Genomic and phenotypic consequences of range expansion and colonisation

Camilla Lo Cascio Sætre

"*If I had to choose, I would rather have birds than airplanes*"
- Charles Lindbergh

Photo by Edward Kluen

Contents

Summary

How species adapt to novel or changed environments is a fundamental question in evolutionary biology. When human activities are the main drivers of the environmental changes, there is an added responsibility to document their effect on contemporary species. Anthropogenic climate change is affecting natural habitats and biodiversity on a massive scale, and many species have responded to climate warming by expanding their range to higher latitudes or altitudes. To counteract the negative impact of habitat destruction and alteration, conservation efforts have also worked on restoring habitat, which in some cases opens up the possibility for colonisation outside of a species' range. Range expansion and colonisation may have major effects on populations in terms of genome-wide differentiation and diversity at range edges, and local adaptation may present itself both phenotypically and genomically.

This thesis explores these processes in the reed warbler *(Acrocephalus scirpaceus)*. The reed warbler has sparked interest in scientific communities for decades, especially its coevolution with the brood-parasitic common cuckoo (*Cuculus canorus*), and in more recent times, its response to climate change. The reed warbler has been described as a climate winner, and its recent range expansion northwards into Fennoscandia is thought to be climate-driven. Interestingly, reed warblers also recently colonised a novel habitat near its southern range edge, following habitat restoration of wetland in Malta. From the beginning of the reed warblers' colonisation of Malta, the birds were monitored and sampled, adding up to a time-series dataset of 19 years. The goal of **Paper I** was to explore this dataset to see how the nascent population developed over time, and to assess the success of the colonisation. In **Paper IA**, we found that the birds had undergone a decrease in body mass through the 19 years, following a trajectory consistent with a population ascending an adaptive peak, an Ornstein–Uhlenbeck process (OU). We corroborated this finding with estimates of individual survival and mean population fitness, suggesting that the reduction in body mass is a local adaptation to the new Maltese environment. Though we only used samples from birds in the breeding season, it is difficult to exclude the possibility of migrants in the dataset, which may inflate body mass estimates. The body mass of resident birds also fluctuates throughout the season. Controlling for possible seasonal sampling effects is therefore important, and in **Paper IB**, we controlled for seasonality in three different ways. With all approaches, the OU-model fits better than the neutral model, further strengthening our original conclusion of adaptive evolution. For this study, we amplified eight microsatellite markers, which we used to estimate effective population size (N_e) and F_{IS}. The molecular N_e estimate corroborated the independent N_e estimate from the OU model, lending support to the model's validity. According to the F_{IS} estimate and compared to published estimates of other populations, the Maltese population shows signs of inbreeding, which could potentially reduce the adaptability of this population in the future. However, our findings document rapid adaptive phenotypic change following colonisation, demonstrating the reed warbler's high potential for adaptation.

To increase our knowledge of the reed warbler's adaptability and population dynamics, and to set the southwards colonisation into context with the range expansion northwards, we wanted to investigate genomic data. As an aid to **Paper III**, and also to provide a tool for numerous future studies, the goal of **Paper II** was to create a high-quality chromosome-length genome assembly of the reed warbler. We utilised three different sequencing methods to achieve this: PacBio, 10x and Hi-C. The genome assembly was carefully curated and annotated, and it will serve as a genomic resource for future studies, thereby taking part in the goal of assembling the genomes of all bird species. The genome assembly enabled us to utilise reference-based mapping of RAD sequences in **Paper III**, and to learn more about specific genes and their function. An interesting additional finding in **Paper II** relates to the synteny between the reed warbler genome and the genomes of closely related species. We found evidence for two independent macrochromosomal fusions in the reed warbler genome which are not present in the great tit (*Parus major*) or garden warbler (*Sylvia borin*) genomes. Further studies should investigate the ubiquity of these fusions and see if they are shared with other members of the *Acrocephalus* genus, and possibly look into the consequences of such fusions.

The goal of **Paper III** was to investigate the genomic consequences of range expansion and colonisation. To achieve this, we used RAD sequencing of 10 populations in Europe from north to south, including both the northern range edge (Finland and Norway) and the southern range edge, using a subset of the Maltese individuals from **Paper I**. We found low levels of population structuring, and no apparent loss in genome-wide diversity at the range edges. We found that isolation by environment is much stronger than isolation by distance amongst the reed warbler populations, and we identified disparity in precipitation variability as the main barrier of gene flow. We found no evidence that the loci involved in population divergence and adaptation in the core populations are the same that are involved in adaptation at the range edges, suggesting partly novel variation at the range edges. Using three genome scan methods to identify signs of selection, we found 49 SNPs putatively under selection, of which 33 were located in introns of 28 unique genes. Most of these are correlated with differences in temperature and precipitation variables. Some genome scan outliers also show signs of being part of nascent selective sweeps. These results suggest that in the reed warbler, contemporary range expansion has had little effect on molecular diversity and has been rapidly followed by local adaptation to climatic conditions, which could further corroborate the rapid pace at which colonisation of novel environments has occurred both northwards and southwards. We found reduced heterozygosity in the Maltese population using microsatellite markers in **Paper I**, indicating inbreeding. We did not find similar results in **Paper III** using RAD-seq, at least not compared to the other populations studied. However, there were high relatedness levels within the population. Continued monitoring is important to ensure the well-being of this recently established population.

