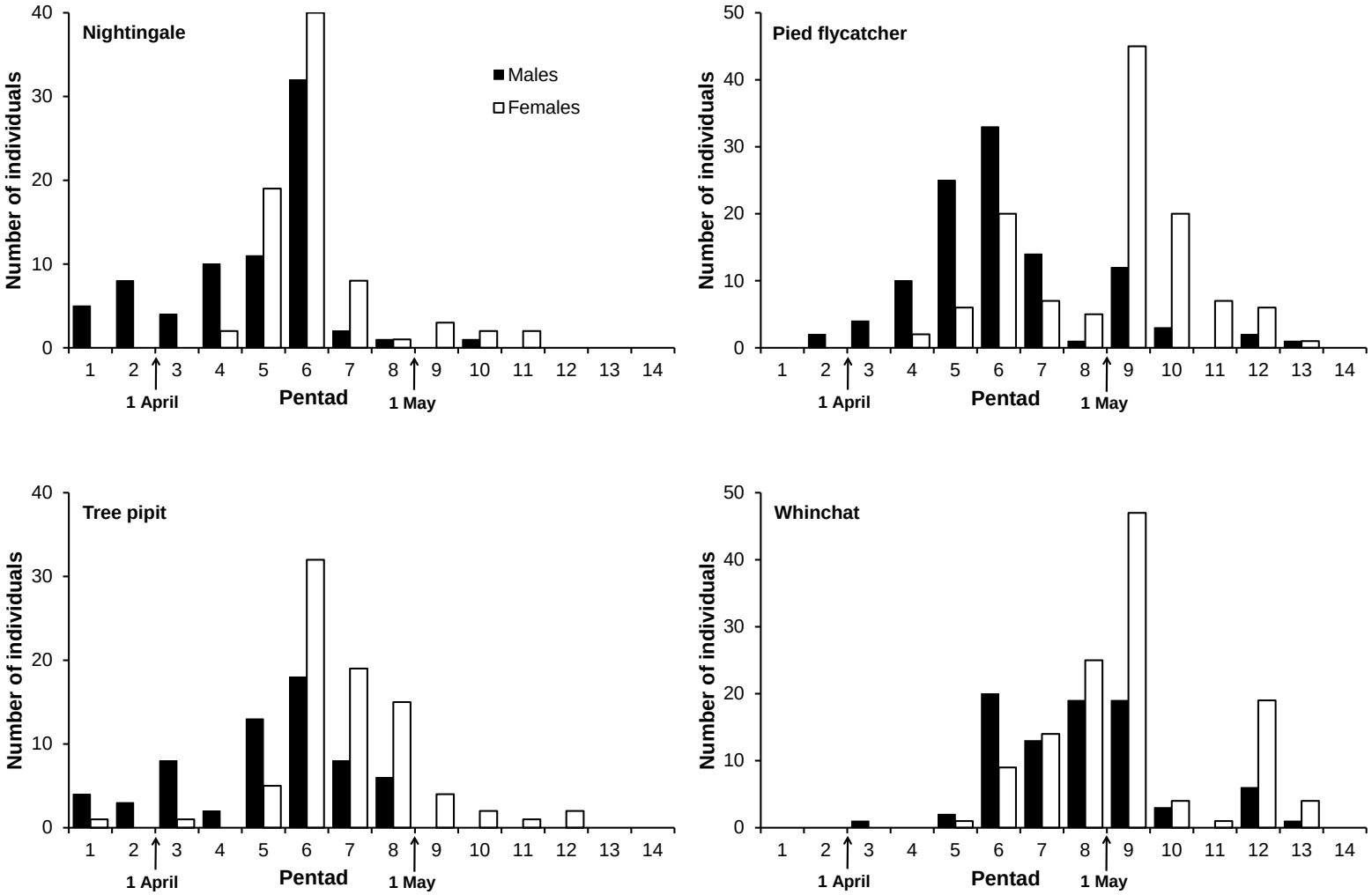


Table SI_1. Linear models of migration date in relation to sex, age, longer *Clock* allele length and wing length. Two-way interactions among sex, age or mean allele length were excluded whenever they did not attain statistical significance. See Table 1 in the main text for sample sizes.

	F	d.f.	*P*	η^2	Estimated marginal means or coefficients (SE)	
Nightingale						
Sex	27.47	1, 144	<0.001	0.160		
Age	16.89	1, 144	<0.001	0.105		
Sex x Age	4.93	1, 144	0.028	0.033		
Longer allele length	5.17	1, 144	0.024	0.035	2.640 (1.161)	
3rd primary length	5.88	1, 144	0.017	0.039	-0.068 (0.028)	
Pied flycatcher						
Sex	56.71	1, 221	<0.001	0.204	F: 118.7 (0.86)	M: 108.8 (0.91)
Age	47.71	1, 221	<0.001	0.178	SY: 118.0 (0.77)	Older:109.4 (0.94)
Longer allele length	0.04	1, 221	0.840	0.000	-0.238 (1.173)	
3rd primary length	4.25	1, 221	0.040	0.019	-0.085 (0.041)	
Tree pipit						
Sex	10.68	1, 138	0.001	0.072	F: 112.7 (0.99)	M: 104.1 (1.15)
Age	7.10	1, 138	0.009	0.049	SY: 110.3 (1.03)	Older: 106.6 (0.94)
Longer allele length	3.25	1, 138	0.073	0.023		
3rd primary length	0.03	1, 138	0.873	0.000	-0.006 (0.035)	
Sex x Longer allele length	12.30	1, 138	0.001	0.082	F: 10.547 (2.592)[a]	M: -3.395 (3.007)
Whinchat						
Sex	10.53	1, 203	0.001	0.049	F: 121.4 (0.99)	M: 116.7 (1.11)
Age	3.03	1, 203	0.083	0.015	SY: 120.5 (0.72)	Older:117.7 (1.38)
Longer allele length	2.23	1, 203	0.137	0.011	3.699 (2.477)	
3rd primary length	0.16	1, 203	0.691	0.001	-0.018 (0.045)	

* Estimated marginal means: Second-year females: 109.9 (1.26), Second-year males: 105.7 (1.18), Older females: 107.5 (1.22), Older males: 97.9 (1.34)
a: $t = 4.07$, $P < 0.001$; b: $t = -1.13$, $P = 0.261$.

Table SI_2. Linear models of migration date in relation to sex, age, shorter *Clock* allele length and wing length. Two-way interactions among sex, age or mean allele length were excluded whenever they did not attain statistical significance. See Table 1 in the main text for sample sizes.

	F	d.f.	*P*	η^2	Estimated marginal means or coefficients (SE)	
Nightingale						
Sex	25.22	1, 144	<0.001	0.149		
Age	16.27	1, 144	<0.001	0.102		
Sex x Age	4.40	1, 144	0.038	0.030		
Shorter allele length	1.01	1, 144	0.316	0.007	-0.798 (0.793)	
3rd primary length	4.81	1, 144	0.030	0.032	-0.062 (0.028)	
Pied flycatcher						
Sex	56.97	1, 221	<0.001	0.205	F: 118.7 (0.86)	M: 108.8 (0.91)
Age	47.78	1, 221	<0.001	0.178	SY: 118.0 (0.77)	Older:109.4 (0.94)
Shorter allele length	0.02	1, 221	0.886	0.000	0.155 (1.084)	
3rd primary length	4.11	1, 221	0.044	0.018	-0.084 (0.041)	
Tree pipit						
Sex	22.73	1, 139	<0.001	0.141	F: 112.2 (1.03)	M: 104.1 (1.20)
Age	5.30	1, 139	0.023	0.037	SY: 109.9 (1.07)	Older: 106.5 (0.98)
Shorter allele length	3.23	1, 139	0.075	0.023	1.752 (0.975)	
3rd primary length	0.13	1, 139	0.718	0.001	-0.013 (0.036)	
Whinchat						
Sex	8.95	1, 203	0.003	0.042	F: 121.2 (0.99)	M: 116.9 (1.11)
Age	3.53	1, 203	0.062	0.017	SY: 120.5 (0.72)	Older:117.6 (1.38)
Shorter allele length	1.57	1, 203	0.212	0.008	0.980 (0.783)	
3rd primary length	0.58	1, 203	0.446	0.003	-0.034 (0.044)	

***** Estimated marginal means: Second-year females: 109.9 (1.28), Second-year males: 105.7 (1.20), Older females: 107.4 (1.23), Older males: 98.1 (1.36)

Table SI_3. Linear models of migration date in relation to sex, age, mean *Adcyap1* allele length and wing length. Two-way interactions among sex, age or mean allele length were excluded as they never attained significance. See Table 3 in the main text for sample sizes.

	F	d.f.	P	η^2	Estimated marginal means or coefficients (SE)	
Nightingale						
Sex	22.27	1, 94	<0.001	0.192	F: 109.8 (1.13)	M: 101.7 (1.17)
Age	12.59	1, 94	0.001	0.118	SY: 108.6 (1.11)	Older:102.9 (1.13)
Mean allele length	0.13	1, 94	0.720	0.001	-0.260 (0.725)	
3rd primary length	3.21	1, 94	0.077	0.033	-0.063 (0.035)	
Pied flycatcher						
Sex	25.66	1, 90	<0.001	0.222	F: 121.6 (1.57)	M: 109.7 (1.53)
Age	7.60	1, 90	0.007	0.078	SY: 118.6 (1.33)	Older:112.6 (1.61)
Mean allele length	1.21	1, 90	0.274	0.013	-0.600 (0.545)	
3rd primary length	2.13	1, 90	0.148	0.023	-0.105 (0.072)	
Tree pipit						
Sex	19.65	1, 92	<0.001	0.176	F: 113.5 (1.27)	M: 103.8 (1.55)
Age	5.42	1, 92	0.022	0.056	SY: 110.8 (1.36)	Older:106.5 (1.23)
Mean allele length	0.13	1, 92	0.725	0.001	0.212 (0.600)	
3rd primary length	0.10	1, 92	0.748	0.001	-0.014 (0.044)	
Whinchat						
Sex	9.43	1, 164	0.002	0.054	F: 121.9 (1.12)	M: 117.0 (1.25)
Age	3.62	1, 164	0.059	0.022	SY: 121.1 (0.82)	Older:117.7 (1.55)
Mean allele length	1.16	1, 164	0.284	0.007	-0.428 (0.398)	
3rd primary length	0.09	1, 164	0.759	0.001	-0.015 (0.050)	

Table SI_4. Linear models of migration date in relation to sex, age, longer *Adcyap1* allele length and wing length. Two-way interactions among sex, age or mean allele length were excluded whenever they did not attain significance. See Table 3 in the main text for sample sizes.

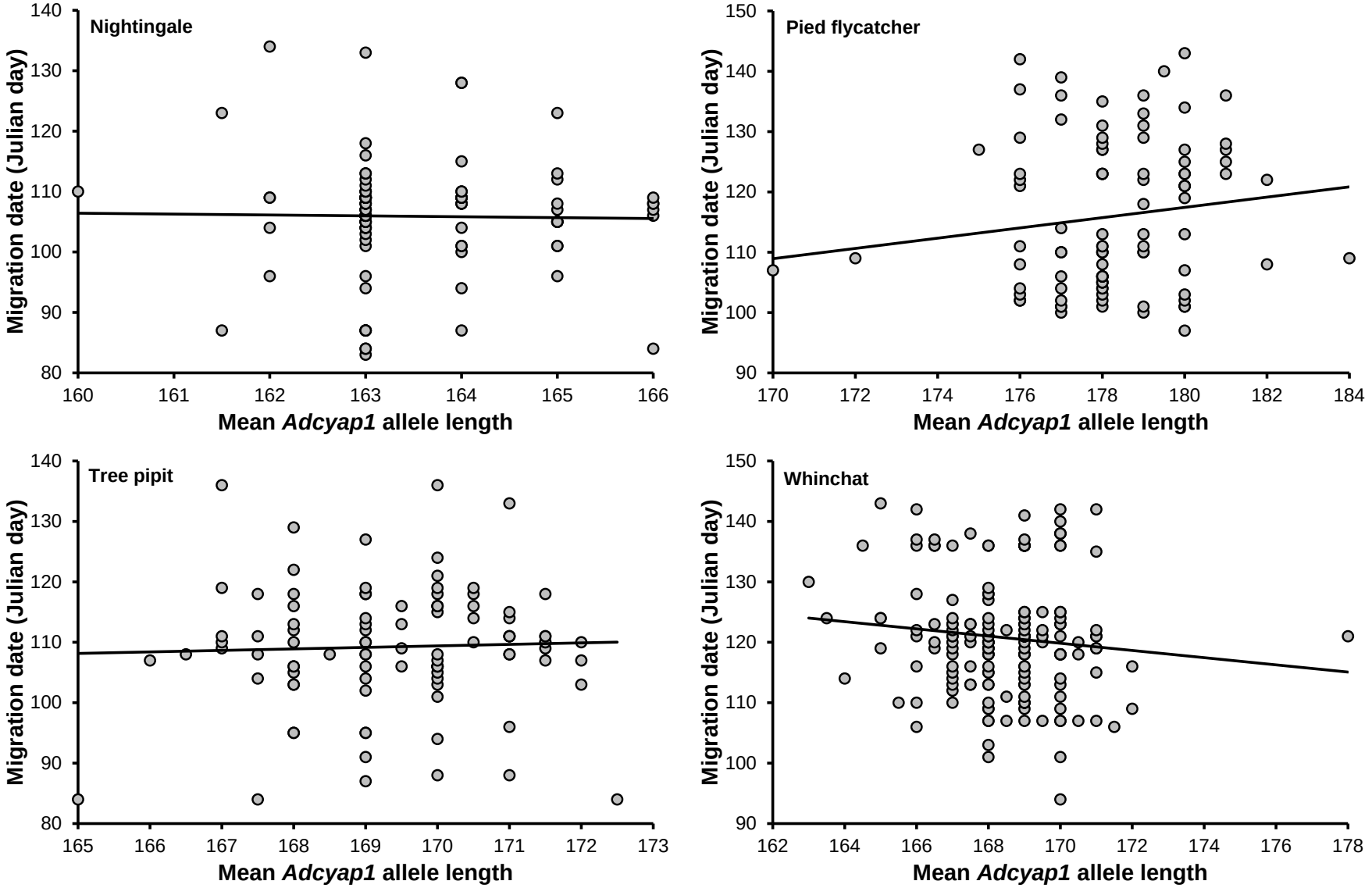
	F	d.f.	*P*	η^2	Estimated marginal means or coefficients (SE)	
Nightingale						
Sex	22.33	1, 94	<0.001	0.192	F: 109.8 (1.13)	M: 101.7 (1.17)
Age	12.42	1, 94	0.001	0.117	SY: 108.6 (1.11)	Older:102.9 (1.13)
Longer allele length	0.05	1, 94	0.823	0.001	-0.100 (0.447)	
3rd primary length	3.21	1, 94	0.076	0.033	-0.063 (0.035)	
Pied flycatcher						
Sex	4.59	1, 89	0.035	0.049	F: 120.9 (1.55)	M: 109.6 (1.50)
Age	7.65	1, 89	0.007	0.079	SY: 118.1 (1.31)	Older:112.3 (1.59)
Longer allele length	0.10	1, 89	0.754	0.001		
3rd primary length	2.21	1, 89	0.140	0.024	-0.105 (0.070)	
Sex x Longer allele Length	4.90	1, 89	0.029	0.052	F: 0.878 (0.713)[a]	M: -1.168 (0.588)[b]
Tree pipit						
Sex	19.16	1, 92	<0.001	0.172	F: 113.4 (1.28)	M: 103.9 (1.56)
Age	5.34	1, 92	0.023	0.055	SY: 110.8 (1.36)	Older: 106.5 (1.23)
Longer allele length	0.06	1, 92	0.800	0.001	0.125 (0.493)	
3rd primary length	0.13	1, 92	0.723	0.001	-0.016 (0.045)	
Whinchat						
Sex	9.84	1, 164	0.002	0.057	F: 121.9 (1.13)	M: 116.8 (1.26)
Age	3.89	1, 164	0.050	0.023	SY: 121.1 (0.83)	Older:117.6 (1.57)
Longer allele length	0.01	1, 164	0.926	0.000	0.032 (0.342)	
3rd primary length	0.08	1, 164	0.783	0.000	-0.014 (0.050)	

a: $t = 1.23$, $P = 0.222$; b: $t = -1.99$, $P = 0.050$.

Table SI_5. Linear models of migration date in relation to sex, age, shorter *Adcyap1* allele length and wing length. Two-way interactions among sex, age or mean allele length were excluded as they never attained significance. See Table 3 in the main text for sample sizes.

	F	d.f.	P	η^2	Estimated marginal means or coefficients (SE)	
Nightingale						
Sex	22.18	1, 94	<0.001	0.191	F: 109.8 (1.13)	M: 101.7 (1.17)
Age	12.53	1, 94	0.001	0.118	SY: 108.6 (1.12)	Older:102.9 (1.13)
Shorter allele length	0.11	1, 94	0.746	0.001	-0.216 (0.665)	
3rd primary length	3.17	1, 94	0.078	0.033	-0.063 (0.035)	
Pied flycatcher						
Sex	25.66	1, 90	<0.001	0.222	F: 121.5 (1.56)	M: 109.7 (1.53)
Age	7.74	1, 90	0.007	0.079	SY: 118.6 (1.33)	Older:112.6 (1.62)
Shorter allele length	1.22	1, 90	0.273	0.013	-0.498 (0.451)	
3rd primary length	2.27	1, 90	0.136	0.025	-0.109 (0.072)	
Tree pipit						
Sex	19.81	1, 92	<0.001	0.177	F: 113.5 (1.28)	M: 103.8 (1.56)
Age	5.41	1, 92	0.022	0.056	SY: 110.8 (1.36)	Older: 106.5 (1.24)
Shorter allele length	0.10	1, 92	0.752	0.001	0.151 (0.478)	
3rd primary length	0.09	1, 92	0.772	0.001	-0.013 (0.045)	
Whinchat						
Sex	9.48	1, 164	0.002	0.055	F: 121.8 (1.12)	M: 116.9 (1.24)
Age	4.18	1, 164	0.043	0.025	SY: 121.2 (0.82)	Older:117.6 (1.54)
Shorter allele length	3.27	1, 164	0.073	0.020	-0.567 (0.314)	
3rd primary length	0.04	1, 164	0.836	0.000	-0.010 (0.049)	

Figure SI_2. Relationships between migration date (1 = January 1st) and mean *Adcyap1* allele length in the four species. The relationship was non-significant in all cases, after controlling for sex, age and wing length (see Table SI_1). Simple linear regression lines are shown to better visualize the trends.



CHAPTER 2

Adcyap1 polymorphism covaries with breeding latitude in a Nearctic migratory songbird, the Wilson's warbler (*Cardellina pusilla*)

Gaia Bazzi, Andrea Galimberti, Quentin R. Hays, Ilaria Bruni, Jacopo G. Cecere, Luca Gianfranceschi, Keith A. Hobson, Yolanda E. Morbey, Nicola Saino, Christopher G. Guglielmo and Diego Rubolini

Ecology and Evolution (2016)

Ecology and Evolution

Open Access

Adcyap1 polymorphism covaries with breeding latitude in a Nearctic migratory songbird, the Wilson's warbler (*Cardellina pusilla*)

Gaia Bazzi[1], Andrea Galimberti[2], Quentin R. Hays[3,4], Ilaria Bruni[2], Jacopo G. Cecere[5], Luca Gianfranceschi[1], Keith A. Hobson[3,6], Yolanda E. Morbey[3], Nicola Saino[1], Christopher G. Guglielmo[3] & Diego Rubolini[1]

[1]Dipartimento di Bioscienze, Università degli Studi di Milano, via Celoria 26, I-20133 Milan, Italy
[2]ZooPlantLab, Dipartimento di Biotecnologie e Bioscienze, Università degli Studi di Milano-Bicocca, Piazza della Scienza 2, I-20126 Milan, Italy
[3]Department of Biology, Advanced Facility for Avian Research, University of Western Ontario, London, Ontario N6A 5B7, Canada
[4]Natural Resources Department, Eastern New Mexico University - Ruidoso, Ruidoso, New Mexico 88345
[5]ISPRA – Istituto Superiore per la Protezione e la Ricerca Ambientale, Via Cà Fornacetta 9, I-40064 Ozzano dell'Emilia (BO), Italy
[6]Environment Canada, 11 Innovation Boulevard, Saskatoon, Saskatchewan S7N 3H5, Canada

Keywords
Candidate genes, *Clock*, deuterium, migration distance, phenology, stable isotopes.

Correspondence
Diego Rubolini, Dipartimento di Bioscienze, Università degli Studi di Milano, via Celoria 26, I-20133 Milan, Italy.
Tel: +390250314718;
Fax: +390250314713;
E-mail: diego.rubolini@unimi.it
and
Christopher G. Guglielmo, Department of Biology, Advanced Facility for Avian Research, University of Western Ontario, 1151 Richmond St. N., London, ON N6A5B7, Canada.
Tel: +1519-661-2111 x84648;
Fax: +1519-661-3935;
E-mail: cguglie2@uwo.ca

Funding Information
Funding was provided by a Discovery Grant to CGG from The Natural Sciences and Engineering Research Council of Canada and by an operating grant from Environment Canada to KAH.

Received: 5 February 2016; Accepted: 12 February 2016

doi: 10.1002/ece3.2053

Abstract

Understanding the genetic background of complex behavioral traits, showing multigenic control and extensive environmental effects, is a challenging task. Among such traits, migration is known to show a large additive genetic component. Yet, the identification of specific genes or gene regions explaining phenotypic variance in migratory behavior has received less attention. Migration ultimately depends on seasonal cycles, and polymorphism at phenological candidate genes may underlie variation in timing of migration or other aspects of migratory behavior. In this study of a Nearctic–Neotropical migratory songbird, the Wilson's warbler (*Cardellina pusilla*), we investigated the association between polymorphism at two phenological candidate genes, *Clock* and *Adcyap1*, and two aspects of the migratory phenotype, timing of spring migration through a stopover site and inferred latitude of the breeding destination. The breeding destination of migrating individuals was identified using feather deuterium ratio (δ^2H), which reliably reflects breeding latitude throughout the species' western breeding range. Ninety-eight percent of the individuals were homozygous at *Clock*, and the rare heterozygotes did not deviate from homozygous migration phenology. *Adcyap1* was highly polymorphic, and allele size was not significantly associated with migration date. However, *Adcyap1* allele size significantly positively predicted the inferred breeding latitude of males but not of females. Moreover, we found a strong positive association between inferred breeding latitude and *Adcyap1* allele size in long-distance migrating birds from the northern sector of the breeding range (western Canada), while this was not the case in short-distance migrating birds from the southern sector of the breeding range (coastal California). Our findings support previous evidence for a role of *Adcyap1* in shaping the avian migratory phenotype, while highlighting that patterns of phenological candidate gene–phenotype associations may be complex, significantly varying between geographically distinct populations and even between the sexes.

Introduction

In many migratory organisms, individuals embark on their first migratory journey without any guidance from conspecifics concerning the timing and the direction of migration and the distance to be covered, and in most cases in the absence of any cue concerning the ecological conditions they subsequently face (Newton 2008). Such an innate spatiotemporal program, which is remarkably consistent at the population level, hints at a strong genetic component of migratory behavior (reviewed in Liedvogel and Lundberg 2014). Indeed, experimental studies of birds focusing on quantitative genetics and heritability of timing and direction of migration and migration distance revealed these traits have a large additive genetic component (Pulido and Berthold 2003). However, identifying which individual genes or gene groups are involved in shaping phenotypic variability in natural populations of migratory species has received less attention (Liedvogel and Lundberg 2014).

Studies of migratory vertebrates have mainly focused on phenological candidate genes (PCG), that is, genes whose allelic variation may explain differences in the photoperiodic responses among populations and individuals (e.g., Johnsen et al. 2007; O'Malley and Banks 2008; O'Malley et al. 2010; Bourret and Garant 2015; Saino et al. 2015). These PCG show short tandem repeats that could affect gene functions and are possibly involved in determining variability in migratory phenotypes (Kashi et al. 1997; Comings 1998; Li et al. 2004; Fondon et al. 2008). Among PCG, *Clock* (*Circadian Locomotor Output Cycles Kaput*) plays a central role within the "core circadian oscillator" (CCO), which is ultimately responsible for the onset and setting of circadian and circannual rhythmicity (Panda et al. 2002; Lincoln et al. 2003; Bell-Pedersen et al. 2005; Ko and Takahashi 2006). Polymorphism at a *Clock* polyglutamine-rich region (Poly-Q) has been reported both among and within populations (e.g., Johnsen et al. 2007; O'Malley and Banks 2008; Liedvogel et al. 2009; Caprioli et al. 2012), and it may play a role in determining phenological responses by differentially affecting the timing of seasonal activities (Hayasaka et al. 2002). At the among-population level, a latitudinal cline in *Clock* allele size, with the frequency of longer alleles increasing along a south–north gradient, has been documented in some migratory bird and fish species (Johnsen et al. 2007; O'Malley and Banks 2008; O'Malley et al. 2010; Lemay and Russello 2014) and probably reflects local adaptation to different photoperiodic regimes (Kyriacou et al. 2008). However, such a latitudinal cline has not been detected in other studies (e.g., Dor et al. 2012; O'Brien et al. 2013).

Within populations, *Clock* polymorphism may affect an individual's response to photoperiod. Intrapopulation variation in *Clock* allele size predicted timing of key life-history events of birds, such as reproduction, molt, and migration (Liedvogel et al. 2009; Caprioli et al. 2012; Saino et al. 2013, 2015; Bazzi et al. 2015; Bourret and Garant 2015), with individuals bearing "shorter" alleles (i.e., with fewer glutamine residues) showing advanced timing, while individuals bearing "longer" alleles showed a delayed phenology. Yet, other studies failed to find any association between *Clock* genotype and timing of phenophases (Liedvogel and Sheldon 2010; Dor et al. 2011; Chakarov et al. 2013; Morbey et al. 2014). Finally, in a Nearctic migratory bird species complex (*Junco hyemalis* and *J. phaeonotus*), *Clock* allele size increased with migration distance within some subspecies groups, but not among populations (Peterson et al. 2013).

A second important PCG, *Adcyap1* (*Adenylate Cyclase-Activating Polypeptide 1*), encodes PACAP (pituitary adenylate cyclase-activating polypeptide), a neuropeptide broadly diffused in the brain and peripheral organs of vertebrates (Nowak and Zawilska 2003; Vaudry et al. 2009; Olano-Marin et al. 2011). PACAP is known to exert several biological functions, many of which could foster physiological and behavioral shifts related to migration (Mueller et al. 2011). PACAP plays a role in circadian and circannual timing, stimulating melatonin synthesis in the pineal gland and conveying light information from the retina to the suprachiasmatic nucleus of the hypothalamus, a key element in the regulation of circadian timing of birds and mammals (Simonneaux et al. 1993; Hannibal et al. 1997; Schwartz and Andrews 2013). PACAP also modulates the expression of the CCO, by directly activating *Clock* and other circadian genes in the chicken pineal gland (Nagy and Csernus 2007; Racz et al. 2008). *Adcyap1* shows microsatellite polymorphism at the 3′-UTR of the gene that has been suggested to modify post-transcriptional processes (Steinmeyer et al. 2009). The first evidence of a link between *Adcyap1* 3′-UTR polymorphism and migratory behavior came from a study of the blackcap (*Sylvia atricapilla*), where longer *Adcyap1* allele sizes were positively associated with the amount of migratory restlessness of caged birds during their first migration toward the wintering areas (Mueller et al. 2011). Hence, allele size variation may modulate migration distance within populations, with birds bearing longer *Adcyap1* alleles migrating farther, because the amount of migratory restlessness is associated with longer migratory flights (Gwinner 1990; Berthold 1996; Maggini and Bairlein 2010). Furthermore, blackcap populations migrating over longer distances had a longer mean *Adcyap1* allele size (Mueller et al. 2011). A study of the dark-eyed junco (*J. hyemalis*) further supported the possible role of *Adcyap1* in affecting migratory restlessness in a migratory population (Peterson et al. 2013). However, *Adcyap1*

polymorphism was not associated with migration distance among junco populations (Peterson et al. 2013).

As PACAP modulates molecular clocks (Nagy and Csernus 2007; Racz et al. 2008), we may also expect *Adcyap1* to predict phenology within populations. Early dispersing juvenile buzzards (*Buteo buteo*) carried longer *Adcyap1* alleles (Chakarov et al. 2013). Moreover, *Adcyap1* allele size predicted laying date of female tree swallows (*Tachycineta bicolor*), although the effect of genotype on phenology varied between local populations (Bourret and Garant 2015). Finally, there was no association between timing of spring migration toward the breeding areas and *Adcyap1* polymorphism in four Palearctic–Afrotropical migrants (Saino et al. 2015), and *Adcyap1* polymorphism predicted timing of spring migration in blackcaps migrating across Europe only among females and in combination with wing morphology (Mettler et al. 2015).

In this study of a Nearctic–Neotropical migratory passerine, the Wilson's warbler (*Cardellina pusilla*), we aimed to assess whether *Clock* and *Adcyap1* polymorphism predicted the timing of spring migration to breeding areas, recorded at a southern stopover site, and the inferred latitude of the breeding destination, as gauged from the stable hydrogen isotope ratio of feathers (Kelly et al. 2002; Paxton et al. 2007; Hobson and Wassenaar 2008). We sampled Wilson's warblers during spring migration along the western flyway at a stopover site in Arizona, USA, halfway between the wintering (from southern Mexico to Panama) and the breeding areas (from California to northern Canada and Alaska) (see Fig. 1). The Wilson's warbler is an obligate migrant which follows a "leapfrog" migration pattern, with birds from northern breeding populations overwintering at more southern latitudes in the wintering range (Central Mexico to Panama) compared to those from southern breeding populations, that overwinter in northern Mexico (Kelly et al. 2002; Clegg et al. 2003; Paxton et al. 2007; Ruegg et al. 2014b). Hence, birds from northern breeding populations perform considerably longer migrations than those from southern breeding ones (Fig. 1).

We took advantage of the well-defined geographical gradient in hydrogen stable isotope ratios (δ^2H hereafter) in precipitation across North America to infer the breeding destination of migrating warblers sampled at the study site (Fig. 1; Kelly et al. 2002; Clegg et al. 2003; Paxton et al. 2007; Hobson and Wassenaar 2008; Hobson et al. 2014).

As outlined above, candidate genes may predict phenotypic variation at two different levels: within and among populations. To investigate within-population effects, we tested the association between timing of migration and genetic polymorphism while statistically removing the confounding effects of breeding latitude (reflecting population of origin; northern populations migrate later, this study and Paxton et al. 2007). On the other hand, to investigate among-population effects and highlight possible adaptive geographic clines in allele frequencies, we analyzed the association between genetic polymorphism and inferred breeding latitude.

In line with previous literature documenting associations between *Clock* polymorphism and phenology, we predicted that within populations (i.e., after controlling for breeding latitude, by including feather δ^2H in the models), individuals bearing shorter *Clock* alleles will migrate early, whereas those with longer alleles will have a delayed migration phenology. Among populations, we predicted birds originating from northern populations to have longer *Clock* alleles than those from southern populations.

On the other hand, we expected *Adcyap1* allele size to increase with latitude of the breeding destination (i.e., decrease with increasing feather δ^2H; Fig. 1), reflecting an among-population effect. We also explored the association between *Adcyap1* and the timing of migration at the within-population level (i.e., after controlling for feather δ^2H), but no clear prediction could be formulated due to the paucity of previous studies (Chakarov et al. 2013; Bourret and Garant 2015; Mettler et al. 2015; Saino et al. 2015).

Materials and Methods

Study species

The Wilson's warbler is a small, sexually dichromatic, widespread Nearctic passerine that breeds in North America and winters in Mexico and Central America (Pyle et al. 1997). The species shows a clear genetic differentiation among geographical populations, with birds from the eastern portion of the breeding range and migrating through eastern North America being well separated from those migrating along the western flyway (Ruegg et al. 2014b). In addition, western breeding populations can be separated into five homogeneous genetic clusters (Ruegg et al. 2014b). Wilson's warblers perform a single complete postbreeding molt while at or near the breeding grounds, before fall migration (Pyle et al. 1997; Paxton et al. 2007). Hence, birds sampled during spring migration will carry tail feathers grown in the breeding areas in the previous breeding season. Feather δ^2H values from spring migrating birds was thus used to infer individual breeding destinations (see below).

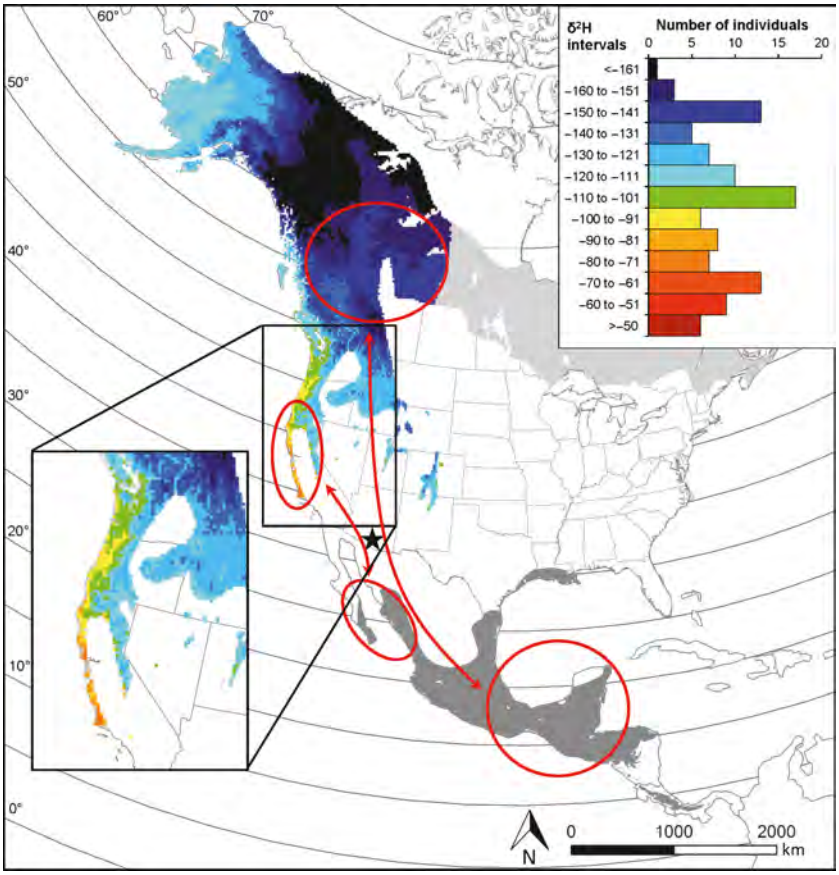


Figure 1. Feather δ^2H isoscape within the breeding range of western North American Wilson's warbler populations. The map (Lambert Conformal Conic projection) shows the expected spatial distribution of feather δ^2H values in the western part of Wilson's warbler breeding range (see Materials and Methods). The eastern portion of the breeding distribution range (light gray) and the entire wintering distribution range (dark gray) of the species is also shown for comparison, and the position of the study site (Buenos Aires National Wildlife Refuge, Arizona) is marked as a black star. Inset histogram: frequency distribution of feather δ^2H values of spring migrating Wilson's warblers at the study site (n = 105). The ellipses denote the approximate breeding destination (based on feather δ^2H thresholds; see Materials and Methods) and wintering distribution (based on Ruegg et al. 2014b) of short-distance migrating southern birds (feather δ^2H $\geq -90‰$, n = 43 individuals) and of long-distance migrating northern birds (feather δ^2H values $< -130‰$, n = 22 individuals) (see Materials and Methods).

