Candidate genes for migration do not distinguish migratory and non-migratory birds

Juan S. Lugo Ramos · Kira E. Delmore · Miriam Liedvogel

Abstract Migratory traits in birds have been shown to have a strong heritable component and several candidate genes have been suggested to control these migratory traits. To investigate if the genetic makeup of one or a set of these candidate genes can be used to identify a general pattern between migratory and non-migratory birds, we extracted genomic sequence data for 25 hypothesised candidate genes for migration from 70 available genomes across all orders of Aves and characterised sequence divergence between migratory and non-migratory phenotypes. When examining each gene separately across all species, we did not identify any genetic variants in candidate genes that distinguished migrants from non-migrants; any resulting pattern was driven by the phylogenetic signal. This was true for each gene analysed independently, but also for concatenated sequence alignments of all candidate genes combined. We also attempted to distinguish between migrant and non-migrants using structural features at four candidate genes that have previously been reported to show association with migratory behaviour but did not pick up a signal for migratory phenotype here either. Finally, a screen for dN/dS ratio across all focal candidate genes to probe for putative features of selection did not uncover a pattern, though this might not be expected given the broad phylogenetic scale used here. Our study demonstrates the potential of public genomic data to test for general patterns of migratory gene candidates in a cross-species comparative context, and raise questions on the applicability of candidate gene approaches in a macro-evolutionary context to understand the genetic architecture of migratory behaviour.

Keywords Migratory traits · Bird migration · Candidate genes · Next-generation sequencing · Genomic data
on genetic variation in specific candidate regions of the genome that have been suggested to directly impact the function of the candidate gene and ultimately the target phenotype in other species as well. In the context of migration, the traits receiving most attention are circadian behaviour and personality traits (e.g. exploratory or anxiety-related behaviour); focal candidate genes include \textit{CLOCK}, \textit{ADCYAP1}, \textit{CREB1} and \textit{NPAS2} for circadian rhythm, and \textit{DRD4} and \textit{SERT} for personality traits (see Müller et al. 2011 as one of the pioneer studies in the context of migration).

\textit{CLOCK} and \textit{ADCYAP1} are among the most studied candidate genes in the field of bird migration. The overall pattern for \textit{CLOCK} variability indicates a latitudinal cline in repeat lengths at the variable region, possibly reflecting local adaptation to seasonal variation at different latitudes. The circadian \textit{CLOCK} gene is highly conserved among birds throughout most of its sequence with the exception of one C-terminal region that contains a variable poly-glutamine(Q) (poly-Q) motif with variability in the number of glutamine repeats both among and within species (e.g. Johnsen et al. 2007; Liedvogel et al. 2009; Bazzi et al. 2016; but also see Liedvogel and Sheldon 2010; Dor et al. 2011). This region has been suggested to influence the transcription activating potential of the protein, potentially altering rhythms in both physiology and behaviour. The aforementioned pattern of a latitudinal cline in lengths polymorphism has been recovered in several species and in a migration context has been suggested to reflect adaptive features related to migration. Similarly for \textit{ADCYAP1}, a neuropeptide-coding gene, one 3′ UTR microsatellite length variation has shown associations with migratory behaviour in blackcaps (Müller et al. 2011; Mettler et al. 2015) with longer alleles associated with higher migratory activity. However, the pattern so far lacks consistency across other avian species (e.g. Peterson et al. 2013) in order to confirm its suggested role as regulatory unit of migratory behaviour.

More recently, Delmore et al. (2015a) used next-generation sequencing data to estimate genomic differentiation between two subspecies groups of \textit{Catharus ustulatus} that exhibit differences in their migratory behaviour (both timing and orientation). They characterised the entire genome to probe for enrichment of candidate genes for migration in areas of elevated differentiation across the genome. The resulting list of candidates included 25 genes that had been identified using a literature search (Ruegg et al. 2014). As predicted, genes from the list of candidates were enriched in areas of elevated differentiation, suggesting selection around these candidates could not only contribute to variation documented in their migratory behaviour, but also help maintain differences between the groups. Delmore et al. (2016) expanded on this work using hybrids and a genome-wide association study to identify one region in particular that is associated with variation in migratory orientation and harbours the \textit{CLOCK} gene.

Despite several clearly species-specific differences in behavioural traits that make up the migratory phenotype, all migratory birds have to meet a general consensus schedule of key adaptations in order to cope with the challenge of migration. Specifically, migratory birds need a set of navigation and orientation mechanisms to migrate successfully, and must complete several key physiological traits such as hyperphagia and feather moulting in time prior to their migratory journey. Migratory individuals also share common morphological and anatomical adaptations, such as elongated and more pointed wings, lighter bone structure, maybe also brain volume (e.g. Lockwood et al. 1998; Fuchs et al. 2014). Given this common set of features among migratory birds and demonstrations of both their heritability and potentially shared genetic basis (e.g. the aforementioned involvement of \textit{CLOCK} and \textit{ADCYAP1} across species), it has been postulated that there may be a shared set of genes for migration among birds (Berthold 1999).

Here, we test the hypothesis that sequence variation at one or a common set of suggested candidate genes for migration has been exploited to adapt to the requirements of migratory behaviour in different avian lineages using a cross-species comparative approach. Specifically, we focused on each of the 25 candidate genes for migration proposed by Ruegg et al. (2014) and compared patterns of genetic variability separately for each candidate gene between migratory and non-migratory species of birds. In addition to analyses per candidate gene, we also analysed concatenated sequence data from all candidate genes. This work benefited from the wealth of genomic data that have been accumulated in the last half-decade, with the assembly of several draft reference genomes for birds and large-scale initiative of the Avian phylogenomics project (http://avian.genomics.cn/en/index.html; more recently expanding to the B10K project). Our full dataset included 17 obligate migratory species, 32 sedentary or non-migratory species and 21 species with an additional intermediate movement phenotype (e.g. dispersive, partial migrant) (Table S1). We compare patterns of evolutionary divergence of each candidate gene using three different approaches: (1) comparing observed topologies for candidate genes to trees built using phylogenetic relationships with and without distinguishing migratory species from non-migratory species; (2) performing a gene-wide and branch-specific dN/dS analysis to identify if selective pressures on these candidate genes play a role related to migration; and (3) focussing on structural features that previous studies have shown to correlate with migratory behaviour (e.g. microsatellites at \textit{ADCYAP1} and \textit{CREB1} 3′ UTR regions and poly-Q regions in \textit{NPAS2}).
Materials and methods

Migratory phenotype characterisation

We classified each of the 70 species included in our study, according to their migratory phenotype. Our classification was based on a careful literature review of bird guides (Svensson et al. 1999), as well as BirdLife (http://birdlife.org) and Handbook of birds of the World (HBW) (http://www.hbw.com). We defined the following categories: clearly non-migratory (resident, sedentary) (0/R; \( n = 32 \)), obligate migrant (2/M; \( n = 17 \)). However, sometimes it is not easy to clearly define a species as either clearly non-migrant or obligate migrant, this is especially true when a migratory trait segregates within a population such as in partial migrants where only some individuals of the population migrate, consequently we added a third category \( n = 21 \). This category includes partial migratory species (i.e. not all individuals of the population migrate), and species that exhibit other kind of migration-independent movement behaviour (e.g. dispersal, homing, foraging flights). For partial migrant species we used additional information of the individual used for generating the reference, such as date and geographical origin of sample collection, in order to clearly define migratory phenotype and grouped that individual/species accordingly whenever possible.

