Evolutionary response to climate change

in migratory pied flycatchers

Barbara Helm[1,2,3,6,*], Benjamin M. Van Doren[4,6], Dieter Hoffmann[5] & Ute Hoffmann[5]

[1] present address: University of Groningen, GELIFES - Groningen Institute for Evolutionary Life Sciences, Nijenborgh 7, Groningen, 9747 AG The Netherlands, b.helm@rug.nl

[2] University of Glasgow, IBAHCM - Institute of Biodiversity, Animal Health and Comparative Medicine, Graham Kerr Building, Glasgow G12 8QQ, UK

[3] Max-Planck-Institut für Ornithologie, Eberhard-Gwinner-Strasse 6a, 82319 Seewiesen, Germany

[4] University of Oxford, Edward Grey Institute, Department of Zoology, Oxford, OX1 3PS, UK, bmvandoren@gmail.com

[5] Hanhofer Straße 35a, Harthausen, 67376, Germany

[6] corresponding authors, who contributed equally

* Lead contact: Barbara Helm

Keywords: migration, circannual, phenology, timing, climate window

SUMMARY

Climate change is rapidly advancing spring phenology [1-3], but at different rates in different species [1, 4]. Whether these advances are solely driven by phenotypic plasticity [2, 5], or also involve evolution, is hotly debated (e.g. [5-7]). In some species, including avian long-distance migrants, plastic responses to early springs may be constrained by inherited circannual timing programmes [8, 9], making evolutionary adjustment the only viable mechanism for keeping pace with shifting phenology [5, 10]. This constraint may be contributing to population declines in migratory species [5, 10-12]. To test whether a migrant's timing programme has evolved [10, 12], we replicated an experimental study of the annual cycle of long-distance migratory pied flycatchers (*Ficedula hypoleuca*) after 21 years of warming. Flycatchers are a model for studying constrained ecological responses to climate change [6, 10, 12, 13]. We show that the phase of the flycatcher circannual clock controlling spring moult, migration, and reproductive timing advanced by 9 days. A nearby wild population mirrored these changes, concurrently advancing egg-laying by 11 days. Furthermore, the time window during which wild flycatcher reproductive timing was most sensitive to ambient temperature advanced by 0.8 d yr^{-1}. These results support a role of phenotypic evolution [14] in changing spring phenology [15, 16]. We suggest that the timing programmes of long-distance migratory birds may have greater adaptive potential than previously thought, leaving some scope for evolutionary rescue in a changing climate.

RESULTS AND DISCUSSION

Replicated experimental study

Changing temperature regimes can impart strong selection pressures on annual cycle timing and migration traits [4, 16-19], which are often heritable [20-23]. However, the extent to which climate adjustment requires evolutionary change depends on an organism's timing strategy [3, 23]. For example, in songbirds, populations that can continuously access information about their reproductive environment (e.g. year-round residents [24]) often show high plasticity. Conversely, many migrants use rigid, inherited circannual programmes to predict suitable conditions over long distances [2, 8-10, 16]. Although these species show some plasticity (e.g., sensitivity to local temperatures [16, 25]), they require evolutionary adjustment of timing and migration traits to keep pace. Currently, it is unclear whether phenotypic evolution, defined as the change in the mean phenotype of a population over successive generations, can match rapid climate change [5, 14, 17, 18, 26]. Evolutionary changes in timing programmes are difficult to detect without experimentation, genetic time series, or longitudinal data from pedigreed populations [5, 17, 19, 22, 24, 26]. Hence, evidence for climate-induced evolution in timing traits is scarce, particularly in vertebrates.

Here, we provide experimental evidence for climate-induced evolution in the annual cycle, from two studies on the pied flycatcher (hereafter "flycatcher") designed by the late Eberhard Gwinner. We investigated the first full annual cycle of flycatchers in replicated studies of cohorts hatched in 1981 [27] and 2002. In this common garden experiment through time [21], nestlings were collected from the same German field site, on the same dates (Figures S1 and S2), and raised in identical captive settings [13, 27]; thus we considered any systematic timing changes between cohorts as evidence of evolutionary

change in the birds' inherited timing programmes. We recorded the timing of annual cycle events and grouped them by season (autumn, winter, spring) [9, 13, 27, 28] (Figure 1).

Based on extensive documentation of climate-induced advances in spring phenology [2, 5, 6, 10, 15, 25, 29], we expected earlier spring timing (end of winter moults, start of migratory restlessness, reproductive activation) in the 2002 cohort compared to 1981. Because flycatchers are protandric and there is a high fitness prime on early male [30] but not necessarily female phenology [18], we expected particularly evident advances in males. In contrast, for autumn, phenological trends and underlying selection pressures are inconsistent for migratory songbirds [31, 32]. We therefore considered advances and delays in autumn phenology (end of post-juvenile body moult and migratory restlessness) to be equally possible, and we did not expect consistent differences between the sexes. Likewise, we had no directional expectation for changes in winter timing (body mass drop, start of winter moults) in either sex. To test these hypotheses, we derived seasonal timing indices by averaging the times of events for each individual. We also examined timing traits individually.

Captive flycatchers showed some evidence of delayed autumn timing. The autumn index averaged later in 2002 by 10 days (95% CI [-1.6,22], χ_1^2 = 2.9, P = 0.087, Figures 2A and 3). This delay was more pronounced in males, although the cohort × sex interaction did not reach statistical significance (χ_1^2 = 2.1, P = 0.15; males: 18 d, 95% CI [2.4,34]; females: 1.2 d, 95% CI [-16,18]). The observed delay was associated with autumn migratory restlessness, which was variable but ended 17 d later in 2002 (95% CI [-7.1,42]; Figures 3 and S3). In contrast, post-juvenile moult, which occurs before migration, ended only slightly later in

2002 (2 d, 95% CI [-1.3,5.4]) and strongly depended on hatch date (0.89 d d^{-1}, 95% CI [0.54,1.3]).

In winter, the sign of phenology changes reversed. The winter timing index averaged 7.7 days earlier in 2002, but the effect was not statistically significant (95% CI [-23,7.8], χ_1^2 = 1.0, P = 0.31, Figures 2B and 3). The observed difference was largely attributable to winter moults, which started 8 to 9 d earlier in 2002 (body moult: -9.1 d, 95% CI [-23,5]; flight feather moult: -8.1 d, 95% CI [-24,7.7]; Figures 3 and S3). In contrast, there was no advance in the timing of body mass drop (0.38 d, 95% CI [-19,20]), which occurs before moult. Winter and autumn timing indices were not correlated (r = 0.18).

Captive flycatchers significantly advanced spring phenology in 2002 relative to 1981 (Figures 2C and 3). The spring timing index averaged 9.3 d earlier in 2002 (95% CI [-16,-2.9], χ_1^2 = 7.3, P = 0.007), and across cohorts males were protandric by 6.4 d (95% CI [-13,-0.18], χ_1^2 = 4, P = 0.045). The spring advance was particularly evident in winter flight feather moult, which terminated 14 days earlier in 2002 (95% CI [-26,-1.8], Figures 3 and S3). End of winter body moult and start of spring migratory restlessness both occurred 4 d earlier in 2002 (body moult: -4.4 d, 95% CI [-13,4.5]; restlessness: -3.8 d, 95% CI [-11,3.6]). The timing of gonadal activation advanced clearly in males (-7.8 d, 95% CI [-15,-0.16]) but not in females (0.34 d, 95% CI [-7.3,8]). There was a correlation between spring and winter timing indices (r = 0.87), but not between spring and autumn (r = 0.16).