This thesis has contributed to the increase of our knowledge of the genomic and phenotypic consequences of range expansion and colonisation, including rapid adaptation to local climate. Further, it has provided a high-quality chromosome-level genome assembly which can be used for various future studies, like studies of comparative genomics, chromosomal rearrangements in birds, or more in-depth studies of the reed warbler.

Acknowledgements

First and foremost, I would like to thank Fabrice, without whom this thesis would not have existed. We started on this path together after you took me on as a research assistant and shared the exciting world of reed warblers with me. I'm so grateful to you! You are always so kind, supportive and positive. Your creativity, enthusiasm and quick and clear thinking are qualities I really admire. I'm also very thankful to my other two supervisors, Kjetill and Ole. Kjetill, you have provided safe and solid support, you are so knowledgeable, and you're always kind and in good humour. My spirits are always lifted after talking with you. Ole, I am in awe of your skills, and you have been a massive help with the bioinformatics parts of the thesis. You also have a wonderfully calming effect on my nerves.

To all my collaborators, I'm deeply grateful, you have significantly improved this thesis. I share the fascination of reed warblers with my Finnish collaborators, Katja, Rose, Nora and Edward. You contributed in a major way with the Finnish reed warblers, and I greatly appreciate all of your invaluable help and input. Katja and Rose, it was wonderful to have you visiting me in Oslo, and the many subsequent Zoom calls with you and Nora have been really constructive. Katja, thanks for all the helpful discussions and the problem-solving we have done together. Rose, your knowledge and brilliant writing skills really made a difference. I'm also very grateful for the other collaborators who provided reed warbler samples and who have been an inspiration in the study of reed warblers: Bård Stokke and Frode Fossøy, Petr Procházka, and the people involved with BirdLife Malta (Charles Coleiro, Martin Austad and Mark Gauci). Special thanks to Petr for your helpful input and kind support. Finally, I'm incredibly thankful to Kjetil Voje and Mark Ravinet for all your help with data analysis. Kjetil, it was wonderful to benefit from your expertise and extensive knowledge of evolution in the early phase of this thesis. Mark, I owe so much of my knowledge of R to you, either directly or by going through your excellent guides, and you have been so generous with your help.

I have absolutely loved being in the field, and I'm extremely lucky to have been part of the Sparrow group's field work. Luckily, there were plenty of reed warblers there as well! I think most biologists can agree field work is the best, but field work in Italy with the Sparrow group is famously "luxury field work". The food, the sun, the beach in the afternoon, and the cosy surroundings relieve any trace of tiredness after a hard day, though the birds are mostly a treat to work with. I'm very grateful to have shared this experience with these amazing people: Glenn, Fabrice, Mark, Melissah, Angélica, Ingvild, Helene, Vitalii and Åsta. I could fill pages and pages with fun anecdotes and inside jokes, but I will limit myself and just allude to one, which always makes me smile: The little bicycle. I'm also very thankful to Severino Vitulano and Matteo Caldarella for making field work possible for us. A special thanks to Matteo for taking us around and being so friendly and open, and to Giuseppe Agnelli for being the best handyman and for speaking Italian very pedagogically with me.

I've been working a lot from home, especially after you-know-what, but looking back to the classes I've taken and taught, and the people I've hung around with or encountered, there are many that deserves to be acknowledged. None mentioned, none forgotten, but I'm very grateful to all co-workers, students, professors, teachers, faculty and others who have helped me or just inspired me. Biologists are a friendly bunch!

To my friends, I'm forever grateful. PP (Linn, Keby, Tonje, Trude and Katrine), I appreciate the group chat more than you know. I can't wait to go on holiday with you! Åsta, I thank the sparrows for bringing us together, and you have become vitally important to me. Helene and Christine, I'm writing this on the train back from Bergen where I've just had the best time with you! You are family, almost entirely literally. Andre, Trym and Karl, you have been so supportive, interested and invested, and extremely fun to hang out with. To all of you, and to my other friends, I appreciate you enormously!