Genome sequences, extraction and alignment

We downloaded genome sequences and annotations for most of the species used in our study from the NCBI database (Supplementary Table S1). We further included genome sequences of five additional migratory species here: Siberian stonechat Saxicola maurus (Van Doren et al. 2017), Swainson’s thrush Catharus ustulatus (Delmore et al. 2015a), European blackcap Sylvia atricapilla (Delmore et al., in preparation), Willow warbler Phylloscopus trochilus (Lundberg et al. 2017, accepted), and Greenish warbler Phylloscopus trochiloides (Irwin et al. 2016).

Once we had sequences data and annotation for all of the species, we used the Bedtools (Quinlan and Hall 2010) getfasta module to extract genomic sequences for each of the 25 candidate genes for every species. Sequences for unpublished genomes or genomes without annotations (for details see Table S1) were generated using Blastn and chicken cDNAs from the Ensembl database as a reference. All genomic sequences were aligned with MAFFT (Katoh and Standley 2013) and manually edited in AliView. Coding (CDS) sequences were also obtained from a multiple alignment of the genomic sequences and Ensemble cDNA sequences (including untranslated regions, UTRs) for the flycatcher, chicken, and Zebra Finch. Only sequences covering 50% or more of the chicken genes were considered for further analysis.

Topological analyses

Evolutionary trees were constructed for each candidate gene using whole genomic sequences and cDNA as reference, using a Neighbour Joining approach in MEGA v5.2 (Tamura et al. 2011). The reliability of the trees was evaluated performing a bootstrap analysis of 1000 replicates with the Kimura 2 Parameters model. To visualise the trees we used Figtree (http://tree.bio.ed.ac.uk/).

We compared the pattern of evolutionary divergence of these gene trees with three different hypothetical scenarios: divergence driven by phylogeny (‘phylogenetic topology’); divergence constrained by migration, i.e. different migratory phenotypes clustering in separate branches, while keeping the evolutionary relationship of the phylogenetic topology within each branch, (‘migratory phenotype topology’); and random divergence (‘random topology’). These comparisons were carried out for (a) the full dataset including all three migratory phenotypes; (b) a restricted dataset only contrasting exclusively obligatory migratory and completely non-migratory (resident) species, and (c) a clade-specific analysis exclusively focusing on the genus of Passeriformes, as this is the only monophyletic clade in our dataset with a sufficiently high number of species for both obligate migrants and non-migratory species, thus allowing for a more fine-tuned assessment on a narrower phylogenetic scale. The clade-specific subset allows us to test if the migratory phenotype might be controlled by a different clade-specific subset of genes. This comparative approach allows us to identify the presence or absence of general patterns, using genetic variation at candidate genes to distinguish between patterns related to phylogenetic relationships and migratory behaviour. The divergence driven by phylogeny (i.e. the gene trees matching the species tree, ‘phylogenetic topology’) was constructed using
the total evidence nucleotide species tree (TENT) phylogeny, published by Jarvis et al. (2014). For species not included in the TENT phylogeny we used timetree (Hedges et al. 2015) divergence times to position these species in our phylogeny. The divergence constrained by migration scenario (‘migratory phenotype topology’) was constructed by clustering each phenotype (once exclusively focusing on migratory versus resident species for the restricted dataset; and also for the full dataset including other movement as a third phenotype category) in one separate branch while keeping the evolutionary relationships of the phylogenetic topology within each branch. Random divergence (‘random tree’) was generated shuffling branches randomly from the gene trees obtained, in order to avoid bias regarding the method of random trees generation by TOPD/fmts that only randomises taxa, but not branches for the statistical comparison. Restricting these analyses to exclusively Passerine species allowed us to analyse the effects of the evolutionary patterns of each candidate genes on a smaller scale. An example of the topologies is illustrated for the candidate gene PER3 in Fig. S1.

Comparisons of these three focal topologies were carried out in TOPD/fmTS (Puigbò et al. 2007) using three different approaches: nodal, splits and disagree from the program. In brief, the ‘nodal approach’ counts the number of nodes that separate two taxa in a given topology and calculates the root mean squared deviation (RSMD) between each pair of trees. For identical topologies RSMD results in a value of zero. To calculate the significance of the RSMD obtained, TOPD/fmts calculates the distance between two contrasted tree pairs and 100 random trees obtaining one standard deviation (SD) confidence interval (CI). Compared topologies are characterised as statistically similar, within noise or different, depending on their distance with respect to CI. Specifically distances below CI denote similar topologies; distances above CI indicate statistical difference (distances around CI are within noise). The ‘disagree method’ characterises how many taxa need to be removed from the compared topology in order to end up with the exact same topologies for both trees (assessed as count of taxa/total taxa). Consequently, a value of 0/total indicates identical topologies. The ‘splits method’ evaluates if there are common branches between both trees, the lower the distance the more branches the tree pair shares.

dN/dS analysis

Accounting for the fact that different species might have found different ways to alter similar phenotypes in the same gene (i.e. different changes in sequence), we also analysed synonymous and non-synonymous mutations of all candidate genes across species. In order to pick up on putative selective pressures on candidate genes for migration, a gene-wide dN/dS analysis was carried out on the Datamonkey server (Pond and Frost 2005). We used three different datasets for each candidate gene: one including all the species, one restricted to migratory, and another restricted to non-migratory species.

Gene-wide dN/dS ratios (w) were estimated by maximum likelihood (ML) methods using a different model for each gene. Each model was obtained from the CMS module of the server. Neighbour Joining (NJ) phylogenies obtained for each candidate gene were used as input to assess likelihood of the tree comparing the neutral null model M1 (w < 1) and the model M2 that allows w > 1. Positive selection was assessed if the likelihood shows a p < 0.05.

To evaluate if lineages with migratory species show a signature of selection, a branch-specific analysis of dN/dS was also carried on Datamonkey with the Branch-Site REL program (Pond et al. 2011). The dataset for each candidate included migratory and non-migratory species. Branches under episodic diversifying selection were identified with a Holm–Bonferroni corrected p ≤ 0.05.