Field data from wild flycatchers

To link the replicated laboratory experiment to responses in wild conspecifics, we analysed data from a 46-year field study of nearby breeding flycatchers [33] (Figure S1), testing for

changes in reproductive timing and sensitivity to local ambient temperature [34] (Figure 4). We expected the degree of advance in the spring phenology of captive birds to be comparable to that of wild conspecifics, although wild birds might show additional phenological plasticity of -1 to -2 d $ºC^{-1}$ [16].

Field data indicated that wild flycatchers also advanced spring phenology (Figure 4A). During the interval between captive studies (1981-2002), wild flycatchers commenced egg-laying progressively earlier (slope: -0.53 d yr^{-1}, 95% CI [-0.73, -0.34]), achieving a 11.2-d advance over those 21 years. Over the entire field time series (1973-2018), laydates changed by -0.31 d yr^{-1} (95% CI [-0.39, -0.24]).

Advances of breeding phenology in wild flycatchers were partly explained by ambient temperature on the breeding grounds (Figure 4B). We identified the time window in which mean temperature was most closely associated with laydate using R package *climwin* [34]. This temperature-sensitive window occurred from March 29 to May 13 (1973-2018; Figure S4; *climwin* randomization $P < 0.001$). Mean temperature during this window increased rapidly, by 0.080 ºC yr^{-1} (95% CI [0.012,0.15]) between the captive experiments, and by 0.063 ºC yr^{-1} (95% CI [0.039,0.088]) from 1973-2018 (Figure 4B).

In addition to temperature sensitivity (i.e., phenological plasticity), flycatcher laydates also showed directional change over time. Between captive experiments (1981-2002), flycatcher laydates covaried with temperature by -1.5 d $ºC^{-1}$ (95% CI [-2.6,-0.32]) while advancing at a rate of -0.41 d yr^{-1} (95% CI [-0.6,-0.21]). From 1973-2018, plasticity was identical (-1.5 d $ºC^{-1}$, 95% CI [-2.3,-0.72]), but the rate of annual change was lower (-0.22 d yr^{-1} (95% CI [-0.3,-0.14]). Hence, our study interval captured a particularly strong directional change during a period of rapid warming. These figures fit well with studies of flycatchers in regions with

strongly increasing spring temperature [10, 12, 16] (Figure S4). Accounting for the effect of year was important; a model including temperature as the sole predictor overestimated plasticity (-2.7 d ºC^{-1}, 95% CI [-3.5,-1.9]).

Beyond advancing egg-laying, wild flycatchers advanced the timing of the temperature-sensitive window itself (Figure 4C). The mean date of the best window advanced by -0.83 d yr^{-1} (bootstrapped 95% CI [-1.1,-0.45]) over 24 years. Laydates early in our time series (e.g. 1973-1995) were best explained by breeding-ground temperatures from mid-April to mid-May, while laydates in the later years (e.g. 1996-2018) were best explained by temperatures from late March to early May.

In summary, our captive experiment revealed advances in the timing of spring events that were not likely attributable to plasticity, since flycatcher cohorts monitored in 1981 and 2002 were raised and studied under replicated laboratory conditions. Spring advances of 9 d in captive birds mirrored advances in the laydates of wild birds of 11 d during the same period. We also detected a potential delay in autumn timing, and a tendency of earlier timing in late winter, in the captive birds.

Changes to the timing programme

Our findings suggest that the circannual timing programme of flycatchers has undergone phenotypic evolution. Circannual clocks are inherited [9] and track the time of year, even under constant experimental conditions. Importantly, they regulate organisms' timing responses to environmental factors, in particular photoperiod and ambient temperature [9, 35]. Rapid microevolutionary change in the circannual programme is feasible in songbirds and has been reported in Eurasian blackcaps (*Sylvia atricapilla*) [36, 37]. Timing changes may

advance or delay the entire annual cycle [29], but in our captive flycatchers shifts were season-specific.

In spring, the clear phenology advance of our flycatchers mirrored widely reported shifts in migratory birds [2, 5, 6, 10, 15, 25, 29], which may partly reflect high selection pressures linked to reproduction [4, 17, 18, 30]. Because early departure from African wintering quarters facilitates early arrival on the breeding grounds [38, 39], it is clear how selection for reproductive timing may translate into earlier preparation to depart Africa. Among the contributing traits, the strong advance of flight feather moult is expected because this moult is largely completed before birds migrate, whereas body moult may overlap with migration (Figure 1). A weaker signal for migratory restlessness timing may be due to the large sampling variance of that trait and our small sample size; migratory restlessness is a proxy for wild migratory behaviour, and its timing is difficult to quantify with high precision [8, 40]. Our findings were robust when this trait was excluded to enable larger sample sizes (see Methods). Lastly, as predicted for our protandric species, we found earlier timing in males during reproductive activation [30].

Climate-associated shifts in autumn migration timing have been reported for many avian species, with variable directionality [31, 32]. In Europe, autumn migration generally advanced in trans-Saharan migrants and single-brooded species, whereas shorter-distance and multi-brooded migrants tended to delay [31]. Flycatchers were among the slightly advancing migrants, but recent observations of increasing late, potentially second, broods (Hoffmann, unpubl. data), may indicate shifts to autumnal delays. For winter, data on changing phenology are scarce, but earlier spring departure dates have been reported, for example for Barn swallows (*Hirundo rustica*) in South Africa [39].

Season-specific changes in phenology in our captive flycatchers imply selective modification of the underlying timing programme [8]. One possible mechanism is a change to the photoperiodic response. In spring, increasing daylengths advance the annual cycle and prompt spring phenology in many bird species; in autumn, photoperiodic responses are reversed [9, 35, 41]. A spring advance could be achieved by heightened photoperiodic sensitivity [19, 22, 23], but individuals showed no correlated spring and autumn responses. Furthermore, pied flycatchers and other migrants are largely insensitive to photoperiod in winter, when the flycatchers' phenology advance began [8, 9, 41]. Instead, it is more likely that flycatchers experienced an evolutionary change to the circannual clock itself. By effectively speeding up the clock over winter, the flycatchers' spring phase, and concurrent environmental sensitivity, were reactivated earlier. Such a change could also explain the advance of the climate window in the wild population.

The selective advance of spring timing also argues against alternative interpretations of the differences between cohorts. Major influences of developmental factors, for example date and conditions during hatching, exclusively on spring phenology are unlikely [42, 43], in particular because within cohorts, we found no effect of hatching date on timing after the juvenile phase. Only delayed manifestation of highly specific developmental effects could explain our finding of season-specific timing shifts. Alternatively, cohort differences might have originated from sampling different subsets of the local population. By keeping collection date constant while laydates advanced, chicks collected in 2002 originated from relatively later-laying parents than those in 1981 (Figure S2). However, this scenario predicts a timing delay in the 2002 cohort instead of the advance we observed [18].

An evolutionary response could have taken several routes seen in other taxa: first, the local population could have experienced selection on existing variation. Selection could have changed allele frequencies of genes involved in circannual rhythms and photoperiodic pathways [23, 44], or modified transgenerational epigenetic effects [43, 45]. Second, the population could have experienced introgression by earlier-timed immigrants [7]. A final possibility is random change due to genetic drift. However, spring timing is linked to fitness in flycatchers [12], and captive data paralleled the climate-linked changes in nearby wild flycatchers. This makes selection a more likely explanation [10, 12, 18, 25], potentially aided by assortative mating for timing [37].