I quite literally could not have done this without my family: Mamma, Pappa, Bestemor, Helene, Tim, Sigurd, Leonie and Henrik. You have supported me in every way possible, been the best at comforting me and backing me up, and you are so fun to hang out with! Pappa, you have inspired me so much and impress me to this day, and I'm so lucky to have you to offer advice, help and support. Mamma, you are the best person I know, and I'm so grateful for everything you have done for me. Helene, I'm so lucky to have you as a sister. I actually can't find the words to say how much I appreciate you (especially without using our sister language). Thanks for the café colloquia and for being my best cheerleader. Tim, thank you for your wonderful support and expert advice.

Jørgen, I have to restrain myself a little before I start to gush about how much you mean to me, but I have to say how extremely grateful I am to have you. Actually, I can thank the PhD for meeting you, and that is without a shadow of doubt the best thing to come out of this PhD. Thank you for being the best emotional support, for being the best desk mate at home, for feeding me delicious food and for making me laugh and smile every single day. Elsker deg for alltid!

A final thanks to the reed warblers, for letting me study you, and for being so interesting!

List of papers

This thesis includes four papers, of which three are published (Paper IA, Paper IB and Paper II) and one is in manuscript form (Paper III). Paper IB is a reply to a correspondence letter, so I will at times refer to Paper IA and Paper IB as Paper I.

Paper IA) Sætre CLC, Coleiro C, Austad M, Gauci M, Sætre GP, Voje KL, Eroukhmanoff F. 2017. Rapid adaptive phenotypic change following colonization of a newly restored habitat. Nature Communications, 8(1), 1-6. doi: https://doi.org/10.1038/ncomms14159

Paper IB) Sætre CLC, Coleiro C, Austad M, Gauci M, Sætre GP, Voje KL, Eroukhmanoff F. 2018. Reply to 'Inconclusive evidence for rapid adaptive evolution'. Nature Communications, 9(1), 1-3. doi: https://doi.org/10.1038/s41467-018-05120-9

Paper II) Sætre CLC, Eroukhmanoff F, Rönkä K, Kluen E, Thorogood R, Torrance J, Tracey A, Chow W, Pelan S, Howe K, Jakobsen KS, Tørresen OK. 2021. A chromosome-level genome assembly of the reed warbler (*Acrocephalus scirpaceus*). Genome Biology and Evolution, 13(9), evab212. doi: https://doi.org/10.1093/gbe/evab212

Paper III) Sætre CLC, Rönkä K, Ravinet M, Vitulano S, Caldarella M, Procházka P, Stokke B, Cuevas A, Bergman N, Thorogood R, Jakobsen KS, Tørresen OK, Eroukhmanoff F. Genomic consequences of range expansion and colonisation in the reed warbler (*Acrocephalus scirpaceus*) (Manuscript)

1   Introduction

1.1   Range expansion

The geographic areas occupied by species naturally change through time, and species may shift, expand or contract their range depending on population dynamics and how they are affected by their surrounding environment. Range expansions have been well documented across taxa through evolutionary times by fossil records, and phylogeographical studies have used molecular markers to trace historical expansions of species. For instance, humans have undergone a massive range expansion starting 45 to 60 thousand years ago from Africa, resulting in human occupation of almost all of the Earth's habitable regions (Henn et al. 2012), and serve as a classical example of range expansion. The reasons for this formidable expansion are not clear, but hypotheses include cultural advances supporting extreme demographic growth and neuro-anatomical changes in the ancestral population, as well as responses to climatic fluctuations at later stages (Henn et al. 2012).

Many studies have focused on population dynamics that took place during the Quaternary and especially after the last glacial maximum (about 20 000 years ago). The ranges of species have shifted repeatedly following major climatic oscillations (Hewitt 2000). During glacial periods when ice sheets and permafrost covered much of Europe and North America, species survived in southern refugia, and in interglacial periods they expanded their ranges northwards (Emerson and Hewitt 2005). Climatic conditions are often crucial in determining where a species is found, and when the opportunity arises that a species' niche becomes available, range expansion may occur.

The ecological niche of a species has been described as an abstract mapping of population dynamics onto an environmental space, where the axes are the abiotic and biotic factors that influence birth and death rates of the species in question (Hutchinson 1978; Holt 2009). A population should persist in a habitat with conditions that are within the species' niche, and may spread to previously uninhabited areas where conditions are similar. Thus, a species can track its original niche, or undergo "niche drift" (Jezkova et al. 2011), but species may also be capable of adapting to a novel niche if driven to it (Mestre et al. 2020; Lustenhouwer and Parker 2022). For example, when climate warmed after the last glacial maximum, it was found that the desert horned lizard *Phrynosoma platyrhinos* expanded its climatic niche towards warmer and drier climates, to enable it to expand its range and persist in a desert habitat (Jezkova et al. 2016).