Structural features and predictors of migration

We used a linear regression analysis to test for a correlation between the genotype of migratory species at each focal candidate gene and both breeding latitude and migratory distance. Models for both predictors were run separately. The genotype used for each gene was the microsatellite length (as number of bases) of the 3’UTR of ADCYAP1 and CREBL1, or poly-Q (as number of predicted glutamine amino acids) on exon 20 of CLOCK and NPAS. For the CLOCK gene we included two separate polymorphic regions with variable poly-Q repeats in our analysis (both variable regions are located in the same exon). The significance of the fit was assessed with a simple linear regression, using a significance threshold p ≤ 0.05.

Within and across population variability in candidate gene sequence

Our comparative analyses focus exclusively on the sequence of one reference genome; inter-individual variation is not taken into account, mostly due to the limitation of available data to examine this level of variation. In order to make an attempt to see if variance within on candidate gene could be higher/lower in a specific migratory phenotype, we focused on CLOCK gene polymorphism, the only candidate gene with a sufficiently large number of individual sequence data available for several species (n = 10), including both migratory (n = 8) and resident (n = 2) species. Here we compare datasets of individually genotyped migratory species: flycatcher Ficedula hypoleuca (Saino et al. 2015; n = 226), willow warbler Phylloscopus
Results

Comparing phylogenetic, migratory phenotype and random topologies

We show constructed gene trees for a select number of candidate genes in Fig. 1, right column. In general, candidate gene trees do not separate migratory from non-migratory birds. Instead, they resemble the phylogenetic topologies expected based on Jarvis et al. (2014). Nonetheless, some groups of birds that comprised exclusively of migrants group together in most of the candidate gene trees. For example, Falconiformes, Accipitriformes and some Passeriformes show a clustering pattern. Nevertheless, this is most likely due to their higher levels of similarity and common ancestry rather than relationships based on migratory phenotype (also see Fig. S2). Note that in addition to separate analyses for each of the 25 candidate genes, we also concatenated sequence data from all genes by species and constructed a combined candidate genes tree (Fig. 2). In this concatenated tree the clustering patterns persist in the aforementioned lineages.

Results from nodal, splits and disagree methods for statistically distinguishing between phylogenetic, migratory phenotype and random trees can be found in Table 1 (full dataset). Our results clearly show that the phylogenetic topology tree provides the best fit to most gene trees. The lack of support for the migratory phenotype topology shows no evidence for a monophyletic origin for a migratory phenotype across avian taxa. This pattern did not change when restricting our analysis to only the extreme phenotypes, i.e. exclusively obligate migrants and clear non-migrants. We further zoomed into just one clade (specifically Passerines). Although overall, we see general patterns either more similar to the speciation phylogeny or not showing differences from a comparison with random topologies (see nodal approach in Table 1), the monophyletic clade analysis on Passeriformes indicates that respective gene topologies for some candidates, specifically HRSP12 and HSPA5, are consistently more similar to the migration topology than the speciation topology (Fig. 3). This clustering pattern is also evident in the tree topologies (see Fig. S2) where most of the migrant species within Passerine tend to cluster in one branch. The other approaches did not show evidence for a general trend towards one or a set of candidate genes being recurrently involved in distinguishing migrants from non-migrants.

Recall that we also concatenated sequence data from all genes by species and constructed a complete gene tree. This concatenated phylogeny shows that most of the taxonomic groups clustering migratory species together are statistically well supported and show larger branch lengths. This could suggest an evolutionary process of acceleration in these lineages for some of the candidate genes.

Selection in candidate genes is not related to migration

To identify selective forces in lineages with migratory species, we performed a gene-wide and branch-specific dN/dS analysis. The gene-wide analysis for selection in candidate genes across all bird species indicates that CLOCK, DRD4, NEK2, HSPA5 and CSNKE1, have been under positive selection at p < 0.05 (Table 2). Nonetheless, and important within the focus of our analysis, this signature does not seem to be linked to the migratory phenotype in any of our comparisons, as independent datasets of separately analysed migratory, partial migratory and non-migratory species generally do not show any signature of selection. The only exception here was the CLOCK gene that showed significant values for selection in partial migrants. Thus, despite the general overall pattern of selection across all avian clade, none of these candidate genes showed lineages under selective pressure.

Association between structural features and predictors of migration

Alignments of the genomic sequences (Fig. S3) of any of the 25 gene candidates did not separate migratory and non-migratory species. ADCYAP1 has variable sequence lengths of a microsatellite with an AG repeat at the 3′ UTR region ranging from 26 to 56 bp for migratory species, and 10–54 bp for non-migratory species. CREB1 also has a microsatellite with a TG/CG repeat motive at the 3′ UTR region, ranging from 12 to 26 bp in migratory and 14–42 bp in non-migratory species, respectively (the longest length variant was exclusively found in tits). NPAS has one variable region of poly-Q repeats, in migratory species; this has length variation between 6 and 10 amino acids (aa), in non-migratory species the repeat length varied between 5 and 12 aa repeats. The CLOCK gene has 2
poly-Q variable regions, both located on exon 20. The first region (R1) shows length variation between 6 and 13 poly-Q repeats in migrants and 4 to 14 repeats in non-migratory species. Length polymorphism in variable region two (R2) varied between 4 and 9 poly-Q repeats in both phenotypes. Lengths variation in neither of the focal candidates showed significant differences between migratory and non-migratory birds.
The intra-specific analysis focusing on individual variability of \textit{CLOCK} gene polymorphism shows a tendency for higher variability in migratory species (Fig. S4); however, given that only two non-migratory species were included (namely great tits and blue tits—great tits being mostly monomorphic, blue tits highly variable), this does not allow us to draw any statistically supported conclusions (Table S3). Nonetheless, this individual based analysis of variability suggests that most common allele lengths (i.e. number of poly-Q repeats) varied considerably between species, as did the overall degree of variability within species (Fig. S4, Table S3). For example, migratory species like \textit{H. rustica} and \textit{F. hypoleuca}, show constrained levels of variability, being mostly monomorphic, while \textit{C. caeruleus}, a non-migrant species, shows a degree of variance comparable to those migrant species. Statistical differences were found between the group of migrant and non-migrant species. Nonetheless when the comparisons were performed between individual species, migrant species and non-migrant species showed statistical differences, but also comparisons between a migrant and a non-migrant species did not show statistical differences (see Table S3).

We used linear models to quantify the correlation of structural repeats (poly-Q in \textit{CLOCK} and \textit{NPAS2}; and allele lengths in \textit{ADCYAP1} and \textit{CREB1}) and both breeding latitude and migratory distance (Fig. 1, left column). The fit of the data to a linear regression model did not show significant associations between genotype and migratory phenotype (Table 3). The best fit, albeit not significant, that was following a similar pattern as earlier work showing a correlation between candidate gene variation and migratory phenotype results from our comparison between \textit{ADCYAP1} genotype variation in relation to migratory distance ($r^2 = 0.2120, p = 0.01132$). Although not significant, here the regression trend is in line with earlier studies, and might be taken as support for earlier findings reporting genotype variation at this locus with relevance to the migratory phenotype on the within-species level.