There is growing evidence of evolutionary change in timing in response to warmer springs [2, 5, 6, 10, 15, 25, 29]. Several studies have also detected components of spring advancement that are not explained by plasticity [2, 15, 16, 25]. In a comparison of long-term breeding data of four UK songbird species [25], flycatchers were the least sensitive and the only species for which the temporal trend in laydate was significantly more extreme than could be explained by plasticity alone [25]. The authors' interpretation, that microevolution may have compensated for imperfect temperature sensitivity, accords with our findings [25]. Our full-annual cycle data from captive flycatchers identify the putative mechanism of these advancements as accelerated circannual timing during winter, before birds prepare for reproduction [3, 8, 9].

It is promising to observe season-specific change in a species whose ability to keep pace with a shifting climate may depend on its capacity for evolutionary change [8, 10, 37]. Long-distance migrants are in decline and face a myriad of anthropogenic threats. As the earth's climate continues to change, the consequences of failing to keep pace with the seasons

have been well demonstrated; flycatchers are declining most strongly where they are the most mistimed relative to the spring peak in food abundance [12]. However, whether evolutionary change will suffice for flycatchers to keep pace with climate change remains to be determined. Further common garden studies over time could shed light on the evolutionary potential of phenology in a changing world.

Acknowledgments

We thank Ebo Gwinner for planning and carrying out this experiment, Helga Gwinner for her support, and Ninon Ballerstaedt for dedicated calibration of the captivity data. We thank Kyle Horton and Mariëlle van Toor for feedback on figures, the Marshall Commission for funding to BVD, and Ben Sheldon, Wesley Hochachka, Albert Phillimore, Marcel Visser, Jelmer Samplonius, and anonymous referees for helpful input.

Author contributions

BH helped the late Ebo Gwinner with the design and execution of the captive study. UH and DH collected the field data. BVD carried out all analyses, with help from BH. BH and BVD wrote the paper.

Declaration of interests

The authors declare no competing interests.

Main text figure legends

Figure 1. Annual cycle events in first-year captive and wild pied flycatchers. Boxplots (central bars indicate median, boxes show interquartile range) show the timing of annual cycle events for 1981 and 2002 cohorts combined, as well as reproductive events in wild birds (inside circle); see also Figures S1 and S2.

Figure 2. Changes in captive flycatchers. Comparison between the 1981 (blue) and 2002 (green) cohorts in annual cycle timing ((A) autumn; (B) winter; (C) spring). The top timeline shows the median time of each event during the year (pooled across study cohorts). Below (A, B, C) are timing indices calculated by averaging across seasonal traits for each individual. Boxplots show the median as a line, the interquartile range as a box, and data within 1.5 times the interquartile range as whiskers; further points are shown as dots. Illustration from Handbook of Birds of the World [46]; see also Figures S1, S2 and S3.

Figure 3. Differences in timing traits between 1981–2002. Shown are means and one standard error of timing of captive flycatchers, estimated from linear mixed-effects models; see also Figures S1 and S2.

Figure 4. Changes in wild flycatchers. Error shading represents 95% confidence limits. (A–B): Solid lines show slopes for the time period between captive studies (delineated by vertical dotted lines); dashed slopes are over the entire study period. (A) Annual mean laydates. (B) Ambient temperature during the climate-sensitive window preceding flycatcher breeding. (C) Timing of the best model-averaged climate window for flycatcher laydates, identified from 23-year subsets centred on the year indicated on the x-axis. Gray circles show the median date of each window. The horizontal dotted lines indicate the start and end of the

overall best climate window identified using the entire dataset. Shading represents 95% confidence limits from 1000 reruns of *climwin* after bootstrapping flycatcher laydates [47]; see also Figures S1 and S4.

STAR * METHODS

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Barbara Helm (b.helm@rug.nl).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Pied flycatchers

The experimental subjects were pied flycatchers (*Ficedula hypoleuca*). All experimental procedures conformed to the relevant regulatory standards under permit by the state of Upper Bavaria, Germany.

METHOD DETAILS

Description of replication study in captivity

Pied flycatchers from southwest Germany were studied in the Max Planck Institute for Ornithology by the late Eberhard Gwinner in research on circannual rhythms and photoperiodism [9, 13, 27]. The original studies in the 1980s established the annual cycle of

hand-raised birds under different photoperiodic cycles [13, 27]. One experiment in 1981 [27] mimicked the natural daylengths experienced by the birds during breeding, on migration, and in their West African wintering area (10° N). Under these conditions, captive young flycatchers underwent post-juvenile moult soon after independence and started autumn migratory restlessness at very young ages (Main text, Figure 1). Autumn migratory restlessness was often biphasic and extended into early winter. The substantial fat reserves deposited in autumn were also maintained until late winter. Thereafter, the flycatchers lost these fat reserves and undertook prenuptial winter moult of body plumage, as well as tertials and some inner secondaries (collectively "flight feathers"). Moult was followed by the start of spring migratory restlessness and gonadal growth.

In 2002, Gwinner, with help from author BH, replicated this study under identically mimicked conditions [13, 27], although sadly he did not live to see the full results. The goal of the replication was to test for evolutionary change in flycatcher timing since 1981. Gwinner collected flycatchers in 2002 from the same area as in 1981 (Figure S1), hand-raised them in the same way, and tested them under identical conditions, using original lighting devices and both, original and new recording methods.

In 2002, birds were collected in Lahr in southern Germany (48.3° N / 7.3° E; elevation 160 m asl; Figure S1) as nestlings and hand-raised as described earlier [48]. For precise replication, nestlings were collected at similar dates and ages (Figure S2). Because the timing of a bird's hatching may influence the timing of subsequent events in its annual cycle [42], we tested whether hatch date was significantly different between cohorts. There was no detectable difference between the 1982 and 2002 cohorts tested by linear model (effect = 1.61 d, t_{30} = 0.97, P = 0.34), nor between males and females (effect = -2.49 d, t_{30} = -1.59, P = 0.12), and

the interaction term was likewise not significant (effect = 2.08 d, $t_{29} = 0.62$, P = 0.54). In 1981, mean hatch date was ordinal (julian) d 148.5; in 2002, it was d 150. Hatch date was also included in all our models and was not a significant predictor of any timing trait, with the exception of the end of postjuvenile moult. Age at collection also did not differ (8.4 d in 2002; 9.0 d in 1981; t-test, P = 0.52).

Once independent, young birds were kept in individual cages (42x23x23cm) in climate-controlled chambers (ca. 20°C) with light provided by 40-W fluorescent bulbs in the daytime (400 lx at perch level) and by 10-W incandescent bulbs at night (ca. 0.01 lx). Birds were exposed to simulated local daylength until the approximate start of autumn migration. Thereafter, they were progressively shifted to the photoperiodic conditions they would naturally experience *en route* and at their West African wintering areas slightly north of the equator, based on information from ringing and field data [38, 49]. Because birds were thereafter kept at a simulated latitude of 10°N for further study, the end of migratory restlessness, gonadal regression, and post-nuptial moult were not analyzed.

Birds were weighed and checked for moult at least weekly, and every 2-3 d during the post-juvenile period, by the Institute team led by the authors of the original flycatcher study [13]. We checked body moult by inspecting the entire bird and scored presence of moult if we detected feather growth in any of 19 defined body areas. Wing moult was scored for each flight feather of the right wing following [50]. In addition, starting in their first winter, birds of both sexes were assessed for the state of their reproductive development (testis diameter in males, diameter of the largest follicle in females) by laparotomy approximately every three weeks [51].