Study site and field methods

We captured Wilson's warblers at the Buenos Aires National Wildlife Refuge (Fig. 1), Pima County, USA (31°33′00″N 111°33′02″W), during March 28–May 24, 2006. All field procedures were approved by the University of Western Ontario Animal Use Sub-Committee (2006-014-02), the US Geological Survey Bird Banding Lab (23423), the US Fish and Wildlife Service (MB121152-0), and the Arizona Game and Fish Department (298399). The study area, located in the Sonoran Desert, provides isolated riparian habitats to a large number of migrating birds which stopover there to refuel during spring migration. Previous studies showed that birds with markedly different breeding destinations (exclusively from the western breeding range) use the area as a stopover site (Paxton et al. 2007; Ruegg et al. 2014b; Fig. 1). We selected a small (approx. 2 ha), isolated trapping site which enabled us to reliably capture most of the foraging individuals at any given time; afternoon resighting efforts indicated that unmarked individuals were rare (QRH, unpublished data). We are therefore confident that almost all the birds were captured on the day of arrival. We used first capture date of each bird as a reliable estimate of timing of migration (hereafter, migration date).

Birds were trapped during the morning hours using mist nets, banded, measured (wing length, to the nearest mm), and sexed according to Pyle et al. (1997). Age was not assessed due to difficulties in discriminating between second year and older individuals based on the differences in feather wear. For each individual, we collected a biological sample including one outer rectrix and three breast contour feathers. Full-grown feathers were collected and stored individually at room temperature for later analyses. Overall, we captured 108 individuals (49 females, 59 males).

Stable isotope analyses and feather isoscape of the breeding destination

Stable hydrogen isotope measurements were made for individual rectrices. δ^2H analyses were performed using the comparative equilibration approach reported in Wassenaar and Hobson (2003) (see Appendix S1 for details). Isotope values were reported in δ notation as parts per thousand ($‰$) deviation from the Vienna

Standard Mean Ocean Water (VSMOW)–Standard Light Antarctic Precipitation (SLAP) scale. We could determine δ^2H for 105 individuals (47 females, 58 males) of the 108 sampled. The remains of analyzed feathers were kept in glass tubes at room temperature until genetic analyses were performed.

To infer the putative breeding destination of migrating Wilson's warblers, we compared the feather δ^2H values with a δ^2H feather isoscape of North America showing the spatial distribution of expected feather δ^2H values on a continental scale (Fig. 1). The δ^2H feather isoscape was derived by combining the most recent available amount-weighted growing season δ^2H precipitation surface (see Terzer et al. 2013 for details and data source) and a feather – precipitation δ^2H transfer function for Neotropical migrant nonground foragers (full details reported in Hobson et al. 2014).

To compare the strength of genotype–phenotype associations between populations differing in migratory behavior (see *Statistical analyses* below), we aimed at separating long-distance migrating "northern" birds from short-distance migrating "southern" ones. Moreover, in these comparisons, we aimed at accounting for the geographical genetic structuring among the species' western population complex previously described by Ruegg et al. (2014b) using high-resolution genetic markers. To this end, we first assigned our feather δ^2H values to 10‰ intervals and plotted the contours of these intervals on the feather isoscape map (Fig. 1). "Southern" birds were identified as those with feather δ^2H values ≥ -90‰ ($n = 43$ individuals), which is consistent with the genetically homogeneous cluster of coastal California breeding birds, migrating south over short-distances to the Pacific coastal areas of northern Mexico ("yellow" cluster in Ruegg et al. 2014b; Fig. 1). "Northern" birds were identified as those with feather δ^2H values < -130‰ ($n = 22$ individuals), which breed throughout Canada and migrate over long distances to winter in southern Mexico and throughout Central America (Kelly et al. 2002; Ruegg et al. 2014b; Fig. 1). These birds belong to a different genetically homogeneous cluster ("violet" cluster in Ruegg et al. 2014b). Intermediate feather δ^2H values (between -130‰ and -91‰) were excluded from these analyses as they may either reflect birds breeding in continental northwestern US states or in central-western Alaska (see Fig. 1), thus including birds from genetically differentiated clusters of populations (Ruegg et al. 2014b). Nevertheless, given that the mean value of the lowest 10th percentile of the feather δ^2H values we recorded in migrating individuals (-150.7‰; see also Fig. 1) corresponds to ca. 59° N on the feather isoscape shown in Fig. 1 (details not shown), we regard the possibility that birds from northern and northwestern Canada and Alaska (i.e., above 60° N, Fig. 1) were included in our sample as unlikely. Indeed, these populations may show a more eastward migration route (Paxton et al. 2007). Finally, Kelly et al. (2002) showed that feather δ^2H of Wilson's warblers sampled throughout their entire western breeding range (from California to northern Alaska) decreases linearly with breeding latitude. Hence, we can safely assume that, in our sample, decreasing feather δ^2H values reflects increasing latitudes of breeding destinations.

Genetic analyses

Total genomic DNA was extracted from feathers (both the rectrix remains and breast contour feathers) using commercial kits (see Appendix S1 for details). Polymorphism at the *Clock* and *Adcyap1* genes (allele size due to the number of sequence repeats) was determined by fragment analysis as described in Caprioli et al. (2012) and Saino et al. (2015) with slight modifications (Appendix S1). Of the 105 birds with feather δ^2H values, we reliably genotyped 102 individuals (*Clock*: 46 females, 56 males; *Adcyap1*: 44 females, 58 males). Although no *Clock* genomic sequence is available for the study species, the alignment of all *Clock* gene sequences of passerine species found in GenBank showed that all length polymorphisms in this order are due to a variable number of glutamine codons (details not shown). Hence, it is safe to assume that the 115- and 118-bp-long alleles we detected (see Results) correspond to Q_8 and Q_9 alleles, respectively (*Clock* alleles are identified according the predicted number of glutamine residues in the mature protein; see Liedvogel et al. 2009).

Statistical analyses

As we found extremely low variability at the *Clock* locus (see Results), to explore the *Clock* genotype-timing of migration association, we compared migration date of the rare heterozygote *Clock* individuals with the phenotypic distribution of migration date of the homozygote ones (as detailed in Bazzi et al. 2015). We thus calculated the 95% nonparametric bootstrap confidence limits (BCa method; see details in Bazzi et al. 2015) of migration date for homozygous individuals, separately for each sex; to account for seasonal variation in the occurrence of birds from different breeding destinations (see Results and Paxton et al. 2007), the migration date values of both the homozygote and the heterozygote individuals were expressed as residuals from the linear regression of migration date on feather δ^2H, separately for each sex. We then compared the heterozygote values (residuals) with the confidence intervals of homozygote individuals (from the distribution of residuals) of their respective sex. No

formal analysis was made to investigate the association between *Clock* genotype and breeding destination.

To test for an association between *Adcyap1* allele size and migration date while controlling for variation in migration date due to variation in breeding destination, we ran linear models of migration date (1 = January 1) as a function of allele size (mean of the two alleles; Mueller et al. 2011; Chakarov et al. 2013; Peterson et al. 2013; Saino et al. 2015), feather δ^2H, and sex (0 = females, 1 = males). Finally, we tested the association between *Adcyap1* genotype and breeding destination by linear models of feather δ^2H as a function of *Adcyap1* allele size and sex. Similar models were run to investigate the association between wing length, another proxy of breeding destination (see Results and Q. R. Hays et al. in prep.), and *Adcyap1* allele size. Two-way interaction terms were included in all initial linear models and removed *en bloc* if statistically nonsignificant ($P > 0.05$).

We also tested whether the strength of the *Adcyap1*–feather δ^2H association varied between the two genetically homogeneous clusters of northern and southern birds by computing the correlation coefficients between *Adcyap1* allele size and feather δ^2H for northern and southern birds, and applying a test for the difference between two correlation coefficients (Sokal and Rohlf 2009).

All the analyses involving *Adcyap1* were repeated for each allele size (short and long) to investigate whether allelic dominance effects occurred, because this is often the case in phenotypic effects of simple sequence repeats (Ross 2002; Fondon et al. 2008; Liedvogel et al. 2009; Saino et al. 2015).

Deviations from Hardy–Weinberg equilibrium (HWE) were tested for the *Adcyap1* locus using the Markov chain method (Guo and Thompson 1992) implemented in GENEPOP (dememorization = 1000, batches = 100, iterations per batch = 1000) to obtain unbiased estimates of the Fisher's exact statistic (Raymond and Rousset 1995). For this locus, we also calculated genetic differentiation by computing the F_{ST} estimate between the northern and southern populations by means of Fstat 2.9.3 software (Goudet 2001).

Means and parameter estimates are reported together with their associated standard error, unless stated otherwise.

Results

Timing of migration and breeding destination

Feather δ^2H values showed a clear temporal decline during the migration period [Fig. 2; linear regression, estimate: −1.08 (0.16) ‰/day, $t_{103} = 6.67$, $P < 0.001$], with birds directed toward northern breeding areas (i.e., with lower feather δ^2H values) migrating later (see also Paxton et al. 2007). Migration date also significantly differed between the sexes, males migrating considerably earlier on average (April 19) than females (May 7) ($t_{106} = 6.76$, $P < 0.001$) (Fig. 2) (see Otahal 1995). Feather δ^2H predicted migration date in a similar way in either sex (see Fig. 2), as indicated by the nonsignificant interaction between sex and feather δ^2H in a linear model ($F_{1,101} = 0.72$, $P = 0.40$). This model also confirmed a statistically significant sex difference in migration date [mean estimated difference in migration date between males and females: 14.33 (2.59) days ($F_{1,102} = 30.68$, $P < 0.001$)] while controlling for the effect of feather δ^2H [−0.21 (0.04) days/‰, $F_{1,102} = 29.76$, $P < 0.001$] (see also above and Fig. 2).

Genetic variation at *Clock* and *Adcyap1*

The *Clock* locus showed very low variability, because only two (one male, one female) of 102 genotyped individuals were heterozygous (Q8/Q9), while all the others were homozygous (Q9/Q9) [mean allele size = 117.97 (0.02) bp; observed heterozygosity $H_o = 0.02$]. On the other hand, *Adcyap1* was highly variable ($H_o = 0.78$) (Table S1). *Adcyap1* genotype frequencies significantly deviated from HWE ($P = 0.030$). However, this was not the case if the test was conducted within southern and northern birds (both P-values > 0.60). This finding

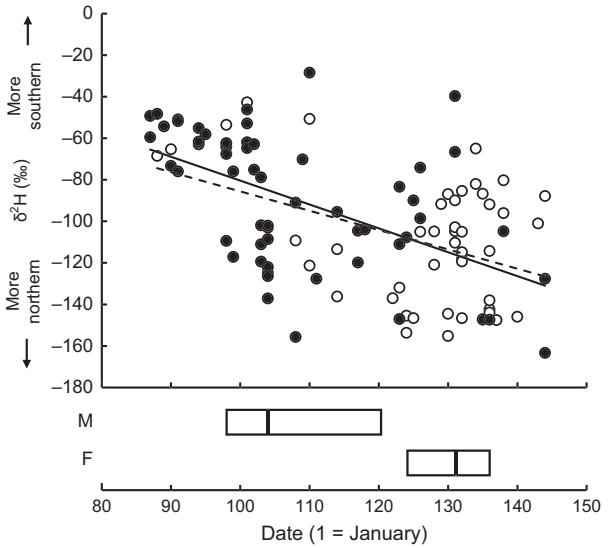


Figure 2. Feather δ^2H decreases during the migration period in male and female Wilson's warblers. Males: filled circles, continuous line; $n = 58$; females: open circles, broken line; $n = 47$. The lines represent simple linear regressions. Boxplots show the median migration date of each sex (boxes: 25th and 75th percentile).

corroborates the hypothesis of reduced gene flow between genetically differentiated southern and northern population clusters (Ruegg et al. 2014b). Indeed, *Adcyap1* genotypes showed a statistically significant differentiation between southern and northern birds (F_{ST} = 0.045, P = 0.001). The estimated F_{ST} for the *Adcyap1* microsatellite was considerably larger than the mean F_{ST} obtained from the reanalysis of the allelic panel of neutral microsatellite loci previously published in Clegg et al. (2003) from birds breeding in the corresponding population clusters (courtesy of S. M. Clegg and T. B. Smith) (see Appendix S2).

However, the mean *Adcyap1* allele size did not significantly differ between southern and northern birds [158.26 (0.19) vs. 158.73 (0.37), respectively; t_{63} = 1.28, P = 0.21]. Results were qualitatively similar for the short and long allele size (details not shown for brevity).

Within-population variation in timing of migration and genotype

Controlling for feather δ^2H, the migration dates of the heterozygous *Clock* male and female fell within the 95% confidence limit of the sex-specific distributions of migration date (Table S2). Hence, the rare heterozygous *Clock* individuals were not phenodeviant with respect to migration timing of homozygotes.

There was no statistically significant association between *Adcyap1* allele size and migration date in linear models controlling for sex and feather δ^2H (Table 1). There were no significant two-way interactions between sex, allele size, and feather δ^2H on migration date (Table 1). Hence, the (nonsignificant) effect of *Adcyap1* genotype on timing of migration was similar irrespective of geographical origin (no statistically significant allele size × δ^2H interaction, Table 1).

Breeding destination and genotype

The two rare heterozygous *Clock* individuals had feather δ^2H values that were at the extremes of the feather δ^2H distribution: The male had a value of −63‰, whereas the female had a value of −145‰. Hence, the male belonged to the southern group, while the female to the northern group (see Materials and Methods) (see also Fig. 1).

In linear models of feather δ^2H as a function of mean *Adcyap1* allele size and sex, allele size marginally and negatively predicted feather δ^2H ($F_{1,99}$ = 4.26, P = 0.042; Table 1, Fig. 3). Moreover, females showed significantly lower feather δ^2H values than males [−110.6‰ (30.2 SD) vs. −89.7‰ (33.0 SD), respectively]. The slope of the relationship between mean allele size and feather δ^2H was similar in the two sexes, as testified by the nonsignificant interaction between allele size and sex (Table 1). However, the relationship was statistically significant in males (r = −0.32, P = 0.015, n = 58) but not in females (r = −0.02, P = 0.88, n = 44) (Fig. 3).

Table 1. Linear models of the effect of *Adcyap1* allele size on migration date, feather δ^2H, and wing length. The interaction terms were removed *en bloc* if nonsignificant (P > 0.05), and statistics for main effects refer to models without the interaction term. Results were qualitatively similar if the short or long allele size was considered instead of the mean allele size (details not shown for brevity).

	Estimate (SE)	df	F	P
Migration date (n = 102)				
Allele size	0.34 (0.86)	1, 98	0.16	0.69
Sex	−14.64 (2.67)	1, 98	30.18	<0.001
δ^2H	−0.21 (0.04)	1, 98	27.93	<0.001
Allele size × sex	0.21 (1.86)	1, 95	0.01	0.91
Allele size × δ^2H	−0.01 (0.03)	1, 95	0.09	0.77
Sex × δ^2H	−0.05 (0.08)	1, 95	0.42	0.52
Feather δ^2H (n = 102)				
Allele size	−4.35 (2.11)	1, 99	4.26	0.042
Sex	21.63 (6.35)	1, 99	11.59	0.001
Allele size × sex	−6.22 (4.32)	1, 98	2.07	0.15
Wing length (n = 102)				
Allele size	0.28 (0.12)	1, 99	5.60	0.020
Sex	1.49 (0.35)	1, 99	17.74	<0.001
Allele size × sex	0.21 (0.24)	1, 98	0.78	0.38

The results were similar if the size of the long *Adcyap1* allele was used instead of mean *Adcyap1* allele size (Table 1) (effect of allele size: $F_{1,99}$ = 5.15, P = 0.025). The correlation coefficient (Pearson's r) between the long allele size and feather δ^2H in males was -0.38 (P = 0.004), whereas in females it was 0.01 (P = 0.97). No statistically significant association emerged between short *Adcyap1* allele size and feather δ^2H (correlation coefficients: males, r = −0.14, P = 0.30; females, r = −0.05, P = 0.74).

Wing length decreased linearly with feather δ^2H, indicating that wing length can be used as a reliable proxy of breeding destination and that more negative feather δ^2H values within the observed range of variation were reflecting northern breeding destinations (see also *Stable isotope analyses and feather isoscape of the breeding destination*) [linear model with sex and feather δ^2H as predictors, effect of feather δ^2H, estimate: −0.03 (0.01), $F_{1,102}$ = 56.63, P < 0.001; effect of sex: $F_{1,102}$ = 55.92, P < 0.001; the sex × δ^2H interaction and the quadratic term of feather δ^2H were not significant (details not shown)] (see also Q. R. Hays et al., in prep.).

Wing length was significantly (positively) predicted by *Adcyap1* mean and long allele size, while controlling for sex differences in wing length (Table 1; effect of long allele size, estimate: 0.20 (0.09), $F_{1,99}$ = 5.27, P = 0.024). The correlation between wing length and mean or long

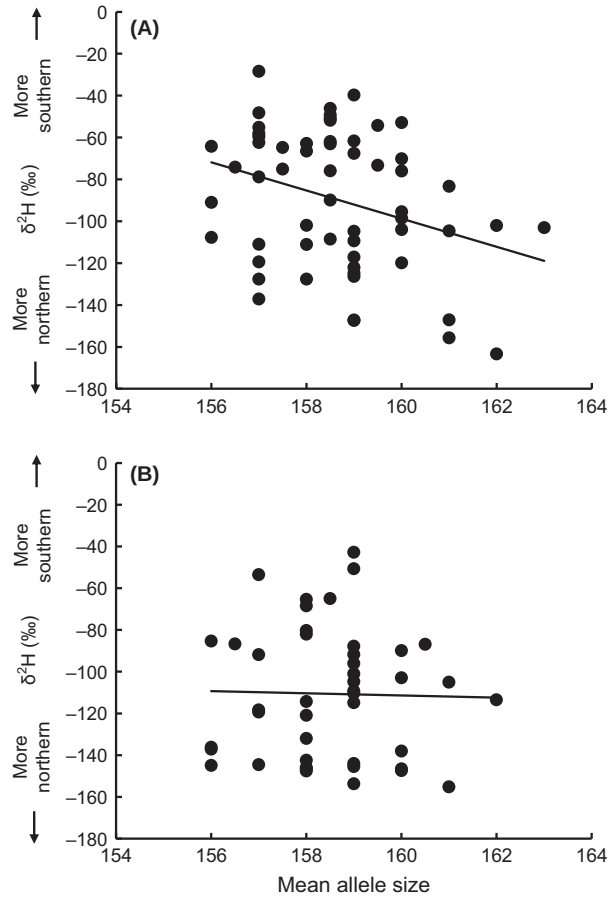


Figure 3. Wilson's warbler feather δ^2H values in relation to *Adcyap1* allele size in (A) males and (B) females. The lines represent simple linear regressions.

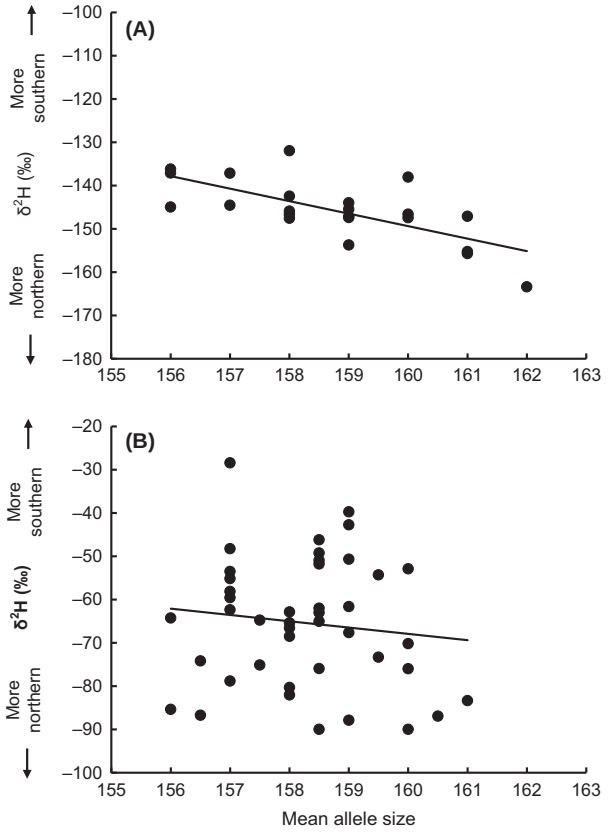


Figure 4. Correlation between *Adcyap1* allele size and feather δ^2H values in (A) northern or (B) southern birds. The same *y*-axis range is shown for ease of comparison. The lines represent simple linear regressions.

Adcyap1 allele size was statistically significant in males (mean allele size: $r = 0.30$, $P = 0.021$; long allele size: $r = 0.33$, $P = 0.011$) but not in females (mean allele size: $r = 0.12$, $P = 0.44$; long allele size: $r = 0.06$, $P = 0.68$). No statistically significant association emerged between wing length and short *Adcyap1* allele size (effect of short allele size: $F_{1,99} = 2.48$, $P = 0.119$; correlation coefficients: males, $r = 0.14$, $P = 0.35$; females, $r = 0.17$, $P = 0.22$). However, controlling for breeding destination, *Adcyap1* allele size (mean, long or short) did not significantly predict wing length anymore (all $P > 0.17$); hence, the association between wing length and *Adcyap1* genotype was a spurious effect of the latitudinal variation of wing length.

The strength of the association between *Adcyap1* mean or long (but not short) allele size and feather δ^2H, as gauged by correlation coefficients, significantly differed between the genetically homogeneous clusters of southern and northern birds (Table 2): *Adcyap1* allele size and feather δ^2H were strongly negatively correlated among northern birds, while this was not the case among southern birds (Table 2, Fig. 4). Results were qualitatively unaltered if these comparisons were based on more restrictive cutoffs for northern (i.e., $< -140‰$) and southern ($\geq -80‰$) birds (Fig. 1 and Table S3).

Table 2. Association between *Adcyap1* allele size and feather δ^2H in southern and northern Wilson's warblers. *Z*-values refer to the test of the difference between the correlation coefficients of southern and northern birds.

	Southern birds (n = 43)	Northern birds (n = 22)	Z	P
Mean allele size	−0.12	−0.69***	2.63	0.009
Short allele size	−0.13	−0.31	0.69	0.490
Long allele size	−0.07	−0.77***	3.44	<0.001

Significance levels of correlation coefficients: ***$P < 0.001$.

To further rule out that the strong association between inferred breeding latitude and *Adcyap1* allele size among northern birds is confounded by geographical variation in

population genetic structure, we reanalyzed previously published data (Clegg et al. 2003) on (presumably) neutral microsatellite loci from Wilson's warbler breeding in range of northern birds (Table S4). We found no statistically significant association between mean allele size of 8 microsatellite markers and either breeding latitude or predicted feather δ^2H values at the breeding sites (Table S4).

Discussion

We investigated whether variation at two PCG, *Clock* and *Adcyap1*, affected the timing of spring migration and breeding destination, as inferred from feather δ^2H values. To our knowledge, this is the first study testing whether interindividual variation in migration distance was predicted by *Clock* and *Adcyap1* genetic polymorphism. We found low variability at the *Clock* gene locus, most individuals being homozygous, and the two rare *Clock* heterozygotes were not deviants compared to homozygotes for timing of migration. Hence, rare genetic variants of this *Clock* gene region are not associated with migratory behavior in Wilson's warbler.

The *Adcyap1* 3′-UTR microsatellite was instead highly polymorphic, and longer *Adcyap1* alleles were significantly associated with northern breeding latitudes, especially among males. The association between breeding latitude and *Adcyap1* polymorphism was considerably strong within the genetically homogenous cluster of long-distance migrating northern birds, while it was nonsignificant within the cluster of short-distance migrating southern birds.

Given the extremely low genetic polymorphism at *Clock*, we will frame our discussion around the *Adcyap1*–phenotype association.

Adcyap1 polymorphism and timing of migration

Relative timing of spring migration through the study site, accounting for phenological differences among populations, was not associated with *Adcyap1* 3′-UTR microsatellite variation. We had no clear predictions on the association between *Adcyap1* polymorphism and timing of migration, as the few previous studies provided mixed or no evidence for links between *Adcyap1* and phenology (see Introduction). However, PACAP is broadly involved in scheduling circadian rhythms (Schwartz and Andrews 2013), and it may activate *Clock*, whose polymorphism covaries with breeding/migration phenology in some species (Liedvogel et al. 2009; Caprioli et al. 2012; Bazzi et al. 2015; Bourret and Garant 2015; Saino et al. 2015). The lack of any statistically detectable association between spring migration phenology and *Adcyap1* polymorphism may possibly indicate that phenology in this species is mostly shaped by other PCG, or by environmental modifications of gene expression via epigenetic mechanisms (Joska et al. 2014). Alternatively, phenology could be under the control of several genes with small additive effects (Manolio et al. 2009), as hypothesized for reproduction and other complex life-history traits (Visser et al. 2010; Liedvogel et al. 2012).

Adcyap1 polymorphism and breeding destination

The observation of a relatively strong association between male *Adcyap1* polymorphism and breeding destination constitutes the first evidence of a link between interindividual variation in migration distance and PCG in wild birds. *Adcyap1* microsatellite polymorphism explained ca. 10% of the variance in male breeding destination, a remarkable value compared to studies analyzing genotype–phenotype associations in the wild (Fondon et al. 2008; Bourret and Garant 2015; Saino et al. 2015).

The association between *Adcyap1* allele size and breeding destination may be the outcome of both inter- and intrapopulation effects. For instance, *Adcyap1* may show a geographical cline, with northern birds having longer alleles than southern ones, but also, within populations, a pattern of individual birds bearing longer alleles migrating over longer distances. The observed difference in the strength of the association between allele size and feather δ^2H between the previously identified genetically homogeneous clusters of southern and northern birds suggests that within-population rather than between-population effects may drive the observed covariation between *Adcyap1* allele size and breeding destination in the entire sample. In line with this, although the two groups were genetically differentiated based on the F_{ST} statistic, the mean *Adcyap1* allele size did not significantly differ between northern and southern birds.

The remarkably strong association between *Adcyap1* polymorphism and breeding destination in long-distance migrating northern birds (ca. 50% of variance explained; Table 2) compared to short-distance migrating southern ones is intriguing. Previous evidence suggested that migratory behavior is under more strict endogenous (genetic) control in long-distance vs. short-distance migrants, because short-distance migrants typically show larger phenotypic variance in migratory behavior than long-distance migrants (Berthold 1996; Pulido and Widmer 2005). This has been hypothesized to result from strong stabilizing selection and environmental canalization

of migratory traits (Pulido and Widmer 2005; Pulido 2007). Our findings suggest that a stronger endogenous control of migratory behavior in long-distance migrants may also involve a tighter genotype–phenotype link for PCG. As these tests were conducted within previously identified genetically homogenous clusters (see Materials and Methods), the distinct spatial genetic structuring of the Wilson's warbler population complex did not confound our findings.

The association between *Adcyap1* genotype and breeding destination differed between the sexes. Using the entire sample of birds, *Adcyap1* allele size significantly predicted breeding destination in males but not in females. PCG may thus control migratory behavior in a sex-specific way, as shown earlier for *Clock* and other genes (Caprioli et al. 2012; Bourret and Garant 2015). We could rule out that the lack of statistically significant genotype–phenotype association among females arose from the larger accidental subsampling of southern females. In fact, even when restricting the analysis to the northern birds ($n = 22$), in spite of the low sample size (6 males and 16 females), the association between *Adcyap1* polymorphism and feather δ^2H was still stronger among males ($r = -0.89$) than females ($r = -0.50$). We could also rule out that feather δ^2H differently predicted actual breeding destination in the two sexes: the covariation between feather δ^2H and wing length was remarkably similar and statistically significant in both sexes (see also Q. R. Hays et al. in prep.).

Finally, we found that the effect of *Adcyap1* genotype on breeding latitude was consistent when we considered both the mean and the longer allele size, while no significant association emerged for the shorter allele size. This suggests a genetic dominance of the longer allele, a feature observed in the *Clock* gene (Liedvogel et al. 2009; Saino et al. 2015) and in other simple sequence repeats (Ross 2002; Fondon et al. 2008).

Potential mechanisms driving the association between *Adcyap1* polymorphism and breeding destination

The mechanisms linking *Adcyap1* polymorphism to migration traits are largely unknown and have been only sparsely addressed (Mueller et al. 2011; Peterson et al. 2013; Saino et al. 2015). We envisage two nonexclusive pathways that may lead to an association between *Adcyap1* polymorphism and breeding destination among long-distance migrating birds. These pathways rest on the assumption that the main determinant of migration distance in passerines is the amount and/or onset of migratory restlessness during their first southward migration (Gwinner 1996; Newton 2008).

First, a latitudinal cline of *Adcyap1* allele size may reflect adaptation to local photoperiodic conditions triggering migratory behavior from the highly seasonal northern breeding quarters at the most appropriate time of the year at different latitudes. In fact, PACAP is involved in conveying light information to the suprachiasmatic nucleus of the hypothalamus and modulates photoperiodic circadian genes (Hannibal et al. 1997; Nagy and Csernus 2007; Racz et al. 2008). *Adcyap1* length polymorphism may thus influence circadian genes' expression (Nagy and Csernus 2007; Racz et al. 2008) and regulate the onset of autumn migratory autumn restlessness. As the duration of autumn migratory restlessness mostly depends on the timing of its onset (Maggini and Bairlein 2010), *Adcyap1* could affect the onset rather than the duration of autumn restlessness in birds breeding at different latitudes. Unfortunately, the association between the onset of migratory restlessness and *Adcyap1* polymorphism was not tested in previous studies (Mueller et al. 2011; Peterson et al. 2013). Evidence from juvenile dispersing common buzzards may support this hypothesis. In this species, juveniles carrying longer *Adcyap1* alleles disperse earlier than those bearing shorter alleles (Chakarov et al. 2013). As both dispersal and migration rely on increased behavioral "restlessness" (Ritchison et al. 1992), we argue that *Adcyap1* polymorphism influences the onset of restlessness rather than its duration.

Second, PACAP may affect migratory restlessness, and hence, the distance to be migrated from different breeding latitudes in autumn, by modulating melatonin secretion (Simonneaux et al. 1993) and lipid utilization, while inhibiting feeding behavior (Tachibana et al. 2003, 2007). PACAP could cause a phase shift of the endogenous oscillators from day-to-night activity through melatonin modulation, a major determinant of migration onset in passerines (Fusani and Gwinner 2005; Mueller et al. 2011), and a fuel shift to lipid metabolism, which may influence nutritional state and night migratory activity (Coppack and Bairlein 2011; Schwartz and Andrews 2013). Longer *Adcyap1* alleles may induce a decrease of melatonin release over longer periods, extending migratory restlessness and thus the distance migrated (Coppack and Bairlein 2011).

Concluding remarks

By taking advantage of feather δ^2H gradients throughout the breeding range of Wilson's warbler in western North America, we discovered the evidence of a very strong latitudinal cline in *Adcyap1* microsatellite polymorphism among the genetically homogeneous cluster of northern breeding birds, that were also genetically differentiated at this PCG from southern breeding ones. We also uncov-

ered that the strength of *Adcyap1*–phenotype association may vary widely among geographical populations and even between the sexes, suggesting that sex-specific selective pressures may affect PCG–phenotype associations. Importantly, between-sex and among-population variation in *Adcyap1*–phenotype associations may contribute to explain the often inconsistent findings of previous studies investigating the link between PCG polymorphism and migratory behavior.

On the whole, results from this study deepen our understanding of the genetic basis of avian migration, as they are consistent with the hypothesis that *Adcyap1* is involved in the regulation of migratory behavior, and may have far-reaching evolutionary implications. For instance, our finding of stronger genetic differentiation for the *Adcyap1* microsatellite than for neutral microsatellite markers between short- and long-distance migratory Wilson's warblers is coherent with a previous study comparing genome-wide patterns of divergence between short- and long-distance migrating populations of the Swainson's thrush (*Catharus ustulatus*) (Ruegg et al. 2014a). That study reported that several migration-linked genes, including *Adcyap1*, showed a stronger pattern of single nucleotide polymorphism (SNP) differentiation between short- and long-distance migrants compared to all the other SNPs from autosomal genes. Coupled with the observation of a strong latitudinal cline of allele size among long-distance migrating northern birds, a significant differentiation of the *Adcyap1* microsatellite between short- and long-distance migrating Wilson's warblers provides support for a role of migration-linked genes and migratory behavior as drivers of population differentiation, promoting reproductive isolation and eventually leading to speciation (Liedvogel et al. 2011; Ruegg et al. 2014a).

Acknowledgments

We thank C. Bitter, N. Donovan, K. Kardynal, and I. Maggini for assistance, and M. Casiraghi and M. Labra for providing laboratory facilities. C. Eizaguirre, R. Inger, and an anonymous referee provided useful comments on a previous draft of the manuscript. We also thank S. Clegg, T. Smith, and R. Harrigan for providing the microsatellite data from their breeding populations. Funding was provided by a Discovery Grant to CGG from The Natural Sciences and Engineering Research Council of Canada and by an operating grant from Environment Canada to KAH.

Conflict of Interest

None declared.

References

Bazzi, G., R. Ambrosini, M. Caprioli, A. Costanzo, F. Liechti, E. Gatti, et al. 2015. *Clock* gene polymorphism and scheduling of migration: a geolocator study of the barn swallow *Hirundo rustica*. Sci. Rep. 5:12443.

Bell-Pedersen, D., V. M. Cassone, D. J. Earnest, S. S. Golden, P. E. Hardin, T. L. Thomas, et al. 2005. Circadian rhythms from multiple oscillators: lessons from diverse organisms. Nat. Rev. Genet. 6:544–556.

Berthold, P. 1996. Control of bird migration. Springer, Berlin 355.

Bourret, A., and D. Garant. 2015. Candidate gene-environment interactions and their relationships with timing of breeding in a wild bird population. Ecol. Evol. 5:3628–3641.

Caprioli, M., R. Ambrosini, G. Boncoraglio, E. Gatti, A. Romano, M. Romano, et al. 2012. *Clock* gene variation is associated with breeding phenology and maybe under directional selection in the migratory barn swallow. PLoS ONE 7:e35140.

Chakarov, N., R. M. Jonker, M. Boerner, J. I. Hoffman, and O. Kruger. 2013. Variation at phenological candidate genes correlates with timing of dispersal and plumage morph in a sedentary bird of prey. Mol. Ecol. 22:5430–5440.

Clegg, S. M., J. F. Kelly, M. Kimura, and T. B. Smith. 2003. Combining genetic markers and stable isotopes to reveal population connectivity and migration patterns in a Neotropical migrant, Wilson's warbler (*Wilsonia pusilla*). Mol. Ecol. 12:819–830.

Comings, D. E. 1998. Polygenic inheritance and micro/minisatellites. Mol. Psychiatry 3:21–31.

Coppack, T., and F. Bairlein. 2011. Circadian control of nocturnal songbird migration. J. Ornithol. 152(Suppl. 1):67–73.

Dor, R., I. J. Lovette, R. J. Safran, S. M. Billerman, G. H. Huber, Y. Vortman, et al. 2011. Low variation in the polymorphic *Clock* gene poly-Q region despite population genetic structure across barn swallow (*Hirundo rustica*) populations. PLoS ONE 6:e28843.

Dor, R., C. B. Cooper, I. J. Lovette, V. Massoni, F. Bulit, M. Liljesthrom, et al. 2012. *Clock* gene variation in *Tachycineta* swallows. Ecol. Evol. 2:95–105.

Fondon, J. W. 3rd, E. A. Hammock, A. J. Hannan, and D. G. King. 2008. Simple sequence repeats: genetic modulators of brain function and behavior. Trends Neurosci. 31:328–334.

Fusani, L., and E. Gwinner. 2005. Melatonin and nocturnal migration. Ann. N. Y. Acad. Sci. 1046:264–270.

Goudet, J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9. 3).

Guo, S. W., and E. A. Thompson. 1992. A Monte Carlo method for combined segregation and linkage analysis. Am. J. Hum. Genet. 51:1111.

Gwinner, E. 1990. Bird migration: physiology and ecophysiology. Springer, Berlin 435.

Gwinner, E. 1996. Circadian and circannual programmes in avian migration. J. Exp. Biol. 199:39–48.