**Discussion**

We used publicly available archived genomic data from non-migratory and migratory bird species to test for the presence or absence of a general clustering pattern in candidate genes for migration. In a cross-species comparative approach, we characterised sequence features in 25 candidate genes in a dataset including birds across all orders of Aves. Our study thus illustrates the potential of public genomic data to test for general patterns of migratory gene candidates in a cross-species comparative context. Despite this powerful dataset we were not able to identify genetic variation that allowed us to distinguish migratory from non-migratory birds based on sequence difference in any of the candidate genes included here.

Most patterns we recover based on candidate gene sequence variation were driven by species phylogeny, and
Table 1  Topology comparison among candidate gene trees and target trees

|                  | RANDOM |             | SPECIATION |             | MIGRATION |             |
|------------------|--------|-------------|------------|-------------|-----------|-------------|
|                  | Dis    | Nod        | Split      | Dis         | Nod       | Split       |
| PASSERINES restricted dataset |        |             |            |             |           |             |
| ADCYAP1          | 12/12  | 2.741378   | 1          | 1/13        | 1.450022  | 0.4         |
| CLOCK            | 12/12  | 2.383566   | 1          | 1/13        | 1.450022  | 0.4         |
| CRY1             | 8/15   | 2.146980   | 0.916667   | 1/15        | 1.751190  | 0.416667    |
| CRY2             | 17/17  | 3.226972   | 1          | 4/17        | 2.711631  | 0.5         |
| DRD4             | 17/17  | 3.391165   | 0          | 0/17        | 0         | 0          |
| HRSP12           | 15/15  | 2.700970   | 1          | 3/15        | 1.825742  | 0.5         |
| HSP90B1          | 14/14  | 2.896522   | 1          | 1/14        | 1.657386  | 0.364656    |
| HSPA5            | 16/16  | 3.345395   | 1          | 6/16        | 2.534758  | 0.4         |
| HSPA8            | 14/14  | 3.275938   | 1          | 1/14        | 2.305879  | 0.435897    |
| NEK2             | 14/14  | 2.604000   | 0.638889   | 1/15        | 1.067071  | 0.272222    |
| NFIL3            | 15/16  | 3.221151   | 2/16       | 2/16        | 2.416335  | 0.5         |
| PARL             | 17/17  | 3.429972   | 5/17       | 2.667892    | 0.642857  | 4/16        |
| PER2             | 13/13  | 2.480695   | 1/14       | 1.500915    | 0.363636  | 1/14        |
| PER3             | 17/17  | 3.564531   | 5/17       | 2.825826    | 0.642857  | 4/17        |
| SLC2A1           | 12/12  | 2.424005   | 2/12       | 1.299174    | 0.309259  | 5/12        |
| AANAT            | 16/16  | 2.966479   | 3/16       | 2.081666    | 0.307692  | 4/16        |
| CSNK1E           | 13/13  | 3.125571   | 4/13       | 2.124340    | 0.5       | 4/13        |
| SLC1A3           | 12/14  | 3.119223   | 1/14       | 2.401755    | 0.465397  | 1/14        |
| HSP90A1          | 13/15  | 3.072418   | 0.823900   | 3/15        | 1.345224  | 0.5         |
| TTR              | 14/14  | 3.034009   | 2/15       | 2.113851    | 0.407280  | 2/15        |
| YPEL1            | 13/13  | 2.916833   | 0.990741   | 3/13        | 1.874818  | 0.5         |
| NPSA2            | 10/10  | 2.653709   | 2/10       | 1.341210    | 0.322046  | 4/10        |
| ARNTL            | 12/12  | 3.286592   | 2/12       | 2.412694    | 0.4       | 2/12        |
| CPNE4            | 13/13  | 3.221151   | 2/14       | 2.416335    | 0.527778  | 2/14        |
| CREB1            | 17/17  | 3.735757   | 4/17       | 2.711631    | 0.5       | 4/17        |
| MIGRANTS versus NON-MIGRANTS restricted dataset |        |             |            |             |           |             |
| ADCYAP1          | 26/26  | 3.478727   | 0.9565     | 5/27        | 2.951652  | 0.375       |
| CLOCK            | 40/40  | 5.640058   | 11/41      | 5.180639    | 0.578947  | 17/41       |
| CRY1             | 39/39  | 5.528725   | 0.9722     | 7/40        | 5.137968  | 0.621622    |
| CRY2             | 41/41  | 5.931387   | 0.9766     | 10/41       | 5.340413  | 0.683058    |
| DRD4             | 41/41  | 5.086807   | 3/17       | 0.496668    | 0.071429  | 12/42       |
| HRSP12           | 42/42  | 4.994887   | 0.9744     | 11/43       | 5.041138  | 0.675       |
| HSP90B1          | 46/47  | 5.639855   | 0.9852     | 16/47       | 5.230941  | 0.640355    |
| HSPA5            | 47/47  | 5.483471   | 12/48      | 5.055016    | 0.622222  | 27/48       |
| HSPA8            | 44/44  | 5.468022   | 13/44      | 4.792124    | 0.670420  | 18/44       |
| NEK2             | 44/44  | 4.559393   | 0.9783     | 16/44       | 4.066146  | 0.476974    |
| NFIL3            | 46/47  | 5.878573   | 0.9852     | 15/47       | 5.185939  | 0.661762    |
| PARL             | 42/42  | 4.882503   | 0.9762     | 17/42       | 4.423420  | 0.525190    |
| PER2             | 42/42  | 5.092866   | 13/43      | 4.831376    | 0.65      | 22/43       |
| PER3             | 46/46  | 4.579834   | 23/46      | 4.538616    | 0.697674  | 15/46       |
| SLC2A1           | 28/28  | 5.259547   | 0.9853     | 13/28       | 5.003874  | 0.646401    |
| AANAT            | 38/38  | 4.836443   | 10/39      | 4.577714    | 0.611111  | 20/39       |
| CSNK1E           | 40/40  | 6.099601   | 20/40      | 5.006534    | 0.729730  | 13/40       |
| SLC1A3           | 46/46  | 5.205493   | 0.9822     | 13/46       | 5.003494  | 0.648874    |
| HSP90A1          | 39/39  | 4.910991   | 0.9556     | 19/39       | 4.893480  | 0.573299    |
| TTR              | 21/22  | 5.185258   | 0.9915     | 12/22       | 4.888213  | 0.665140    |
| YPEL1            | 47/48  | 4.874233   | 0.9928     | 13/48       | 4.553259  | 0.395       |

 Springer
did not show indications separating migratory and non-migratory birds. This could suggest that candidate genes do not play a general role in controlling migration across the full avian clade, or that patterns cannot be evident in a wide phylogenetic analysis. In addition to analysing each candidate gene individually, we also analysed concatenated sequence data from all candidate genes, which also did not allow clear separation of migratory and non-migratory birds. The latter finding might not be too surprising, as this scenario would assume divergence of candidate genes among lineages in parallel to species divergence, a rather unlikely scenario.