To quantify the timing of migratory restlessness, we measured activity continuously to identify phases of nocturnal activity [40]. Activity was recorded throughout the study period via microswitches attached to the perches. We then derived the number of 30-min intervals showing any activity during the night (i.e., during the lights-off period, discounting immediate effects of switching on and off of the lights). We analysed the resulting time series of nocturnal activity with a changepoint algorithm that defines the start and end of migratory restlessness [40].

Because the point of our experiment was to investigate whether flycatchers had changed their behaviour compared to the original captivity experiment 21 years ago, we took particular care to ascertain that in 2002 we quantified the birds' behaviour in the same ways as in 1981, and that no systematic measurement bias occurred between replicates. In 2002, microswitch data were collected electronically for all birds by computer-based event recorders. In 1981, the microswitches were attached to an inkwriter (Esterline Angus, Washington USA). The inkwriter recorded activity onto time-charted paper rolls, after which the ink marks were hand-counted by an observer. For each 30 min interval on the recording paper that showed an ink mark during night hours, a bird was scored as "active" for that interval.

In order to minimize differences between the 1981 and 2002 replicates, we carried out two calibration steps of recording methods. The first involved comparing activity recording by inkwriters to those of electronic event recorders. In 2002, in parallel to electronic event recorders, we recorded activity with two Esterline-Angus inkwriters from the original stock, which we moved between cages during the entire recording period. In each cage, birds were recorded simultaneously by both methods for one week, and then the inkwriters were

moved to the next bird, so that 2-3 weeks of comparative data were available for all birds. We then hand-counted the ink recordings for comparison with the parallel electronic recordings. Using a linear mixed-effects model (n = 328 nights of paired recordings), we quantified the methods' repeatability and the mean difference between them: repeatability was high (0.951), and the mean difference was 0.75 (95% CI [0.56,0.95]).

Additionally, we calibrated our hand-counting in 2002 against hand-counting in 1981 using the original ink paper rolls of 5 birds from the 1981 experiment. Our new counts were compared against those noted in the original scoring sheets from 1981 for the same birds. The repeatability (quantified as above) was 0.952 (n = 590 recounted nights). The recounting slightly overestimated activity compared to the original count (mean = 1.01, 95% CI [0.93,1.10]).

Thus, the calibration data indicated close correspondence between the methods. The slight deviations in both steps are expected to partially offset each other. The original observer of ink counts had counted somewhat more conservatively, but the new electronic method, in turn, was slightly more conservative than the inkwriter. Remaining small mean differences between methods were not expected to affect outcomes because we generated bird-specific estimates for start and end of migratory restlessness by changepoint analysis, which uses relative differences in time series [40]. Thus, we are confident that we measured behaviour equivalently in the two replicates.

Description of field study

We obtained field information from three sites, which, like the origin of the captive population, were all located in the Upper Rhine valley (Figure S1; see there for distances).

One site is an active study location of free-living flycatchers [33]. The remaining two sites are weather stations, which framed the flycatcher sites to the north and south within the Rhine valley.

Breeding phenology

To assess changes in local pied flycatcher breeding phenology in the wild during the study period, we used a 46-year dataset from Harthausen near Speyer, Germany (49.3° N / 8.4° E; elevation 105 m asl; Figure S1). From 1973-2018, authors DH and UH collected information on the timing of clutch initiation (laydate), hatching, and breeding success of a population of flycatchers, monitoring 55±14 nests per year, of which we obtained laydate information from 40±15 per year. Data were gathered as part of a ringing study in a nest-box population situated in a mixed coniferous/deciduous woodland at 100 m asl [33]. First arrival of birds was in the first ten days of April (range: 1 to 9 April; data from 15 years). Mean clutch size was 6 eggs, mean incubation period 12 days (12.4±1.73 d; n = 49 nests from 2 years; Hoffmann, unpubl.), and on rare occasions birds were double-brooded. To focus on changes at the start of the breeding season, we followed [16] by only including clutches initiated within 30 days of the mean laydate of the first five nests in a given year. In total, we analysed laydates from 1,834 clutches over 46 years (998 of which occurred in the 21 years spanning the captive studies). In our phenology analyses, we used the mean laydate for each year.

Local ambient spring temperature

We obtained local hourly ambient temperature data from two weather stations in southwest Germany (German weather service, ftp://ftp-

cdc.dwd.de/pub/CDC/observations_germany/; Figure S1): Mannheim (station ID 5906; 49.47ºN, 8.50ºE) and Freiburg (station ID 1443; 48.02ºN, 7.83ºE) from 1973 to 2018. Ambient temperatures of the two stations were closely correlated during the study period (r = 0.96). We averaged the temperature data from these two stations to develop a single regional temperature measure relevant for our flycatcher studies. For missing hourly data points (0.03% of data), we used an exponentially weighted moving average to replace the missing temperature values.

QUANTIFICATION AND STATISTICAL ANALYSIS

Analysis of captivity data

Overall, we compared data from 11 birds from 6 families in 1981 (5 females and 6 males), and 22 birds from 8 families in 2002 (11 females and 11 males). In spring, data were missing for one 2002 bird, and for autumn migratory restlessness, data were missing for three 2002 birds.

Annual cycle timing traits

We compiled data on the timing of moults, migratory restlessness, body mass, and reproductive activation (Main text, Figure 1). The timing of migratory restlessness was quantified from nightly activity profiles as described above. The timing of body mass changes was also quantified using changepoint analysis [40] to determine the date at which a bird shifted from high (winter) to low (spring) body mass states. Moult timing traits were dates of start and end. For the body and flight feather moults, we defined start as the first date on which a given moult was recorded, and end as the last date of recording this moult.

We quantified variation in reproductive timing with weighted averages, weighting each measuring date by gonad size on that date. Thus, birds that showed enlarged gonads earlier in the season were assigned an earlier date, and *vice versa*. We did not include the declining phase of the reproductive cycle.

Because we had season-specific predictions, we analysed timing traits in seasonal blocks. In autumn, our measures included only the end date of post-juvenile body moult and the end date of autumn migratory restlessness. We did not use the start dates because on several occasions these events may have started before data collection began. In winter, we used the start dates of winter moult of body plumage and flight feathers, and the start date of the winter drop in mass. Finally, in spring, we examined the end dates of winter body plumage and flight feather moult, the start date of spring migratory restlessness, and the weighted mean date of gonadal activation as described above.

Model construction and evaluation

We used linear mixed-effects models (lme4 package in R [52]) to test for a difference between cohorts in timing traits during autumn, winter and spring. Because our hypotheses were structured by season and all of our predictors (traits) were in the same units (d, days), we first derived seasonal timing indices by averaging across seasonal traits for each individual. We thus obtained autumn, winter, and spring mean timings for each bird. We could not compare seasonal means for individuals missing data in any trait in a season, so we excluded individuals with missing data. We retained 30 birds in autumn (11 from 1981, 19 from 2002) and 28 in winter (10 from 1981, 18 from 2002). In spring, we had 23 individuals with complete data (8 from 1981, 15 from 2002); an additional 5 did not show any spring migratory restlessness or were not monitored. Therefore, we calculated two

versions of the spring index, one including migratory restlessness but fewer (23) birds, and another version that excluded migratory restlessness but included 28 birds (10 from 1981, 18 from 2002). Both versions produced highly similar results in our analysis. The spring index without migratory restlessness, including the five additional individuals that were missing data, averaged 8.5 d earlier in 2002 (95% CI [-17,-0.49]), compared to 9.3 d earlier (95% CI [-16,-2.9], χ_1^2 = 7.3, P = 0.007) based on the 23 individuals with complete data (see Main text).