Hannibal, J., J. M. Ding, D. Chen, J. Fahrenkrug, P. J. Larsen, M. U. Gillette, et al. 1997. Pituitary adenylate cyclase-activating peptide (PACAP) in the retinohypothalamic tract: a potential daytime regulator of the biological clock. J. Neurosci. 17:2637–2644.

Hayasaka, N., S. I. LaRue, and C. B. Green. 2002. In Vivo Disruption of *Xenopus* CLOCK in the retinal photoreceptor cells abolishes circadian melatonin rhythmicity without affecting its production levels. J. Neurosci. 22:1600–1607.

Hobson, K. A., and L. I. Wassenaar. 2008. Tracking animal migration with stable isotopes. Academic Press, London 144.

Hobson, K. A., S. L. Van Wilgenburg, J. Faaborg, J. D. Toms, C. Rengifo, A. L. Sosa, et al. 2014. Connecting breeding and wintering grounds of Neotropical migrant songbirds using stable hydrogen isotopes: a call for an isotopic atlas of migratory connectivity. J. Field Ornithol. 85:237–257.

Johnsen, A., A. E. Fidler, S. Kuhn, K. L. Carter, A. Hoffmann, I. R. Barr, et al. 2007. Avian *Clock* gene polymorphism: evidence for a latitudinal cline in allele frequencies. Mol. Ecol. 16:4867–4880.

Joska, T. M., R. Zaman, and W. J. Belden. 2014. Regulated DNA methylation and the circadian clock: implications in cancer. Biology 3:560–577.

Kashi, Y., D. King, and M. Soller. 1997. Simple sequence repeats as a source of quantitative genetic variation. Trends Genet. 13:74–78.

Kelly, J. F., V. Atudorei, Z. D. Sharp, and D. M. Finch. 2002. Insights into Wilson's Warbler migration from analyses of hydrogen stable-isotope ratios. Oecologia 130:216–221.

Ko, C. H., and J. S. Takahashi. 2006. Molecular components of the mammalian circadian clock. Hum. Mol. Genet. 15: R271–R277.

Kyriacou, C. P., A. A. Peixoto, F. Sandrelli, R. Costa, and E. Tauber. 2008. Clines in clock genes: fine-tuning circadian rhythms to the environment. Trends Genet. 24:124–132.

Lemay, M. A., and M. A. Russello. 2014. Latitudinal cline in allele length provides evidence for selection in a circadian rhythm gene. Biol. J. Linn. Soc. 111:869–877.

Li, Y. C., A. B. Korol, T. Fahima, and E. Nevo. 2004. Microsatellites within genes: structure, function, and evolution. Mol. Biol. Evol. 21:991–1007.

Liedvogel, M., and M. Lundberg. 2014. The genetics of migration. Pp. 219–231 *in* L. A. Hansson, S. Åkesson, eds. Animal movements across scales. Oxford University Press, Oxford.

Liedvogel, M., and B. C. Sheldon. 2010. Low variability and absence of phenotypic correlates of *Clock* gene variation in a great tit *Parus major* population. J. Avian Biol. 41:543–550.

Liedvogel, M., M. Szulkin, S. C. Knowles, M. J. Wood, and B. C. Sheldon. 2009. 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.

Liedvogel, M., S. Akesson, and S. Bensch. 2011. The genetics of migration on the move. Trends Ecol. Evol. 26:561–569.

Liedvogel, M., C. K. Cornwallis, and B. C. Sheldon. 2012. Integrating candidate gene and quantitative genetic approaches to understand variation in timing of breeding in wild tit populations. J. Evol. Biol. 25:813–823.

Lincoln, G. A., H. Andersson, and A. Loudon. 2003. Clock genes in calendar cells as the basis of annual timekeeping in mammals - a unifying hypothesis. J. Endocrinol. 179:1–13.

Maggini, I., and F. Bairlein. 2010. Endogenous rhythms of seasonal migratory body mass changes and nocturnal restlessness in different populations of Northern Wheatear *Oenanthe oenanthe*. J. Biol. Rhythms 25:268–276.

Manolio, T. A., F. S. Collins, N. J. Cox, D. B. Goldstein, L. A. Hindorff, D. J. Hunter, et al. 2009. Finding the missing heritability of complex diseases. Nature 461:747–753.

Mettler, R., G. Segelbacher, and H. M. Schaefer. 2015. Interactions between a candidate gene for migration (*ADCYAP1*), morphology and sex predict spring arrival in blackcap populations. PLoS ONE 10:e0144587.

Morbey, Y. E., E. L. Jensen, and M. A. Russello. 2014. Time scale matters: genetic analysis does not support adaptation-by-time as the mechanism for adaptive seasonal declines in kokanee reproductive life span. Ecol. Evol. 4:3714–3722.

Mueller, J. C., F. Pulido, and B. Kempenaers. 2011. Identification of a gene associated with avian migratory behaviour. Proc. R. Soc. B Biol. Sci. 278:2848–2856.

Nagy, A. D., and V. J. Csernus. 2007. The role of PACAP in the control of circadian expression of clock genes in the chicken pineal gland. Peptides 28:1767–1774.

Newton, I. 2008. The migration ecology of birds. Academic Press, London.

Nowak, J. Z., and J. B. Zawilska. 2003. PACAP in avians: origin, occurrence, and receptors-pharmacological and functional considerations. Curr. Pharm. Des. 9:467–481.

O'Brien, C., L. Unruh, C. Zimmerman, W. E. Bradshaw, C. M. Holzapfel, and W. A. Cresko. 2013. 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.

Olano-Marin, J., J. C. Mueller, and B. Kempenaers. 2011. Heterozygosity and survival in blue tits (*Cyanistes caeruleus*): contrasting effects of presumably functional and neutral loci. Mol. Ecol. 20:4028–4041.

O'Malley, K. G., and M. A. Banks. 2008. 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.

O'Malley, K. G., M. J. Ford, and J. J. Hard. 2010. *Clock* polymorphism in Pacific salmon: evidence for variable selection along a latitudinal gradient. Proc. R. Soc. B Biol. Sci. 277:3703–3714.

Otahal, C. D.. 1995. Sexual differences in Wilson's Warbler migration. J. Field Ornithol., 66:60–69.

Panda, S., J. B. Hogenesch, and S. A. Kay. 2002. Circadian rhythms from flies to human. Nature 417:329–335.

Paxton, K. L., C. Van Riper, T. C. Theimer, and E. H. Paxton. 2007. Spatial and Temporal Migration Patterns of Wilson's Warbler (*Wilsonia Pusilla*) in the Southwest as Revealed by Stable Isotopes. Auk 124:162.

Peterson, M. P., M. Abolins-Abols, J. W. Atwell, R. J. Rice, B. Mila, and E. D. Ketterson. 2013. Variation in candidate genes CLOCK and ADCYAP1 does not consistently predict differences in migratory behavior in the songbird genus *Junco*. F1000Research, 2:115.

Pulido, F. 2007. The genetics and evolution of avian migration. Bioscience 57:165–174.

Pulido, F., and P. Berthold. 2003. Quantitative genetic analysis of migratory behaviour. Pp. 53–77 in P. Berthold, E. Gwinner and E. Sonnenschein, eds. Avian migration. Springer, Berlin.

Pulido, F., and M. Widmer. 2005. Are long-distance migrants constrained in their evolutionary response to environmental change? Causes of variation in the timing of autumn migration in a blackcap (*S. atricapilla*) and two garden warbler (*Sylvia borin*) populations. Ann. N. Y. Acad. Sci. 1046:228–241.

Pyle, P., S. N. G. Howell, R. P. Yunick, and D. F. DeSante. 1997. Identification guide to North American passerines. Slate Creek Press, Bolinas, CA.

Racz, B., G. Horvath, N. Faluhelyi, A. D. Nagy, A. Tamas, P. Kiss, et al. 2008. Effects of PACAP on the circadian changes of signaling pathways in chicken pinealocytes. J. Neurosci. 36:220–226.

Raymond, M., and F. Rousset. 1995. An exact test for population differentiation. Evolution 49:1280–1283.

Ritchison, G., J. R. Belthoff, and E. J. Sparks. 1992. Dispersal restlessness: evidence for innate dispersal by juvenile eastern screech-owls? Anim. Behav. 43:57–65.

Ross, C. A. 2002. Polyglutamine pathogenesis: emergence of unifying mechanisms for Huntington's disease and related disorders. Neuron 35:819–822.

Ruegg, K., E. C. Anderson, J. Boone, J. Pouls, and T. B. Smith. 2014a. A role for migration-linked genes and genomic islands in divergence of a songbird. Mol. Ecol. 23:4757–4769.

Ruegg, K. C., E. C. Anderson, K. L. Paxton, V. Apkenas, S. Lao, R. B. Siegel, et al. 2014b. Mapping migration in a songbird using high-resolution genetic markers. Mol. Ecol. 23:5726–5739.

Saino, N., M. Romano, M. Caprioli, M. Fasola, R. Lardelli, P. Micheloni, et al. 2013. Timing of molt of barn swallows is delayed in a rare *Clock* genotype. PeerJ 1:e17.

Saino, N., G. Bazzi, E. Gatti, M. Caprioli, J. G. Cecere, C. D. Possenti, et al. 2015. Polymorphism at the *Clock* gene predicts phenology of long-distance migration in birds. Mol. Ecol. 24:1758–1773.

Schwartz, C., and M. T. Andrews. 2013. Circannual transitions in gene expression: lessons from seasonal adaptations. Pp. 247–273 *in* A. E. Rougvie, M. B. O'Connor, eds. Developmental timing. Elsevier, Philadelphia, PA.

Simonneaux, V., A. Ouichou, and P. Pévet. 1993. Pituitary adenylate cyclase-activating polypeptide (PACAP) stimulates melatonin synthesis from rat pineal gland. Brain Res. 603:148–152.

Sokal, R. R., and F. J. Rohlf. 2009. Introduction to biostatistics. Dover Publications, New York.

Steinmeyer, C., J. C. Mueller, and B. Kempenaers. 2009. Search for informative polymorphisms in candidate genes: clock genes and circadian behaviour in blue tits. Genetica 136:109–117.

Tachibana, T., S. Saito, S. Tomonaga, T. Takagi, E.-S. Saito, T. Boswell, et al. 2003. Intracerebroventricular injection of vasoactive intestinal peptide and pituitary adenylate cyclase-activating polypeptide inhibits feeding in chicks. Neurosci. Lett. 339:203–206.

Tachibana, T., D. Oikawa, N. Adachi, T. Boswell, and M. Furuse. 2007. Central administration of vasoactive intestinal peptide and pituitary adenylate cyclase-activating polypeptide differentially regulates energy metabolism in chicks. Comp. Biochem. Physiol. A 147:156–164.

Terzer, S., L. I. Wassenaar, L. J. Araguás-Araguás, and P. K. Aggarwal. 2013. Global isoscapes for $\delta^{18}O$ and $\delta^{2}H$ in precipitation: improved prediction using regionalized climatic regression models. Hydrol. Earth Syst. Sci. 17:4713–4728.

Vaudry, D., A. Falluel-Morel, S. Bourgault, M. Basille, D. Burel, O. Wurtz, et al. 2009. Pituitary adenylate cyclase-activating polypeptide and its receptors: 20 years after the discovery. Pharmacol. Rev. 61:283–357.

Visser, M. E., S. P. Caro, K. van Oers, S. V. Schaper, and B. Helm. 2010. Phenology, seasonal timing and circannual rhythms: towards a unified framework. Philos. Trans. R. Soc. B Biol. Sci. 365:3113–3127.

Wassenaar, L. I., and K. A. Hobson. 2003. Comparative equilibration and online technique for determination of non-exchangeable hydrogen of keratins for use in animal migration studies. Isot. Environ. Health Stud. 39:211–217.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Determination of feather $\delta^{2}H$ and genetic analyses.

Appendix S2. Comparison of F_{ST} of neutral microsatellites (from Clegg et al. 2003) between southern and northern Wilson's warblers.

Table S1. Wilson's warbler *Adcyap1* allele frequencies, observed heterozygosity (H_o) and mean (SE) allele size.

Table S2. Phenotypic distribution of migration date of male and female Wilson's warblers homozygotes for *Clock*.

Table S3. Variation in the strength of the association between *Adcyap1* allele size and feather δ^2H value among southern and northern Wilson's warblers based on different thresholds to identify southern and northern birds.

Table S4. Correlation between mean allele size at 8 neutral microsatellite markers (from Clegg et al. 2003) and breeding latitude or predicted feather δ^2H of the breeding site.

Supporting Information

Adcyap1 polymorphism covaries with breeding latitude in a Nearctic migratory songbird, the Wilson's warbler (*Cardellina pusilla*)

Gaia Bazzi, Andrea Galimberti, Quentin R. Hays, Ilaria Bruni, Jacopo G. Cecere, Luca Gianfranceschi, Keith A. Hobson, Yolanda E. Morbey, Nicola Saino, Christopher G. Guglielmo, Diego Rubolini

List of items:

Appendix S1. Determination of feather δ^2H and genetic analyses.

Appendix S2. Comparison of F_{ST} of neutral microsatellites (from Clegg *et al.* 2003) between southern and northern Wilson's warblers.

Table S1. Wilson's warbler *Adcyap1* allele frequencies, observed heterozygosity (H_o) and mean (SE) allele size.

Table S2. Phenotypic distribution of migration date of male and female Wilson's warblers homozygotes for *Clock*.

Table S3. Variation in the strength of the association between *Adcyap1* allele size and feather δ^2H among southern and northern Wilson's warblers based on different thresholds to identify southern and northern birds.

Table S4. Correlation between mean allele size at 8 neutral microsatellite markers (from Clegg *et al.* 2003) and breeding latitude or predicted feather δ^2H of the breeding site.

Appendix S1. Determination of feather δ^2H and genetic analyses

Determination of feather δ^2H

Tail feathers were first washed in a 2:1 chloroform/methanol solution to remove surface oils, rinsed and air dried. A small sample (ca. 0.35 mg) was cut from the distal vane of the feather and transferred to a silver capsule. In order to avoid uncontrolled exchange between non carbon-bound hydrogen in the feathers and ambient water vapor, δ^2H analyses were performed using the comparative equilibration approach reported in Wassenaar and Hobson (2003). Feather and keratin standard δ^2H values were analysed on H_2 gases produced by flash pyrolysis under He flow. Within each feather run, we measured 5 replicates of the Environment Canada keratin reference materials CBS (-197‰), SPK (-121.6‰) and KHS (-54.1‰). Based on replicate measurements of standards, measurement error was estimated as ± 2‰.

Genetic analyses

DNA was extracted from feathers using the Dneasy® Blood & Tissue Kit (Qiagen, Hilden, Germany) following manufacturer's instructions (i.e. Purification of Total DNA from Animal Tissues protocol), except for pre-treatment and elution steps. The basal portion (i.e., the portion of feather calamus from the tip to 2-3 mm above the superior umbilicus) of rectrices and breast contour feathers was isolated and incubated for at least six hours in sterile NaCl solution (0.9%) in agitation at 4°C. After rehydration, the calamus was chopped in small pieces on a sterile glass slide by using sterile instruments and incubated overnight in lysis buffer. After purification steps, DNA was eluted in 50 µl of deionized water (molecular biology grade) and DNA quality and concentration were measured fluorometrically with a NanoDrop™ 1000 Spectrophotometer (Thermo Scientific, USA). Polymorphism at the *Clock* and *Adcyap1* genes (allele size due to the number of sequence repeats) were determined by fragment analysis as described in Caprioli *et al.* (2012) and Saino *et al.* (2015), with some modifications. Differently from these two previous studies, given the lower DNA concentration of the collected samples, PCRs were conducted starting from 10 ng of DNA by using puReTaq Ready-To-Go PCR beads (Amersham Bioscience, Freiburg, Germany) in a 25 µl reaction according to the manufacturer's

instructions. Moreover, for both markers, the final elongation step (72°C) of PCR thermal profile was prolonged up to 30 minutes. PCR products (1 μl) were mixed with 12 μl of formamide and 0.2 μl of LIZ-500 size standard (Applied Biosystems, Foster City, CA, USA) and then analyzed on an ABI 3130 automated sequencer (Applied Biosystems).

The allele sizes for each locus were identified using Genemapper 4.0 software (Applied Biosystems). Genotyping (i.e. amplification and fragment electrophoresis) was repeated in samples with low or missing peaks at least two times before excluding them from the dataset.

Appendix S2. Comparison of F_{ST} of neutral microsatellites (from Clegg *et al.* 2003) between southern and northern Wilson's warblers.

We reanalysed microsatellite data previously published by Clegg *et al.* (2003) to estimate F_{ST} at (presumably) neutral loci from breeding birds belonging to the same northern and southern population clusters of our study. We considered the data from the populations labelled JUN, WA and BC in Table 1 of Clegg *et al.* (2003) as belonging to the northern population cluster (n = 42 individuals), whereas the population labelled CA was considered as belonging to the southern population cluster (n = 17 individuals). As in Clegg *et al.* (2003), for estimating F_{ST} we discarded data from three out eight microsatellites because of a significant heterozygote deficit. The estimated mean F_{ST} value across all 5 loci was -0.008 (95% bootstrap confidence intervals: -0.019 to 0.001) and not significantly different from 0 ($P = 0.05$). Hence, the F_{ST} point estimate for the *Adcyap1* microsatellite (0.045), besides being significantly different from 0, was considerably larger (outside the 95% confidence intervals) than the mean F_{ST} of the other five neutral microsatellites. Such a low level of genetic differentiation based on neutral microsatellites is in line with the lack of detectable divergence among western breeding populations of the Wilson's warbler originally reported in Clegg *et al.* (2003). We note that significant patterns of genetic divergence among western breeding populations was later uncovered by Ruegg *et al.* (2014) by adopting a genomic approach.

References

Caprioli, M., Ambrosini, R., Boncoraglio, G., Gatti, E., Romano, A., Romano, M., Rubolini, D., Gianfranceschi, L. & Saino, N. (2012) *Clock* gene variation is associated with breeding phenology and maybe under directional selection in the migratory barn swallow. *PLoS ONE,* **7,** e35140.

Clegg, S.M., Kelly, J.F., Kimura, M. & Smith, T.B. (2003) Combining genetic markers and stable isotopes to reveal population connectivity and migration patterns in a Neotropical migrant, Wilson's warbler (*Wilsonia pusilla*). *Molecular Ecology,* **12,** 819-830.

Ruegg, K.C., Anderson, E.C., Paxton, K.L., Apkenas, V., Lao, S., Siegel, R.B., DeSante, D.F., Moore, F. & Smith, T.B. (2014) Mapping migration in a songbird using high-resolution genetic markers. *Molecular Ecology,* **23,** 5726-5739.

Saino, N., Bazzi, G., Gatti, E., Caprioli, M., Cecere, J.G., Possenti, C.D., Galimberti, A., Orioli, V., Bani, L., Rubolini, D., Gianfranceschi, L. & Spina, F. (2015) Polymorphism at the *Clock* gene predicts phenology of long-distance migration in birds. *Molecular Ecology,* **24,** 1758-1773.

Wassenaar, L.I. & Hobson, K.A. (2003) Comparative equilibration and online technique for determination of non-exchangeable hydrogen of keratins for use in animal migration studies. *Isotopes in Environmental and Health Studies,* **39,** 211-217.

Table S1. Wilson's warbler *Adcyap1* allele frequencies, observed heterozygosity (H_o) and mean (SE) allele size.

	154	156	157	158	159	160	161	162	164	H_o	Mean allele size
Adcyap1	0.010	0.225	0.118	0.240	0.020	0.211	0.005	0.147	0.025	0.775	158.58 (0.16)

Table S2. Phenotypic distribution of migration date of male and female Wilson's warblers homozygotes for *Clock*. Migration date is expressed as residuals of a linear regression of migration date on δ^2H, separately for each sex, in order to control for the effect of geographical origin on migration date (see Fig. 3). Mean values are shown together with SD (round brackets) and non-parametric bootstrap 95% confidence limits (see Methods) (square brackets). The values for the two heterozygote individuals are shown on the rightmost column: both values fall within the 95% confidence limits of the homozygous males and females, respectively.

	Homozygotes	Heterozygote
Males (n = 55)	0.01 (12.83) [-2.98 to 3.74]	-0.30
Females (n = 45)	0.05 (12.12) [-3.68 to 3.43]	-2.31

Table S3. Variation in the strength of the association between *Adcyap1* allele size and feather δ^2H among southern and northern Wilson's warblers based on different thresholds to identify southern (S) and northern birds (N) [N: feather $\delta^2H < -130$ ‰ or < -140 ‰; S: feather $\delta^2H \geq -80$ ‰ or ≥ -90 ‰; see Methods for details]. Numbers in parentheses denote sample sizes. Pairs of letters highlight statistically significant ($P < 0.01$) Z-tests of the difference between correlation coefficients of southern and northern birds, with letters indicating different correlations (superscripts). Significance levels of correlation coefficients: * = $P < 0.05$; ** = $P < 0.01$; *** = $P < 0.001$.

	S (≥ -80 ‰)[a]	S (≥ -90 ‰)[b]	N (< -130 ‰)[c]	N (< -140 ‰)[d]	Z-test
Mean allele size	0.07 (35)	-0.12 (43)	-0.69 (22)***	-0.71 (17)**	ac-ad-bc-bd
Short allele size	-0.05 (35)	-0.13 (43)	-0.31 (22)	-0.40 (17)	
Long allele size	0.05 (35)	-0.07 (43)	-0.77 (22)***	-0.73 (17)**	ac-ad-bc-bd

Table S4. Correlation between mean allele size at 8 neutral microsatellite markers (from Clegg *et al.* 2003) and breeding latitude or predicted feather δ^2H of the breeding site. Microsatellite data refer to the breeding populations labeled JUN, WA and BC in Table 1 of Clegg *et al.* (2003). These populations occur in the latitude range spanning between 48°N and 58°N, corresponding to predicted feather δ^2H spanning between -123 ‰ and -157 ‰ based on our feather isoscape (see Fig. 1). Hence, the range of these three populations entirely encompasses the 'northern breeding populations' as defined in this study (δ^2H < -130 ‰; see Methods). As the original microsatellite data from these three populations were not matched to individual-based isotope data (see Clegg *et al.* 2003), for each breeding population we calculated the predicted mean feather δ^2H from a buffer of 50 km around each breeding site from the raster of the feather isoscape depicted in Fig. 1. Correlation coefficients and sample size (number of individuals; in parentheses) are shown. None of the correlations is statistically significant (all $P > 0.09$).

Locus	Latitude	Feather δ^2H
*Wp*C6	0.120 (41)	-0.127 (41)
*Wp*D23	0.055 (42)	-0.052 (42)
*Wp*D30	0.133 (41)	-0.126 (41)
*Wp*D4	0.248 (42)	-0.240 (42)
*Dpμ*01	0.212 (41)	-0.198 (41)
*Dpμ*03	-0.085 (42)	0.084 (42)
*Dpμ*05	-0.119 (39)	0.121 (39)
*Dpμ*16	-0.257 (41)	0.265 (41)

CHAPTER 3

Sex-specific associations between *Clock*, *Npas2* and *Creb1* allelic variation, timing of spring migration and moult speed in the willow warbler (*Phylloscopus trochilus*)

Gaia Bazzi, Stefano Podofillini, Emanuele Gatti, Luca Gianfranceschi, Jacopo G. Cecere, Fernando Spina, Nicola Saino and Diego Rubolini

Manuscript to be submitted

Photo: Andrea Galimberti

Sex-specific associations between *Clock*, *Npas2* and *Creb1* allelic variation, timing of spring migration and moult speed in the willow warbler (*Phylloscopus trochilus*)

Gaia Bazzi[1], Stefano Podofillini[1], Emanuele Gatti[1], Luca Gianfranceschi[1], Jacopo G. Cecere[2], Fernando Spina[2], Nicola Saino[1] and Diego Rubolini[1]

1. Dipartimento di Bioscienze, Università degli Studi di Milano, via Celoria 26, I-20133, Milan, Italy

2. ISPRA – Istituto Superiore per la Protezione e la Ricerca Ambientale, via Cà Fornacetta 9, I-40064 Ozzano dell'Emilia (BO), Italy

Keywords: *Adcyap1*, candidate genes, phenology, ptilochronology

Abstract

The timing of major life-history events, such as migration and moult, is set by endogenous circadian and circannual clocks, that have been well characterized at the molecular level. Conversely, the genetic sources of variation in phenology and in other behavioural traits have been sparsely addressed. It has been proposed that inter-individual variability in the timing of seasonal events may arise from allelic polymorphism at phenological candidate genes involved in the signalling cascade of the endogenous clocks. In this study of a long-distance migratory passerine bird, the willow warbler (*Phylloscopus trochilus*), we investigated whether allelic variation at four candidate genes (*Adcyap1*, *Clock*, *Creb1* and *Npas2*), predicted two major components of the annual schedule, timing of spring migration across the central Mediterranean sea and moult speed, the latter gauged from ptilochronological analyses of tail feathers moulted in the African winter quarters. *Npas2* and *Clock* allele size predicted spring migration date, with individuals bearing longer alleles migrating significantly earlier than those bearing shorter alleles. The allele size-phenology association was statistically significant for males only for *Npas2* gene and for females (but not for males) for *Clock* gene. Moreover, *Creb1* allele size significantly predicted male (but not female) moult speed, longer alleles being associated with faster moult. All other genotype-phenotype associations were statistically non-significant. These findings provide new evidence for a role of candidate genes in modulating the phenology of different circannual activities in long-distance migratory birds, and for the occurrence of sex-specific candidate gene effects.

Introduction

The annual schedule of migratory birds is controlled by an endogenous program, which is synchronized with seasonal changes primarily by daily changes in photoperiod (e.g. Gwinner 1986, Gwinner 2003, Sharp 2005, Pulido 2007, Visser et al. 2010). The endogenous clock that modulates circadian and circannual rhythmicity has been extensively studied in several organisms, from prokaryotes to vertebrates, and the genes controlling such mechanisms have been well characterized (Bell-Pedersen et al. 2005). Conversely, the genetic bases of phenotypic variation in the timing of seasonal events are poorly understood, and only a few genes have been rather firmly linked to phenological variability in wild organisms (e.g. the Clock gene Liedvogel et al. 2009, Caprioli et al. 2012, Saino et al. 2015a).

It has been suggested that differences in the timing of life-history events among individuals could arise from polymorphism at genes involved in the signalling cascade of the endogenous clock (Visser et al. 2010). Studies of among-individuals and among-populations phenological variability in vertebrate species have mainly focused on length polymorphism at four candidate genes, namely *Adcyap1*, *Clock*, *Creb1* and *Npas2* (e.g. Liedvogel et al. 2009, O'Malley et al. 2010, Caprioli et al. 2012, Chakarov et al. 2013, Bourret and Garant 2015). *Clock* and its paralogue *Npas2* are two components of the endogenous clock and show a polymorphic polyglutamine (poly-Q) repeat sequence in their exonic regions (Fidler and Gwinner 2003, Steinmeyer et al. 2009). On the other hand, short tandem repeat sequences at 3' UTR have been reported for *Creb1*, a transcription factor involved in the light-induced clock entrainment (Gau et al. 2002, Tischkau et al. 2003), and *Adcyap1*, encoding for PACAP, a neurotransmitter with several biological functions related to the circadian and circannual rhythmicity (Simonneaux et al. 1993, Hannibal et al. 1997, Nagy and Csernus 2007, Racz et al. 2008, Schwartz and Andrews 2013).

In birds, allele size variations at *Clock* and *Npas2* have been linked with differences in the timing of breeding among individuals, longer alleles being associated with delayed reproduction

and in some species with shorter incubation periods (Liedvogel et al. 2009, Caprioli et al. 2012, Bourret and Garant 2015, but see Liedvogel and Sheldon 2010, Dor et al. 2012). Moreover, timing of migration (Bazzi et al. 2015, Saino et al. 2015a) and winter moult (Saino et al. 2013) was delayed among individuals bearing longer *Clock* alleles in some long-distance migratory bird species. Polymorphism at *Adcyap1* and *Creb1* genes was found to be associated with juvenile dispersal behaviour in buzzards (*Buteo buteo*): individuals dispersing earlier carried longer *Adcyap1* alleles and shorter *Creb1* alleles than those dispersing later (Chakarov et al. 2013). Furthermore *Creb1* allele size was related to incubation duration in male tree swallows (*Tachycineta bicolor*), though in combination with spring temperatures only (Bourret and Garant 2015). Finally, *Adcyap1* allele size was associated with laying date in female tree swallows; however, the direction of the association varied with latitude, being negative at lower latitudes but becoming positive at higher latitudes (Bourret and Garant 2015). Although other studies did not find any relationship between candidate genes and phenology (e.g. Liedvogel and Sheldon 2010, Dor et al. 2012), there is evidence that polymorphism at some of such genes is associated with other behavioural traits that may be indirectly linked to circannual rhythms and/or photoperiodic response, such as migratory restlessness and migration distance (Mueller et al. 2011, Peterson et al. 2013, Bazzi et al. 2016). Moreover, a latitudinal cline in the frequency of alleles of different length has been reported for *Clock* and *Adcyap1* in a few species (Johnsen et al. 2007, Bazzi et al. 2016, but see Kuhn et al. 2013); allele size of both candidate genes increased northwards, hinting a possible role of polymorphism in the adaptation to different photoperiodic regimes or to the timing of breeding season, that is delayed and shorter at higher latitudes (Gwinner 1986, Berthold 1996, Johnsen et al. 2007, Bazzi et al. 2016).

On the whole, these findings suggest that candidate genes polymorphism may underlie variability in the timing of life-history events through the whole life-cycle of birds and at different life stages, but the general picture is still patchy. In this study of the willow warbler (*Phylloscopus trochilus*) we aimed at assessing whether length polymorphism at five candidate genes (the

98 previously studied *Adcyap1*, *Clock*, *Creb1* and *Npas2* genes and a newly identified polymorphic
99 region of *Clock* gene, *Clock3*) predicted the timing of spring migration and the speed of winter
100 moult, as assessed by measuring the growth rate of tail feathers by means of ptilochronological
101 techniques (Grubb 2006). We assumed that a larger feather growth rate corresponds to a faster
102 moult (Bensch and Grahn 1993, de La Hera et al. 2009). The willow warbler, a small (ca. 10 g)
103 trans-Saharan migratory passerine that breeds in Eurasia at medium-high latitudes and overwinters
104 in sub-Saharan Africa, is among the few species that perform two complete annual moults, one of
105 which occurs during winter, while the birds are in Africa (Underhill et al. 1992). Birds leave for
106 spring migration in late February-March and reach the breeding grounds in mid-March to late May
107 (Cramp 1998), and were sampled during spring migration across the central Mediterranean sea. We
108 expected migration date to be delayed among birds with longer *Clock* and *Npas2* alleles.
109 Conversely, due to the variable genotype-phenotype associations detected by previous studies, we
110 had no clear predictions on the allele size-phenology or feather growth rate association for the other
111 candidate genes.
112
113
114 **Materials and Methods**
115
116 **Field methods**
117 Willow warblers were sampled at Ventotene (40°48'N – 13°25'E), a small island located in the
118 central Mediterranean Sea, ca. 50 km off the Italian coast. We captured birds during the period
119 March 22 -May 27 2013; the sampling period encompassed the entire spring migration period of the
120 study species at Ventotene (Spina et al. 1993, Messineo et al. 2001, Saino et al. 2010). Birds were
121 trapped using mist-nets following standard capture protocols and banded with metal rings (Spina et
122 al. 1993, Saino et al. 2010). Wing length (length of the 8th outermost primary feather, Jenni and
123 Winkler 1994) and tarsus length were recorded for all birds. Wing length and tail length (but not

124 tarsus length) can be used as proxies of breeding destination among willow warblers breeding in
125 Fennoscandia (Bensch et al. 1999): both wing and tail length strongly increased with breeding
126 latitude ($r^2 > 0.58$). Since the willow warblers migrating through the central Mediterranean are
127 directed mostly towards Fennoscandia (Jonzén et al. 2006a, Jonzén et al. 2006b), we used wing
128 length and tail feather length (see Ptilochronological analyses; wing and tail feather length were
129 strongly positively correlated in our sample of birds: males, $r = 0.83$; females, $r = 0.87$) as proxies
130 of breeding destination (separately for each sex as males are significantly larger than females; see
131 Cramp 1998 and Results). Birds usually rest on Ventotene for a few hours before resuming their
132 travel towards breeding quarters (Goymann et al. 2010, Tenan and Spina 2010). We considered
133 only first capture dates (i.e. we excluded recaptures of birds previously ringed at the study site
134 during the same season). We assumed that the distribution of first capture dates reflects the
135 phenology of species's timing of spring migration at Ventotene (see Saino et al. 2010, Saino et al.
136 2015a).
137 We aimed at sampling ca. 100 individuals, evenly distributed along the whole spring
138 migration season. According to the number of willow warblers captured during the previous years
139 (2006-2011), we sampled one every eight individuals (see Saino et al. 2010, Saino et al. 2015a). For
140 each individual we collected a small blood (ca. 10-30 µl, collected in heparinized capillary tubes
141 and stored at -20 °C) or feather (3/4 undertail coverts, stored in 99% ethanol at room temperature)
142 sample as a source of DNA. Moreover, the fourth outermost rectrix (Jenni and Winkler 1994;
143 hereafter R_4) was collected and stored in individual bags for ptilochronological analyses.
144
145 **Ptilochronological analyses**
146 Moult speed was indirectly assessed by measuring growth bar width (GBW) on R_4 (see de La Hera
147 et al. 2009). A single feather growth bar consists of one light band and one dark band, which
148 correspond to the portion of the feather grown during a single night-day cycle (Brodin 1993). Wider
149 growth bars reflect faster feather growth (Grubb 2006). Although moult speed depends on the

number of feather growing simultaneously, as well as on the individual feather growth rate, it has been shown that individuals with high feather growth rates moult many feathers at the same time (Bensch and Grahn 1993). Hence, we can assume that GBW could be safely used to analyze moult speed (see de La Hera et al. 2009). We measured the width of six bars, three on either side of a point located at two thirds of the feather length (modified from Grubb 2006 and de la Hera et al. 2009 according to the number of growth bars clearly recognizable on willow warblers' R_4). Total bars width was measured with a digital caliper (in mm, 0.01 mm resolution) on the dorsal surface of the vane. GBW was expressed as total bars width/6. In order to avoid any bias, all feather measures were taken by the same person (SP). Repeatability of GBW, as assessed on a sample of feathers measured twice, was very high (n = 20, $r = 0.96$, $P < 0.001$).

After measuring GBW, feathers were taped to tracing paper across the shaft, and scanned; tail feather length (in mm; 0.01 mm resolution) was measured on the resulting images using the 'segmented line' tool of ImageJ 1.46r software (rsbweb.nih.gov) (Saino et al. 2015b). Individuals whose feather tips were broken, for which feather length could thus not be measured, were excluded from moult speed analyses. We obtained GWB from 119 individuals.

Genetic analyses

Genomic DNA was extracted from blood samples by means of alkaline lysis of 6 µl of blood in 100 µl of a 50 mM NaOH solution at 100°C for 20 minutes. Extracted DNA was quantified using a spectrophotometer and diluted to a final concentration of 50–100 ng/µL. Genomic DNA from feathers was extracted using a commercial kit (5 PRIME, ArchivePure DNA purification kit, Hilden, Deutschland). The procedure is described in details in Saino et al. (2015a).