### Table 1 continued

|          | RANDOM |          | SPECIATION |          | MIGRATION |          |
|----------|--------|----------|------------|----------|-----------|----------|
|          | Dis    | Nod      | Split      | Dis      | Nod       | Split    |
|          |        |          |            |          |           |          |
| NPAS2    | 40/40  | 5.479820 | 1          | 19/40    | 4.936480  | 0.660727 |
| ARNTL    | 36/36  | 4.673910 | 0.9848     | 17/36    | 4.342727  | 0.566613 |
| CPNE4    | 38/39  | 5.878573 | 0.9852     | 18/39    | 5.185939  | 0.661762 |
| CREB1    | 40/48  | 6.699674 | 1          | 24/48    | 5.973568  | 0.644444 |
| ADCYAP1  | 30/34  | 3.794169 | 0.954854   | 8/35     | 3.906039  | 0.468750 |
| CLOCK    | 59/59  | 5.931323 | 1          | 20/60    | 5.510288  | 0.684211 |
| CRY1     | 56/60  | 5.894122 | 0.982456   | 23/62    | 5.500493  | 0.679966 |
| CRY2     | 61/61  | 5.690368 | 0.991729   | 19/61    | 5.419998  | 0.676801 |
| DRD4     | 60/60  | 5.837537 | 1          | 3/17     | 0.469668  | 0.071429 |
| HRSP12   | 62/62  | 5.237258 | 1          | 27/63    | 5.349055  | 0.701754 |
| HSP90B1  | 67/67  | 4.944896 | 0.966574   | 25/67    | 4.556093  | 0.604968 |
| HSPA5    | 67/67  | 6.240832 | 1          | 18/68    | 5.691938  | 0.630769 |
| HSPA8    | 62/63  | 5.620202 | 0.984544   | 22/63    | 5.125876  | 0.663344 |
| NEK2     | 65/66  | 6.254055 | 0.983173   | 25/66    | 5.649675  | 0.651162 |
| NFIL3    | 67/67  | 5.970355 | 0.991729   | 21/67    | 5.528167  | 0.637798 |
| PARL     | 66/66  | 5.523274 | 0.987992   | 20/66    | 5.203310  | 0.632417 |
| PER2     | 59/59  | 5.391185 | 1          | 22/60    | 5.182543  | 0.666667 |
| PER3     | 66/66  | 5.478672 | 1          | 30/67    | 5.203342  | 0.671875 |
| SLC2A1   | 41/41  | 5.860697 | 0.981719   | 15/41    | 5.275366  | 0.663248 |
| AANAT    | 54/55  | 5.090447 | 1          | 14/55    | 4.810622  | 0.653846 |
| CSNK1E   | 51/55  | 5.950073 | 0.961538   | 25/55    | 4.951618  | 0.629308 |
| SLC1A3   | 65/65  | 4.862746 | 0.967742   | 16/65    | 4.708164  | 0.576480 |
| HSP90AA1 | 52/53  | 5.206538 | 0.972647   | 17/53    | 4.972273  | 0.610309 |
| TTR      | 22/22  | 6.378516 | 0.961444   | 11/22    | 5.565637  | 0.666466 |
| YPEL1    | 48/48  | 5.687568 | 0.994152   | 16/48    | 5.453279  | 0.687392 |
| NPAS2    | 56/56  | 5.182652 | 0.980130   | 20/56    | 4.961226  | 0.619555 |
| ARNTL    | 55/55  | 5.949312 | 0.974225   | 17/56    | 5.253574  | 0.667268 |
| CPNE4    | 54/54  | 5.503541 | 0.991195   | 19/54    | 5.202493  | 0.676264 |
| CREB1    | 61/67  | 6.806959 | 0.968750   | 29/67    | 6.179656  | 0.640625 |

To contrast topology patterns we compared single gene trees derived for each candidate gene with three target trees: (1) randomly generated trees based on randomly permutating branches of single gene tree topology 'random topology'; (2) genes based on overall phylogeny (Jarvis et al. 2014), thus representing evolutionary divergence among all the species included in the study 'phylogenetic topology', and (3) topology classified by migratory behaviour, i.e. an artificial tree clustering migratory species and non-migratory species in two different branches 'topology based on migratory phenotype'.

**Bold** characters highlight the most similar topology for each candidate gene tree for the three methods.

Tables shown for three datasets: Monophyletic analysis on Passerines; exclusively contrasting obligate migrants versus clear residents; and the full dataset including the third phenotype category with partial migrants, dispersive species.

* and ** Denote significance below and above one SD of 100 randomly generated trees, respectively.
In order to address the fact that we cannot exclude that the relevance and contribution of certain traits may differ between avian lineages in a way that different traits may be controlled by different mutations in the same genes, and/or varying patterns of epistatic interactions with other genes, we also screened for putative features of selection within these candidate genes. However, our dN/dS approach did not allow identifying patterns of selection at either the gene-wide or branch-sites level for the migratory lineages. This result does not provide support for lineage-specific selection on the gene candidates.

Indicative patterns might get lost if not approached at the appropriate phylogenetic scale. Our more focused analyses on a restricted data assembly exclusive to the Passeriformes allow aims to address this issue at least partly allowing us to potentially recover general trends towards some or all candidate genes being recurrently involved in the control of migration. Analyses limited to this family show improved resolution and come close to the margin of randomness in nodal comparisons. However, even on this more focal scale only a few of these comparisons show better values than the phylogeny species tree, suggesting that even on a more constrained phylogenetic scale, the genes keep following

Fig. 3 Topologies on a constrained phylogenetic scale. Clade-specific analysis exclusively focusing on all species within the genus Passerines. The candidate gene trees obtained from the Neighbour Joining analysis for HRSP12 and HSPA5, two candidate genes that follow clustering patterns consistently more similar to the migration topology than the speciation topology. *Colouring scheme* and node support as in Fig. 1
### Table 2: Gene-wide dN/dS analysis of migratory gene candidates