The response variables were the seasonal timing indices. We included a random intercept of brood ID (sibgroup) to account for any similarities in timing due to genetic similarities among siblings. The fixed effects were cohort (1981 or 2002), sex, a cohort × sex interaction, and hatch date (to account for any effect of the timing of hatching on subsequent annual cycle timing). To maximize the precision of our estimates given a small sample size, we removed non-cohort fixed effects if they were weakly supported (P > 0.15). We report effect sizes, 95% confidence intervals, and likelihood ratio test P-values for remaining fixed effects. If there was evidence for a cohort × sex interaction, we report separate effects for males and females.

After testing seasonal indices, we repeated the above procedure for each individual timing trait and present effect sizes and confidence intervals for the effect of cohort on these traits. Our goal here was to better understand the drivers of seasonal differences while fully utilizing all data.

Analysis of field data

We tested for change in laydate ($d_{lay)}$ with linear models. For nests where hatchdate (d_{hatch}) but not laydate was recorded, we estimated laydate with the following formula:

$d_{lay} = d_{hatch} - (N_{egg} - 1) - 12$

where N_{egg} is the number of eggs in the complete clutch. The constant 12 reflects the local incubation period.

We used the R package *climwin* [47] to identify the absolute spring time window ("climate window") in which mean ambient temperature at the breeding site most closely predicted breeding phenology (Figure S4). We searched all climate windows of one week or longer in duration, up to 90 days before the last recorded laydate in our dataset (4 June). Searching a large number of climate windows increases the likelihood of a false positive result. Therefore, we used the *randwin* function to create 100 randomized datasets and used the *pvalue* function to determine the probability of discovering the relationship we observed by chance. We determined overall start and end dates by taking an average across models, weighted by the Akaike model weights provided by *climwin*.

We also calculated climate windows for direct comparison with another recent study of phenology in flycatchers [16]. Samplonius et al. (see Figure S4) restricted the length of study years and set their reference date to the average of annual mean laydates (May 2 in our case), searching all possible climate windows at least 15 days in duration between 0-60 days before this date. This was in contrast to the wider search interval (90 d) and the later reference date (June 4) in our analysis, which yielded a larger number of possible windows.

Once we identified these biologically-relevant time windows, we calculated the mean temperature for each year during the window and regressed these values against year to determine the change in temperature during the study period (21 years) and over the entire time series (46 years). We then constructed linear models where the response variable was mean annual laydate. In one model, we included temperature as the sole predictor. In a

second model, we included both temperature and year; this allowed us to test whether there was a significant effect of year while accounting for plasticity in response to temperature, and vice versa. In both models, we weighted each observation by the square root of the number of nests monitored in that year.

Because we expected an advance in the birds' spring phenology, we speculated that the birds' climate-sensitive window itself may have also advanced. We explored the possibility of a shifting window by searching for climate windows across different subsets of study years. Specifically, we used subsets that were 23 years in duration (50% of the years in the study), incremented by one year. For example, we started with the 23-year subset from 1973-1995, then 1974-1996, then 1975-1997, etc., until the final subset of 1996-2018. Therefore, in total we tested 24 different subsets. In this manner, we investigated robustness of the climate window approach to small changes in the choice of study years and searched for any longitudinal trends in identified climate windows. We calculated the slope of change in the median date of the window over time and performed 1000 bootstrapped reanalyses with *climwin* to assess the robustness of the slope to variation in sampled nests.

DATA AND CODE AVAILABILITY

All biological data used in the analysis are available within the article and on the Mendeley Data repository (http://dx.doi.org/10.17632/6n38vwnwc7.1). The weather data are publicly available from the German weather service, ftp://ftp-cdc.dwd.de/pub/CDC/observations_germany).

KEY RESOURCES TABLE

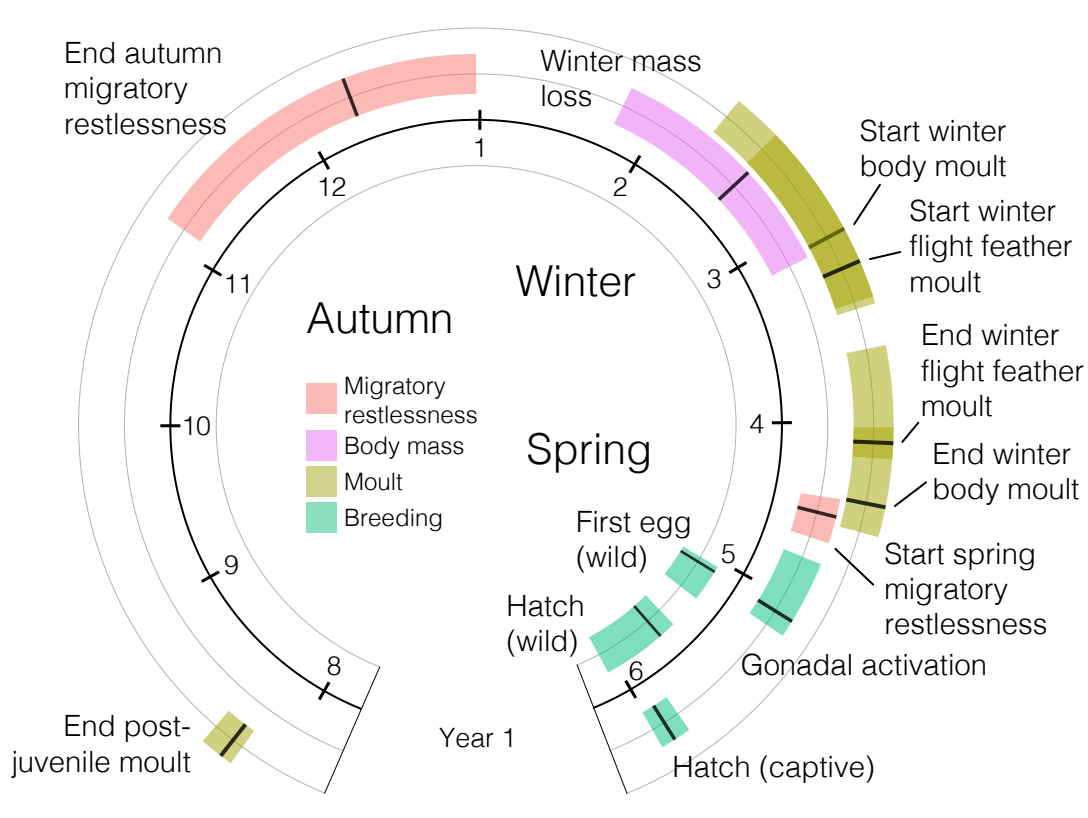
Attached as a separate file.