Willow warblers are sexually size dimorphic (males are larger than females) but sexually monochromatic (Cramp 1998), and sex cannot be determined in the field. Hence, sex was determined using CHD1 primers (for DNA extracted from blood samples; details on primers and PCR amplification in Saino et al. 2015a). As PCR amplification performed on DNA extracted from

feathers with CHD1 primers did not produce reliable results, we designed a new set of primers on *Passer montanus* CHD gene (Sequence ID in GenBank: gb|GU370350.1|): PassexF 5'-GAGAAACTGTGCAAAACAGG-3' and PassexR 5'-GAGTCACTATCAGATCCAGARTATC-3'. PCR amplification were performed in a final volume of 15 µl, with 6 µl DNA solution, 1× PCR buffer (Promega), 1.5 mM of Mg2+, 0.3 µl of each primer (stock 10 mM), 1.5 µl of dNTPs (stock 2 mM) and 1 U Taq DNA polymerase (Promega). PCR amplification profile was as follows: 95 °C for 3 min, 35 cycles at 95 °C for 45 s, 55 °C for 45 s and 72 °C for 50 s and further extension at 72 °C for 5 min. PCR products were then separated on 2.5% agarose gel and visualized after ethidium bromide staining.

We assessed polymorphism at *Adcyap1, Creb1* and *Npas2* genes and at two polymorphic *Clock* gene regions [*Clock1* and *Clock3*; *Clock1* was the region investigated in Johnsen et al. (2007) and subsequent studies on *Clock* gene polymorphism, while *Clock3* was a newly identified polymorphic region; see below] by means of PCR amplification followed by fragment analysis. Primers for *Adcyap1* PCR amplifications were taken from Saino et al. (2015a), whereas *Clock1* primers are described in Caprioli et al. (2012). Finally, *Creb1* and *Npas2* primers correspond to those described in Steinmeyer et al. (2009) [with the slight modifications proposed by Bourret and Garant (2015) for the *Creb1* gene]. The *Clock3* locus was identified by aligning all *Clock* avian gene sequences available in GenBank and searching for polymorphic regions that vary in number of CAG repeats; *Clock3* primers (*Clock3*F 5'-TCTGCTGCTTTCCCACTACA-3' and *Clock3*R 5'-ATCAGTCATCTTGTCAGTTCTGTG-3') were designed *ex novo*.

PCR amplification was performed using a commercial kit (Quiagen, Multiplex PCR Kit) in a final volume of 25/26 µl with 12.5 µl 2x QUIAGEN Multiplex PCR Master Mix, 2.5 µl 10x primer mix (0.5 µl for each primer), 2 µl RNase-free water (for genomic DNA extracted from blood only), 5 µl 5x Q-Solution and 3 µl of DNA solution (6 µl for DNA extracted from feather samples). PCR amplification profile was: 95 °C for 15 min, 35 cycles at 94 °C for 30 s, 56 °C for 90 s, 72 °C for 60 s and a final extension at 60 °C for 30 min. PCR products were labelled with 6-FAM (*Clock1* and

202 *Creb1*), HEX (*Clock*3 and *Npas2*), or TAMRA (*Adcyap1*). Polymorphism at candidate genes was
203 determined using fragment analysis (Macrogen Inc., Seoul, Republic of Korea) (see Caprioli et al.
204 2012, Bazzi et al. 2015, Saino et al. 2015a).
205
206 **Statistical analyses**
207 First, we tested for deviations from Hardy–Weinberg equilibrium (HWE) for the five loci using the
208 Markov chain method (Guo and Thompson 1992) implemented in GENEPOP (dememorization =
209 1000, batches = 100, iterations per batch = 1000). Moreover, we explored genetic differentiation
210 between the sexes for the five loci separately and for the combination of all loci by estimating F_{ST}
211 between males and females using Fstat 2.9.3 software (Goudet 2001).
212 Since length polymorphisms detected at the two *Clock* polymorphic regions could affect
213 phenotype simultaneously, we also considered in the models the sum of *Clock1* and *Clock3* allele
214 sizes (*sumClock* hereafter). To investigate the association between candidate genes allele size and
215 migration date, while controlling for variation in migration timing due to sex, we ran six separate
216 linear models of migration date (1 = January 1) as a function of sex (0 = females, 1 = males) and
217 *Adcyap1*, *Clock1*, *Clock3*, *sumClock*, *Creb1*, or *Npas2* allele size (mean of long and short allele
218 size). Each model tested the effect of a single candidate gene on migration date. Since any possible
219 association between candidate genes may arise from a latitudinal cline of allele size, we accounted
220 for the possible confounding effects of wing length (that increases with latitude across Europe in
221 several passerine bird species; Cramp 1998, Bensch et al. 1999, Peiró 2003, Evans et al. 2009,
222 Tarka et al. 2010), by including wing length as a covariate. Since previous studies highlighted a
223 dominance of the longer allele in shaping phenology and other phenotypic traits for *Clock* and other
224 genes with poly-Q polymorphism (Ross 2002, Fondon et al. 2008, Liedvogel et al. 2009, Saino et
225 al. 2015a), we ran five separate models similar to those described above (for *Adcyap1*, *Clock1*,
226 *Clock3*, *Creb1* and *Npas2*), in which the long instead of the mean allele size was included as a
227 predictor.

Finally, we tested whether candidate genes allele size predicted GBW. Since GBW and feather length were strongly correlated ($r = 0.51$, $P < 0.001$), to control for the effect of feather length on GBW we computed the residuals of a linear regression of GBW on feather length (feather growth rate, FGR hereafter). We then ran six separate linear models of FGR as a function of sex and *Adcyap1*, *Clock1*, *Clock3*, *sumClock*, *Creb1*, or *Npas2* allele size. We also tested for the allelic dominance of the long allele by including long, instead of mean, allele size in a further similar set of models.

Results

Migration phenology and morphology

The willow warbler is a highly protandrous species, with mean migration date of males [99.6 (11.0 SD), n = 63] being much earlier than that of females [117.0 (12.5 SD), n = 60; $t_{117.4} = -8.24$, $P < 0.001$; see also Saino *et al.* 2010]. Males were significantly larger than females for all biometrics [wing length, males: 53.3 mm (1.9 SD), females: 49.6 mm (1.9 SD); tail feather length, males: 56.5 mm (1.8 SD), females: 52.5 (2.2 SD); tarsus length, males: 19.7 mm (0.7 SD), females: 18.6 mm (0.6 SD); all $t > 9.39$, all $P < 0.001$] (see also Cramp 1998).

Wing and tail feather length of males did not significantly vary with migration date [wing length, estimate: -0.20 (0.22 SE), $t_{61} = -0.92$, $P = 0.36$; tail feather length, estimate: -0.001 (0.002 SE), $t_{61} = -0.02$ $P = 0.99$]. On the other hand, wing and tail length of females significantly declined with migration date [wing length, estimate: -0.50 (0.19 SE), $t_{58} = -2.60$, $P = 0.012$; tail feather length, estimate: -0.006 (0.002), $t_{58} = -2.61$, $P = 0.011$]. Tarsus length did not significantly vary with migration date in both sexes (males, estimate: -0.12 (0.08 SE), $t_{60} = -1.60$, $P = 0.12$; females, estimate: -0.09 (0.06 SE), $t_{58} = -1.51$, $P = 0.14$). Hence, these analyses suggest that, at least among

253 females, birds from northern populations migrate earlier across the central Mediterranean than those
254 from southern populations.
255
256 **Candidate genes variation**
257 We successfully genotyped 93 to 112 individuals, depending on locus (Table 1). Polymorphism
258 broadly varied among candidate genes: the *Clock3* locus showed rather low variability, with two
259 alleles only (108 and 111), one of which accounted for the 90.2% of allelic variability (Table 1); on
260 the other hand, the *Adcyap1* locus was high variable (Table 1), with 10 alleles detected. The other
261 candidate genes showed intermediate observed levels of heterozygosity (Table 1). The *Creb1* locus
262 significantly deviated from the Hardy-Weinberg equilibrium ($P < 0.001$), while this was not the
263 case for the other loci ($P > 0.21$). Allele frequencies of males and females were similar for all loci,
264 as indicated by the small F_{ST} values (*Adcyap1*: $F_{ST} = 0.001$, $P = 0.20$; *Clock1*: $F_{ST} = -0.002$, $P =$
265 0.75; *Clock3*: $F_{ST} = -0.009$, $P = 0.70$; *Creb1*: $F_{ST} = 0.015$, $P = 0.30$; *Npas2*: $F_{ST} = -0.005$, $P = 0.70$;
266 all loci pooled: $F_{ST} = 0.008$, $P = 0.25$).
267
268 **Candidate genes, timing of migration and morphology**
269 *Npas2* allele size predicted migration date in a different way according to sex, as indicated by the
270 statistically significant sex × *Npas2* interaction ($F_{1,88} = 6.95$, $P = 0.010$) in a linear model including
271 *Npas2* allele size, sex, wing length and the *Npas2* × sex interaction as predictors (Table 2). Indeed,
272 the *Npas2*-migration date association was statistically significant for males but not for females
273 [males, *Npas2* allele size, estimate: -6.90 (1.96 SE), $F_{1,44} = 12.47$, $P = 0.001$; wing length, estimate:
274 -0.09 (0.09 SE), $F_{1,44} = 1.11$, $P = 0.30$; females, *Npas2* allele size, estimate: 0.01 (1.47 SE), $F_{1,43} =$
275 0.001, $P = 0.99$; wing length, estimate: -0.13 (0.09 SE), $F_{1,43} = 1.85$, $P = 0.18$]. Conversely, none of
276 the other candidate genes was statistically significantly associated with migration date (Table 2).
277 Results were similar if tail feather length was used instead of wing length as a predictor and for
278 models in which we did not account for morphology (details not shown for brevity). Results for

279 *Adcyap1*, *Clock3*, *Creb1* and *Npas2* were qualitatively similar when using long allele size instead of
280 mean allele size (datails not shown for brevity).

281 On the other hand, in a linear model of migration date as a function of long *Clock1* allele size,
282 sex and wing length, long *Clock1* allele size significantly and negatively predicted migration date in
283 combination with sex, as indicated by the statistically significant long *Clock1* allele size × sex
284 interaction [allele size × sex, estimate: -2.52 (1.12 SE), $F_{1,116} = 5.08$, $P = 0.026$; other model details
285 not shown for brevity]. The effect of long *Clock* allele size on migration date was statistically
286 significant for females, but not for males in sex-specific models including wing length as a
287 covariate [females, long *Clock1* allele size, estimate: -1.89 (0.87 SE), $F_{1,56} = 4.79$, $P = 0.033$; wing
288 length, estimate: -0.21 (0.08 SE), $F_{1,56} = 7.35$, $P = 0.009$; males, long *Clock1* allele size, estimate:
289 0.15 (1.03 SE), $F_{1,59} = 0.02$, $P = 0.89$; wing length, estimate: -0.07 (0.08), $F_{1,59} = 0.82$, $P = 0.37$]
290 (Fig. 2). All these results were qualitatively unaltered of tail feather length was included in the
291 models instead of wing length and if we did not account for the possible confounding effects of
292 morphology (details not shown).

293 Wing and tail feather length did not significantly covary with allele size of any locus in
294 models separated by either sex (all $P > 0.08$).

295

296 **Candidate genes and moult**

297 FGR did not significantly differ between the sexes [males: 0.02 (0.17 SD), n = 59; females: 0.01
298 (0.18 SD), n = 58; $t_{115} = -0.96$; $P = 0.34$]. FGR was significantly predicted by *Creb1* allele size in
299 combination with sex, as gauged by the statistically significant interaction between *Creb1* allele size
300 and sex (Table 3; Fig. 3). FGR significantly increased with *Creb1* allele size in males but not in
301 females (Table 3; Fig. 3), indicating that males bearing longer *Creb1* alleles grew their feathers
302 faster than those bearing shorter *Creb1* alleles. Allele size polymorphism at other candidate genes
303 did not significantly predict FGR (Table 3). Results were qualitatively unaltered if long, instead of
304 mean, alleles were included in the models (details not shown for brevity).

Discussion

We investigated whether allelic variation at five candidate genes (*Adcyap1*, *Clock1*, *Clock3*, *Creb1* and *Npas2*) predicted the timing of two important life-history activities, namely timing of spring migration across the central Mediterranean sea and speed of tail feather moult in the African winter quarters, in the long-distance migrating willow warbler. Allelic variation broadly differed between candidate genes, from the low values of observed heterozygosity shown by *Clock3* locus to the high variability of *Adcyap1* gene. *Npas2* allele size affected migration date in male (but not female) willow warblers, but the relationship was opposite to our expectations based on previous research, with individuals bearing shorter alleles migrating later through the study site compared with those bearing longer alleles. Furthermore, long *Clock1* allele size significantly predicted migration date in females, but not in males; again, the relationship was opposite to expectations, the genotype-phenology pattern of association being similar to that detected for *Npas2*. Moreover, *Creb1* allele size significantly predicted FGR, a proxy of overall moult speed, faster moult being associated with longer *Creb1* alleles in male (but not female) willow warblers.

The delayed migration of males bearing shorter *Npas2* alleles was opposite to expectations. According to the few studies investigating the association between *Npas2* gene polymorphism and phenology (Chakarov et al. 2013, Bourret and Garant 2015) and the hypothesis that *Npas2* could overtake *Clock* gene functions, representing an alternative or additional source of adaptive poly-Q variation for the regulation of timing of seasonal events (Debruyne 2008, Steinmeyer et al. 2009), we expected *Npas2* allele size to increase with migration date. In addition, the long *Clock* allele size-migration date association found among females was opposite to our expectations based on previous research (e.g. Bazzi et al. 2015, Saino et al. 2015a), with individuals bearing longer alleles migrating earlier compared to those bearing shorter alleles.

A possible explanation for this finding is that different willow warbler populations that have diverged for *Npas2* and *Clock1* migrate through the study site at different times. Since we found

that late migrating female willow warblers had shorter wings and tails (but not tarsus length), this is indeed a possibility. Interestingly, this pattern of wing length decrease during the migration season, that emerged for other long-distance migratory passerines sampled at the same study site (Saino et al. 2015a), is different from expectations. In fact, wing length increases with latitude across Europe in several passerine species (including the willow warbler; Bensch et al. 1999) and northern populations usually migrate later than southern ones (see e.g.Cramp 1998, Rubolini et al. 2005, Conklin et al. 2010). The negative association between *Npas2* and long *Clock1* and migration date could thus originate because of geographic differentiation in *Npas2* and *Clock1* combined with an unusual differential migration pattern of willow warblers from different populations. However, wing length did not covary with *Npas2* and *Clock1* allele size, and the relationship between *Npas2* genotype and male migration, and between long *Clock1* allele size and female phenology, emerged when controlling for wing length differences (see Results). Hence, the explanation for a negative association between *Npas2* and *Clock1* allele size and migration date remains elusive.

The association between *Clock1* genotype and migration date emerged for long allele size only, being statistically nonsignificant for the mean allele size. This finding further corroborates the hypothesis of a dominance of the long allele, that has been suggested in previous studies of *Clock* and other genes that show poly-Q polymorphism (Ross 2002, Fondon et al. 2008, Liedvogel et al. 2009, Saino et al. 2015a).

Feather growth rate was predicted by *Creb1* allele size, longer *Creb1* alleles being associated with faster moult. Again, the relationship differed between sexes, as gauged by the statistically significant interaction between sex and *Creb1* allele size. Sampling birds during migration allowed us to make inferences about the duration of moult by means of ptilochronological analysis of tail feathers (Grubb 2006); nevertheless, we could not obtain information about the timing of the moult, i.e. the date of the onset or of the end of plumage moult. Previous research mainly focused on the onset of life-history events, or their timing at a given stage (e.g. Caprioli et al. 2012, Chakarov et al. 2013, Saino et al. 2013, Saino et al. 2015a), while only a few studies investigated the association

between candidate genes and the duration of circannual activities (Liedvogel et al. 2009; Bourret and Garant 2015; see also Liedvogel and Sheldon 2010; Mueller et al. 2011; Dor et al. 2012; Peterson et al. 2013). However, studying moult duration in relation to candidate genes polymorphism could improve our understanding of the genetic regulation of annual scheduling. Moult requires considerable amounts of resources, and overlap between moult and other circannual activities is largely avoided by most species (Jenni and Winkler 1994, Hemborg and Lundberg 1998). Hence, in winter moulting species, such as the willow warbler, moult duration may constrain the timing of spring migration (Hedenström et al. 2007, Møller et al. 2011). Indeed, comparative studies of trans-Saharan migrants with different moult strategies, showed that species performing a complete moult during wintering migrate later than those moulting in Europe before autumn migration (Rubolini et al. 2005).

We had no a priori expectation on the possible effect of candidate genes allele size on moult speed, since the single previous study investigating the relationship between genotype and moult phenology focused on *Clock* gene only (Saino et al. 2013). However, Chakarov et al. (2013) found that longer *Creb1* alleles were associated with delayed juvenile dispersal in buzzards. Hence, the *Creb1* allele size-moult speed association we detected may arise from a delayed onset of plumage moult among individuals bearing longer *Creb1* alleles. A delayed timing of moult might constrain its duration, leading to faster feather growth, as demonstrated in small migratory passerines experimentally subjected to shorter moulting periods by altering photoperiod (e.g. Dawson et al. 2000, Hall and Fransson 2000). Alternatively, we might speculate that *Creb1* allele size may directly affect moult speed through its involvement in the melanin synthesis pathway, that has been demonstrated at least in mammals (see e.g. Park et al. 2011). Indeed, it is known that melanin granules increase feather resistance to abrasion at low metabolic costs, and that non-melanic feathers should be much more thicker than melanic ones in order to resist to the same amount of wear (Bonser 1995). Improved melanin synthesis may allow birds to grow feather faster without reducing their resistance to wear.

383 Our results showed that *Npas2, Clock1* and *Creb1* genes had sex-specific phenotypic effects.
384 Sex-specific effects of candidate genes have been previously highlighted for different life-history
385 events by several studies (Caprioli et al. 2012, Bourret and Garant 2015, Saino et al. 2015a, Bazzi et
386 al. 2016). Sex-specific effects may originate because of sex-specific selective pressures on timing of
387 life-history events. For instance, in proterandrous migratory species, males are subjected to stronger
388 selective pressures for early arrival at the breeding grounds than females (e.g. Morbey and
389 Ydenberg 2001, Spottiswoode et al. 2006, Newton 2008, Reudink et al. 2009, Spottiswoode and
390 Saino 2010). Proximately, sex-specific genotype-phenotype associations may arise because of sex-
391 specific genetic architecture. For instance, the autosomal genome is shared by both sexes, but gene
392 expression and regulation is often sexually dimorphic, leading to genotype-sex interactions in
393 genotype-phenotype association studies (review in Ellegren and Parsch 2007, Ober et al. 2008). An
394 alternative possibility is that males and females migrating at Ventotene originated from different
395 breeding populations and that the observed sex-specific genotype-phenotype associations may
396 instead originate because of population-specific candidate gene effects. However, the lack of
397 genetic differentiation at candidate genes between the sexes (both for single loci and for the
398 combination of the five loci) argues against this possibility.
399 To conclude, our study provides novel insights into avian migratory phenotype-genotype
400 associations for a large set of candidate genes. Our findings suggest that different candidate genes
401 may contribute to regulating different life-history events in a sex-specific fashion, and that
402 candidate gene polymorphism underlies among-individuals variation in phenology throughout the
403 annual cycle. Intriguingly, the association between *Creb1*, a candidate gene which constitutes a key
404 element for the light entrainment of the endogenous clock, and moult speed, a life-history event that
405 occurs at equatorial latitudes, may suggest that daylength plays a role in the synchronization of
406 circadian and circannual rhythms of birds even where daily changes in photoperiod are small.
407

Acknowledgements – We thank M. Caprioli and C.D. Possenti for assistance during laboratory and field work, and A. Galimberti for the support with statistical analyses. We thank the Riserva Naturale Isole di Ventotene e Santo Stefano for the logistic support and the field assistants and ringers that helped collecting the data on the field.

References

Bazzi, G., Ambrosini, R., Caprioli, M., Costanzo, A., Liechti, F., Gatti, E., Gianfranceschi, L., Podofillini, S., Romano, A., Romano, M., Scandolara, C., Saino, N. and Rubolini, D. 2015. *Clock* gene polymorphism and scheduling of migration: a geolocator study of the barn swallow *Hirundo rustica*. -Sci Rep, 5: 12443.

Bazzi, G., Galimberti, A., Hays, Q. R., Bruni, I., Cecere, J. G., Gianfranceschi, L., Hobson, K. A., Morbey, Y. E., Saino, N., Guglielmo, C. G. and Rubolini, D. 2016. Adcyap1polymorphism covaries with breeding latitude in a Nearctic migratory songbird, the Wilson's warbler (Cardellina pusilla). -Ecol. Evol.

Bell-Pedersen, D., Cassone, V. M., Earnest, D. J., Golden, S. S., Hardin, P. E., Thomas, T. L. and Zoran, M. J. 2005. Circadian rhythms from multiple oscillators: lessons from diverse organisms. -Nat. Rev. Genet., 6: 544-56.

Bensch, S., Andersson, T. and Akesson, S. 1999. Morphological and molecular variation across a migratory divide in willow warblers, *Phylloscopus trochilus*. -Evolution, 53: 1925-1935.

Bensch, S. and Grahn, M. 1993. A new method for estimating individual speed of molt. -Condor, 95: 305-315.

Berthold, P. 1996. Control of bird migration, Springer Science & Business Media.

Bonser, R. H. C. 1995. Melanin and the abrasion resistance of feathers. -Condor, 97: 590-591.

Bourret, A. and Garant, D. 2015. Candidate gene-environment interactions and their relationships with timing of breeding in a wild bird population. -Ecol. Evol., 5: 3628-41.

433 Brodin, A. 1993. Radio-ptilochronology tracing radioactively labelled food in feathers. -Ornis
434 Scand., 24: 167-173.
435 Caprioli, M., Ambrosini, R., Boncoraglio, G., Gatti, E., Romano, A., Romano, M., Rubolini, D.,
436 Gianfranceschi, L. and Saino, N. 2012. *Clock* gene variation is associated with breeding
437 phenology and maybe under directional selection in the migratory barn swallow. -PLoS
438 One, 7: e35140.
439 Chakarov, N., Jonker, R. M., Boerner, M., Hoffman, J. I. and Kruger, O. 2013. Variation at
440 phenological candidate genes correlates with timing of dispersal and plumage morph in a
441 sedentary bird of prey. -Mol. Ecol., 22: 5430-40.
442 Conklin, J. R., Battley, P. F., Potter, M. A. and Fox, J. W. 2010. Breeding latitude drives individual
443 schedules in a trans-hemispheric migrant bird. -Nat. Comm., 1: 67.
444 Cramp, S. 1998. The Complete Birds of the Western Palearctic on CD-ROM, Oxford, Oxford
445 University Press.
446 Dawson, A., Hinsley, S. A., Ferns, P. N., Bonser, R. H. and Eccleston, L. 2000. Rate of moult
447 affects feather quality: a mechanism linking current reproductive effort to future survival. -
448 Proc. R. Soc. B, 267: 2093-8.
449 De La Hera, I., Pérez-Tris, J. and Tellería, J. 2009. Migratory behaviour affects the trade-off
450 between feather growth rate and feather quality in a passerine bird. -Biol. J. Linn. Soc., 97:
451 98-105.
452 Debruyne, J. P. 2008. Oscillating perceptions: the ups and downs of the CLOCK protein in the
453 mouse circadian system. -Journal of Genetics, 87: 437-446.
454 Dor, R., Cooper, C. B., Lovette, I. J., Massoni, V., Bulit, F., Liljesthrom, M. and Winkler, D. W.
455 2012. Clock gene variation in *Tachycineta* swallows. -Ecol. Evol., 2: 95-105.
456 Ellegren, H. and Parsch, J. 2007. The evolution of sex-biased genes and sex-biased gene expression.
457 -Nat. Rev. Genet., 8: 689-98.

458 Evans, K. L., Gaston, K. J., Sharp, S. P., Mcgowan, A. and Hatchwell, B. J. 2009. The effect of
459 urbanisation on avian morphology and latitudinal gradients in body size. -Oikos, 118: 251-
460 259.
461 Fidler, A. E. and Gwinner, E. 2003. Comparative analysis of Avian BMAL1 and CLOCK protein
462 sequences: a search for features associated with owl nocturnal behaviour. -Comp. Biochem.
463 Physiol. B, 136: 861-874.
464 Fondon, J. W., 3rd, Hammock, E. A., Hannan, A. J. and King, D. G. 2008. Simple sequence
465 repeats: genetic modulators of brain function and behavior. -Trends Neurosci., 31: 328-34.
466 Gau, D., Lemberger, T., Von Gall, C., Kretz, O., Le Minh, N., Gass, P., Schmid, W., Schibler, U.,
467 Korf, H. W. and Schutz, G. 2002. Phosphorylation of CREB Ser142 regulates light-induced
468 phase shifts of the circadian clock. -Neuron, 34: 245-53.
469 Goudet, J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices
470 (version 2.9. 3).
471 Goymann, W., Spina, F., Ferri, A. and Fusani, L. 2010. Body fat influences departure from stopover
472 sites in migratory birds: evidence from whole-island telemetry. -Biol. Lett., 6: 478-81.
473 Grubb, T. C. 2006. Ptilochronology. Feather time and the biology of birds, New York, Oxford
474 University Press.
475 Guo, S. W. and Thompson, E. A. 1992. A Monte Carlo method for combined segregation and
476 linkage analysis. -Am. J. Hum. Genet., 51: 1111-1126.
477 Gwinner, E. 1986. Circannual rhythms, Berlin, Springer-Verlag.
478 Gwinner, E. 2003. Circannual rhythms in birds. -Curr. Opin. Neurobiol., 13: 770-778.
479 Hall, K. S. S. and Fransson, T. 2000. Lesser Whitethroats under time-constraint moult more rapidly
480 and grow shorter wing feathers. -J. Avian Biol., 31: 583-587.
481 Hannibal, J., Ding, J. M., Chen, D., Fahrenkrug, J., Larsen, P. J., Gillette, M. U. and Mikkelsen, J.
482 D. 1997. Pituitary adenylate cyclase-activating peptide (PACAP) in the retinohypothalamic
483 tract: a potential daytime regulator of the biological clock. -J. Neurosci., 17: 2637-2644.

484 Hedenström, A., Barta, Z., Helm, B., Houston, A. I., Mcnamara, J. M. and Jonzén, N. 2007.
485 Migration speed and scheduling of annual events by migrating birds in relation to climate
486 change. -Clim. Res., 35: 79-91.
487 Hemborg, C. and Lundberg, A. 1998. Costs of overlapping reproduction and moult in passerine
488 birds: an experiment with the pied flycatcher. -Behav. Ecol. Sociobiol., 43: 19-23.
489 Jenni, L. and Winkler, R. R. 1994. Moult and Ageing of European Passerines, London, Academic
490 Press.
491 Johnsen, A., Fidler, A. E., Kuhn, S., Carter, K. L., Hoffmann, A., Barr, I. R., Biard, C.,
492 Charmantier, A., Eens, M., Korsten, P., Siitari, H., Tomiuk, J. and Kempenaers, B. 2007.
493 Avian Clock gene polymorphism: evidence for a latitudinal cline in allele frequencies. -Mol.
494 Ecol., 16: 4867-80.
495 Jonzén, N., Linden, A., Ergon, T., Knudsen, E., Vik, J. O., Rubolini, D., Piacentini, D., Brinch, C.,
496 Spina, F., Karlsson, L., Stervander, M., Andersson, A., Waldenstrom, J., Lehikoinen, A.,
497 Edvardsen, E., Solvang, R. and Stenseth, N. C. 2006a. Rapid advance of spring arrival dates
498 in long-distance migratory birds. -Science, 312: 1959-61.
499 Jonzén, N., Piacentini, D., Andersson, A., Montemaggiori, A., Stervander, M., Rubolini, D.,
500 Waldenström, J. and Spina, F. 2006b. The timing of spring migration in trans-Saharan
501 migrants: a comparison between Ottenby, Sweden and Capri, Italy. -Ornis Svec., 16: 27-33.
502 Kuhn, K., Schwenk, K., Both, C., Canal, D., Johansson, U. S., Van Der Mije, S., Topfer, T. and
503 Packert, M. 2013. Differentiation in neutral genes and a candidate gene in the pied
504 flycatcher: using biological archives to track global climate change. -Ecol. Evol., 3: 4799-
505 814.
506 Liedvogel, M. and Sheldon, B. C. 2010. Low variability and absence of phenotypic correlates of
507 *Clock* gene variation in a great tit *Parus major* population. -J. Avian Biol., 41: 543-550.

508 Liedvogel, M., Szulkin, M., Knowles, S. C., Wood, M. J. and Sheldon, B. C. 2009. Phenotypic
509 correlates of *Clock* gene variation in a wild blue tit population: evidence for a role in
510 seasonal timing of reproduction. -Mol. Ecol., 18: 2444-56.
511 Messineo, A., Grattarola, A. and Spina, F. 2001. Dieci anni di Progetto Piccole Isole. -Biologia e
512 Conservazione della Fauna, 106: 1-244.
513 Møller, A. P., Nuttall, R., Piper, S. E., Szép, T. and Vickers, E. J. 2011. Migration, moult and
514 climate change in barn swallows *Hirundo rustic*a in South Africa. -Clim. Res., 47: 201-205.
515 Morbey, Y. E. and Ydenberg, R. C. 2001. Protandrous arrival timing to breeding areas: a review. -
516 Ecol. Lett., 4: 663-673.
517 Mueller, J. C., Pulido, F. and Kempenaers, B. 2011. Identification of a gene associated with avian
518 migratory behaviour. -Proc. R. Soc. B, 278: 2848-56.
519 Nagy, A. D. and Csernus, V. J. 2007. The role of PACAP in the control of circadian expression of
520 clock genes in the chicken pineal gland. -Peptides, 28: 1767-74.
521 Newton, I. 2008. The migration ecology of birds, London, Academic Press.
522 O'malley, K. G., Ford, M. J. and Hard, J. J. 2010. *Clock* polymorphism in Pacific salmon: evidence
523 for variable selection along a latitudinal gradient. -Proc. R. Soc. B, 277: 3703-14.
524 Ober, C., Loisel, D. A. and Gilad, Y. 2008. Sex-specific genetic architecture of human disease. -
525 Nat. Rev. Genet., 9: 911-22.
526 Park, S. Y., Jin, M. L., Kim, Y. H., Kim, Y. and Lee, S. J. 2011. Aromatic-turmerone inhibits alpha-
527 MSH and IBMX-induced melanogenesis by inactivating CREB and MITF signaling
528 pathways. -Arch. Dermatol. Res., 303: 737-44.
529 Peiró, G. 2003. Intraspecific variation in the wing shape of the long-distance migrant reed warbler
530 *Acrocephalus scirpaceus*: effects of age and distance of migration. -Ardeola, 50: 31-37.
531 Peterson, M. P., Abolins-Abols, M., Atwell, J. W., Rice, R. J., Mila, B. and Ketterson, E. D. 2013.
532 Variation in candidate genes CLOCK and ADCYAP1 does not consistently predict
533 differences in migratory behavior in the songbird genus Junco. -F1000Res., 2: 115.

534 Pulido, F. 2007. The Genetics and Evolution of Avian Migration. -Bioscience, 57: 165.
535 Racz, B., Horvath, G., Faluhelyi, N., Nagy, A. D., Tamas, A., Kiss, P., Gallyas, F., Jr., Toth, G.,
536 Gaszner, B., Csernus, V. and Reglodi, D. 2008. Effects of PACAP on the circadian changes
537 of signaling pathways in chicken pinealocytes. -J. Neurosci., 36: 220-6.
538 Reudink, M. W., Marra, P. P., Kyser, T. K., Boag, P. T., Langin, K. M. and Ratcliffe, L. M. 2009.
539 Non-breeding season events influence sexual selection in a long-distance migratory bird. -
540 Proc. R. Soc. B, 276: 1619-26.
541 Ross, C. A. 2002. Polyglutamine Pathogenesis. -Neuron, 35: 819-822.
542 Rubolini, D., Spina, F. and Saino, N. 2005. Correlates of timing of spring migration in birds: a
543 comparative study of trans-Saharan migrants. -Biol. J. Linn. Soc., 85: 199-210.
544 Saino, N., Bazzi, G., Gatti, E., Caprioli, M., Cecere, J. G., Possenti, C. D., Galimberti, A., Orioli,
545 V., Bani, L., Rubolini, D., Gianfranceschi, L. and Spina, F. 2015a. Polymorphism at the
546 Clock gene predicts phenology of long-distance migration in birds. -Mol. Ecol., 24: 1758-
547 73.
548 Saino, N., Romano, M., Caprioli, M., Fasola, M., Lardelli, R., Micheloni, P., Scandolara, C.,
549 Rubolini, D. and Gianfranceschi, L. 2013. Timing of molt of barn swallows is delayed in a
550 rare *Clock* genotype. -PeerJ, 1: e17.
551 Saino, N., Romano, M., Romano, A., Rubolini, D., Ambrosini, R., Caprioli, M., Parolini, M.,
552 Scandolara, C., Bazzi, G., Costanzo, A. and Perez-Tris, J. 2015b. White tail spots in
553 breeding Barn Swallows *Hirundo rustica* signal body condition during winter moult. -Ibis,
554 157: 722-730.
555 Saino, N., Rubolini, D., Serra, L., Caprioli, M., Morganti, M., Ambrosini, R. and Spina, F. 2010.
556 Sex-related variation in migration phenology in relation to sexual dimorphism: a test of
557 competing hypotheses for the evolution of protandry. -J. Evol. Biol., 23: 2054-2065.

558 Schwartz, C. and Andrews, M. T. 2013. Circannual Transitions in Gene Expression: Lessons from
559 Seasonal Adaptations. In: ROUGVIE, A. E. and O'CONNOR, M. B. (eds.) Developmental
560 timing. Academic Press.
561 Sharp, P. J. 2005. Photoperiodic regulation of seasonal breeding in birds. -Ann. N. Y. Acad. Sci.,
562 1040: 189-99.
563 Simonneaux, V., Ouichou, A. and Pévet, P. 1993. Pituitary adenylate cyclase-activating polypeptide
564 (PACAP) stimulates melatonin synthesis from rat pineal gland. -Brain Res., 603: 148-152.
565 Spina, F., Massi, A., Montemaggiori, A. and Baccetti, N. 1993. Spring Migration Across Central
566 Mediterranean: General Results from the "Progetto Piccole Isole". -Vogelwarte, 37: 1-94.
567 Spottiswoode, C. and Saino, N. 2010. Sexual selection and climate change. In: MØLLER, A. P.,
568 FIEDLER, W. and BERTHOLD, P. (eds.) Effects of Climate Change in Birds. Oxford
569 University Press, Oxford, pp. 169-189.
570 Spottiswoode, C. N., Tøttrup, A. P. and Coppack, T. 2006. Sexual selection predicts advancement
571 of avian spring migration in response to climate change. -Proc. R. Soc. B, 273: 3023-9.
572 Steinmeyer, C., Mueller, J. C. and Kempenaers, B. 2009. Search for informative polymorphisms in
573 candidate genes: clock genes and circadian behaviour in blue tits. -Genetica, 136: 109-17.
574 Tarka, M., Åkesson, M., Beraldi, D., Hernández-Sánchez, J., Hasselquist, D., Bensch, S. and
575 Hansson, B. 2010. A strong quantitative trait locus for wing length on chromosome 2 in a
576 wild population of great reed warblers. -Proc. R. Soc. B, 277: 2361-9.
577 Tenan, S. and Spina, F. 2010. Timing and condition-related effects on recapture probability, mass
578 change and stopover length of spring migrating songbirds on a small mediterranean island. -
579 Ardeola, 57: 121-132.
580 Tischkau, S. A., Mitchell, J. W., Tyan, S. H., Buchanan, G. F. and Gillette, M. U. 2003.
581 Ca2+/cAMP response element-binding protein (CREB)-dependent activation of *Per1* is
582 required for light-induced signaling in the suprachiasmatic nucleus circadian clock. -J. Biol.
583 Chem., 278: 718-23.

584 Underhill, L. G., Prys-Jones, R. P., Dowsett, R. J., Herroelen, P., Johnson, D. N., Lawn, M. R.,
585 Norman, S. C., Pearson, D. J. and Tree, A. J. 1992. The biannual primary moult of Willow
586 Warblers *Phylloscopus trochilus* in Europe and Africa. -Ibis, 134: 286-297.
587 Visser, M. E., Caro, S. P., Van Oers, K., Schaper, S. V. and Helm, B. 2010. Phenology, seasonal
588 timing and circannual rhythms: towards a unified framework. -Phil. Trans. R. Soc. Lond. B,
589 365: 3113-3127.