| Gene | Phenotype | lnL.M1 | lnL.M2 | LRT    | P value |
|------|-----------|--------|--------|--------|---------|
| ARNTL | ALL       | −12014.300 | −12014.300 | 1.30E−04 | 1.000   |
|       | Migratory | −2443.540  | −2443.540  | 8.43E−05 | 1.000   |
|       | Partial migrants | −5751.000  | −5883.250  | 3.42E−03 | 1.000   |
|       | Residents | −9022.375  | −9022.055  | 1.17E+00 | 0.623   |
| CLOCK | ALL       | −17203.100 | −17199.300 | 7.49E+00 | **0.0236713** |
|       | Migratory | −2784.523  | −2784.444  | 2.89E−04 | 0.899   |
|       | Partial migrants | −6084.900  | −6801.840  | 6.13E+00 | **0.0467** |
|       | Residents | −12321.600 | −12321.600 | 1.27E−04 | 1.000   |
| PER3  | ALL       | −41720.300 | −41718.500 | 3.49E+00 | 0.175   |
|       | Migratory | −2124.250  | −2124.250  | 2.34E−06 | 1.000   |
|       | Partial migrants | −5883.250  | −5883.250  | 6.48E−05 | 1.000   |
|       | Residents | −13206.766 | −13206.727 | 7.11E−02 | 0.969   |
| CRY1  | ALL       | −11297.000 | −11297.000 | 6.54E−05 | 1.000   |
|       | Migratory | −2618.571  | −2618.509  | 5.14E−01 | 0.949   |
|       | Partial migrants | −6543.055  | −6543.055  | 2.94E−05 | 0.999   |
|       | Residents | −7465.354  | −7465.354  | 4.34E−05 | 1.000   |
| CRY2  | ALL       | −12140.600 | −12140.600 | −5.26E−04 | 1.000   |
|       | Migratory | −2718.098  | −2718.030  | 2.57E−01 | 0.869   |
|       | Partial migrants | −5137.220  | −5134.740  | 4.96E+00 | **0.084** |
|       | Residents | −11345.504 | −11345.444 | 1.11E−01 | 0.952   |
| NEK2  | ALL       | −15013.300 | −15008.900 | 8.86E+00 | **0.0119051** |
|       | Migratory | −3123.490  | −3123.430  | 1.03E+00 | 0.798   |
|       | Partial migrants | −6882.490  | −6882.400  | 2.48E−06 | 1.000   |
|       | Residents | −13729.370 | −13729.286 | 1.55E−01 | 0.932   |
| NPAS2 | ALL       | −9773.415  | −9773.415  | 3.58E−07 | 1.000   |
|       | Migratory | −2788.055  | −2787.988  | 6.78E−04 | 1.000   |
|       | Partial migrants | −6711.240  | −6711.240  | 5.43E−06 | 1.000   |
|       | Residents | −8384.493  | −8384.493  | 4.83E−05 | 0.998   |
| PER2  | ALL       | −34024.900 | −34024.900 | 2.68E−04 | 1.000   |
|       | Migratory | −2703.313  | −2703.248  | 3.57E−04 | 0.975   |
|       | Partial migrants | −5601.859  | −6548.185  | −1.29E−04 | 1.000   |
|       | Residents | −14439.024 | −14438.968 | 1.02E−01 | 0.955   |
| DRD4  | ALL       | −9382.370  | −9377.070  | 1.06E+01 | **0.00499358** |
|       | Migratory | −3260.480  | −3257.910  | 5.14E−00 | **0.077** |
|       | Partial migrants | −5392.430  | −6882.400  | 5.45E−06 | 1.000   |
|       | Residents | −6979.373  | −6979.033  | 1.25E+00 | 0.599   |
| AANAT | ALL       | −5698.000  | −5698.000  | 2.84E−05 | 1.000   |
|       | Migratory | −2699.671  | −2699.604  | 6.46E−02 | 0.914   |
|       | Partial migrants | −6545.730  | −6545.730  | 2.02E−05 | 0.999   |
|       | Residents | −4721.560  | −4721.560  | −7.37E−05 | 1.000   |
| CPNE4 | ALL       | −11228.606 | −11228.588 | 3.60E−02 | 1.000   |
|       | Migratory | −2452.620  | −2452.575  | 5.76E−04 | 1.000   |
|       | Partial migrants | −5633.000  | −6539.900  | 3.20E−04 | 0.999   |
|       | Residents | −9385.890  | −9385.890  | 3.73E−04 | 0.999   |
| HSPA5 | ALL       | −16331.700 | −16328.200 | 6.92E+00 | **0.0314173** |
|       | Migratory | −2707.037  | −2706.969  | 3.26E−02 | 0.796   |
|       | Partial migrants | −5595.625  | −6566.695  | 1.26E−03 | 0.993   |
|       | Residents | −12404.581 | −12404.509 | 1.33E−01 | 0.942   |
Table 2 continued

| Gene     | Phenotype  | lnL M1    | lnL M2    | LRT      | P value |
|----------|------------|-----------|-----------|----------|---------|
| HSPA8    | ALL        | -11781.944| -11781.943| 2.92E-03 | 1.000   |
|          | Migratory  | -2703.351 | -2703.283 | 1.29E-01 | 0.866   |
|          | Partial migrants | -5594.500 | -6547.055 | -5.42E-05| 1.000   |
|          | Residents   | -7568.080 | -7568.080 | 8.25E-05 | 1.000   |
| HSP90B1  | ALL        | -16126.900| -16126.900| -9.01E-05| 1.000   |
|          | Migratory  | -2780.990 | -2780.900 | 4.24E-04 | 1.000   |
|          | Partial migrants | -5603.844 | -6540.379 | 2.81E-04 | 0.998   |
|          | Residents   | -15017.600| -15017.300| 5.54E-01 | 0.758   |
| HSP90AA1 | ALL        | -14902.200| -14902.200| 2.91E-05 | 1.000   |
|          | Migratory  | -13307.269| -13307.209| 1.11E-01 | 0.951   |
|          | Partial migrants | -5596.750 | -6586.335 | 2.21E-03 | 0.990   |
|          | Residents   | -13025.049| -13024.987| 1.14E-01 | 0.950   |
| SLC2a1   | ALL        | -3747.905 | -3747.905 | -1.05E-03| 1.000   |
|          | Migratory  | -2703.354 | -2703.286 | 1.93E-01 | 0.832   |
|          | Partial migrants | -6535.469 | -6535.469 | 2.20E-05 | 0.977   |
|          | Residents   | -2483.450 | -2483.450 | 3.48E-04 | 0.954   |
| SLC1a3   | ALL        | -9229.518 | -9229.518 | 2.19E-06 | 1.000   |
|          | Migratory  | -2681.245 | -2681.178 | 5.17E-04 | 0.943   |
|          | Partial migrants | -6547.055 | -6547.055 | -2.15E-04| 1.000   |
|          | Residents   | -5723.150 | -5722.510 | 2.34E-00 | 0.246   |
| PARL     | ALL        | -5321.898 | -5321.143 | 1.51E-00 | 0.621   |
|          | Migratory  | -2695.975 | -2695.908 | 2.57E-01 | 0.859   |
|          | Partial migrants | -5602.781 | -6543.055 | 1.18E-03 | 0.999   |
|          | Residents   | -3519.020 | -3519.020 | -4.90E-06| 1.000   |
| TTR      | ALL        | -10543.983| -10543.914| 1.38E-01 | 0.999   |
|          | Migratory  | -2699.665 | -2699.597 | 1.45E-01 | 0.873   |
|          | Partial migrants | -5611.500 | -6504.243 | -2.44E-05| 1.000   |
|          | Residents   | -7637.370 | -7637.370 | 2.34E-04 | 0.989   |
| YPEL1    | ALL        | -9350.850 | -9350.062 | 1.58E+00 | 0.532   |
|          | Migratory  | -2692.279 | -2692.213 | 1.29E-01 | 0.733   |
|          | Partial migrants | -6545.289 | -6545.289 | 8.76E-06 | 0.989   |
|          | Residents   | -6270.698 | -6270.538 | 5.86E-01 | 0.811   |
| HRSP12   | ALL        | -6547.120 | -6547.110 | -2.40E-03| 1.000   |
|          | Migratory  | -2745.243 | -2745.243 | 3.24E-06 | 0.987   |
|          | Partial migrants | -5604.125 | -6545.289 | 8.76E-06 | 0.989   |
|          | Residents   | -4936.372 | -4936.012 | 1.32E+00 | 0.576   |
| CREB1    | ALL        | -5330.580 | -5330.580 | 6.44E-05 | 1.000   |
|          | Migratory  | -2786.289 | -2786.289 | 6.08E-05 | 0.999   |
|          | Partial migrants | -6504.243 | -6504.243 | -1.06E-04| 1.000   |
|          | Residents   | -3287.980 | -3287.980 | 7.82E-04 | 1.000   |
| CSNKE1   | ALL        | -5495.620 | -5490.040 | 1.12E+01 | 0.00376182** |
|          | Migratory  | -2784.055 | -2784.055 | 1.80E-04 | 0.921   |
|          | Partial migrants | -6546.958 | -6546.958 | -4.89E-04| 1.000   |
|          | Residents   | -4149.594 | -4149.514 | 2.93E-01 | 0.906   |
| ADCYAP1  | ALL        | -1903.400 | -1903.540 | -2.82E-01| 1.000   |
|          | Migratory  | -2781.379 | -2781.379 | 2.14E-05 | 0.859   |
|          | Partial migrants | -6555.992 | -6555.992 | 5.65E-06 | 1.000   |
|          | Residents   | -2028.490 | -2028.490 | 3.43E-05 | 1.000   |
the trends of phylogeny or speciation rather than clustering according to migratory phenotype.