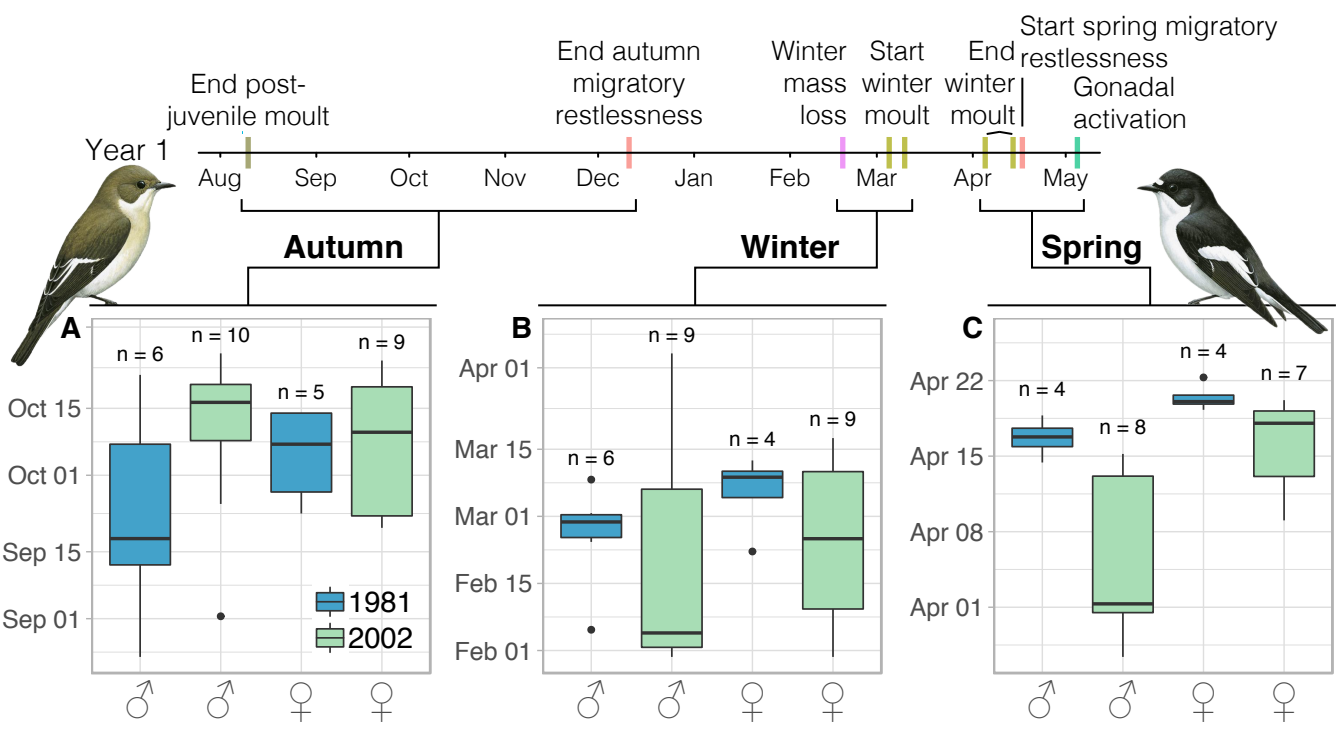
REFERENCES

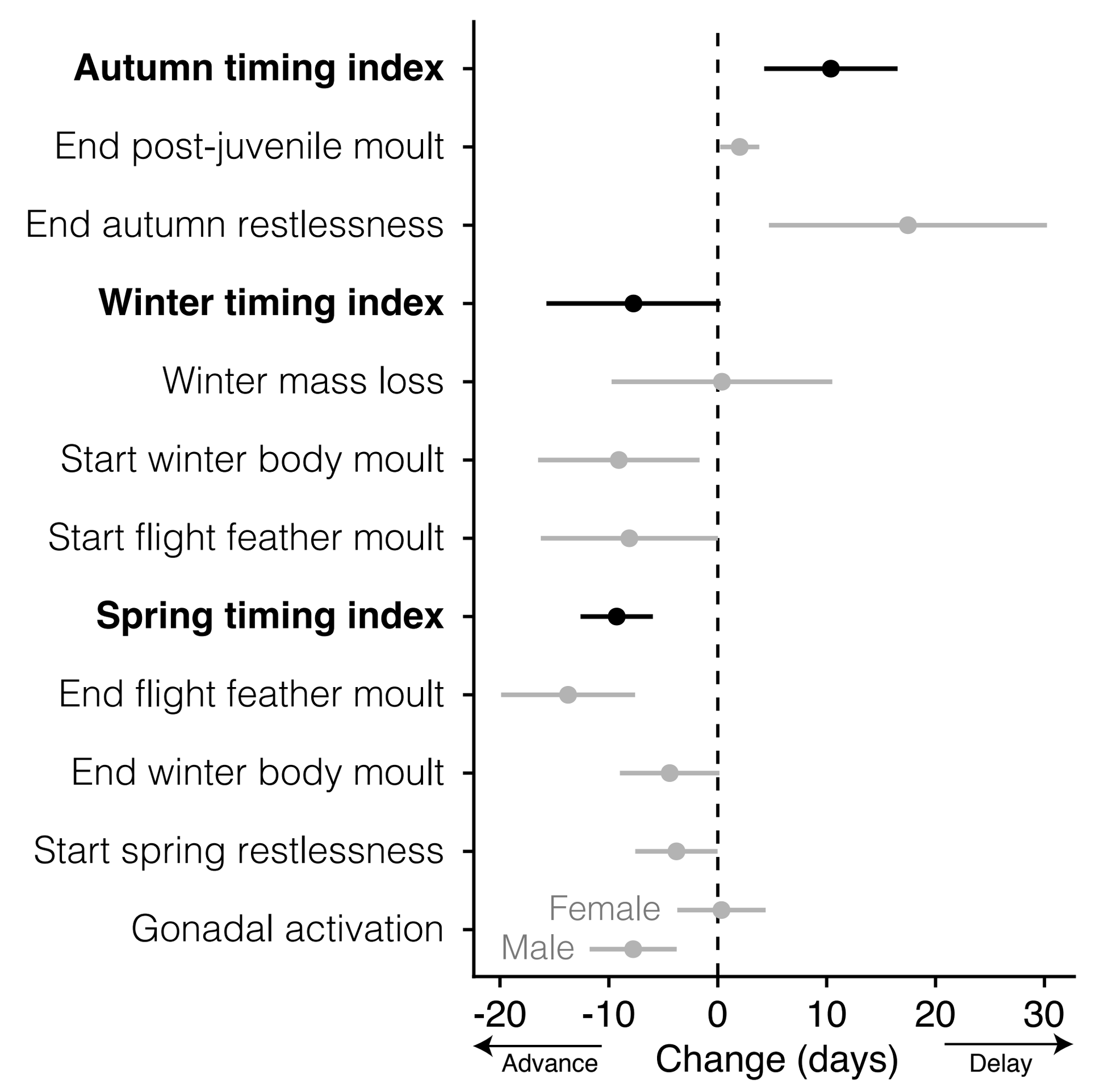
1. Thackeray, S.J., Henrys, P.A., Hemming, D., Bell, J.R., Botham, M.S., Burthe, S., Helaouet, P., Johns, D.G., Jones, I.D., Leech, D.I., et al. (2016). Phenological sensitivity to climate across taxa and trophic levels. Nature *535*, 241-245.
2. Usui, T., Butchart, S.H.M., and Phillimore, A.B. (2017). Temporal shifts and temperature sensitivity of avian spring migratory phenology: a phylogenetic meta-analysis. J Anim Ecol *86*, 250-261.
3. Visser, M.E., Caro, S.P., van Oers, K., Schaper, S.V., and Helm, B. (2010). Phenology, seasonal timing and circannual rhythms: towards a unified framework. Philos Trans R Soc Lond B *365*, 3113-3127.
4. Visser, M.E., and Gienapp, P. (2019). Evolutionary and demographic consequences of phenological mismatches. Nature Ecol Evol *3*, 879–885.
5. Charmantier, A., and Gienapp, P. (2014). Climate change and timing of avian breeding and migration: evolutionary versus plastic changes. Evol Appl *7*, 15-28.
6. Jonzén, N., Lindén, A., Ergon, T., Knudsen, E., Vik, J.O., Rubolini, D., Piacentini, D., Brinch, C., Spina, F., Karlsson, L., et al. (2006). Rapid Advance of Spring Arrival Dates in Long-Distance Migratory Birds. Science *312*, 1959-1961.
7. Visser, M.E. (2008). Keeping up with a warming world; assessing the rate of adaptation to climate change. Proc R Soc Lond B *275*, 649-659.
8. Akesson, S., Ilieva, M., Karagicheva, J., Rakhimberdiev, E., Tomotani, B., and Helm, B. (2017). Timing avian long-distance migration: from internal clock mechanisms to global flights. Philos Trans R Soc Lond B *372*, 1734.
9. Gwinner, E. (1996). Circannual clocks in avian reproduction and migration. Ibis *138*, 47-63.
10. Both, C., and Visser, M.E. (2001). Adjustment to climate change is constrained by arrival date in a long-distance migrant bird. Nature *411*, 296-298.
11. Lameris, T.K., van der Jeugd, H.P., Eichhorn, G., Dokter, A.M., Bouten, W., Boom, M.P., Litvin, K.E., Ens, B.J., and Nolet, B.A. (2018). Arctic Geese Tune Migration to a Warming Climate but Still Suffer from a Phenological Mismatch. Curr Biol. *28*, 2467-2473.e2464.
12. Both, C., Bouwhuis, S., Lessells, C.M., and Visser, M.E. (2006). Climate change and population declines in a long-distance migratory bird. Nature *441*, 81-83.
13. Gwinner, E., and Schwabl-Benzinger, I. (1982). Adaptive temporal programming of molt and migratory disposition in two closely related long-distance migrants, the pied flycatcher (*Ficedula hypoleuca*) and the collared flycatcher (*Ficedula albicollis*). In Avian Navigation, F. Papi and H. Wallraff, eds. (Berlin und Heidelberg: Springer), pp. 75-89.
14. Schlichting, C., and Pigliucci, M. (1998). Phenotypic evolution. A reaction norm perspective, (Sunderland MA: Sinauer).
15. Buskirk, J., Mulvihill, R.S., and Leberman, R.C. (2012). Phenotypic plasticity alone cannot explain climate-induced change in avian migration timing. Ecol Evol *2*, 2430-2437.