590

591

592

Table 1. List of candidate genes, sample size (N), number of the alleles observed at each locus (K), range of alleles length (alleles range, in bp), mean allele size (in bp, with associated standard error, in brackets) and observed heterzygosity (H_o).

Candidate gene	N	K	Alleles range	Mean allele size (SE)	H_o
Adcyap1	112	10	160-176	170.21 (0.22)	0.83
Clock1	121	5	114-126	120.03 (0.14)	0.47
Clock3	97	2	108-111	108.29 (0.06)	0.15
Creb1	92	4	271-277	274.01 (0.10)	0.50
Npas2	93	5	166-178	172.37 (0.11)	0.38

Table 2. Linear models of the effect of candidate genes allele size on migration date (1 = January 1). The interaction terms were removed *en bloc* if nonsignificant ($P > 0.05$), and statistics for main effects refer to models without the interaction term.

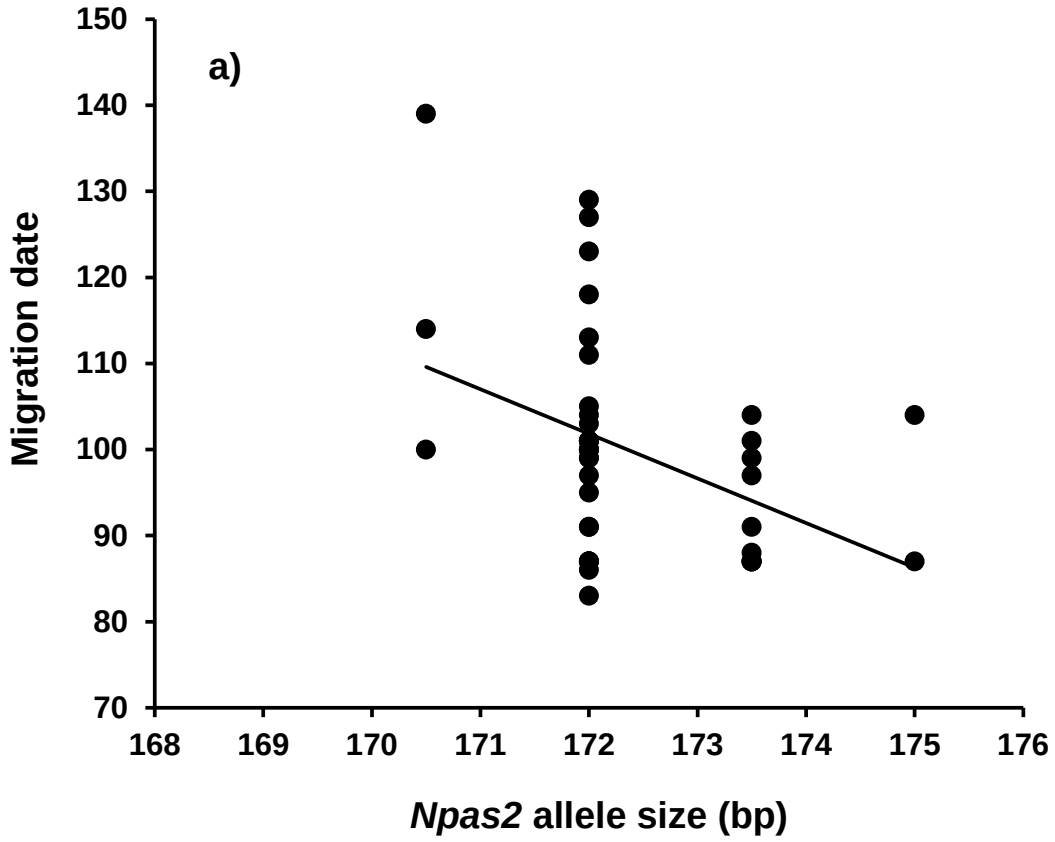
Variable	Estimate (SE)	df	*F*	*P*
Adcyap1 allele size	0.19 (0.50)	1, 104	0.15	0.70
Clock1 allele size	-0.80 (0.68)	1, 115	1.38	0.24
Clock3 allele size	0.13 (1.93)	1, 89	0.01	0.95
sumClock allele size	-0.65 (0.79)	1, 89	0.67	0.41
Creb1 allele size	0.96 (1.28)	1, 85	0.56	0.46
Npas2 allele size	-5.51 (1.86)	1, 84	4.88	0.020
Sex	-983.19 (416.12)	1, 84	5.58	< 0.001
Sex × *Npas2* allele size	5.77 (2.41)	1, 84	5.73	0.019

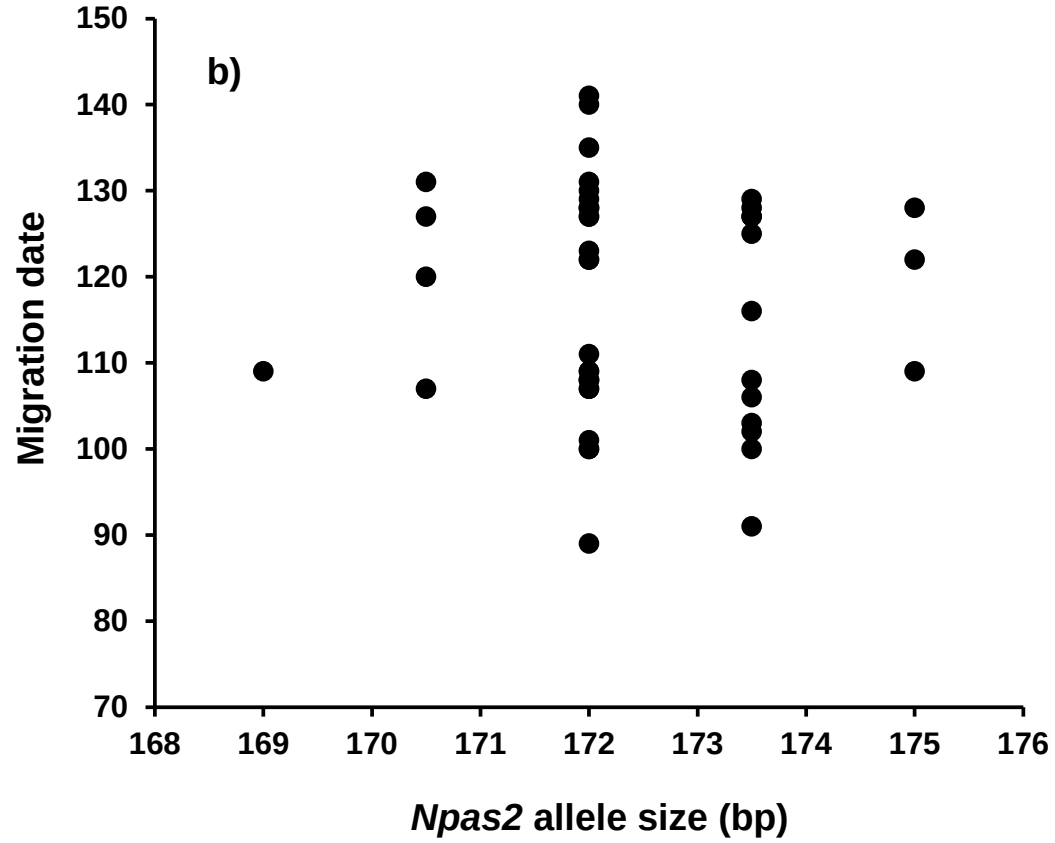
Table 3. Linear models of the effect of candidate genes allele size on feather growth rate (residuals of a regression of GBW on feather length; see Methods). The interaction terms were removed *en bloc* if nonsignificant ($P > 0.05$), and statistics for main effects refer to models without the interaction term.

Variable	Estimate (SE)	df	*F*	*P*
Adcyap1 allele size	0.006 (0.007)	1, 104	0.65	0.42
Clock1 allele size	0.003 (0.010)	1, 112	0.07	0.80
Clock3 allele size	0.001 (0.027)	1, 90	0.01	0.99
sumClock allele size	0.012 (0.011)	1, 88	1.04	0.31
Creb1 allele size	0.054 (0.022)	1, 85	0.12	0.73
Sex	26.04 (10.70)	1, 85	5.93	0.017
Creb1 allele size × sex[a]	-0.10 (0.04)	1, 85	5.93	0.017
Npas2	0.011 (0.017)	1, 86	0.43	0.51

a) estimates from linear models of migration date as a function of *Creb1* allele size separated for either sex; males, estimate: 0.054 (0.020 SE), $F_{1,45} = 7.45$, $P = 0.009$; females, estimate: -0.041 (0.035 SE), $F_{1,40} = 1.33$, $P = 0.26$

Figure 1. Migration date (1 = January 1) in relation to *Npas2* allele size in (a) male and (b) female willow warblers. The line represents simple linear regression.





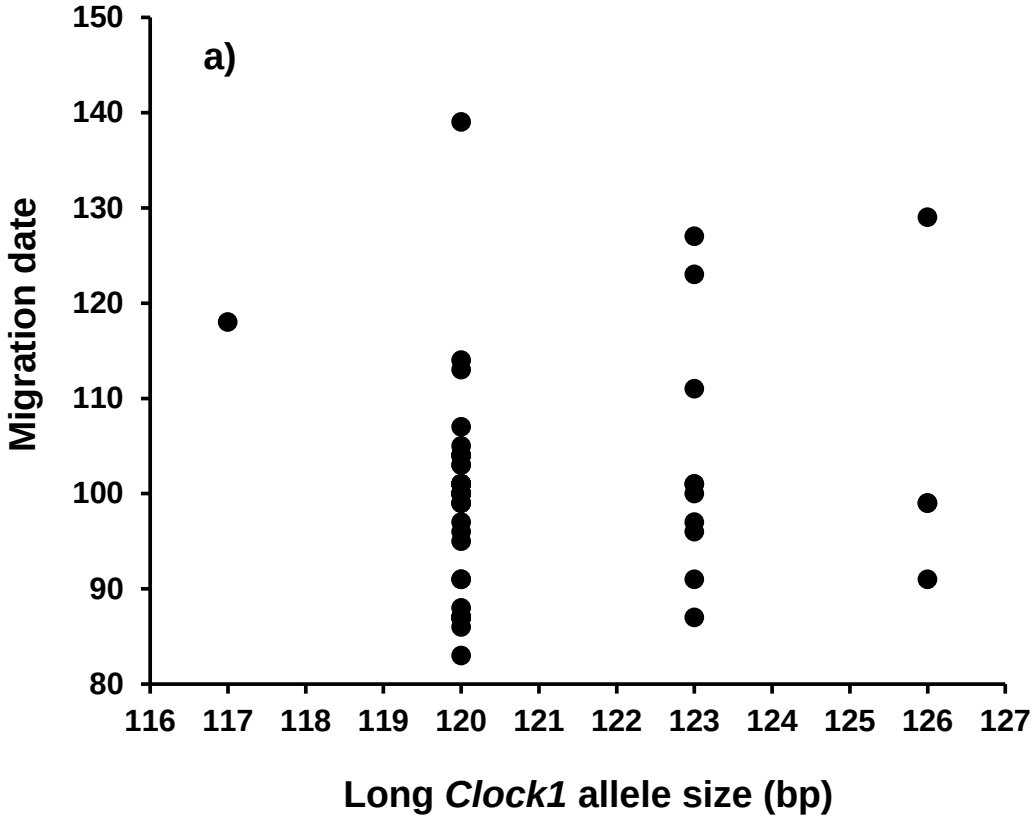
618 **Figure 2.** Migration date (1 = January 1) in association with long *Clock1* allele size in (a) males and

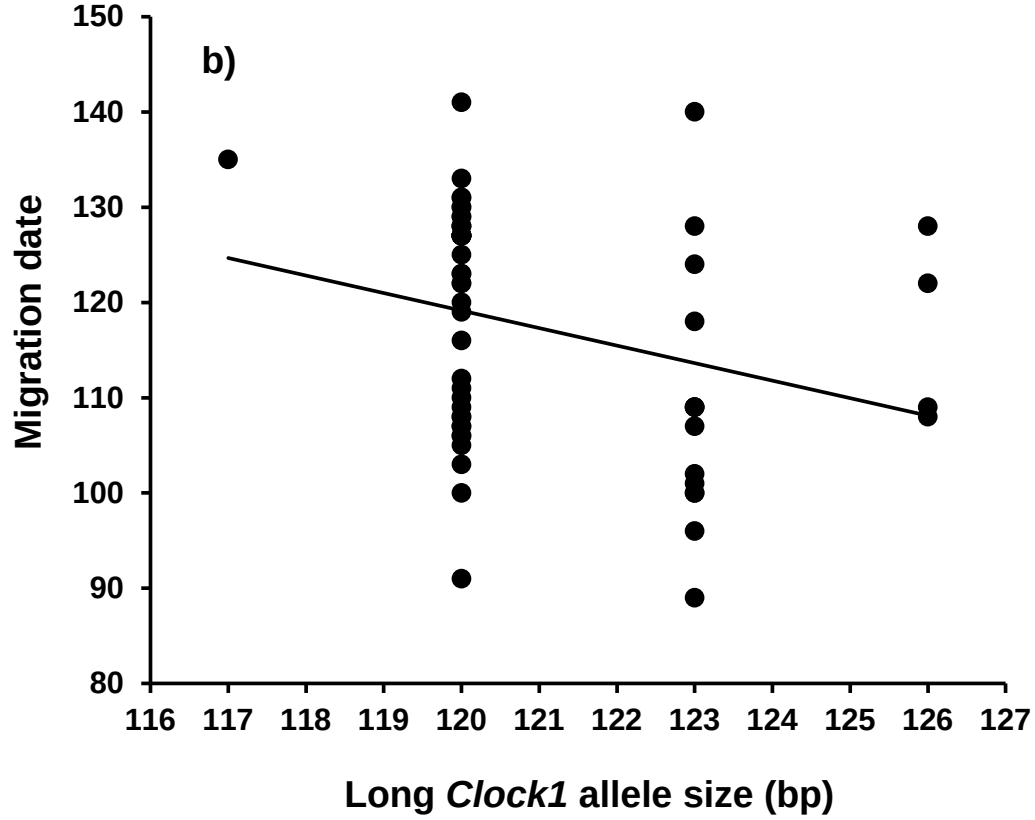
619 (b) females. The line represents simple linear regression.

620

621

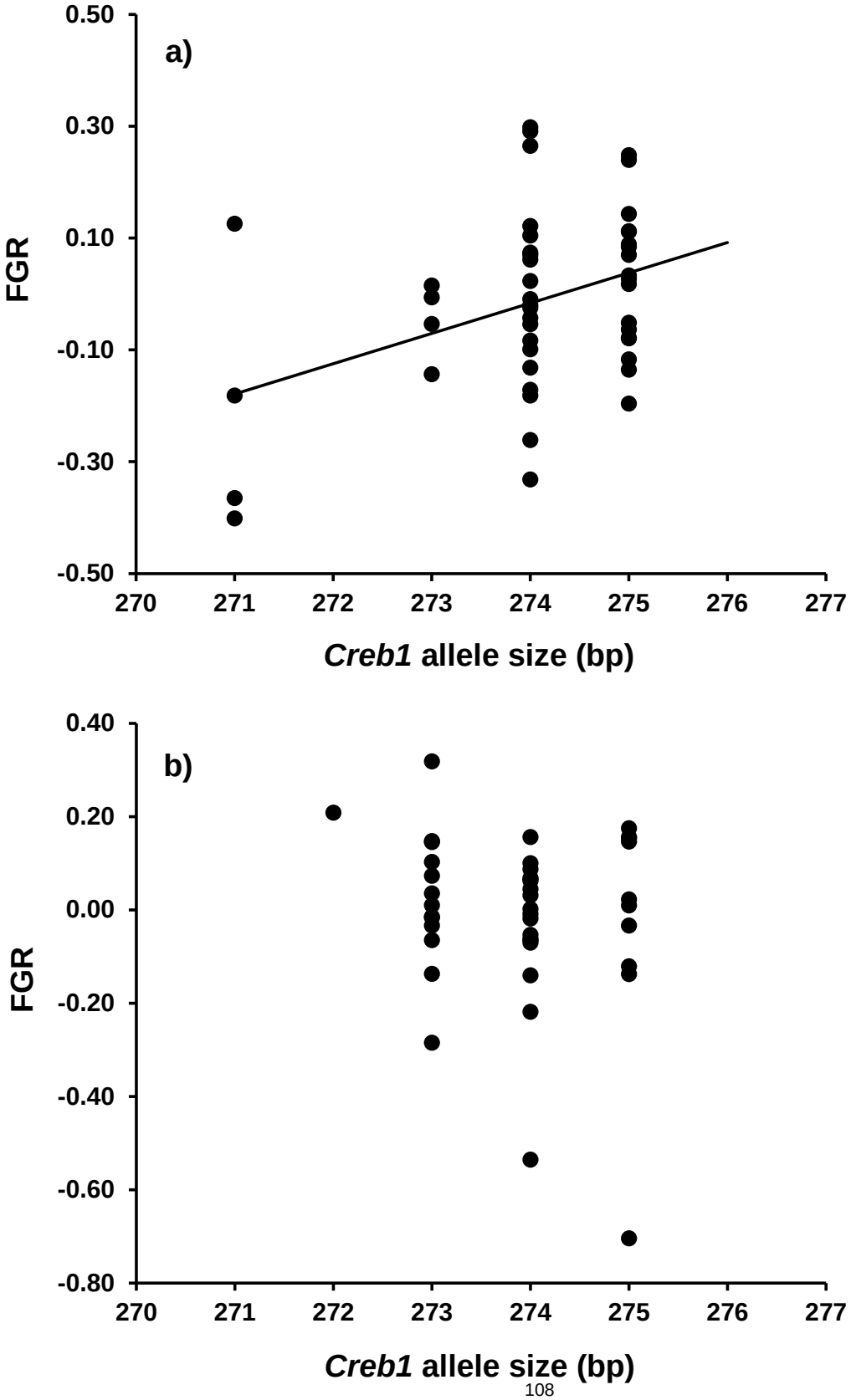
622





623 **Figure 3.** Feather growth rate (FGR; residuals of a regression of GBW on feather length; see
624 Methods) *vs Creb1* allele size in (a) males and (b) females. High FGR values are assumed to reflect
625 faster moult. The line represents simple linear regression. The result for females was similar after
626 removing the outliers from the model (details not show for brevity).

627
628
629
630



PART 2

CHAPTER 4

Clock gene polymorphism and scheduling of migration: a geolocator study of the barn swallow *Hirundo rustica*

Gaia Bazzi, Roberto Ambrosini, Manuela Caprioli, Alessandra Costanzo, Felix Liechti, Emanuele Gatti, Luca Gianfranceschi, Stefano Podofillini, Andrea Romano, Maria Romano, Chiara Scandolara, Nicola Saino and Diego Rubolini

Scientific Reports **5** (2015)

Photo: Francesco Renzi

SCIENTIFIC REPORTS

OPEN

Clock gene polymorphism and scheduling of migration: a geolocator study of the barn swallow *Hirundo rustica*

Received: 30 January 2015
Accepted: 26 June 2015
Published: 22 July 2015

Gaia Bazzi[1], Roberto Ambrosini[2], Manuela Caprioli[1], Alessandra Costanzo[1], Felix Liechti[3], Emanuele Gatti[1], Luca Gianfranceschi[1], Stefano Podofillini[1], Andrea Romano[1], Maria Romano[1], Chiara Scandolara[1,3], Nicola Saino[1] & Diego Rubolini[1]

Circannual rhythms often rely on endogenous seasonal photoperiodic timers involving 'clock' genes, and *Clock* gene polymorphism has been associated to variation in phenology in some bird species. In the long-distance migratory barn swallow *Hirundo rustica*, individuals bearing the rare *Clock* allele with the largest number of C-terminal polyglutamine repeats found in this species (Q_8) show a delayed reproduction and moult later. We explored the association between *Clock* polymorphism and migration scheduling, as gauged by light-level geolocators, in two barn swallow populations (Switzerland; Po Plain, Italy). Genetic polymorphism was low: 91% of the 64 individuals tracked year-round were Q_7/Q_7 homozygotes. We compared the phenology of the rare genotypes with the phenotypic distribution of Q_7/Q_7 homozygotes within each population. In Switzerland, compared to Q_7/Q_7, two Q_6/Q_7 males departed earlier from the wintering grounds and arrived earlier to their colony in spring, while a single Q_7/Q_8 female was delayed for both phenophases. On the other hand, in the Po Plain, three Q_6/Q_7 individuals had a similar phenology compared to Q_7/Q_7. The Swiss data are suggestive for a role of genetic polymorphism at a candidate phenological gene in shaping migration traits, and support the idea that *Clock* polymorphism underlies phenological variation in birds.

The timing of seasonal activities has major fitness consequences and is subjected to intense natural selection[1,2]. In species experiencing seasonal peaks of resource availability, typically occurring at medium-high latitudes, proper matching of life-cycle events to such resource pulses is fundamental for achieving a high fitness. The timing of seasonal activities is often set by genetically controlled endogenous circadian and circannual rhythms[3], which are synchronized with ecological conditions through external drivers[4], among which photoperiod plays a pivotal role[5]. This seasonal photoperiodic timer involves 'circadian clock' genes, which are responsible for the onset and setting of circadian and circannual rhythmicity[6–8]. The circadian clock relies on an auto-regulated negative feedback loop, the 'core circadian oscillator' (CCO)[9,10]. In birds and mammals, the CLOCK transcription factor, encoded by the *Circadian Locomotor Output Cycles Kaput* (*Clock*) gene, plays a central role within the CCO, both by acting as a transcriptional activator of the CCO itself and by activating the transcription of several output genes[11,12]. The C-terminal domain of the *Clock* gene contains a series of polyglutamine residues (Poly-Q) which may vary in length among species as well as among individuals within populations[13–18].

[1]Dipartimento di Bioscienze, Università degli Studi di Milano, via Celoria 26, I-20133 Milano, Italy. [2]Dipartimento di Scienze dell'Ambiente e del Territorio e di Scienze della Terra (DISAT), Università di Milano-Bicocca, Piazza della Scienza 1, I-20126 Milano, Italy. [3]Swiss Ornithological Institute, Seerose 1, CH-6204 Sempach, Switzerland. Correspondence and requests for materials should be addressed to D.R (email: diego.rubolini@unimi.it)

Latitudinal increase in the number of Poly-Q repeats across populations has been suggested to reflect adaptation to local regimes of circannual photoperiodic variation[14,16,19]. Similarly, Poly-Q polymorphism may control variation in photoperiodic responses among individuals from the same population, causing variation in phenology at the individual level[15,17,20,21]. Though the evidence for an association between *Clock* polymorphism and timing of seasonal events at the population level is scant, two recent studies conducted on a long-distance migratory bird, the barn swallow *Hirundo rustica*, have shown that a very rare *Clock* genotype, the heterozygous Q_7/Q_8 (c.a. 1.5% of the population), which contains the largest Poly-Q stretch observed in this species (Q_8), has a delayed timing compared to the most frequent (94%) Q_7/Q_7 genotype[17,20]. These findings are consistent with the idea that larger Poly-Q stretches are associated with a delayed phenology[15,17,20,21], and suggest that the Q_7/Q_8 genotype shows a delayed annual routine. Specifically, the first study, carried out on an Italian barn swallow population, showed that the onset of laying by Q_7/Q_8 females was delayed by 13 days relative to Q_7/Q_7 counterparts, corresponding to ca. 0.7 s.d. later than the mean of Q_7/Q_7 females[17]. In the second study, carried out on moulting barn swallows in the sub-Saharan wintering areas, wing feather moult of three Q_7/Q_8 individuals was found to be significantly delayed, by ca. 8/9 days, than Q_7/Q_7 birds[20]. In both studies, the phenology of the other rare genotype Q_6/Q_7 (4.8%)[17], with a smaller Poly-Q stretch, did not differ from that of the common genotype Q_7/Q_7[17,20].

To date, the analysis of individual variation in scheduling of migration events across the entire annual cycle in small-sized birds has been hampered by the technical difficulties of tracking large numbers of individuals from the same population. Here we exploit data retrieved from 64 individual barn swallows tracked along the entire annual migration cycle using miniaturized light-level geolocators[22] to explore the association between *Clock* genotype polymorphism and scheduling of migration events in two barn swallow populations breeding in Southern Switzerland (Magadino) and Northern Italy (Po Plain). We expected that rare individuals bearing the heterozygous Q_7/Q_8 genotype should show a delayed scheduling of migration and wintering events compared to the phenotypic distribution of phenological traits of Q_7/Q_7 birds, complementing previous evidence concerning delayed timing of breeding and moult.

Results and Discussion

Most (90.6%) of the 64 individuals tracked for an entire annual cycle (Switzerland, $n = 26$; Po Plain, $n = 38$) were homozygous Q_7/Q_7, as expected. In the Swiss population, we detected two Q_6/Q_7 males and a single Q_7/Q_8 female, while in the Po Plain we detected three Q_6/Q_7 individuals (two males, one female) and no Q_7/Q_8 birds.

The migration and wintering phenology of Q_7/Q_7 birds showed a broad overlap between the two study populations, with the single exception of spring arrival date to the breeding colony, that was ca. 10 days later among Swiss birds compared to Po Plain ones (Fig. 1; Supplementary Table S1, Fig. S1; see also Liechti *et al.*[22]).

In the Swiss population, the single Q_7/Q_8 female showed a consistently delayed phenology and was outside the 95% confidence interval (CI) compared to the phenotypic distribution of Q_7/Q_7 for all traits (Fig. 1), both when actual or within-group centered values (thus removing any between-year and between-sex variation in phenology) were considered, the single exception being departure date from the breeding colony for actual values (Fig. 1a; Supplementary Table S2). On the other hand, the two Q_6/Q_7 males were consistently delayed and outside the 95% CI compared to the phenotypic distribution of Q_7/Q_7 individuals for departure from the breeding colony, but were consistently early and outside the 95% CI compared to Q_7/Q_7 for traits related to pre-breeding migration (Fig. 1c,d; Supplementary Table S2). They were instead not consistently advanced for arrival date to wintering area (Fig 1b; Supplementary Table S2). Results were similar if the analyses were separated by sex (Supplementary Table S3) and if we accounted for the effect of individual variation in wintering latitude and longitude (stationary wintering positions, SWP; see Methods) on arrival date to the breeding colony (Supplementary Table S2).

In the Po Plain, the lack of Q_7/Q_8 individuals allowed only a partial evaluation of the association between *Clock* polymorphism and phenology. Although being somewhat delayed in some cases (Supplementary Fig. S1, Table S2), the scheduling of migration and wintering events of the three Q_6/Q_7 was not consistently different from the phenotypic distribution of Q_7/Q_7 birds, with the exception of colony departure, that was delayed in all three individuals compared to the 95% CI of Q_7/Q_7 birds, similarly to the Swiss population (Supplementary Fig. S1, Table S2).

It may be worth noting that the two Swiss Q_6/Q_7 males and the Q_7/Q_8 female were also outside the CIs (advanced and delayed, respectively) of the Q_7/Q_7 Po Plain birds for departure from the wintering area and arrival to the breeding colony (cf. Fig. 1 and Supplementary Fig. S1; Table S2).

For the Swiss Q_7/Q_8 female, delayed arrival could lead to a delayed reproduction[17]. This female started egg laying on May 28th, 2012 (day 148), but unfortunately we have no other breeding data from geolocator females in the same year (nor from the same female in the previous year) for comparison. However, data from Q_7/Q_7 geolocator females from the previous year (spring 2011) revealed a mean first egg laying date that was ca. 20 days earlier (day 126, CI 120 to 136, $n = 5$) than the Q_7/Q_8 female. We are confident that the delayed laying of the Q_7/Q_8 female compared to Q_7/Q_7 females in the previous year was robust to between-year differences in egg laying dates: in fact, the Q_7/Q_8 female was a delayed phenodeviant for arrival date (Table 1; Fig. 1), and over the entire sample of geolocator females (both Swiss and Po Plain individuals) there was a strict association between arrival date and breeding date ($r = 0.90$, $n = 16$;

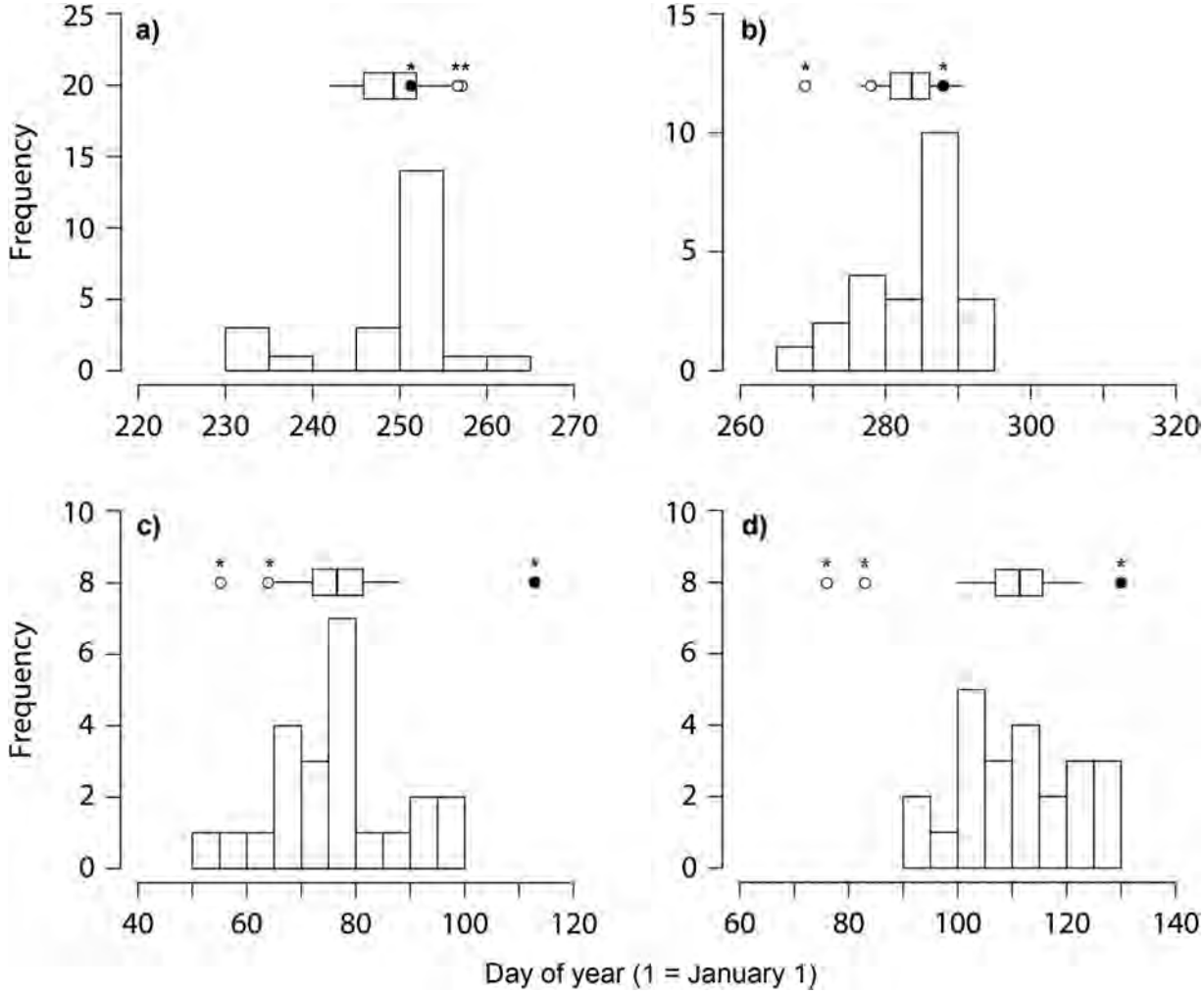


Figure 1. Frequency distribution of phenotypic values of the Q_7/Q_7 individuals and phenology of the Q_6/Q_7 and Q_7/Q_8 individuals in the Swiss population. (**a**) Departure from the breeding colony; (**b**) arrival to the wintering area; (**c**) departure from the wintering area; (**d**) arrival to the breeding colony. The horizontal boxes above histograms show the mean and 95% non-parametric bootstrap confidence interval (CI) of the Q_7/Q_7 phenotypic distribution, while the horizontal lines show the standard deviation. Dots show the phenotypic values of the rare genotypes (open dot = Q_6/Q_7; filled dot = Q_7/Q_8). Histograms and dots refer to actual data, while asterisks above dots denote that the Q_6/Q_7 or Q_7/Q_8 values were outside the 95% CI of a given phenological trait for centred values (removing any between-year and between-sex variation in phenology; see Supplementary Table S2 and Methods for details).

$Zr = 1.47$, CI 0.93 to 2.01; Fig. 2). On the other hand, the two Swiss Q_6/Q_7 males did not have a clearly advanced breeding date compared to Q_7/Q_7 birds (details not shown for brevity). Indeed, the correlation between breeding date (start of egg laying by the social mate) and arrival date in geolocator males was weaker than in females ($r = 0.31$, $n = 37$; $Zr = 0.32$, CI −0.02 to 0.66), suggesting that timing of reproduction is mainly under female control[17].

These findings, based on the first individual year-round tracking data for the barn swallow, complement previous evidence that the rare Q_7/Q_8 barn swallow genotype has a delayed timing[17,20]. Importantly, this evidence originates from the analysis of a further independent dataset, as no individuals from the Swiss population had been included in previous studies[17,20]. The delay of the Swiss Q_7/Q_8 individual was most pronounced for phenological variables related to timing of spring migration (departure from the wintering area and arrival date to the breeding colony) (Fig. 1).

Intriguingly, the two Swiss Q_6/Q_7 males showed an apparently advanced timing of spring migration compared to Q_7/Q_7 birds in the same population. In previous studies of the barn swallow, no statistically significant differences in timing emerged between Q_6/Q_7 and Q_7/Q_7 birds[17,20]. The Po Plain data show instead an inconsistent tendency towards a delayed pre-breeding migration phenology of the three Q_6/Q_7 birds compared to Q_7/Q_7 ones. Albeit not conclusively due to the small sample size of the 'rare'

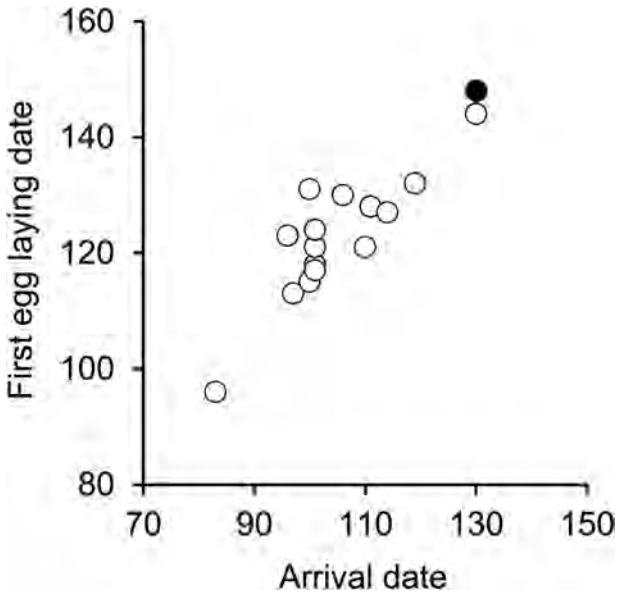


Figure 2. Arrival date predicts onset of reproduction among female barn swallows equipped with geolocators (day 1 = January 1st). The fitted regression equation is: first egg laying date = 0.88 (0.12 s.e.m.) × arrival date + 30.48 (12.51 s.e.m.) ($R^2 = 0.80$). The filled dot denotes the single Q_7/Q_8 female from the Swiss population.

genotypes, the picture emerging from this and our earlier work may suggest that 'clock' genes differently affect phenology in different barn swallow populations, as previously documented by an analysis of the songbird genus *Junco*[18]. The most likely explanation for such among-population variation in 'clock' gene-phenology associations has been proposed to be variation in genetic background among different populations of the same species, with possibly other genes related to phenology variably interacting with *Clock* in different geographical populations[18].