Most studies of candidate genes look for correlates of length polymorphism in a limited number of species. Our approach aims to extend this framework to assessing sequence variability in candidate genes to a broader evolutionary scale. However, our data do not uncover statistical differences between species with different phenotypes, neither do they recover correlations with specific migratory features, specifically distance and breeding latitude. **ADCYAP1** did show a tendency for positive, though non-significant, correlation with migratory distance in our dataset, supporting earlier findings in species-specific analyses. As the length polymorphism in this gene is located in a 3′ UTR this could suggest that a specific pattern of gene expression is playing a role for this gene, and this feature might be relevant in long-distance migrants.

Our focus is put on general sequence differences in candidate genes for migration across species, but we also note that there is a lot of variation in the level of individual and inter-population diversity across avian species in many of the candidate genes included here, such as **CLOCK** poly-Q. Within-species variability at polymorphic loci often is high in some, but certainly not all species, irrespective of migratory phenotype. Previous work has suggested that variability in lengths polymorphism of **CLOCK** related to the timing of migration is enhanced in migratory birds (Saino et al. 2015). Our analyses testing for a role of inter-individual variation in **CLOCK** lengths polymorphism between migratory and non-migratory species falls short in strengthening this suggestion. We detect similarly high levels of polymorphisms as well as scenarios with almost monomorphic genotypes for various species, irrespective of migratory phenotype. This could mean that the observed variability in **CLOCK** is more tightly linked to other parameters that shape the variability of the migratory phenotype, such as breeding latitudes, distances and timing regimes, rather than specifically controlling timing traits in a migration context.

But even when leaving inter-individual variability aside and focusing on just one reference sequence, we expect general patterns that allow separating migrants from non-migratory species of suggested candidate genes for migration to show up if they were in fact under parallel selection. Our results do not support this hypothesis and thus highlight limitations and question the usefulness of a candidate gene approach in the context of understanding migratory behaviour. Consequently, we call for caution and highlight the limitation of candidate gene approaches in macro-evolutionary contexts, as in most cases functionality cannot be easily inferred. Also, most candidate gene studies do not allow for a clear distinction from genetic drift, come with uncertainty about the amount of loci involved in the trait and do not allow controlling for possible effect of linkage disequilibrium. Our study further points out how this approach can erroneously simplify a highly complex phenotype that most probably is a multilocus trait.

One further point we feel is especially important to keep in mind when using a candidate gene approach in the context of migration: many candidate genes have been identified in genetic model organisms—none of that migrate or show any correlate to a migratory phenotype, and functional validation or relevance is mostly lacking. In many of these studies, the subset of candidate genes for migration were selected based on their assumed effect on anticipated candidate traits that feed into the complex phenomenon of migration. Importantly, one possibility explaining the lack of pattern might be the limitation of the starter set of ad

---

**Table 3 Predictors of linear models**

| Predictor          | df | Res err | F    | p    |
|--------------------|----|---------|------|------|
| ADCYAP1 vs distance| 11 | 2228    | 2.962| 0.1132|
| ADCYAP1 vs breeding| 11 | 7.299   | 0.05144| 0.8252|
| CREB1 vs distance  | 13 | 2093   | 0.8929| 0.3619|
| CREB1 vs breeding  | 13 | 6.126   | 0.2411| 0.6316|
| CLOCK R1 vs distance| 23 | 2675 | 1.125 | 0.3   |
| CLOCK R1 vs breeding| 23 | 12.8  | 0.01369| 0.9079|
| CLOCK R2 vs distance| 21 | 2726 | 0.3082 | 0.5846|
| CLOCK R2 vs breeding| 21 | 17.73 | 0.1876 | 0.6693|
| NPAS2 vs distance  | 17 | 2701  | 0.647 | 0.432 |
| NPAS2 vs breeding  | 17 | 11.29 | 0.0788 | 0.7825|

For every comparison df degree of freedom, res residual error, F statistic and significance levels are shown

Significance at p < 0.05

**Table 2 continued**

| Gene   | Phenotype | lnL M1 | lnL M2 | LRT   | P value |
|--------|-----------|--------|--------|-------|---------|
| NFIL3  | ALL       | −9638.910 | −9638.910 | 1.55E−04 | 1.000 |
|        | Migratory | −2788.055 | −2788.055 | 5.14E−01 | 0.949 |
|        | Partial migrants | −5565.000 | −6711.240 | −2.83E−05 | 1.000 |
|        | Residents  | −8583.364 | −8583.304 | 1.11E−01 | 0.952 |

Significance of positive selection was assessed with p < 0.1 indicated with bold font and p < 0.05 indicated with **
hoc hypothesised candidate genes we investigate and may not be completely unexpected. Once more genomic data for migratory and non-migratory species will become available, the genomic toolbox allowing to investigate the genetic basis of migratory traits will grow and allow for an increasing-ly more informed list of de novo identified candidates to be tested in migratory birds.