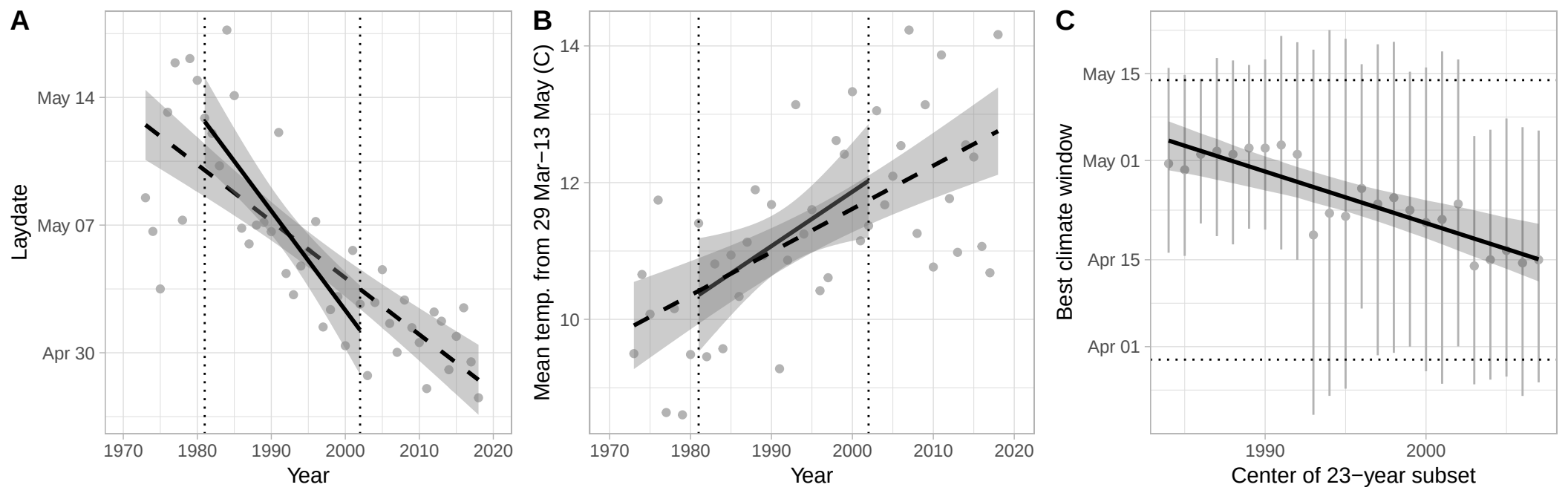
16. Samplonius, J.M., Bartosova, L., Burgess, M.D., Bushuev, A.V., Eeva, T., Ivankina, E.V., Kerimov, A.B., Krams, I., Laaksonen, T., Magi, M., et al. (2018). Phenological sensitivity to climate change is higher in resident than in migrant bird populations among European cavity breeders. Glob Change Biol *24*, 3780-3790.
17. Hansen, M.M., Olivieri, I., Waller, D.M., and Nielsen, E.E. (2012). Monitoring adaptive genetic responses to environmental change. Mol Ecol *21*, 1311-1329.
18. Visser, M.E., Gienapp, P., Husby, A., Morrisey, M., de la Hera, I., Pulido, F., and Both, C. (2015). Effects of Spring Temperatures on the Strength of Selection on Timing of Reproduction in a Long-Distance Migratory Bird. PLoS Biol *13*, e1002120.
19. Bradshaw, W.E., and Holzapfel, C.M. (2001). Genetic shift in photoperiodic response correlated with global warming. Proc Natl Acad Sci *98*, 14509-14511.
20. Pulido, F., and Berthold, P. (2003). Quantitative genetic analysis of migratory behavior. In Avian Migration, P. Berthold, E. Gwinner and E. Sonnenschein, eds. (Berlin Heidelberg: Springer-Verlag Berlin Heidelberg), pp. 53-77.
21. Franks, S.J., Weber, J.J., and Aitken, S.N. (2014). Evolutionary and plastic responses to climate change in terrestrial plant populations. Evol Appl *7*, 123-139.
22. van Asch, M., Salis, L., Holleman, L.J.M., van Lith, B., and Visser, M.E. (2013). Evolutionary response of the egg hatching date of a herbivorous insect under climate change. Nature Clim Change *3*, 244-248.
23. Bradshaw, W.E., and Holzapfel, C.M. (2007). Evolution of animal photoperiodism. Ann Rev Ecol Evol Syst *38*, 1-25.
24. Nussey, D.H., Postma, E., Gienapp, P., and Visser, M.E. (2005). Selection on heritable phenotypic plasticity in a wild bird population. Science *310*, 304-306.
25. Phillimore, A.B., Leech, D.I., Pearce-Higgins, J.W., and Hadfield, J.D. (2016). Passerines may be sufficiently plastic to track temperature-mediated shifts in optimum lay date. Glob Change Biol *22*, 3259-3272.
26. Merilä, J., and Hendry, A.P. (2014). Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. Evol Appl *7*, 1-14.
27. Gwinner, E. (1989). Einfluß der Photoperiode auf das circannuale System des Halsbandschnäppers (*Ficedula albicollis*) und des Trauerschnäppers (*F. hypoleuca)*. J Ornithol *130*, 1-13.
28. Gwinner, E. (1986). Circannual rhythms (Heidelberg, Berlin: Springer).
29. Winkler, D.W., Gandoy, F.A., Areta, J.I., Iliff, M.J., Rakhimberdiev, E., Kardynal, K.J., and Hobson, K.A. (2017). Long-Distance Range Expansion and Rapid Adjustment of Migration in a Newly Established Population of Barn Swallows Breeding in Argentina. Curr Biol *27*, 1080-1084.
30. Samplonius, J.M., and Both, C. (2017). Competitor phenology as a social cue in breeding site selection. J Anim Ecol *86*, 615-623.
31. Jenni, L., and Kéry, M. (2003). Timing of autumn bird migration under climate change: advances in long-distance migrants, delays in short-distance migrants. Proc R Soc Lond B *270*, 1467-1471.
32. Buskirk, J.v., Mulvihill, R., and Leberman, R.C. (2009). Variable shifts in spring and autumn migration phenology in North American songbirds associated with climate change. Glob Change Biol *15*, 760-771.
33. Hoffmann, D., Hoffmann, U., and Flörchinger, M. (2012). Änderungen in der Brutphänologie des Trauerschnäppers (Ficedula hypoleuca). Ergebnisse einer Langzeitstudie (1973-2012) an Höhlenbrütern in Rheinland-Pfalz. Fauna Flora Rheinland-Pfalz *12*, 523-537.
34. Van de Pol, M., Bailey, L.D., McLean, N., Rijsdijk, L., Lawson, C.R., and Brouwer, L. (2016). Identifying the best climatic predictors in ecology and evolution. Methods Ecol Evol *7*, 1246-1257.