The proximate mechanisms leading to an association between length of Poly-Q stretches and phenology remain unknown, but may involve a different response to photoperiod or to other astronomical cues while in Africa. Despite photoperiodic changes around the Equator are minor, birds are able to detect them and tune the timing of their seasonal activities (including timing of moult) accordingly[23]. Moreover, it has recently been suggested that tropical birds can use variation in time of sunrise and sunset to adjust the timing of life-history events along the annual cycle[24]. Our findings suggest that the previously documented reproductive delay in Q_7/Q_8 females was likely to be mainly a consequence of carry-over effects of events occurring during wintering (such as moult) and at the onset of spring migration. Elucidating whether the *Clock*-phenology association causally stems from variable photoperiodic responses while in the African wintering grounds, triggering variable onset of migration activities with carry-over effects on timing of reproduction, would be a challenge for future research.

The fact that the association between *Clock* genotype and timing of autumn migration was somewhat less clear may be due to the difficulties in identifying the onset of autumn migration. Geolocator data allow only an estimate of the timing of departure from the breeding colony, which may not be indicative of variation in onset of post-breeding migration. In fact, the timing of departure from the breeding colony may be heavily affected by involvement in reproductive activities (e.g. nest failure, laying of multiple clutches and extent of post-fledging parental care) besides timing of reproduction. In addition, departures are very concentrated and apparently highly synchronized with stochastic events such as rainfall bouts, which may vary locally between study colonies[22].

Why such rare genotypes persist in barn swallow populations is open to speculation. It is known that occasional spring cold spells, occurring after arrival of barn swallows to their breeding areas, can cause large mortality[25,26]. Clearly, the rare, delayed Q_7/Q_8 individuals may accrue a survival advantage under such intense natural selection episodes, and this may contribute to maintain *Clock* polymorphism in barn swallow populations. On the other hand, early arriving Q_6/Q_7 males may accrue a high fitness advantage by mating earlier if environmental conditions are favourable, even if the association between timing of spring arrival and reproduction was not as strong as in females in the study populations, since timing of reproduction predicts seasonal fitness in barn swallows[25,27]. Finally, we wish to point out that both here and in our previous studies we focused on the association between *Clock* gene polymorphism and phenology, but future studies will need to address whether the huge phenological variance observed among the most common Q_7/Q_7 genotype is mostly due to environmental effects or whether it can be explained by genetic variation at other phenological candidate genes.

To conclude, capitalizing on a unique dataset of individual year-round tracking data of a small long-distance migratory bird obtained by miniaturized light-level geolocators, the Swiss data are suggestive for a role of genetic polymorphism at a candidate phenological gene in shaping variation in migration traits. Despite the very low genetic variation and the consequently small sample of 'rare' genotypes, our findings foster the idea that *Clock* gene polymorphism may underlie phenological variation in avian species.

Methods

Study species, study sites and field procedures. The barn swallow is a small (ca. 20 g) Holarctic semi-colonial insectivorous passerine bird. European populations are largely migratory[28]. Birds move southwards in September-October, and winter in sub-Saharan Africa, where they undergo their annual complete plumage moult. During wintering, birds are mostly stationary[22]. They then leave for spring migration in February-March and reach their breeding sites mid-March to late May[22]. Females lay one to three clutches (3–7 eggs) per season[28]. In Europe, barn swallows mostly breed in small colonies settled in rural buildings (farms, cowsheds, stables)[29].

The study was carried out at three study sites, one in southern Switzerland, in an Alpine valley floor (Magadino, 46°09' N, 8°55' E, 211 m a.s.l.) and two in the Po Plain (northern Italy; Lombardy, 45°19' N, 9°40' E, 60 m a.s.l.; Piedmont, 45°33' N, 8°44' E, 160 m a.s.l.) (details in Scandolara *et al.*[29,30] and Liechti *et al.*[22]). We captured barn swallows at several colonies during two breeding seasons (2010 and 2011). A very large sample (640 individuals; 2010, n = 310; 2011, n = 330) of adult birds (i.e. at least 1 year old) was equipped with miniaturized light-level geolocators weighting on average 0.68–0.77 g (3.74–4.14% of body mass)[30]. During the subsequent breeding season we retrieved 124 geolocators (2010, n = 78; 2011, n = 46)[22]. We collected a small blood sample from birds that successfully returned to their breeding colony with the geolocator. Blood was collected from the brachial vein into heparinized capillary tubes, stored in a cool bag and subsequently kept at −20/−80 °C. Birds were marked with individual colour rings upon capture and were assigned to their nests by direct observation[17,30]. Laying date (Julian date of first egg laying) was determined for most individual by means of frequent inspections of nest content (details in Scandolara *et al.*[30]).

All procedures were performed in accordance with the Swiss and Italian regulations concerning scientific investigations on bird species in the wild, and approved by the Office fédéral de l'environnement (OFEV, Division Espèces, écosystèmes, paysages; Switzerland) (permit n. F044-0799), by Regione Lombardia (auth. n. 329 issued on January 21, 2009 and n. 2141 issued on March 9, 2011) and by the Provincia di Novara (auth. n. 905 issued on March 21, 2011).

Geolocator data analysis. Deriving accurate geographical positions from light-level data obtained from geolocators constitutes a major methodological challenge[31]. The full details and rationale of the analyses of barn swallow geolocator data are reported in Liechti *et al.*[22]. Hereafter we briefly summarize the interpretation of the phenological traits of interest for defining individual variation in the timing of the annual routine. First, based on standardized analyses of daily changes in sunrise and sunset time derived from light-level data (latitudinal estimates were not calculated during migration because the barn swallow migration periods coincide with the period around the Equinoxes, when data derived from geolocators do not allow to accurately calculate latitudes[22,31]), we assigned each daily position to a stationary or movement (migration) period. We then calculated the following phenological traits at the individual level: departure from the breeding colony (day of departure from the breeding colony, determined by means of visual inspection of variation in daily light-level profiles); arrival to the wintering area (day of the first stationary period south of the Sahara, i.e. south of 23.5°N); departure from the wintering area (day of the last stationary period south of the Sahara); arrival to the breeding colony (date of arrival to the breeding colony, determined by means of visual inspection of variation in daily light-level profiles). Among the 124 individuals that returned with the geolocator (see above), we could obtain complete phenological information (i.e. we could obtain phenological data until spring arrival date to the breeding colony) for 68 individuals only (2010, n = 37; 2011, n = 31), due to total or partial failure of the geolocators (e.g. battery failure), which did not allow recovering the entire migration path (see Table 1 in Liechti *et al.*[22]). For each of these complete tracks we determined the stationary wintering position (SWP) as the centre of density (mode) of longitude and latitude of all the daily positions, taking into account the stationary periods south of the Sahara[22].

Genetic analyses. Genomic DNA extraction was performed by alkaline lysis using 6 μl of blood in 100 μl of a 50 mM NaOH at 100 °C for 20 minutes. DNA was quantified using a spectrophotometer and diluted to a final concentration of 50–100 ng/μl. PCR amplification was performed using primers designed on the barn swallow genomic sequence: forward primer (5′-labelled with 6-FAM dye) [6FAM]GGGACAGGTGGTGACAGCTTATC and reverse primer CTGCTGATGGTCCTGCTGACT (Sigma-Aldrich)[17]; PCR fragments were screened for length polymorphism at *Clk*polyQcds by resolution and detection on a conventional DNA sequencing machine using a commercial fragment analysis service (Macrogen Inc., Seoul, Republic of Korea)[17]. The reliability of molecular data for the barn swallow *Clk*polyQcds locus has previously been confirmed by independently repeating the fragment analysis of several individuals[17,20]. Alleles were named according to the number of glutamine repeats predicted in

the mature protein as Q_6 to Q_8, after sequencing the most common allele (Q_7, 112 bp long, coding for a CLOCK protein containing a stretch of 7 glutamine repeats)[32]. We genotyped the vast majority (64 out of 68 individuals; 2010, n = 34; 2011, n = 30) of the birds for which we obtained complete migration tracks (missing genotypes were due to failed genotyping or missing samples).

Statistical analyses. Since the genotype-phenology associations may differ between populations[18], tests of associations between *Clock* polymorphism and phenology were conducted within population. First, we calculated the 95% non-parametric bootstrap CIs (*BCa* method, based on 5000 replicates; R-package *boot*, ver. 1.3–9[33]) of each phenological trait of Q_7/Q_7 individuals. We then checked whether the phenological values of the Q_6/Q_7 and Q_7/Q_8 individuals were included or not within the CIs. If a phenological value was outside the Q_7/Q_7 CI of a given trait, that individual was considered phenodeviant for that trait (i.e. it had a phenotype that significantly deviated from the phenotypic distribution of the reference genotype). Male and female data were pooled for analyses as there were no statistically significant differences between the sexes for most phenological traits, with the exception of departure from the breeding colony, that was 2 days later on average in males compared to females[22]. In addition, to completely remove between-year and between-sex differences[22], we adopted a within-group centring procedure, subtracting from a given value the mean of its corresponding group (i.e. 2010 males, 2010 females, 2011 males, 2011 females). CIs were calculated for centred values as well. Data from the Lombardy and Piedmont study areas were pooled because both were in the same geographical area (Po Plain) and because of the small sample of rare genotypes at each study site (see Results and Discussion).

References

1. Clutton-Brock, T. H. Reproductive success: studies of individual variation in contrasting breeding systems. University of Chicago Press (1988).
2. Verboven, N. & Visser, M. E. Seasonal variation in local recruitment of great tits: the importance of being early. *Oikos* **81,** 511–524 (1998).
3. Visser, M. E., Caro, S. P., van Oers, K., Schaper, S. V. & Helm, B. Phenology, seasonal timing and circannual rhythms: towards a unified framework. *Philos. Trans. R. Soc. Lond. B* **365,** 3113–3127 (2010).
4. Merrow, M., Spoelstra, K. & Roenneberg, T. The circadian cycle: daily rhythms from behaviour to genes. *EMBO Reports* **6,** 930–935 (2005).
5. Gwinner, E. Circannual rhythms in birds. *Curr. Opin. Neurobiol.* **13,** 770–778 (2003).
6. Bartell, P. A. & Gwinner, E. A separate circadian oscillator controls nocturnal migratory restlessness in the songbird Sylvia borin. *J. Biol. Rhythms* **20,** 538–549 (2005).
7. Bell-Pedersen, D. *et al.* Circadian rhythms from multiple oscillators: lessons from diverse organisms. *Nat. Rev. Genet.* **6,** 544–556 (2005).
8. Lincoln, G. A., Andersson, H. & Loudon, A. Clock genes in calendar cells as the basis of annual timekeeping in mammals - a unifying hypothesis. *J. Endocrinol.* **179,** 1–13 (2003).
9. Darlington, T. K. *et al.* Closing the circadian loop: CLOCK-induced transcription of its own inhibitors *per* and *tim*. *Science* **280,** 1599–1603 (1998).
10. Sharp, P. J. Photoperiodic regulation of seasonal breeding in birds. *Ann. N. Y. Acad. Sci.* **1040,** 189–199 (2005).
11. Reppert, S. M. & Weaver, D. R. Coordination of circadian timing in mammals. *Nature* **418,** 935–941 (2002).
12. Rutter, J., Reick, M. & McKnight, S. L. Metabolism and the control of circadian rhythms. *Annu. Rev. Biochem.* **71,** 307–331 (2002).
13. Fidler, A. E. & Gwinner, E. Comparative analysis of Avian BMAL1 and CLOCK protein sequences: a search for features associated with owl nocturnal behaviour. *Comp. Biochem. Physiol. B* **136,** 861–874 (2003).
14. Johnsen, A. *et al.* Avian Clock gene polymorphism: evidence for a latitudinal cline in allele frequencies. *Mol. Ecol.* **16,** 4867–4880 (2007).
15. Liedvogel, M., Szulkin, M., Knowles, S. C., Wood, M. J. & 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).
16. 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. Lond. B* **277,** 3703–3714 (2010).
17. 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).
18. 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. *F1000Research* **2,** 115 (2013).
19. 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. Lond. B* **275,** 2813–2821 (2008).
20. Saino, N. *et al.* Timing of molt of barn swallows is delayed in a rare *Clock* genotype. *PeerJ* **1,** e17 (2013).
21. Saino, N. *et al.* Polymorphism at the *Clock* gene predicts phenology of long-distance migration in birds. *Mol. Ecol.* **24,** 1758–1773 (2015).
22. Liechti, F. *et al.* Timing of migration and residence areas during the non-breeding period of barn swallows *Hirundo rustica* in relation to sex and population. *J. Avian Biol.* **46,** 254–265 (2015).
23. Hau, M., Wikelski, M. & Wingfield, J. C. A neotropical forest bird can measure the slight changes in tropical photoperiod. *Proc. R. Soc. Lond. B* **265,** 89–95 (1998).
24. Goymann, W., Helm, B., Jensen, W., Schwabl, I. & Moore, I. T. A tropical bird can use the equatorial change in sunrise and sunset times to synchronize its circannual clock. *Proc. R. Soc. Lond. B* **279,** 3527–3534 (2012).
25. Møller, A. P. Sexual selection and the barn swallow. Oxford University Press (1994).
26. Brown, C. R. & Brown, M. B. Natural selection on tail and bill morphology in Barn Swallows *Hirundo rustica* during severe weather. *Ibis* **141,** 652–659 (1999).
27. Ambrosini, R., Ferrari, R. P., Martinelli, R., Romano, M. & Saino, N. Seasonal, meteorological, and microhabitat effects on breeding success and offspring phenotype in the barn swallow, *Hirundo rustica*. *Ecoscience* **13,** 298–307 (2006).
28. Turner, A. K. The barn swallow. A & C Black (2006).
29. Scandolara, C. *et al.* Context-, phenotype-, and kin-dependent natal dispersal of barn swallows (*Hirundo rustica*). *Behav. Ecol.* **25,** 180–190 (2014).

30. Scandolara, C. *et al.* Impact of miniaturized geolocators on barn swallow *Hirundo rustica* fitness traits. *J. Avian Biol.* **45,** 417–423 (2014).
31. Lisovski, S. *et al.* Geolocation by light: accuracy and precision affected by environmental factors. *Methods Ecol. Evol.* **3,** 603–612 (2012).
32. Dor, R. *et al.* Clock gene variation in *Tachycineta* swallows. *Ecol. Evol.* **2,** 95–105 (2012).
33. R Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna (2013).

Acknowledgements

We are grateful to farm owners and to many field assistants. We thank the Fondazione Bolle di Magadino for support to C.S. and the 'Ente di gestione delle aree protette del Ticino e del Lago Maggiore' for help with fieldwork. The study was funded by the EU INTERREG program (project ID 15 7624065), Fondazione Cariplo (grant UNIAGI 13357 to N.S.), University of Milano (grant 2009-ATE-0015 to D.R.), and University of Milano-Bicocca (grant 2011-ATE-0272 to R.A.). The Swiss federal office for environment contributed financial support for the development of the data loggers (UTF-Nr. 254, 332, 363, 400).

Author Contributions

N.S. and D.R. conceived and designed the experiment. G.B., D.R. and N.S. wrote the paper. G.B., R.A., M.C., L.G., E.G., S.P., F.L. and D.R. analysed the data. G.B., R.A., M.C., A.C., A.R., M.R., C.S., N.S. and D.R. collected the data.

Additional Information

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Bazzi, G. *et al. Clock* gene polymorphism and scheduling of migration: a geolocator study of the barn swallow *Hirundo rustica*. *Sci. Rep.* **5**, 12443; doi: 10.1038/srep12443 (2015).

Supplementary Information

Clock gene polymorphism and scheduling of migration: a geolocator study of the barn swallow *Hirundo rustica*

Gaia Bazzi[1], Roberto Ambrosini[2], Manuela Caprioli[1], Alessandra Costanzo[1], Felix Liechti[3], Emanuele Gatti[1], Luca Gianfranceschi[1], Stefano Podofillini[1], Andrea Romano[1], Maria Romano[1], Chiara Scandolara[1,3], Nicola Saino[1] & Diego Rubolini[1]

1. Dipartimento di Bioscienze, Università degli Studi di Milano, via Celoria 26, I-20133 Milano, Italy

2. Dipartimento di Biotecnologie e Bioscienze, Università di Milano-Bicocca, Piazza della Scienza 2, I-20126 Milano, Italy

3. Swiss Ornithological Institute, Seerose 1, CH-6204 Sempach, Switzerland

Table S1. Descriptive statistics of phenological traits (day 1 = January 1st) for Q_7/Q_7 individuals from the Swiss and Po Plain population. Mean, sample size (round brackets), and non-parametric bootstrap 95% confidence interval (CI) (square brackets) are shown. Results were qualitatively similar if we controlled for differences among years in linear models with year and population as two-level fixed factors (see also Ref. 22), and if we include data for the Q_6/Q_7 and Q_7/Q_8 individuals (details not shown for brevity). In the Po Plain, we could not estimate departure from the breeding colony and arrival to the wintering area for two Q_7/Q_7 individuals.

Phenological trait	**Switzerland**	**Po Plain**
Departure from the breeding colony	249.3 (23) [245.4; 251.8]	251.2 (33) [248.8; 253.2]
Arrival to the wintering area	283.8 (23) [280.8; 286.2]	281.5 (33) [278.4; 285.3]
Departure from the wintering area	76.7 (23) [72.3; 81.4]	71.5 (35) [67.7; 75.9]
Arrival to the breeding colony	111.3 (23) [107.0; 115.7]	101.1 (35) [97.0; 104.6]

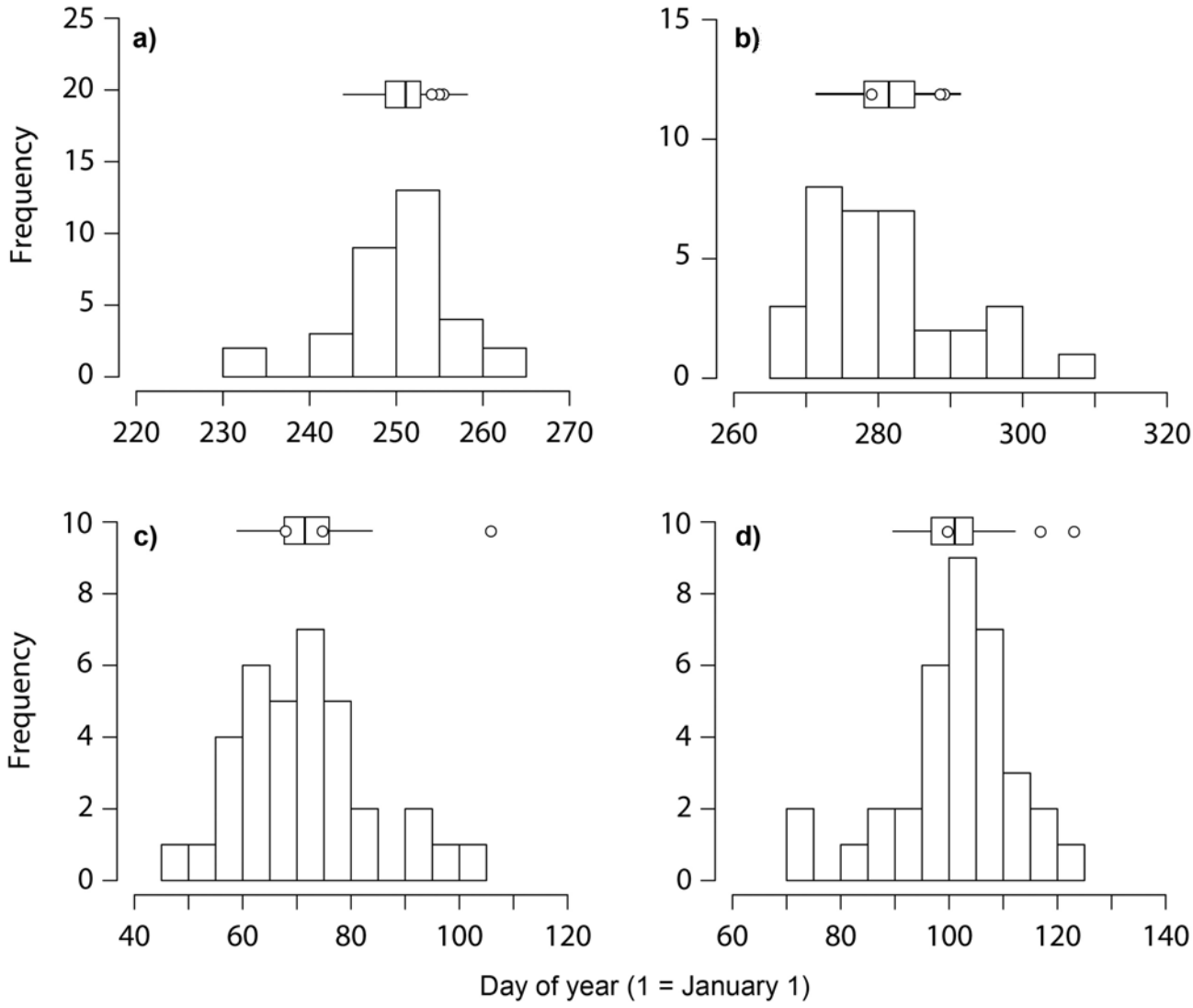
Table S2. Mean, sample size (round brackets), and non-parametric bootstrap 95% confidence interval (CI) (square brackets) of phenological traits for Q_7/Q_7 individuals from the Swiss and Po Plain population. Data for the Q_6/Q_7 individuals (mean value, with individual values in square brackets) and for the Swiss Q_7/Q_8 female are also shown. Data were centred within-groups to remove any between-year and between-sex variation (see Methods) and are expressed in days (difference from the mean). For the Swiss population, statistics of residuals from the linear regression between centred arrival date to the breeding colony values and centred latitude and longitude of the SWP (see Methods), thus removing the effects of individual variation in wintering positions on arrival date to the breeding colony (see Ref. 22), are also shown. In the Po Plain, we could not estimate departure from the breeding colony and arrival to the wintering area for two Q_7/Q_7 individuals.

Phenological trait	**Q_6/Q_7**	**Q_7/Q_7**	**Q_7/Q_8**
Switzerland			
Departure from the breeding colony	2.33 [2.33; 2.33]	-0.48 (23) [-3.70; 1.74]	6.33
Arrival to the wintering area	-5.50 [-10.00; -1.00]	0.35 (23) [-2.01; 2.67]	3.00
Departure from the wintering area	-11.50 [-16.00; -7.00]	0.10 (23) [-5.01; 4.47]	20.67
Arrival to the breeding colony	-21.06 [-24.56; -17.56]	1.57 (23) [-2.51; 5.79]	6.00
Arrival to the breeding colony (residuals)	-14.77 [-20.43; -9.10]	1.03 (23) [-2.65; 4.61]	5.86
Po Plain			
Departure from the breeding colony	3.65 [1.80; 4.08; 5.08]	-0.33 (33) [-2.53; 1.36]	-
Arrival to the wintering area	4.10 [-1.40; 6.85; 6.85]	-0.37 (33) [-3.32; 3.61]	-
Departure from the wintering area	8.16 [-6.07; -1.40; 31.93]	-0.70 (35) [-4.50; 3.54]	-
Arrival to the breeding colony	11.16 [0.00; 11.93; 17.93]	-0.85 (35) [-4.92; 2.59]	-

Table S3. Mean, sample size (round brackets), and non-parametric bootstrap 95% confidence interval (CI) (square brackets) of phenological traits (day 1 = January 1st) for Q_7/Q_7 males and females from the Swiss population. Data for the two Q_6/Q_7 males are shown in the leftmost column (mean value, with the two individual values in square brackets), while data for the Q_7/Q_8 female are shown in the rightmost column.

Phenological trait	**Q_6/Q_7**	**Q_7/Q_7 males**	**Q_7/Q_7 females**	**Q_7/Q_8**
Departure from the breeding colony	257.0 [257; 257]	250.2 (14) [245.6; 253.0]	248.0 (9) [240.5; 252.4]	251
Arrival to the wintering area	273.5 [269; 278]	284.6 (14) [280.4; 287.4]	282.6 (9) [278.2; 286.2]	288
Departure from the wintering area	59.5 [55; 64]	74.6 (14) [68.4; 79.6]	80.1 (9) [73.2; 88.6]	113
Arrival to the breeding colony	79.5 [76; 83]	110.3 (14) [105.4; 115.3]	112.8 (9) [104.8; 120.9]	130

Figure S1. Frequency distribution of phenotypic values of the Q_7/Q_7 individuals and phenology of the Q_6/Q_7 individuals in the Po Plain. a) Departure from the breeding colony; b) arrival to the wintering area; c) departure from the wintering area; d) arrival to the breeding colony. The horizontal boxes above histograms shows the mean and 95% non-parametric confidence interval (CI) of the Q_7/Q_7 phenotypic distribution, while the horizontal lines show the standard deviation. Dots show the phenotypic values of the Q_6/Q_7 individuals. The x- and y-axis are on the same scale as those of the corresponding panels of Fig. 1, to facilitate comparisons of phenology between populations.



PART 3

CHAPTER 5

Clock gene polymorphism and interspecific variation in migratory behaviour: a comparative study of trans-Saharan migratory birds

Gaia Bazzi, Jacopo G. Cecere, Manuela Caprioli, Emanuele Gatti, Luca Gianfranceschi, Stefano Podofillini, Cristina D. Possenti, Roberto Ambrosini, Nicola Saino, Fernando Spina, Diego Rubolini

Submitted to *Molecular Ecology*

Photo: Museo Civico Lentate sul Seveso, Andrea Galimberti, don Amedeo Folladori, Matteo Fransci

Clock gene polymorphism and interspecific variation in migratory behaviour: a comparative study of trans-Saharan migratory birds

Gaia Bazzi[1], Jacopo G. Cecere[2], Manuela Caprioli[1], Emanuele Gatti[1], Luca Gianfranceschi[1], Stefano Podofillini[1], Cristina D. Possenti[1], Roberto Ambrosini[3], Nicola Saino[1], Fernando Spina[2], Diego Rubolini[1]*

1. Dipartimento di Bioscienze, Università degli Studi di Milano, via Celoria 26, I-20133, Milan, Italy

2. ISPRA – Istituto Superiore per la Protezione e la Ricerca Ambientale, via Cà Fornacetta 9, I-40064 Ozzano dell'Emilia (BO), Italy

3 Dipartimento di Scienze dell'Ambiente e del Territorio e di Scienze della Terra (DISAT), Università degli Studi di Milano Bicocca, Piazza della Scienza 1, I-20126 Milan, Italy

Keywords: *Adcyap1*, breeding latitude, candidate genes, long-distance migration, phenology, phylogenetic path analysis

*** Correspondence should be addressed to:**
Diego Rubolini, Dipartimento di Bioscienze, Via Celoria 26, 20133 Milan, Italy,
email: diego.rubolini@unimi.it

Running head: Candidate genes and long-distance migration

Author's names with initials: Gaia Bazzi (GB), Jacopo G. Cecere (JGC), Manuela Caprioli (MC), Emanuele Gatti (EG), Luca Gianfranceschi (LG), Stefano Podofillini (SP), Cristina D. Possenti (CDP), Roberto Ambrosini (RA), Nicola Saino (NS), Fernando Spina (FS), Diego Rubolini (DR)

Abstract

Migratory behaviour is controlled by endogenous circannual rhythms that are synchronized by external cues, such as photoperiod. Investigations on the genetic basis of circannual rhythmicity in vertebrates have highlighted that variation at 'circadian clock' genes may play a major role in regulating photoperiodic responses and timing of life cycle events, such as reproduction and migration. In this comparative study of 23 trans-Saharan migratory bird species we investigated the relationship between genetic variation at two candidate genes, *Clock* and *Adcyap1*, and migratory traits, including timing of spring migration across the central Mediterranean Sea. Coherently with previous evidence showing latitudinal clines in 'circadian clock' genes, *Clock* allele size increased with breeding latitude across species. However, there was no evidence that *Clock* allele size predicted the timing of spring migration. *Clock* (but not *Adcyap1*) gene diversity was significantly lower among species migrating over longer distances, and among those showing delayed spring migration and smaller phenotypic variance in migration timing. Phylogenetic confirmatory path analysis suggested that migration date and distance were the most important variables directly affecting *Clock* gene diversity. Hence, our novel comparative study supports the hypothesis that *Clock* affects migratory behaviour, possibly via adaptation to photoperiodic conditions in the breeding areas. Moreover, it suggests for the first time that long-distance migration is associated with lower gene diversity at candidate genes for migration, likely due to strong stabilizing selection on timing of life cycle events and to the time constraints imposed by a late migration and delayed life cycle.

Introduction

Many bird species breeding in the Northern Hemisphere respond to seasonal variation in ecological conditions by moving to more favourable areas of the tropics during the non-breeding (winter) period. This long-distance migration strategy has likely evolved in species that progressively extended their breeding range northward during the post-glacial phase to exploit favourable summer ecological conditions in temperate seasonal habitats, and were then subjected to selective pressures gradually promoting migratoriness to maximize survival during the non-breeding period (Milá *et al.* 2006; Salewski & Bruderer 2007). Migration between the breeding and non-breeding areas is accomplished by physiological and behavioural shifts that take place at the appropriate time of the year and are triggered by endogenous circannual rhythms (Visser *et al.* 2010). Such endogenous rhythms have a strong genetic component (Visser *et al.* 2010) and are synchronized by external cues (Merrow *et al.* 2005). Variation in photoperiod has been suggested among the primary external cues promoting the synchronization of circannual rhythms (e.g. Gwinner 1986; Gwinner 2003; Sharp 2005; Pulido 2007). Translating photoperiodic signals into appropriate behavioural and physiological changes takes place through complex neurophysiological cascades that involve specific receptors and their encoding genes, among which those of the 'clock' gene family play a major role (Panda *et al.* 2002; Lincoln *et al.* 2003; Bell-Pedersen *et al.* 2005; Ko & Takahashi 2006). Specifically, research on migratory birds has focused on the *Clock* (*Circadian Locomotor Output Cycles Kaput*) gene, that regulates the core circadian oscillator, which is ultimately responsible for the onset and setting of circadian and circannual rhythms (Panda *et al.* 2002; Lincoln *et al.* 2003; Bell-Pedersen *et al.* 2005; Ko & Takahashi 2006). The avian *Clock* exonic region shows a C-terminal polyglutamine (Poly-Q) repeat sequence that varies in extent both among populations and among individuals within populations (e.g. Fidler & Gwinner 2003; Johnsen *et al.* 2007; Liedvogel *et al.* 2009; Caprioli *et al.* 2012; Peterson *et al.* 2013). It has been suggested that this allelic length polymorphism may modulate gene function (Gekakis *et al.* 1998; Hayasaka *et al.*

2002). Among-population comparisons in size of the *Clock* Poly-Q (number of repeat CAG sequences) have highlighted that *Clock* Poly-Q allele size increases at northern latitudes in the blue tit (*Parus caeruleus*) (Johnsen *et al.* 2007). Thus, specific *Clock* genotypes may be adapted to latitudinal differences in the photoperiod that migratory birds experience in their breeding areas. Since the timing of spring seasonal activities becomes increasingly delayed while moving polewards, longer *Clock* Poly-Q alleles should be associated with delayed timing of reproduction and pre-nuptial migration. In line with this hypothesis there is some evidence that, within-populations, individuals bearing longer *Clock* alleles show a delayed timing through the entire life cycle (Liedvogel *et al.* 2009; O'Malley *et al.* 2010; Caprioli *et al.* 2012; Saino *et al.* 2013; Bazzi *et al.* 2015; Bourret & Garant 2015; Saino *et al.* 2015, but see Liedvogel & Sheldon; Dor *et al.* 2011; Chakarov *et al.* 2013).

Hence, both theoretical and empirical arguments suggest that the candidate gene *Clock* is involved in the control of phenology and migratory behaviour. The polymorphism of candidate genes may represent the outcome of adaptation to natural selection pressures favouring specific traits of migratory birds that are at least partly controlled by those genes. If this hypothesis is true, interspecies differences in candidate gene polymorphism are expected to explain interspecies differences in migratory behaviour, such as migration timing or migration distance. Other candidate genes have been suggested to contribute to shaping migratory behaviour. For instance, *Adcyap1*, another candidate gene whose encoded polypeptide (PACAP) is also involved in the physiological regulation of circadian and circannual rhythms (Simonneaux *et al.* 1993; Hannibal *et al.* 1997; Nagy & Csernus 2007; Racz *et al.* 2008; Schwartz & Andrews 2013), shows a 3'-UTR microsatellite polymorphism that may explain variation in migratory restlessness and migration distance (Mueller *et al.* 2011; Peterson *et al.* 2013; Bazzi *et al.* 2016).

In this study, we assessed the degree of interspecific variation in *Clock* and *Adcyap1* gene polymorphism in a set of 23 trans-Saharan migratory birds that share the same broad selective pressures, ecology and migratory behaviour (Table S1). All these species breed in Eurasia and

winter in sub-Saharan Africa (Cramp 1998). We then assessed how genetic variation at these two candidate genes for migration covaries with interspecific variation in migratory traits. Despite the species considered in our study share the same general migration strategy (i.e. they all breed in Eurasia, winter south of the Sahara and cross the desert and central Mediterranean sea during spring migration), they show large among-species variation in migratory traits (see e.g. Rubolini *et al.* 2005). For instance, several species migrate to the Mediterranean to breed, while others go further to northern Europe. Moreover, some species overwinter just south of the Sahara, in the Sahel region, while others extend their winter distribution to southern Africa (Rubolini *et al.* 2005). As a consequence, conspicuous interspecies differences in the timing of spring migration towards the breeding areas exist, as assessed by variation spring migration date across the Mediterranean (Spina *et al.* 1993; Pilastro & Spina 1997; Pilastro *et al.* 1998; Rubolini *et al.* 2004; Rubolini *et al.* 2005; Saino *et al.* 2007; Saino *et al.* 2010a; Robson & Barriocanal 2011). Moreover, large interspecific differences also exist in the variance of migration dates, with some species migrating over several weeks while others migrate over a few days only (Spina *et al.* 1993).

We first tested whether long *Clock* allele size predicted relatively late spring migration across the Mediterranean at the species level. This is expected because *Clock* genotypes with larger poly-Q stretches were found to be associated with a delayed phenology in some previous intraspecific studies (Liedvogel *et al.* 2009; Caprioli *et al.* 2012; Saino *et al.* 2013; Bazzi *et al.* 2015; Bourret & Garant 2015; Saino *et al.* 2015). Moreover, we tested the hypothesis that species migrating to more northern latitudes have longer *Clock* alleles. If *Clock* controls the rhythms of seasonal activities in the breeding areas (Johnsen *et al.* 2007; Liedvogel *et al.* 2009; Caprioli *et al.* 2012; Bourret & Garant 2015), species migrating to northern breeding areas are expected to show longer *Clock* alleles.