Acknowledgements Open access funding provided by Max Planck Society. Funding was provided by the Max Planck Society and the International Max Planck Research School (IMPRS) for Evolutionary Biology.

Compliance with ethical standards

Conflict of interest The authors confirm that they do not have any conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

Bazzi G, Cecere JG, Caprioli M, Gatti E, Gianfranceschi L, Podolfini S, Possenti CD, Ambrosini R, Saino N, Spina F, Rubolini D (2016) Clock gene polymorphism, migratory behaviour and geographic distribution: a comparative study of trans-Saharan migratory birds. Mol Ecol. doi:10.1111/mec.13913

Berthold P (1999) A comprehensive theory for the evolution, control and adaptability of avian migration. Ostrich 70:1–11

Berthold P, Helfig AJ, Mohr G, Querner U (1992) Rapid microevolution of migratory behaviour in a wild bird species. Nature 360:668–670. doi:10.1038/360668a0

Chernetsov N, Kishkinov D, Mourtis M (2008) A long-distance avian migrant compensates for longitudinal displacement during spring migration. Curr Biol 18:188–190. doi:10.1016/j.cub.2008.01.018

Delmore KE, Liedvogel M (2016) Investigating factors that generate and maintain variation in migratory orientation: a primer for recent and future work. Front Behav Neurosci 10:3. doi:10.3389/fnbeh.2016.00003

Delmore KE, Hüblner S, Kane Schuster R, Andrew RL, Câmara F, Guigó R, Irwin DE (2015a) Genomic analysis of a migratory divide reveals candidate genes for migration and implicates selective sweeps in generating islands of differentiation. Mol Ecol 24:1873–1888

Delmore KE, Kenyon HL, Germain RR, Irwin DE (2015b) Phenotypic divergence during speciation is inversely associated with differences in seasonal migration. Proc R Soc B Biol Sci 282:20151921

Delmore KE, Toews DPL, Germain RR, Owens GL, Irwin DE (2016) The genetics of seasonal migration and plumage color. Curr Biol 26:2167–2173

Dor R, Lovette IJ, Safran RJ, Billerman SM, Huber GH, Vortman Y 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(12):e28843

Fuchs R, Winkler H, Bingman VP, Ross JD, Bernroeder G (2014) Brain geometry and its relation to migratory behavior in birds. J Adv Neurosci Res 1:1–9

Hedges SB, Marin J, Suleski M, Paymer M, Kumar S (2015) Tree of life reveals clock-like speciation and diversification. Mol Biol Evol 32:835–845

Irwin DE, Alcaide M, Delmore K, Irwin JH, Owens GL (2016) Recurrent selection explains parallel evolution of genomic regions of high relative but low absolute differentiation in a ring species. Mol Ecol 25:4488–4507

Jarvis ED, Aberer AJ, Li B, Houde P, Li Ho SYW, Faircloth BC, Nabholz B, Howard JT, Suth A, Weber CC, Li J, Zhang F, Li H, Zhou L, Narula N, Liu L, Ganapathy G, Boussau B (2014) Whole-genome analyses resolve early branches in the tree of life of modern birds. Science 346:1126–1138

Johnsen A, Fidler AE, Kuhn S, Carter KL, Hoffmann A, Barr IR, Biard C, Charmantier A, Eens M, Korsten P, Sitiari H, Tomiuk J, Kempenaers B (2007) Avian Clock gene polymorphism: evidence for a latitudinal cline in allele frequencies. Mol Ecol 16:4867–4880. doi:10.1111/j.1365-294X.2007.03552.x

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30:772–780

Liedvogel M, Sheldon BC (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, Szulkin M, Knowles SCL, 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. Mol Ecol 18:2444–2456

Liedvogel M, Åkesson S, Bensch S (2011) The genetics of migration on the move. TREE 26:561–569. doi:10.1093/treeceli.2011.07.009

Lockwood R, Swaddle JP, Rayner JMV (1998) Avian wing-tip shape reconsidered: wingtip shape indices and morphological adaptations to migration. J Avian Biol 29:273–292. doi:10.2307/3677110

Lundberg M, Liedvogel M, Larson K, Sigerman H, Grahn M, Wright A, Åkesson S, Bensch S (2017) Genetic differences between willow warbler migratory phenotypes are few and cluster in large haplotype blocks. Evol Lett (accepted manuscript)

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(12):e0144587. doi:10.1371/journal.pone.0144587

Müller JC, Pulido F, Kempenaers B (2011) Identification of a gene associated with avian migratory behaviour. Proc R Soc B Biol Sci 278:2848–2856. doi:10.1098/rspb.2010.2567

Perdeck AC (1958) Two types of orientation in migrating starlings, Sturnus vulgaris L., and chaffinches, Fringilla coelebs L., as revealed by displacement experiments. Ardea 46:1–37

Peterson MP, Abolins-Abols M, Atwell JW et al (2013) Variation in candidate genes CLOCK and ADCYP1 does not consistently predict differences in migratory behavior in the songbird genus Junco. PLoS One 8(11):e70668. doi:10.1371/journal.pone.0070668

Pond SLK, Frost SDW (2005) Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. Bioinformatics 21:2531–2533

Pond SLK, Murrell B, Fourment M, Frost SDW, Delport W, Scheffler K (2011) A random effects branch-site model for detecting episodic diversifying selection. Mol Biol Evol. doi:10.1093/molbev/msr125

Puigbò P, Garcia-Valvè S, Mccrenery JO (2007) TOPD/FMTS: a new software to compare phylogenetic trees. Bioinf 23:1556–1558

© Springer
Pulido F, Berthold P (2010) Current selection for lower migratory activity will drive the evolution of residency in a migratory bird population. PNAS 107:7341–7346. doi:10.1073/pnas.0910361107

Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. Bioinf 26:841–842

Ruegg KC, Anderson EC, Paxton KL, Apkenas V, Lao S, Siegel RB, DeSante DF, Moore F, Smith TB (2014) Mapping migration in a songbird using high-resolution genetic markers. Mol Ecol 23:5726–5739. doi:10.1111/mec.12977

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. Mol Ecol 24:1758–1773. doi:10.1111/mec.13159

Svensson L, Grant PJ, Mullarney K, Zetterstroem D (1999) Der neue Kosmos-Vogelführer—Alle Arten Europas, Nordafrikas und Vorderasiens. Franckh-Kosmos Verlags-GmbH & Co., Stuttgart. ISBN 3-440-07720-9

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetic analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28:2731–2739

Van Doren BM, Campagna L, Helm B, Illera JC, Lovette IJ, Liedvogel M (2017) Correlated patterns of genetic diversity and differentiation across an avian family. Mol Ecol. doi:10.1111/mec.14083