35. Helm, B., Schwabl, I., and Gwinner, E. (2009). Circannual basis of geographically distinct bird schedules. J. Exp. Biol. *212*, 1259-1269.
36. Pulido, F., and Berthold, P. (2010). Current selection for lower migratory activity will drive the evolution of residency in a migratory bird population. Proc Natl Acad Sci *107*, 7341-7346.
37. Bearhop, S., Fiedler, W., Furness, R.W., Votier, S.C., Waldron, S., Newton, J., Bowen, G.J., Berthold, P., and Farnsworth, K. (2005). Assortative mating as a mechanism for rapid evolution of a migratory divide. Science *310*, 502-504.
38. Ouwehand, J., and Both, C. (2017). African departure rather than migration speed determines variation in spring arrival in pied flycatchers. J Anim Ecol *86*, 88-97.
39. Altwegg, R., Broms, K., Erni, B., Barnard, P., Midgley, G.F., and Underhill, L.G. (2012). Novel methods reveal shifts in migration phenology of barn swallows in South Africa. Proc R Soc Lond B *279*, 1485-1490.
40. Van Doren, B.M., Liedvogel, M., and Helm, B. (2017). Programmed and flexible: long-term Zugunruhe data highlight the many axes of variation in avian migratory behaviour. J Avian Biol *48*, 155-172.
41. Coppack, T., Tindemans, I., Czisch, M., Van der Linden, A., Berthold, P., and Pulido, F. (2008). Can long-distance migratory birds adjust to the advancement of spring by shortening migration distance? The response of the pied flycatcher to latitudinal photoperiodic variation. Glob Change Biol *14*, 2516-2522.
42. Ouwehand, J., Burger, C., and Both, C. (2017). Shifts in hatch dates do not provide pied flycatchers with a rapid ontogenetic route to adjust offspring time schedules to climate change. Funct Ecol *31*, 2087-2097.
43. Ryu, T., Veilleux, H.D., Donelson, J.M., Munday, P.L., and Ravasi, T. (2018). The epigenetic landscape of transgenerational acclimation to ocean warming. Nature Clim Change *8*, 504-509.
44. Prentice, M.B., Bowman, J., Lalor, J.L., McKay, M.M., Thomson, L.A., Watt, C.M., McAdam, A.G., Murray, D.L., and Wilson, P.J. (2017). Signatures of selection in mammalian clock genes with coding trinucleotide repeats: Implications for studying the genomics of high-pace adaptation. Ecol Evol *7*, 7254-7276.
45. Wilschut, R.A., Oplaat, C., Snoek, L.B., Kirschner, J., and Verhoeven, K.J.F. (2016). Natural epigenetic variation contributes to heritable flowering divergence in a widespread asexual dandelion lineage. Mol Ecol *25*, 1759-1768.
46. del Hoyo, J., Elliott, A., Christie, D.A., and Kirwan, G.M. eds. (2018). Handbook of the Birds of the World Alive, retrieved from http://www.hbw.com/ on 25/07/2018 Edition (Barcelona: Lynx ediciones).
47. Bailey, L.D., and van de Pol, M. (2016). climwin: An R Toolbox for Climate Window Analysis. PLOS ONE *11*, e0167980.
48. Gwinner, E., Neusser, V., Engl, E., Schmidl, D., and Bals, L. (1987). Haltung, Zucht und Eiaufzucht afrikanischer und europäischer Schwarzkehlchen *Saxicola torquata*. Gefiederte Welt *111*, 118-120.
49. Ouwehand, J., Ahola, M.P., Ausems, A.N.M.A., Bridge, E.S., Burgess, M., Hahn, S., Hewson, C.M., Klaassen, R.H.G., Laaksonen, T., Lampe, H.M., et al. (2016). Light-level geolocators reveal migratory connectivity in European populations of pied flycatchers Ficedula hypoleuca. J Avian Biol *47*, 69-83.
50. Newton, I. (1966). The moult of the bullfinch *Pyrrhula pyrrhula*. Ibis *108*, 41-67.
51. Gwinner, E. (1975). Die circannuale Periodik der Fortpflanzungsaktivität beim Star (Sturnus vulgaris) unter dem Einfluß gleich- und andersgeschlechtiger Artgenossen. Z. Tierpsychol. *38*, 34-43.
52. Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. J Statist Software *67*, 1-48.


End autumn migratory restlessness
Winter mass loss
Start winter body moult
Start winter flight feather moult
End winter flight feather moult
End winter body moult
Start spring migratory restlessness
Gonadal activation
Hatch (captive)
Hatch (wild)
First egg (wild)
End post-juvenile moult
Year 1
Autumn
Winter
Spring
Migratory restlessness
Body mass
Moult
Breeding
1
2
3
4
5
6
8
9
10
11
12


End post-juvenile moult
End autumn migratory restlessness
Winter mass loss
Start winter moult
End winter moult
Start spring migratory restlessness
Gonadal activation
Year 1
Aug
Sep
Oct
Nov
Dec
Jan
Feb
Mar
Apr
May
Autumn
Winter
Spring
A
B
C
n = 6
n = 10
n = 5
n = 9
Oct 15
Oct 01
Sep 15
Sep 01
1981
2002
n = 6
n = 9
n = 4
n = 9
Apr 01
Mar 15
Mar 01
Feb 15
Feb 01
n = 4
n = 8
n = 4
n = 7
Apr 22
Apr 15
Apr 08
Apr 01
♂
♂
♀
♀


Autumn timing index
End post-juvenile moult
End autumn restlessness
Winter timing index
Winter mass loss
Start winter body moult
Start flight feather moult
Spring timing index
End flight feather moult
End winter body moult
Start spring restlessness
Gonadal activation
Female
Male
-20
-10
0
10
20
30
Advance
Change (days)
Delay


A
Laydate
May 14
May 07
Apr 30
1970
1980
1990
2000
2010
2020
Year
B
Mean temp. from 29 Mar–13 May (C)
14
12
10
1970
1980
1990
2000
2010
2020
Year
C
Best climate window
May 15
May 01
Apr 15
Apr 01
1990
2000
Center of 23−year subset