Finally, species differ broadly in the extent of candidate genes polymorphism. Some species show very low polymorphism, especially at *Clock* , while others show a broad diversity of genotypes (Liedvogel *et al.* 2009; Liedvogel & Sheldon 2010; Dor *et al.* 2011; Caprioli *et al.* 2012;

Chakarov *et al.* 2013; Kuhn *et al.* 2013; Peterson *et al.* 2013; Saino *et al.* 2013; Bazzi *et al.* 2015; Bourret & Garant 2015; Saino *et al.* 2015; Bazzi *et al.* 2016). To our knowledge, no study has ever assessed whether such variability among species is associated with migratory behaviour. Several arguments suggest that this may be the case. For instance, species migrating over longer distances generally show reduced phenotypic variance in migratory behaviour compared to those migrating over shorter distances (Berthold 1996, 2001; Pulido & Widmer 2005). The former are likely under strong selection favouring those phenotypes and genotypes that performed better throughout a tightly scheduled annual cycle. Indeed, long-distance migrants arrive in late spring at their breeding sites (e.g. Hagan *et al.* 1991; Rubolini *et al.* 2010), have a rapid juvenile development, and leave for their non-breeding areas earlier than short-distance migrants (e.g. Gwinner 1986; Berthold 1996). This is achieved by a strong endogenous regulation of migratory behaviour (e.g. Gwinner 1971; Gwinner 1972; Hagan *et al.* 1991; Gwinner & Helm 2003). By contrast, species migrating over shorter distances have less tight annual routines and show a more labile endogenous control, resulting in larger phenotypic variance (e.g. Gwinner 1972; Pulido & Widmer 2005). We might thus predict that species migrating over longer distances, showing a delayed migration and a reduced variance in timing of migration, would show reduced candidate gene polymorphism. We built several hypothetical models relating migration distance, spring migration timing (date and variance) to both *Clock* and *Adcyap1* gene polymorphism, and used phylogenetic path analysis (von Hardenberg & Gonzalez-Voyer 2013; Gonzalez-Voyer & von Hardenberg 2014) to try to infer the possible causal relationships among variables.

Materials and Methods

Study species and field methods

We included in our study 23 small- to medium-sized long-distance migratory bird species belonging to 6 orders (Columbiformes, Caprimulgiformes, Coraciiformes, Bucerotiformes, Piciformes and Passeriformes) (Table S1; Fig. S1). The study was conducted at Ventotene (40°48'N – 13°25'E), a small island located in the central Mediterranean sea, ca. 50 km off the Italian coast. Birds were trapped during spring migration using mist-nets, following standard protocols with constant capture effort, and individually marked with metal rings (Spina *et al.* 1993; Saino *et al.* 2010a). We considered in this study all those trans-Saharan migratory species (i.e. species that breed in Eurasia and mainly or partly overwinter in sub-Saharan Africa, according to Cramp 1998) that regularly stopover at the study site during spring migration and that we could sample in sufficiently large numbers to allow reliable estimates of genetic variation (see below). During spring migration, most birds arrive at Ventotene directly from North Africa, after crossing the Sahara Desert and the Mediterranean Sea (Pilastro *et al.* 1995; Pilastro & Spina 1997; Pilastro *et al.* 1998; Grattarola *et al.* 1999; Saino *et al.* 2010b). Most of the study species are known to perform only very short stopovers at Ventotene, and in most cases birds only rest on the island for a few hours before resuming their migration towards their breeding quarters (Goymann *et al.* 2010; Tenan & Spina 2010). Only first capture dates (i.e. excluding recaptures of birds previously ringed during the same season) were considered in this study. As in previous studies, we assumed that the distribution of first capture date reliably reflects the phenology of spring migration across the Mediterranean Sea of the target species (Saino *et al.* 2010a; Saino *et al.* 2015).

Samples for genetic analyses were taken during March-May 2002, 2008, 2013 and 2014. Each species was sampled in a single study year, the only exception being the European nightjar (*Caprimulgus europaeus*) that was sampled in both 2013 and 2014 to increase sample size. The

trapping period encompassed the whole spring migration period of the study species at Ventotene (Spina *et al.* 1993; Messineo *et al.* 2001; Saino *et al.* 2010a). We aimed at genotyping at least 30 individuals per species, evenly distributed throughout the migration period of each species. Hence, we sampled every *i*-th individual of a given species, and species-specific *i* values were decided based on the capture data of previous years (see Saino *et al.* 2010a; Saino *et al.* 2015). We also included data from four species (*Anthus trivialis*, *Ficedula hypoleuca, Luscinia megarhynchos, Saxicola rubetra* and *Phylloscopus trochilus*) that had been genotyped in larger numbers for other purposes (Saino *et al.* 2015; our unpubl. data) (see Table S2). For each individual we took a small blood (ca. 10-30 µl; collected in heparinized capillary tubes, on sterile tissue paper or on glass slides) or feather sample (3/4 undertail coverts) for genetic analyses. Samples were stored at -20 °C (blood collected in capillary tubes) or at room temperature (blood collected on tissue paper, preserved in buffer; feathers; blood collected on ethanol-fixed and dried glass slides).

Life-history traits

For each species, we recorded several life-history traits, including distribution, population size and phenological variables. All data are reported in Table S1. Information on distribution was obtained from the most recent distribution maps available (BirdLife International & NatureServe 2015) and distribution traits were calculated using ArcMap 9.3 (ArcGis® software by ESRI). Considering that birds were sampled at Ventonene during spring migration towards northern breeding areas, we excluded from the calculations the breeding range south of Ventotene (i.e. south of 40°48’ N). Moreover, we excluded all breeding ranges extending east of the Urals (i.e. east of 60°00’ E). Ring recoveries from the central Mediterranean corroborate our decision to exclude those populations (see Spina & Volponi 2008a, b). Similarly, we excluded from the calculations the European sector of the wintering range north of the latitude of Ventotene (some species are establishing wintering populations in southern Europe, e.g. the hoopoe *Upupa epops* and the wryneck *Jynx torquilla*)

(Cramp 1998; van Wijk *et al.* 2013) and included only information from the African winter distribution and from those regions of the European winter distribution that lye south of the latitude of Ventotene.

For each species we recorded the following distribution variables: 1) northernmost breeding latitude (hereafter breeding latitude, expressed in decimal degrees; the northernmost breeding latitude indicates of how much a species' distribution spreads northwards); 2) breeding range size (in km^2); 3) latitudinal width of breeding and of wintering distribution ranges (breeding and wintering range widths, hereafter), calculated as the difference between the northernmost and the southernmost latitude of breeding and wintering distribution ranges, respectively (in decimal degrees); 4) migration distance, calculated as the distance (in km) between the centroids of breeding and wintering ranges obtained as outlined above (Table S1).

We computed two components of species-specific phenology at Ventotene: the median spring migration date (day 1 = January 1) (mean value of yearly median migration dates for each species; migration date hereafter), and the spread of migration dates (migration spread hereafter; calculated as the mean value of the yearly difference between the 90^{th} and the 10^{th} percentile of migration dates, expressed in days). To do this, we used data collected in the years 2006-2012 (sampling period March 1 to May 31, encompassing the entire migration period of trans-Saharan migrants in the Mediterranean; see also Saino *et al.* 2010a). Both timing of spring migration and migration spread were significantly repeatable within species (migration date, repeatability: $R = 0.87$, $F_{22,138} = 49.88$, $P < 0.001$; migration spread, repeatability: $R = 0.49$, $F_{22,138} = 8.09$, $P < 0.001$; N = 23 species and N = 161 species × year combinations) (see also Rubolini *et al.* 2005).

Genetic analyses and candidate gene polymorphism

Blood genomic DNA was extracted by alkaline lysis. Approximatively 6 µl of blood (from capillary tubes, tissue paper or grass slides) were boiled at 100°C in 70-100 µl (depending on the quantity of

blood available) of a 50mM NaOH solution for 20 min. Extracted DNA was then quantified using a spectrophotometer and diluted to a final concentration of 50-100 ng/µl. Genomic DNA from feathers was extracted using a commercial kit, as described in Saino *et al.* (2015) (ArchivePure DNA purification kit, 5 PRIME, Hilden, Deutschland).

Polymorphisms at *Clock* poly-Q and *Adcyap1* 3'-UTR were determined by fragment analysis after PCR amplification as in Caprioli *et al.* (2012) and Saino *et al.* (2015). PCR products were labelled with HEX (*Clock*), 6-FAM (*Clock* and *Adcyap1*) or TAMRA (*Adcyap1*) dyes and PCR products of 2 individuals or 2 loci were mixed in a single well in most cases.

Clock allele size was expressed for each species as the mean of individuals' mean or long *Clock* allele sizes. The use of the long (besides mean) allele size is justified because dominance of the longer allele in shaping the phenotype has been suggested for *Clock* (Liedvogel *et al.* 2009; Saino *et al.* 2015), similarly to other genes that show poly-Q polymorphism (Ross 2002; Fondon *et al.* 2008). As *Clock* allele length polymorphism is due to a variable number of glutamine codons sequences, we can safely compare *Clock* allele sizes of different species. Conversely, *Adcyap1* allele size could not be compared among species. The *Adcyap1* 3'-UTR length polymorphism is mainly due to a compound microsatellite sequence variation, but alleles differing by just one nucleotide occur (Saino *et al.* 2015; Bazzi *et al.* 2016), proving that the observed polymorphism could be due to any type of nucleotide variation, among species. Hence, *Adcyap1* allele size differences among species could not be attributed to short tandem repeats only, making variation in allele size of this candidate gene not useful for comparative purposes.

Finally, as species (besides differing in allele length) broadly varied in allele and genotype diversity, we calculated an index of gene diversity across species for both the *Clock* and *Adcyap1* loci (hereafter *Clock* and *Adcyap1* gene diversity, respectively). Gene diversity ($\hat{h}$) was calculated as the allelic diversity for a given locus according to the formula: $\hat{h} = 2m\,(1 - \sum x_i^2) / (2m - 1)$, where m is the number of individuals for each species and x_i the frequency of the i-th allele (see Nei & Roychoudhury 1974; formula 12.33 in Nei & Kumar 2000).

Gene diversity of both *Clock* and *Adcyap1* loci was unrelated to sample size, population size ($\log_{10}$-transformed estimated number of breeding pairs in Europe) (from the BirdLife International website www.birdlife.org), or breeding range size (Table S1). Hence, these variables did not affect candidate gene diversity estimates, and were not further considered.

On the whole, we successfully genotyped 1400 individuals for the *Clock* locus and 1088 individuals for *Adcyap1* (Table S1).

Phylogenetic comparative analyses

In order to statistically control for phylogenetic relationships among species, as required in comparative studies (Felsenstein 1985; Harvey & Pagel 1991; Freckleton *et al.* 2002), we downloaded from the BirdTree website (www.birdtree.org) (Jetz *et al.* 2012) 2000 trees from both the 'Ericson' and the 'Hackett' phylogenies of our set of 23 species (following Rubolini *et al.* 2015; details in Jetz *et al.* 2012). Both the 'Ericson' and the 'Hackett' tree sets were summarized into consensus trees (50% majority-rule consensus tree), following the procedure described in Rubolini *et al.* (2015) , using the SumTrees program, part of DendroPy, a Python library for phylogenetic computing (Sukumaran & Holder 2010). Since the 'Hackett' consensus tree showed a basal polytomy between the European nightjar and the turtle dove (*Streptopelia turtur*), it could not be used in phylogenetic analyses. Hence, all subsequent analyses were based on the 'Ericson' phylogeny only. The consensus tree is shown in Fig. S1.

To investigate the associations between *Clock* allele size or *Clock* and *Adcyap1* gene diversity and migratory traits, we relied on phylogenetically corrected generalized least-square regression models (PGLS). PGLS models were ran using the 'pgls' function of the R library caper (ver. 0.5.2) (Orme *et al.* 2011), with the λ parameter (Freckleton *et al.* 2002) estimated using maximum likelihood (ML).

We first tested whether *Clock* allele size predicted phenology by means of PGLS models of migration date as a function of *Clock* allele size. Since interspecific variation in timing of spring migration of trans-Saharan migrants across the Mediterranean is strongly related to wintering latitude, with species wintering at more southern latitudes in Africa migrating later at Ventotene (Rubolini *et al.* 2005), we accounted for the possible confounding effects of migration distance on migration date by including migration distance as a further predictor in the PGLS models. On the other hand, breeding latitude was not a significant predictor of migration date at Ventotene neither in this (details not shown) nor in previous studies of timing spring migration at Ventotene (Rubolini *et al.* 2005) while controlling for migration distance. Secondly, we ran PGLS models of *Clock* allele size as a function of breeding latitude. All models were run by considering either mean and long *Clock* allele size.

Finally, to investigate how migration-related traits affected gene diversity and species' traits, we ran four separate PGLS models in which gene diversity (both *Clock* and *Adcyap1*) was the dependent variable and breeding range width, wintering range width, migration distance and migration date were the predictors. We relied on univariate models since the different species' traits were intrinsically intercorrelated (details not shown). If candidate gene diversity reflects an adaptation to phenologically and ecologically variable environmental conditions, we may expect gene diversity to increase with increasing (breeding and wintering) range width. On the other hand, we expected gene diversity to decrease with increasing migration distance and later migration date (see Introduction). Moreover, if low gene diversity at candidate genes imposes constraints on phenotypic variance, we could expect low gene diversity to result in a smaller spread of migration dates. Hence we ran a further PGLS model in which migration spread was the dependent variable and gene diversity (both *Clock* and *Adcyap1*) was the predictor.

The robustness of PGLS models was checked with respect to phylogenetic uncertainty by means of the model-averaging method recommended by Garamszegi and Mundry (2014). BirdTree tree sets include a sample of Markow Chain Monte Carlo (MCMC) trees sampled in proportion to

their posterior probability (Jetz *et al.* 2012), and individual trees might vary either in topology or branch length. To this end, we ran a PGLS model for each of the 2000 trees of our tree set. Parameter estimates were then averaged across the models according to an information-theoretic (IT) approach (Burnham & Anderson 2002), with each model weighted according to its fit to the data, as evaluated using the Akaike's information criterion (AIC), and reported with their associated 95% confidence intervals (CIs).

To disentangle the causal relationships behind the associations between migratory traits and gene diversity, we carried out a phylogenetic confirmatory path analysis, following von Hardenberg and Gonzalez-Voyer (2013), Gonzalez-Voyer and von Hardenberg (2014). The causal hypotheses, described in the Results section, were represented with directed acyclic graphs (DAGs), that can be mathematically expressed as a set of structural equations. The hypothesized causal models were tested by means of the d-sep test applied to PGLS models (von Hardenberg & Gonzalez-Voyer 2013). The Fisher's *C* statistic was computed for each causal model (Shipley 2000), and the best-fitting models were selected by ranking them according to the *C*-statistic Information Criterion (CIC) (von Hardenberg & Gonzalez-Voyer 2013). If models were equally well-supported (i.e. they had a $\Delta CIC < 2$ from the best-fitting model; von Hardenberg & Gonzalez-Voyer 2013), path coefficients (calculated on standardized variables) were averaged according to their CICc weights (Symonds & Moussalli 2011). Standardized path coefficients are reported with their associated 95% CIs. All analyses were run using R 3.0.1 (R Core Team 2013).

Results

Clock allele size, phenology and distribution

We observed a strong phylogenetic signal in *Clock* allele size, with estimated λ values varying between 0.75 and 0.93 for long and mean allele size, respectively (ML λ estimates derived from intercept-only PGLS models). Hence, closely related species show more similar *Clock* allele sizes than distantly related ones due to phylogenetic descent.

Interspecific variation in spring migration date was not significantly affected by *Clock* allele size (either mean or long) across species (Table 1). The results were similar when controlling for the effect of migration distance on migration date [mean *Clock* allele size, estimate: -0.24 (0.28 SE), $t = -0.85$, $P = 0.40$; effect of migration distance, estimate: 65.14 (13.94 SE), $t = 4.67$, $P < 0.001$; results for the long *Clock* allele were qualitatively similar, details not shown for brevity].

On the other hand, interspecific variation in long *Clock* allele size was significantly and positively predicted by breeding latitude (Table 1, Fig. 2), while the association was weaker and non-significant for the mean *Clock* allele size (Table 1). The alternative scenario, i.e. that longer *Clock* alleles might allow species to colonize more northern latitudes, was also supported [effect of *Clock* allele size on breeding latitude; mean allele size, PGLS estimate: 0.42 (0.18 SE), $t = 2.40$, $P = 0.026$; long allele size, estimate: 0.48 (0.16 SE), $t = 2.93$, $P = 0.008$]. Model-averaged parameter estimates confirmed in all cases the results from PGLS models ran on the consensus tree (see Table 1; other details not shown for brevity for the alternative scenario).

Gene diversity, phenology and distribution

We observed a strong phylogenetic signal in *Clock* (ML estimate of $\lambda = 1$) but not *Adcyap1* gene diversity ($\lambda = 0$), implying that closely related species show more similar *Clock* (but not *Adcyap1*) gene diversity than distantly related ones due to common ancestry.

Clock gene diversity was negatively and significantly predicted by migration distance and migration date in univariate PGLS models (Table 2; Fig. 3). Moreover, *Clock* gene diversity significantly and positively predicted migration spread [PGLS, estimate: 9.71 (4.42 SE), $t = 2.19$, $P = 0.039$; Fig. 3]. Conversely, neither the width of the breeding range nor that of the wintering range significantly predicted *Clock* gene diversity (Table 2). In contrast, in similar univariate PGLS models, *Adcyap1* gene diversity was not significantly affected by any of the phenological or distribution variables we considered (Table 2), and it did not significantly predict migration spread (PGLS, estimate = 1.94 (6.51 SE), $t = 0.30$, $P = 0.77$). All results from the univariate PGLS models of gene diversity run on the consensus tree were confirmed by the model-averaged parameter estimates (see Table 2; details not shown for gene diversity vs. migration spread).

Phylogenetic path analysis of the association between *Clock* gene diversity, migration distance, migration date and migration spread

To limit complexity, causal models were built based on the results derived from the exploratory univariate PGLS models described in the previous paragraph. We considered only those variables that were significantly associated with gene diversity. Hence, we did not run any path analysis involving *Adcyap1* gene diversity and built a set of 12 causal models relating migration distance (M), migration date (D) and migration spread (S) to *Clock* gene diversity (H) (Fig. 1). Migration distance, migration date and migration spread can be regarded as a part of the same 'migratory syndrome' and are tightly related to each other (see also Dingle 2006). Based on previous evidence (e.g. Francis & Cooke 1986; Hagan *et al.* 1991; Kissner *et al.* 2003; Rubolini *et al.* 2005), in all causal models we assumed that species migrating over longer distances arrive later in spring (late

migration date), and that species that arrive late in spring also show smaller migration spread, either directly as a consequence of longer migrations (M09, M10 and M12) or indirectly as a consequence of later migration (M02-M08). As explained above, we assumed that migration distance and migration date affected *Clock* gene diversity, while the latter was assumed to affect migration spread. M01 was *a priori* deemed as highly unlikely (it implied that migration distance and migration date independently and directly affected *Clock* gene diversity) and was tested as a 'control' model.

Most causal models fitted the data relatively well, as gauged by the non-significant values of the Fisher's *C* statistic (Table 3). M03 and M02 were identified as the best causal models ($\Delta CIC < 2$) (Table 3). These models envisaged that both migration distance and migration date directly affect *Clock* gene diversity (Fig. 1). Moreover, migration date was affected by migration distance, which in turn affected migration spread (Fig. 1). Standardized path coefficients from the best fitting models suggested that migration date and distance both directly affected *Clock* gene diversity with a similar intensity (Table 4). Finally, the other direct causal path showed that migration distance positively and significantly predicted median migration date, and that migration spread was negatively and significantly predicted by migration date (Table 4).

Discussion

In this comparative study of 23 trans-Saharan migratory bird species we investigated the association between *Clock* and *Adcyap1* polymorphism and life-history traits related to migration and geographic distribution. We found a significant relationship between breeding latitude and *Clock* allele size, with species having breeding ranges that extend to more northern latitudes that bear on average longer alleles. Conversely, *Clock* allele size was not related to migration date through the central Mediterranean. *Clock* gene diversity was predicted by several traits related to migration:

longer migration distances, delayed spring migration across the central Mediterranean sea, and reduced variance of migration dates were all significantly associated with lower *Clock* gene diversity. The possible causal relationships between *Clock* gene diversity and migratory traits were investigated using phylogenetic confirmatory path analysis (von Hardenberg & Gonzalez-Voyer 2013; Gonzalez-Voyer & von Hardenberg 2014). No association between *Adcyap1* genetic diversity and life-history traits emerged: we will henceforth focus on *Clock* gene polymorphism only.

Life-history and interspecific variation in *Clock* allele size

We found *Clock* allele size to increase with the northernmost latitude reached by the breeding range of a species, and the relationship was stronger for the long rather than the mean *Clock* allele size. This finding corroborates the hypothesis that length polymorphism could reflect an adaptation to different photoperiodic conditions experienced by birds breeding at different latitudes (Johnsen *et al.* 2007). The timing of the breeding season broadly varies across Europe along a southwest-northeast axis; birds directed towards northern latitudes generally have delayed and shorter reproductive periods and experience much more marked daily differences in photoperiod compared to their southern breeding counterparts in their breeding areas (e.g. Gwinner 1986; Berthold 1996). *Clock* alleles with more CAG repeats could provide an adaptive advantage for migrants that must delay the timing of breeding and modulate the reproductive season (and other events of the annual life cycle, including timing of migration) accordingly. Since photoperiod acts as a major external cue to synchronize circannual rhythms (e.g. Gwinner 1986; Gwinner 2003; Sharp 2005; Pulido 2007), the photoperiodic *Clock* gene could mainly schedule the timing of life-history events of migrants while they are at the breeding grounds (Liedvogel *et al.* 2009; Caprioli *et al.* 2012; Bourret & Garant 2015), where the seasonal differences in day length are more pronounced. Hence, natural selection could have favoured the occurrence of longer alleles where birds need to cope with sharp daily changes of photoperiod. In both cases, natural selection may have led to a higher frequency of

such alleles in species breeding at northern latitudes compared to those species with a more southerly distribution. These findings are in accordance with latitudinal clines in *Clock* allele size that have been previously reported among blue tit populations (Johnsen *et al.* 2007) and in allelic polymorphism of 'clock' family genes among *Drosophila* populations (Costa *et al.* 1992; Hut *et al.* 2013).

Finally, the strength of the allele size-breeding latitude association was larger for the long compared to the mean *Clock* allele size: this hints a dominance of the longer allele, as previously suggested for *Clock* and other candidate genes for migration (Liedvogel *et al.* 2009; Saino *et al.* 2015; Bazzi *et al.* 2016). This would be consistent with the notion that many phenotypic traits (including pathological states, such as the Huntington's disease) originating from tandem repeat sequences show dominance of the longer allele (Ross 2002; Fondon *et al.* 2008).

On the other hand, we did not find any association between timing of spring migration and *Clock* polymorphism. Such an association was expected because *Clock* polymorphism has been previously shown to predict phenology of spring migration at the individual level in some bird species (Bazzi *et al.* 2015; Saino *et al.* 2015). The lack of significant associations between *Clock* allele size and migration date was confirmed also when removing the confounding effect of migration distance, that is a strong determinant of timing of spring migration across species (Francis & Cooke 1986; Hagan *et al.* 1991; Rubolini *et al.* 2005). It might be speculated that *Clock* allele size might be more strongly tied with the phenology of other seasonal life cycle events, such as timing of breeding or departure for autumn migration, rather than timing of spring migration across the Mediterranenan. Indeed, timing of breeding and autumn migration are expressed in highly seasonal environments, and are expected to be more strongly linked to photoperiodic cues that can affect *Clock* gene expression than timing of spring migration from equatorial Africa to Eurasia.

458 *Clock* gene diversity and migratory traits

459

460 Long-distance migratory birds typically show less phenotypic variance in migratory traits compared
461 to those migrating over shorter distances (Berthold 1996, 2001; Pulido & Widmer 2005). It has
462 been suggested that the low phenotypic variance observed among long-distance migratory species
463 could arise from a tighter endogenous control of migration, resulting in environmental canalization
464 of migratory behaviour (Debat & David 2001; Pulido & Widmer 2005), rather than from depleted
465 additive genetic variation (Pulido & Widmer 2005). Long-distance migrants leave the wintering
466 grounds thousands of km far from their breeding quarters, and weeks or months before their arrival,
467 with very limited (if any) cues about the ecological conditions at their goal areas (Saino &
468 Ambrosini 2007). Thus, the low genetic variation exhibited by such species, along with
469 environmental canalization, could be explained by the occurrence of a strong stabilizing selection
470 towards the few phenotypes (and genotypes) that allow birds to reach the critical fuel load required
471 for successfully completing their long journey across ecological barriers of the Sahara Desert and
472 Mediterranean Sea, to migrate during the short window of optimal conditions *en route* and to reach
473 the breeding sites just in time to exploit the peak of food resources that are required for achieving a
474 successful reproduction (e.g. Berthold 1995; Merilä 1999; Both & Visser 2001). Although our study
475 included only trans-Saharan migratory bird species, interspecies differences in migration distances
476 were huge, varying from less than 3000 km in Mediterranean breeding species wintering in the
477 Sahel (e.g. *Sylvia cantillans*) to more than 7000 km in species breeding in north-central Europe and
478 wintering in southern Africa (e.g. *Hippolais icterina*). Similar huge differences occurred in the
479 timing of spring migration across the Mediterranean, with early trans-Saharan migratory species
480 arriving at Ventotene as early as the beginning of April and late species as late as mid-May (Table
481 S1 and Rubolini *et al.* 2005). Our finding showed that such differences in migratory behaviour are
482 associated with interspecies differences in *Clock* gene diversity. The phylogenetic path analysis and
483 the analysis of standardized path coefficients indicated that both migration date and migration

484 distance significantly and directly affected *Clock* gene diversity. The best supported models
485 envisaged also direct effects of migration distance on migration date and migration spread, with
486 species migrating over longer distances arriving late at Ventotene and having a smaller migration
487 spread. Hence, the path analysis suggests that both migration distance and the time constraints
488 imposed by a delayed migration negatively affected *Clock* gene diversity. Late migration implies a
489 shorter breeding season, rapid nestling development and early departure for autumn migration: all
490 these traits require a relative inflexibility to external cues and tight endogenous regulation of life
491 cycle events. Hence, depleted genetic variation at *Clock*, and possibly other candidate genes as well
492 (though not *Adcyap1*), may be a consequence of such tight life cycle constraints that are ultimately
493 due to migrating over longer distances.
494
495 Concluding remarks
496
497 This study demonstrates that a comparative approach may prove extremely useful to highlight the
498 role of candidate genes in the evolution of migratory traits, and vice-versa. To our knowledge, such
499 a broad, inter-species approach has never been previously applied in studies of candidate genes
500 variation. By highlighting that species directed to northern latitudes had a higher frequency of
501 longer *Clock* alleles, we provide novel evidence for a possible role of *Clock* gene polymorphism in
502 adaptation to local photoperiodic during the breeding season. Moreover, we showed that the extent
503 of *Clock* gene diversity was related to migratory behaviour. Hence, the *Clock* gene may be of
504 pivotal evolutionary importance for long-distance migrants. The populations of these species, and in
505 particular of those migrating over longer distances, are heavily negatively affected by climate
506 changes (Sanderson *et al.* 2006; Møller *et al.* 2008; Saino *et al.* 2011), partly because they appear to
507 be constrained in their response to environmental changes by reduced phenotypic plasticity (Pulido
508 & Widmer 2005; Rubolini *et al.* 2010; Knudsen *et al.* 2011). Our results suggest that the weaker
509 response of these species to climate change may at least partly result from loss of genetic variation

at candidate genes controlling migratory behaviour. Importantly, depleted diversity at candidate genes for migration may provide a mechanisms contributing, besides environmental canalization, to the smaller phenotypic variance of migration-related traits among long-distance migrants.

Acknowledgements – We thank many people for help with fieldwork and data collection at Ventotene during the years, and especially E. Mancuso, S. Fusetti, A. Mazzoleni, M. Morganti, P. Fusi, V. Petricola and A. de Gioia. We warmly thank the Riserva Naturale Isole di Ventotene e Santo Stefano for logistic support. We also thank the many volunteers ringers and field assistants that helped collecting the data on Ventotene. Results from the Progetto Piccole Isole (INFS-ISPRA): study No. XXX (to be updated at proof stage).

References

Bazzi G, Ambrosini R, Caprioli M*, et al.* (2015) Clock gene polymorphism and scheduling of migration: a geolocator study of the barn swallow *Hirundo rustica*. *Scientific Reports* **5**, 12443.

Bazzi G, Galimberti A, Hays QR*, et al.* (2016) *Adcyap1* polymorphism covaries with breeding latitude in a Nearctic migratory songbird, the Wilson's warbler (*Cardellina pusilla*). *Ecology and Evolution*.

Bell-Pedersen D, Cassone VM, Earnest DJ*, et al.* (2005) Circadian rhythms from multiple oscillators: lessons from diverse organisms. *Nature Reviews Genetics* **6**, 544-556.

Berthold P (1995) Microevolutionary aspects of bird migration based on experimental result. *Israel Journal of Zoology* **41**, 377-385.

Berthold P (1996) *Control of bird migration* Springer Science & Business Media.

Berthold P (2001) *Bird migration: a general survey* Oxford University Press, Oxford.

BirdLife International, NatureServe (2015) Bird species distribution maps of the world, Cambridge.

Both C, Visser ME (2001) Adjustment to climate change is constrained by arrival date in a long-distance migrant bird. *Nature* **411**, 296-298.

Bourret A, Garant D (2015) Candidate gene-environment interactions and their relationships with timing of breeding in a wild bird population. *Ecology and Evolution* **5**, 3628-3641.

Burnham KP, Anderson DR (2002) *Model selection and multimodel inference: A practical information-theoretic approacapproach*, 2nd edn. Springer, New York.

Caprioli M, Ambrosini R, Boncoraglio G*, et al.* (2012) *Clock* gene variation is associated with breeding phenology and maybe under directional selection in the migratory barn swallow. *PLoS ONE* **7**, e35140.

546 Chakarov N, Jonker RM, Boerner M, Hoffman JI, Kruger O (2013) Variation at phenological
547 candidate genes correlates with timing of dispersal and plumage morph in a sedentary bird
548 of prey. *Molecular Ecology* **22**, 5430-5440.
549 Costa R, Peixoto AA, Barbujani G, Kyriacou CP (1992) A latitudinal cline in a *Drosophila* clock
550 gene. *Proceedings of the Royal Society B: Biological Sciences* **250**, 43-49.
551 Cramp S (1998) *The Complete Birds of the Western Palearctic on CD-ROM* Oxford University
552 Press, Oxford.
553 Debat V, David P (2001) Mapping phenotypes: canalization, plasticity and developmental stability.
554 *Trends in Ecology & Evolution* **16**, 555-561.
555 Dingle H (2006) Animal migration: is there a common migratory syndrome? *Journal of*
556 *Ornithology* **147**, 212-220.
557 Dor R, Lovette IJ, Safran RJ*, et al.* (2011) Low variation in the polymorphic *Clock* gene poly-Q
558 region despite population genetic structure across barn swallow (*Hirundo rustica*)
559 populations. *PLoS ONE* **6**, e28843.
560 Felsenstein J (1985) Phylogenies and the comparative method. *American Naturalist* **125**, 1-15.
561 Fidler AE, Gwinner E (2003) Comparative analysis of Avian BMAL1 and CLOCK protein
562 sequences: a search for features associated with owl nocturnal behaviour. *Comparative*
563 *Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* **136**, 861-874.
564 Fondon JW, 3rd, Hammock EA, Hannan AJ, King DG (2008) Simple sequence repeats: genetic
565 modulators of brain function and behavior. *Trends in Neurosciences* **31**, 328-334.
566 Francis CM, Cooke F (1986) Differential timing of spring migration in wood warblers (Parulinae).
567 *Auk* **103**, 548-556.
568 Freckleton RP, Harvey PH, Pagel m (2002) Phylogenetic analysis and comparative data: a test and
569 review of evidence. *American Naturalist* **160**, 712-726.
570 Garamszegi LZ, Mundry R (2014) Multimodel-inference in comparative analyses. In: *Modern*
571 *Phylogenetic Comparative Methods and Their Application in Evolutionary Biology: Theory*
572 *and Practice* (ed. Garamszegi LZ), pp. 305-331. Springer-Verlag, Berlin.
573 Gekakis N, Staknis D, Nguyen HB*, et al.* (1998) Role of the CLOCK Protein in the Mammalian
574 Circadian Mechanism. *Science* **280**, 1564-1569.
575 Gonzalez-Voyer A, von Hardenberg A (2014) An introduction to phylogenetic path analysis. In:
576 *Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology*,
577 pp. 201-229. Springer Berlin Heidelberg, Berlin.
578 Goymann W, Spina F, Ferri A, Fusani L (2010) Body fat influences departure from stopover sites in
579 migratory birds: evidence from whole-island telemetry. *Biology Letters* **6**, 478-481.
580 Grattarola A, Spina F, Pilastro A (1999) Spring migration of the garden warbler (*Sylvia borin*)
581 across the Mediterranean Sea. *Journal of Ornithology* **140**, 419-430.
582 Gwinner E (1971) A comparative study of circannual rhythms in warblers. In: *Biochronometry* (ed.
583 Menaker M), pp. 405–427. National Academy of Sciences, Washington, D. C.
584 Gwinner E (1972) Adaptive functions of circannual rhythms in warblers. In: *XVth International*
585 *Ornithological Congress*, pp. 218-236.
586 Gwinner E (1986) *Circannual rhythms* Springer-Verlag, Berlin.
587 Gwinner E (2003) Circannual rhythms in birds. *Current Opinion in Neurobiology* **13**, 770-778.
588 Gwinner E, Helm B (2003) Circannual and circadian contributions to the timing of avian migration.
589 In: *Avian migration* (eds. Berthold P, Gwinner E, Sonnenschein E), pp. 81-95. Springer-
590 Verlag, Berlin.
591 Hagan JM, Lloyd-Evans L, Alwood JL (1991) The relationship between latitude and the timing of
592 spring migration of North American landbirds. *Ornis Scandinavica* **22**, 129-136.
593 Hannibal J, Ding JM, Chen D*, et al.* (1997) Pituitary adenylate cyclase-activating peptide (PACAP)
594 in the retinohypothalamic tract: a potential daytime regulator of the biological clock.
595 *Journal of Neuroscience* **17**, 2637-2644.

Harvey Ph, Pagel MD (1991) *The comparative method in evolutionary biology* Oxford University Press, Oxford.
Hayasaka N, LaRue SI, Green CB (2002) In Vivo Disruption of Xenopus CLOCK in the Retinal Photoreceptor Cells Abolishes Circadian Melatonin Rhythmicity without Affecting Its Production Levels. *Journal of Neuroscience* **22**, 1600-1607.
Hut RA, Paolucci S, Dor R, Kyriacou CP, Daan S (2013) Latitudinal clines: an evolutionary view on biological rhythms. *Proceedings of the Royal Society B: Biological Sciences* **280**, 20130433.
Jetz W, Thomas GH, Joy JB, Hartmann K, Mooers AO (2012) The global diversity of birds in space and time. *Nature* **491**, 444-448.
Johnsen A, Fidler AE, Kuhn S*, et al.* (2007) Avian Clock gene polymorphism: evidence for a latitudinal cline in allele frequencies. *Molecular Ecology* **16**, 4867-4880.
Kissner KJ, Weatherhead PJ, Francis CM (2003) Sexual size dimorphism and timing of spring migration in birds. *Journal of Evolutionary Biology* **16**, 154-162.
Knudsen E, Linden A, Both C*, et al.* (2011) Challenging claims in the study of migratory birds and climate change. *Biological Reviews* **86**, 928-946.
Ko CH, Takahashi JS (2006) Molecular components of the mammalian circadian clock. *Human Molecular Genetics* **15 Spec No 2**, R271-277.
Kuhn K, Schwenk K, Both C*, et al.* (2013) Differentiation in neutral genes and a candidate gene in the pied flycatcher: using biological archives to track global climate change. *Ecology and Evolution* **3**, 4799-4814.
Liedvogel M, Sheldon BC (2010) Low variability and absence of phenotypic correlates of *Clock* gene variation in a great tit *Parus major* population. *Journal of Avian Biology* **41**, 543-550.
Liedvogel M, Szulkin M, Knowles SC, Wood MJ, Sheldon BC (2009) Phenotypic correlates of *Clock* gene variation in a wild blue tit population: evidence for a role in seasonal timing of reproduction. *Molecular Ecology* **18**, 2444-2456.
Lincoln GA, Andersson H, Loudon A (2003) Clock genes in calendar cells as the basis of annual timekeeping in mammals - a unifying hypothesis. *Journal of Endocrinology* **179**, 1-13.
Merilä J (1999) Genetic architecture of fitness and nonfitness traits: empirical patterns and development of ideas. *Heredity* **83**, 103.
Merrow M, Spoelstra K, Roenneberg T (2005) The circadian cycle: daily rhythms from behaviour to genes. *EMBO Reports* **6**, 930-935.
Messineo A, Grattarola A, Spina F (2001) Dieci anni di Progetto Piccole Isole. *Biologia e Conservazione della Fauna* **106**, 1-244.
Milá B, Smith TB, Wayne RK (2006) Postglacial Population Expansion Drives the Evolution of Long-Distance Migration in a Songbird. *Evolution* **60**, 2403-2409.
Møller AP, Rubolini D, Lehikoinen E (2008) Populations of migratory bird species that did not show a phenological response to climate change are declining. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 16195-16200.
Mueller JC, Pulido F, Kempenaers B (2011) Identification of a gene associated with avian migratory behaviour. *Proceedings of the Royal Society B: Biological Sciences* **278**, 2848-2856.
Nagy AD, Csernus VJ (2007) The role of PACAP in the control of circadian expression of clock genes in the chicken pineal gland. *Peptides* **28**, 1767-1774.
Nei M, Kumar s (2000) *Molecular evolution and phylogenetics* Oford University Press, Oxford.
Nei M, Roychoudhury AK (1974) Genic variation within and between the three major races of man, Caucasoids, Negroids, and Mongoloids. *American Journal of Human Genetics* **26**, 421-443.
O'Malley KG, Ford MJ, Hard JJ (2010) *Clock* polymorphism in Pacific salmon: evidence for variable selection along a latitudinal gradient. *Proceedings of the Royal Society B: Biological Sciences* **277**, 3703-3714.

Orme CDL, Freckleton RP, Thomas GH, Petzoldt T, Fritz SA (2011) Caper: Comparative Analyses of Phylogenetics and Evolution in R. R Package Version 0.5.2. (http://cran.r-project.org/web/packages/caper/index.html).
Panda S, Hogenesch JB, Kay SA (2002) Circadian rhythms from flies to human. *Nature* **417**, 329-335.
Peterson MP, Abolins-Abols M, Atwell JW*, et al.* (2013) Variation in candidate genes CLOCK and ADCYAP1 does not consistently predict differences in migratory behavior in the songbird genus Junco. *F1000Research* **2**, 115.
Pilastro A, Baccetti N, Massi A*, et al.* (1995) Direction of migration and fat consumption rate estimates in spring migrating garden warblers (*Sylvia borin*). *Ric Biol Selvag* **12**, 435-445.
Pilastro A, Macchio S, Massi A, Montemaggiori A, Spina F (1998) Spring migratory routes of eight trans-Saharan passerines through the central and western Mediterranean; results from a network of insular and coastal ringing sites. *Ibis* **140**, 591-598.
Pilastro A, Spina F (1997) Ecological and morphological correlates of residual fat reserves in passerine migrants at their spring arrival in southern Europe. *Journal of Avian Biology* **28**, 309-318.
Pulido F (2007) The Genetics and Evolution of Avian Migration. *Bioscience* **57**, 165.
Pulido F, Widmer M (2005) Are long-distance migrants constrained in their evolutionary response to environmental change? Causes of variation in the timing of autumn migration in a blackcap (*S. atricapilla*) and two garden warbler (*Sylvia borin*) populations. *Annals of the New York Academy of Sciences* **1046**, 228-241.
Racz B, Horvath G, Faluhelyi N*, et al.* (2008) Effects of PACAP on the circadian changes of signaling pathways in chicken pinealocytes. *Journal of Neuroscience* **36**, 220-226.
Robson D, Barriocanal C (2011) Ecological conditions in wintering and passage areas as determinants of timing of spring migration in trans-Saharan migratory birds. *Journal of Animal Ecology* **80**, 320-331.
Ross CA (2002) Polyglutamine Pathogenesis. *Neuron* **35**, 819-822.
Rubolini D, Liker A, Garamszegi LZ, Møller AP, Saino N (2015) Using the BirdTree.org website to obtain robust phylogenies for avian comparative studies: A primer *Current Zoology* **61**, 959-965.
Rubolini D, Saino N, Møller AP (2010) Migratory behaviour constrains the phenological response of birds to climate change. *Climate Research* **42**, 45-55.
Rubolini D, Spina F, Saino N (2004) Protandry and sexual dimorphism in trans-Saharan migratory birds. *Behavioral Ecology* **15**, 592-601.
Rubolini D, Spina F, Saino N (2005) Correlates of timing of spring migration in birds: a comparative study of trans-Saharan migrants. *Biological Journal of the Linnean Society* **85**, 199-210.
Saino N, Ambrosini R (2007) Climatic connectivity between Africa and Europe may serve as a basis for phenotypic adjustment of migration schedules of trans-Saharan migratory birds. *Global Change Biology* **14**, 250-263.
Saino N, Ambrosini R, Rubolini D*, et al.* (2011) Climate warming, ecological mismatch at arrival and population decline in migratory birds. *Proceedings of the Royal Society B: Biological Sciences* **278**, 835-842.
Saino N, Bazzi G, Gatti E*, et al.* (2015) Polymorphism at the Clock gene predicts phenology of long-distance migration in birds. *Molecular Ecology* **24**, 1758-1773.
Saino N, Romano M, Caprioli M*, et al.* (2013) Timing of molt of barn swallows is delayed in a rare *Clock* genotype. *PeerJ* **1**, e17.
Saino N, Rubolini D, Jonzén N*, et al.* (2007) Temperature and rainfall anomalies in Africa predict timing of spring migration in trans-Saharan migratory birds. *Climate Research* **35**, 123-134.

Saino N, Rubolini D, Serra L*, et al.* (2010a) Sex-related variation in migration phenology in relation to sexual dimorphism: a test of competing hypotheses for the evolution of protandry. *Journal of Evolutionary Biology* **23**, 2054-2065.
Saino N, Rubolini D, von Hardenberg J*, et al.* (2010b) Spring migration decisions in relation to weather are predicted by wing morphology among trans-Mediterranean migratory birds. *Functional Ecology* **24**, 658-669.
Salewski V, Bruderer B (2007) The evolution of bird migration--a synthesis. *Naturwissenschaften* **94**, 268-279.
Sanderson FJ, Donald PF, Pain DJ, Burfield IJ, van Bommel FPJ (2006) Long-term population declines in Afro-Palearctic migrant birds. *Biological Conservation* **131**, 93-105.
Schwartz C, Andrews MT (2013) Circannual Transitions in Gene Expression: Lessons from Seasonal Adaptations. In: *Developmental timing* (eds. Rougvie AE, O'Connor MB). Academic Press.
Sharp PJ (2005) Photoperiodic regulation of seasonal breeding in birds. *Annals of the New York Academy of Sciences* **1040**, 189-199.
Shipley B (2000) *Cause and correlation in biology: a user's guide to path analysis, structural equations and causal inference* Cambridge University Press, Cambridge, UK.
Simonneaux V, Ouichou A, Pévet P (1993) Pituitary adenylate cyclase-activating polypeptide (PACAP) stimulates melatonin synthesis from rat pineal gland. *Brain Research* **603**, 148-152.
Spina F, Massi A, Montemaggiori A, Baccetti N (1993) Spring Migration Across Central Mediterranean: General Results from the "Progetto Piccole Isole". *Vogelwarte* **37**, 1-94.
Spina F, Volponi S (2008a) *Atlante della migrazione degli uccelli in Italia. Vol. I non-passeriformi.* ISPRA - MATTM, Roma.
Spina F, Volponi S (2008b) *Atlante della migrazione degli uccelli in Italia. Vol. II passeriformi* ISPRA - MATTR.
Sukumaran J, Holder MT (2010) DendroPy: a Python library for phylogenetic computing. *Bioinformatics* **26**, 1569-1571.
Symonds MRE, Moussalli A (2011) A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion. *Behavioral Ecology and Sociobiology* **65**, 13-21.
Tenan S, Spina F (2010) Timing and condition-related effects on recapture probability, mass change and stopover length of spring migrating songbirds on a small mediterranean island. *Ardeola* **57**, 121-132.
van Wijk RE, Schaub M, Tolkmitt D*, et al.* (2013) Short-distance migration of Wrynecks *Jynx torquilla* from Central European populations. *Ibis* **155**, 886-890.
Visser ME, Caro SP, van Oers K, Schaper SV, Helm B (2010) Phenology, seasonal timing and circannual rhythms: towards a unified framework. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**, 3113-3127.
von Hardenberg A, Gonzalez-Voyer A (2013) Disentangling evolutionary cause-effect relationships with phylogenetic confirmatory path analysis. *Evolution* **67**, 378-387.

Data Accessibility

The data on which the present study is based are all included in the Supporting Information.

Author contributions

GB, NS, and DR designed the study. GB, DR, CDP, NS, FS, and JGC collected the data. GB, MC, SP, CDP, LG and EG performed the lab work. DR, GB and RA analyzed the data. DR and GB wrote the paper.

749 **Table 1.** PGLS regression models testing the effect of *Clock* allele size (species-specific mean of
750 mean or long allele size) on migration date (above) and the effect of breeding latitude on *Clock*
751 allele size (below) of 23 trans-Saharan migratory bird species. Model-averaged parameters and their
752 95% CIs (accounting for phylogenetic uncertainty) are also shown (see Methods).
753
754
755

Trait	**Estimate (SE)**	***t***	***P***	**Model-averaged parameter [95% CIs]**
Migration date vs. Clock *allele size*				
Mean allele size	-0.19 (0.34)	-0.55	0.59	-0.08 [-0.69; 0.53]
Long allele size	-0.12 (0.32)	-0.37	0.72	-0.02 [-0.60; 0.54]
Clock *allele size vs. breeding latitude*				
Mean allele size	0.23 (0.14)	1.61	0.12	0.22 [-0.05; 0.50]
Long allele size	0.42 (0.17)	2.55	0.018	0.43 [0.10; 0.76]

756
757
758

Table 2. Univariate PGLS regression models testing the effect of distribution and phenology on *Clock* and *Adcyap1* gene diversity of 23 trans-Saharan migratory bird species. Migration distance is expressed in km × 1000. Model-averaged parameters and their 95% CIs (accounting for phylogenetic uncertainty) are also shown (see Methods).

Trait	**Estimate (SE)**	***t***	***P***	**Model-averaged parameter [95% CIs]**
Clock *gene diversity*				
Breeding range width	-0.006 (0.005)	-1.11	0.28	-0.006 [-0.015; 0.004]
Wintering range width	-0.003 (0.003)	-1.02	0.32	-0.003 [-0.008; 0.003]
Migration distance	-0.086 (0.033)	-2.62	0.016	-0.081 [-0.145; -0.016]
Migration date	-0.011 (0.004)	-2.60	0.017	-0.010 [-0.018; -0.002]
Adcyap1 *gene diversity*				
Breeding range width	0.008 (0.006)	1.46	0.17	0.008 [-0.003; 0.020]
Wintering range width	0.001 (0.003)	0.48	0.64	0.001 [-0.004; 0.007]
Migration distance	0.004 (0.036)	0.11	0.91	0.004 [-0.067; 0.075]
Migration date	0.001 (0.004)	0.22	0.83	-0.001 [-0.007; 0.009]

768 **Table 3.** List of the 12 hypothesized causal models (see Fig. 1 for graphical representation) and
769 their statistics. Fisher's *C* statistic is reported together with its associated degrees of freedom, p-
770 value and CIC value. Models are sorted according to their CIC values (lowest values represent the
771 best-fitting models). The best models (i.e. those with ΔCIC < 2) are highlighted in boldface.
772

773
774
775

Model	*C*	df	*P*	CICc	ΔCIC
M03	**5.18**	**6**	**0.52**	**26.64**	**0.00**
M02	**5.93**	**6**	**0.43**	**27.39**	**0.75**
M05	3.26	4	0.52	29.55	2.90
M06	3.61	4	0.46	29.90	3.25
M07	3.68	4	0.45	29.97	3.32
M11	10.80	6	0.09	32.26	5.62
M08	1.01	2	0.60	32.86	6.22
M04	11.95	6	0.06	33.41	6.77
M09	9.53	4	0.049	35.82	9.18
M12	9.60	4	0.048	35.89	9.25
M10	17.87	6	0.007	39.34	12.69
M01	25.72	6	< 0.001	47.19	20.55

776
777

778 **Table 4.** Standardized path coefficients (with lower, LCL, and upper, UCL, 95% CI) from the two
779 best-fitting models identified in Table 3 (M03 and M02; see Fig. 1).

780

Path	Coefficient (SE)	LCL	UCL
Direct causal effects on H			
D → H	-0.43 (0.17)	-0.76	-0.11
M → H	-0.39 (0.15)	-0.68	-0.10
Other direct causal paths			
M → D	0.59 (0.18)	0.25	0.94
D → S	-0.59 (0.18)	-0.93	-0.24

781

782

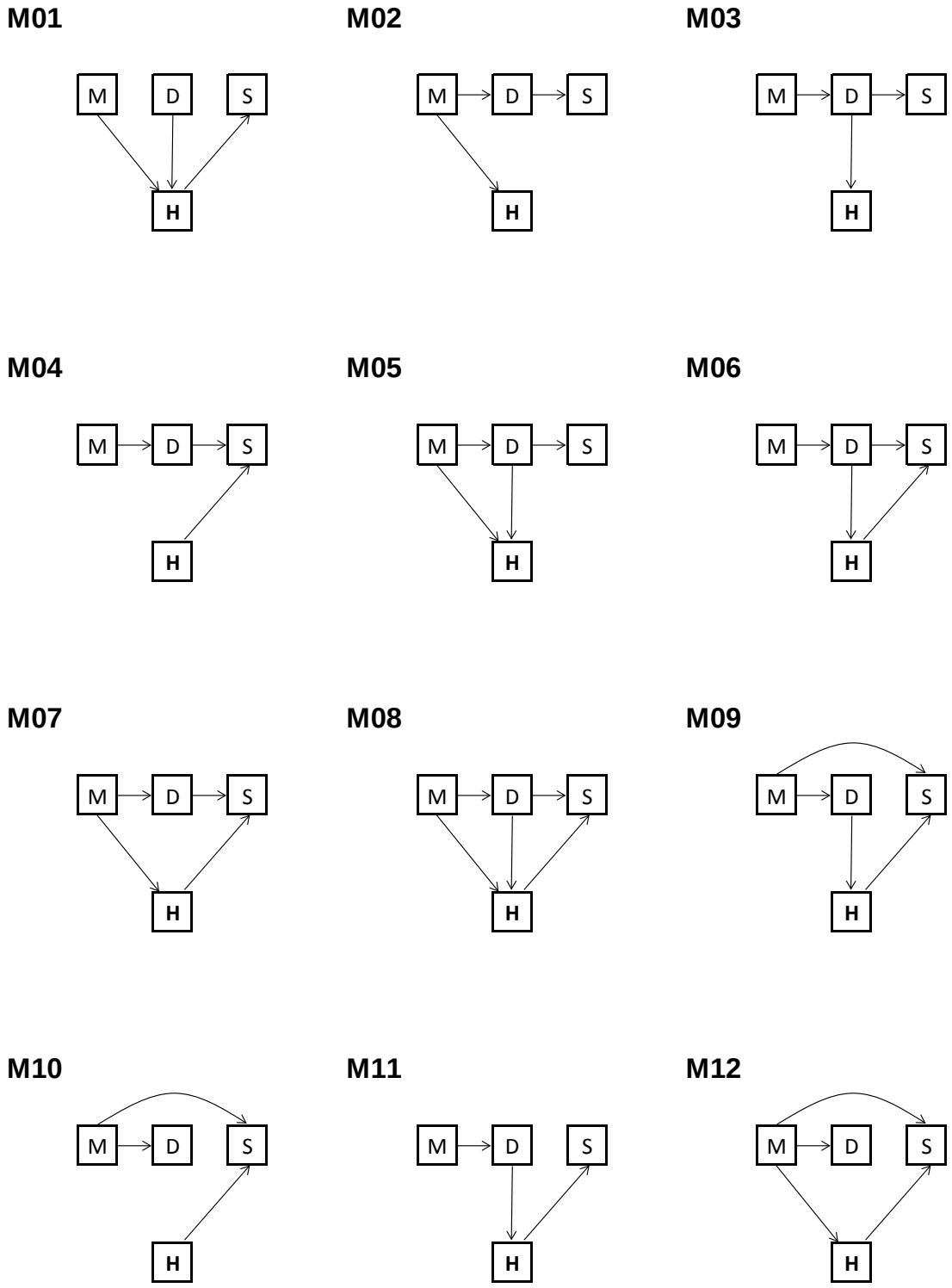
783

784 **Figure 1.** Directed acyclic graphs (DAG) representing causal hypotheses of the relationships
785 between *Clock* gene diversity (H), migration distance (M), migration date (D) and spread of
786 migration dates (S).

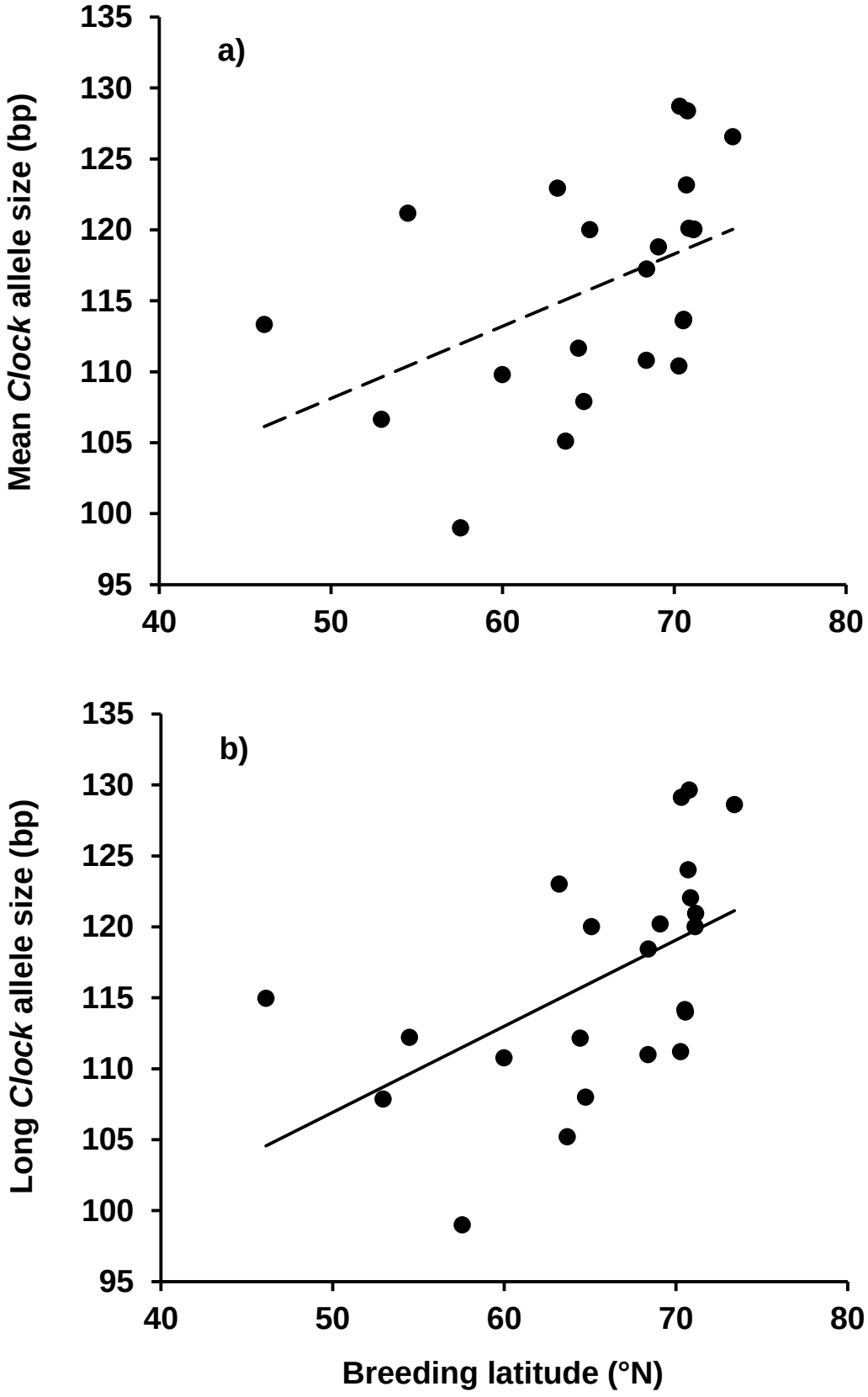
787

788

789



790 **Figure 2.** Mean (a) and long (b) *Clock* allele size (species-specific mean values) vs. breeding
791 latitude. The lines represent phylogenetically regression equations from PGLS models reported in
792 Table 1. The broken line in a) denotes non-significant regression line (see Table 1 for statistics).
793
794



795 **Figure 3.** *Clock* gene diversity vs. a) migration distance and b) migration date (1 = January 1), and
796 migration spread vs. c) *Clock* gene diversity. The lines represent phylogenetically corrected
797 regression equations from PGLS models.
798

799

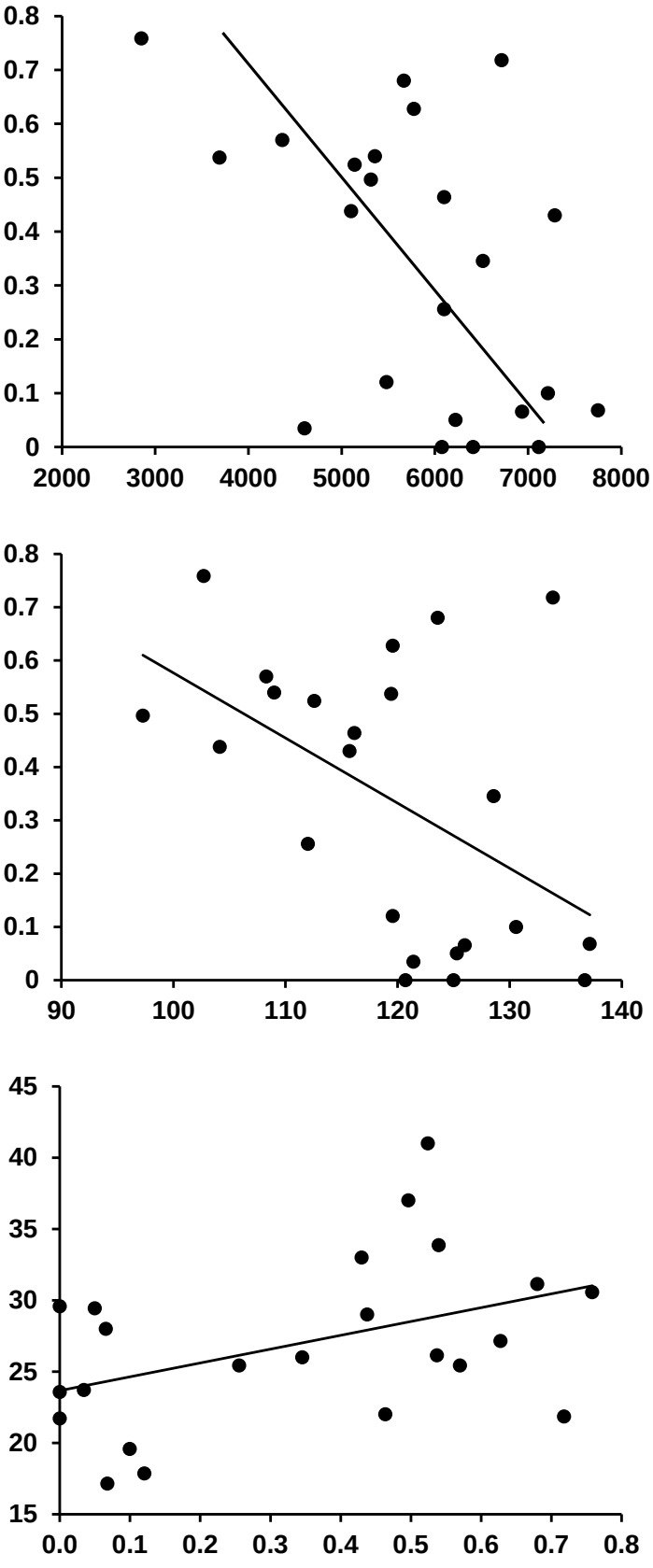


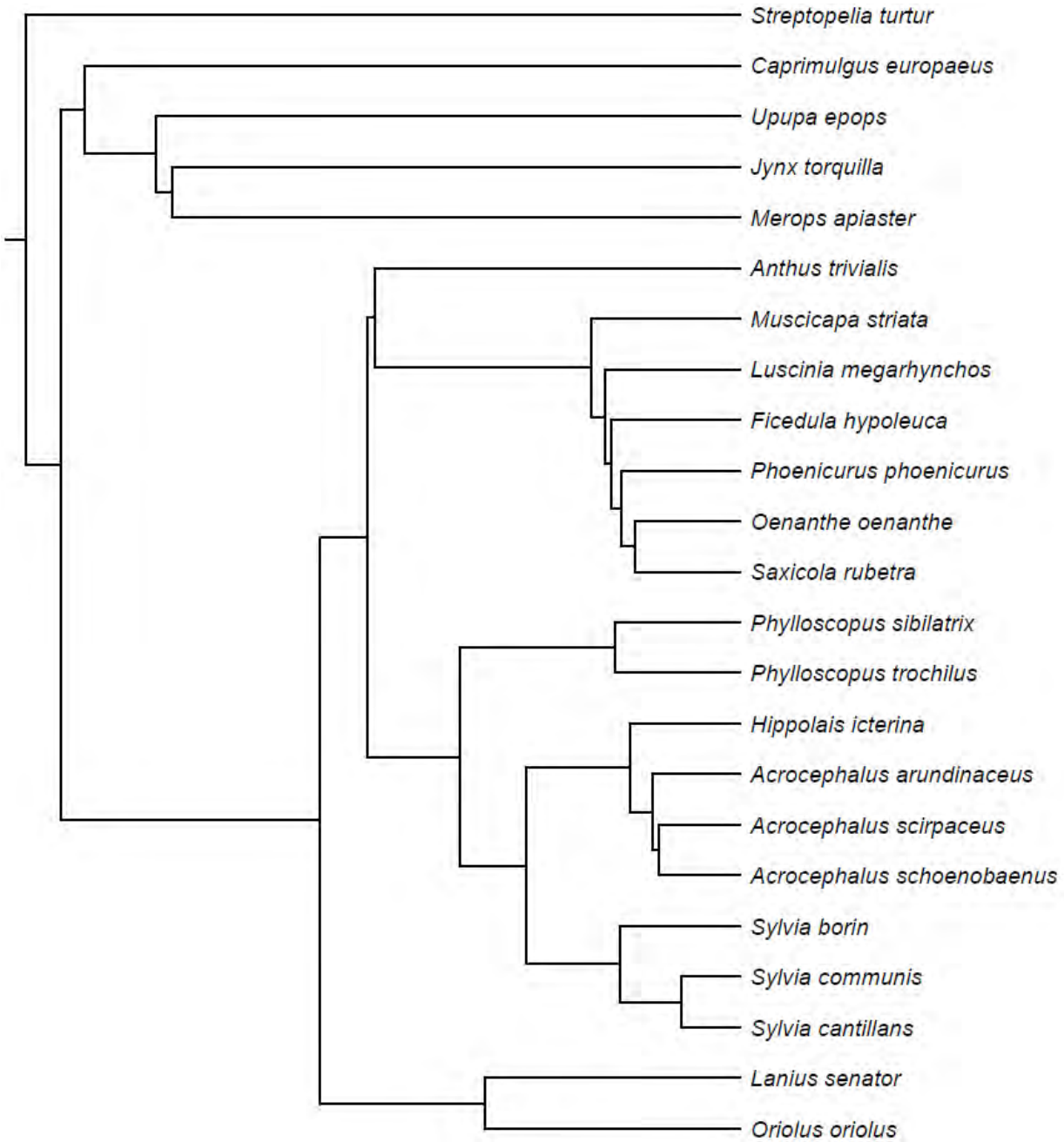
Table S1. List of the 23 trans-Saharan migratory bird species, information on sample size ($N_{2006-2012}$: mean number of birds captured at Ventotene during 2006-2012) and life-history traits considered in the analyses (see Methods): migration date (1 = January 1; mean value of yearly median migration dates for each species); migration spread (mean value of the yearly difference of the 90th and the 10th percentile of migration dates); breeding latitude (northernmost breeding latitude); breeding and wintering range width (differences of the northernmost and the southernmost breeding and wintering latitudes, respectively). Population size (from the BirdLife International website www.birdlife.org) is also shown. Species are sorted in alphabetical order.

Species	Migration date	Migration spread (d)	$N_{2006-2012}$	Migration distance (km)	Breeding latitude (°N)	Breeding range width (°)	Wintering range width (°)	Population size (pairs)
Acrocephalus arundinaceus	125.29	29.43	59	6223	63.20	22.72	49.90	3640000
Acrocephalus schoenobaenus	125.00	29.57	217	6411	71.12	30.64	51.81	5135000
Acrocephalus scirpaceus	136.71	23.57	127	6078	65.09	24.61	53.70	3000000
Anthus trivialis	112.00	25.43	289	6100	70.53	30.05	42.42	32500000
Caprimulgus europaeus	128.57	26.00	23	6516	64.42	23.94	53.09	857000
Ficedula hypoleuca	116.14	22.00	977	6100	70.72	30.24	18.91	15950000
Hippolais icterina	137.14	17.14	2243	7751	68.38	26.57	33.08	5110000
Jynx torquilla	104.14	29.00	83	5102	70.27	29.79	43.38	1137000
Lanius senator	119.43	26.14	71	3691	52.94	12.46	23.61	2520000
Luscinia megarhynchos	108.29	25.43	321	4364	54.49	14.01	23.85	15500000
Merops apiaster	120.71	21.71	65	7115	57.55	17.07	45.15	3925000
Muscicapa striata	130.57	19.57	796	7215	70.56	30.08	49.98	18800000
Oenanthe oenanthe	109.00	33.86	104	5357	73.40	32.92	54.54	10540000
Oriolus oriolus	126.00	28.00	150	6938	63.67	23.19	42.82	6195000
Phoenicurus phoenicurus	112.57	41.00	420	5140	70.77	30.29	39.50	12315000
Phylloscopus sibilatrix	119.57	27.14	1304	5775	69.08	28.60	15.27	9080000
Phylloscopus trochilus	115.71	33.00	1616	7289	71.16	29.78	50.28	79650000
Saxicola rubetra	119.57	17.86	851	5483	70.31	29.83	30.94	8585000
Streptopelia turtur	121.43	23.71	239	4604	64.73	24.25	13.43	4545000
Sylvia borin	133.86	21.86	3957	6716	70.85	30.37	46.24	21800000
Sylvia cantillans	102.71	30.57	1876	2853	46.12	5.64	21.78	4375000
Sylvia communis	123.57	31.14	1652	5669	68.40	27.92	47.83	22550000
Upupa epops	97.29	37.00	55	5314	59.99	19.51	76.01	2030000

Table S2. List of the 23 trans-Saharan migratory bird species and genetic variables (species-specific mean of mean or long *Clock* allele sizes, in bp; *Clock* and *Adcyap1* gene diversity) considered in the analyses (see Methods). Species are sorted in alphabetic order. *Clock* (N_{Clock}) and *Adcyap1* ($N_{Adcyap1}$) sample sizes are also shown.

Species	**Mean *Clock* allele size (bp)**	**Long *Clock* allele size (bp)**	***Clock* gene diversity**	N_{Clock}	***Adcyap1* gene diversity**	$N_{Adcyap1}$
Acrocephalus arundinaceus	122.93	123.00	0.10	20	0.32	21
Acrocephalus schoenobaenus	120.00	120.00	0	30	0.59	30
Acrocephalus scirpaceus	120.00	120.00	0	24	0.94	24
Anthus trivialis	113.59	114.16	0.42	153	0.94	97
Caprimulgus europaeus	111.65	112.15	0.51	39	0.91	39
Ficedula hypoleuca	123.16	124.01	0.67	226	0.87	95
Hippolais icterina	110.79	111.00	0.13	29	0.83	29
Jynx torquilla	110.40	111.20	0.64	30	0.84	31
Lanius senator	106.65	107.85	0.75	20	0.88	18
Luscinia megarhynchos	121.16	112.20	0.72	162	0.69	99
Merops apiaster	99.00	99.00	0	35	0.35	36
Muscicapa striata	113.69	114.00	0.19	29	0.85	29
Oenanthe oenanthe	126.55	128.60	0.71	30	0.80	30
Oriolus oriolus	105.10	105.20	0.13	30	0.80	30
Phoenicurus phoenicurus	128.37	129.63	0.74	43	0.90	35
Phylloscopus sibilatrix	118.80	120.20	0.83	30	0.73	29
Phylloscopus trochilus	120.04	120.94	0.68	121	0.94	112
Saxicola rubetra	128.71	129.12	0.23	208	0.92	169
Streptopelia turtur	107.90	108.00	0.07	29	0.45	30
Sylvia borin	120.10	122.03	0.83	31	0.86	31
Sylvia cantillans	113.32	114.97	0.87	31	0.86	22
Sylvia communis	117.24	118.44	0.82	25	0.93	26
Upupa epops	109.80	110.76	0.64	25	0.73	26

Figure S1. Phylogenetic hypothesis used for the comparative analyses. The tree is a 50% majority-rule consensus tree (see Methods for details) obtained from 2000 random trees of the 'Ericson' phylogeny, downloaded from the BirdTree.org website (www.birdtree.org).



Appendix

List of publications in ISI-ranked journals by Gaia Bazzi

The list includes all the publications that I have co-authored during my PhD.

7. **Bazzi G**, Galimberti A, Hays QR, Bruni I, Cecere JG, Gianfranceschi L, Hobson KA, Morbey YE, Saino N, Guglielmo CG, Rubolini D (2016) *Adcyap1* polymorphism covaries with breeding latitude in a Nearctic migratory songbird, the Wilson's warbler (*Cardellina pusilla*). *Ecology and Evolution*.

6. Romano A, **Bazzi G**, Caprioli M, Corti M, Costanzo A, Rubolini D, Saino N (2016) Nestling sex and plumage color predict food allocation by barn swallow parents. *Behavioral Ecology*, arw040.

5. **Bazzi G**, Ambrosini R, Caprioli M, Costanzo A, Liechti F, Gatti E, Gianfranceschi L, Podofillini S, Romano A, Romano M, Scandolara C, Saino N, Rubolini D (2015) *Clock* gene polymorphism and scheduling of migration: a geolocator study of the barn swallow *Hirundo rustica*. *Scientific Reports* **5**, 12443.

4. Saino N, Romano M, Romano A, Rubolini D, Ambrosini R, Caprioli M, Parolini M, Scandolara C, **Bazzi G**, Costanzo A (2015) White tail spots in breeding Barn Swallows *Hirundo rustica* signal body condition during winter moult. *Ibis* **157**, 722-730.

3. **Bazzi G**, Foglini C, Brambilla M, Saino N, Rubolini D. (2014) Habitat management effects on Prealpine grassland bird communities. *Italian Journal of Zoology* **82**, 251-261.

2. Saino N, **Bazzi G**, Gatti E, Caprioli M, Cecere JG, Possenti CD, Galimberti A, Orioli V, Bani L, Rubolini D, Gianfranceschi L, Spina F (2015) Polymorphism at the *Clock* gene predicts phenology of long-distance migration in birds. *Molecular Ecology* **24**, 1758-1773.

1. Saino N, Romano M, Rubolini D, Ambrosini R, Romano A, Caprioli M, Costanzo A, **Bazzi G** (2014) A trade-off between reproduction and feather growth in the barn swallow (*Hirundo rustica*). *PLoS ONE* **9**, e96428.

Photo: Museo Civico Lentate sul Seveso