Factors like population growth, population density and intraspecific competition may drive range expansions (Skellam 1951; McLeod and MacDonald 1997; Excoffier et al. 2009; Van Petegem et al. 2017). High population density and limited breeding opportunities should increase reproductive competition, and drive expansion (Williams et al. 1994). Individual traits such as aggressive behaviour may then also facilitate range expansion (Hudina et al. 2014). Relative brain size is another individual trait which may facilitate expansion (Fristoe

et al. 2017). Interestingly, the evolution of larger brains was found to pre-date and facilitate the colonisation of novel habitats rather than the other way around in several bird species (Fristoe et al. 2017).

The scale and speed of range expansions may differ greatly. Classic diffusion theory depicts a symmetrical expansion from a central point, and the speed of the expansion is determined by the reproductive rate and dispersal in the population (Skellam 1951). However, changing environmental conditions and complex landscapes may create more complicated expansion dynamics (e.g., Urban et al. 2008; Fraser et al. 2015). The range expansion of American mink (*Neovison vison*) in Scotland was found to vary spatially and temporally related to landscape heterogeneity, and suitable habitats were colonised much more quickly than less suitable habitats (Fraser et al. 2015). The speed of the post-glacial range expansion across Europe and North America differed between species and was probably much slower in some regions, particularly in the tropics (Hewitt 2000).

Recent human activities, in particular human-induced climate change, is having a major impact on contemporary ecosystems, and has increased the speed of many range expansions (Chen et al. 2011). Many species are threatened with extinction, and either have to adapt *in situ* or shift their range to persist. Across taxa, species are shifting their range to higher latitudes or altitudes (e.g., Parmesan and Yohe 2003; Thomas 2010), largely matching the documented global warming (Chen et al. 2011). Climate may affect range boundaries directly as births and deaths may be affected by variables such as cold, heat or drought, or indirectly through changes in species interactions or habitat structure (Thomas 2010). However, some species do not follow the expected direction in range expansion due to climate change, which may be due to physiological constraints in the species or different drivers of change such as habitat loss (Chen et al. 2011). One study found that the association between climate and the distribution of 68 of 100 European bird species is to some extent random (Beale et al. 2008). However, the influence of climate was greater in northern areas, and the lack of association in Beale et al. (2008) may reflect a greater human impact on habitats in more southern regions. Range shifts to higher latitudes have been shown to be greatest where the climate has warmed the most (Chen et al. 2011).

To counteract the negative impact of habitat destruction and alteration due to human activities conservation efforts have been made to restore habitat. For example, many pollinators are at risk from habitat loss and the use of pesticides, but restoring habitat and planting flowering shrubs and hedgerows have been found to increase pollinator species richness and abundance (M'Gonigle et al. 2015). Habitat restoration does not always lead to successful colonisations, and habitat quality and proximity to source populations may be important factors in determining colonisation success (WallisDeVries and Ens 2010). However, in some cases habitat restoration open the possibility of small-scale range expansions.

1.2   What limits range expansion?

Range expansion requires dispersal of individuals to occupy new locations, which then have to successfully survive and reproduce there. In some cases, a species range is limited by geographic barriers and may rapidly invade a novel habitat when they are introduced (Sakai et al. 2001; Van Kleunen et al. 2016; Winkler et al. 2019). For example, the cane toad (*Bufo marinus*) was introduced to Australia in 1935 to control crop pests and it subsequently spread throughout tropical and subtropical areas (Phillips et al. 2007). The toads are highly toxic, and Australian predators such as snakes have no evolutionary history with these toxins and die if they attempt to eat them (Smith and Phillips 2006), which likely facilitated their spread. Species may spread into novel habitats through intentional or unintentional human activities (for example through ballast water; Sakai et al. 2001), but normally they spread by natural means.

Dispersal propensities vary greatly between species. An individual's ability to disperse depends on its morphology, size and endurance, but dispersal propensity may also be related to behavioural aspects (Saastamoinen et al. 2018), such as exploratory behaviour, foraging behaviour and migratory restlessness (Zera and Brisson 2012). Dispersal propensities are also determined by population dynamics and species characteristics. A comparison of a variety of bird species found several trends which may explain increased dispersal (Paradis et al. 1998). Species living in wet habitats disperse further than species living in dry habitats, which may be explained by greater patchiness of wet habitats. Migrant species disperse further than resident ones, and dispersal distances are lower among abundant species and among species with large geographical ranges. Within a species, population density may be a major factor in determining dispersal. Density-dependent dispersal may both be positive, if competition induces emigration, or negative if for instance social crowding effects impede free movement (Matthysen 2005). High site fidelity may also impose competition between individuals for limited resources, and if site fidelity is strong, it can prevent the metapopulation from occupying all potential available sites (Matthiopoulos et al. 2005). On the other hand, high site fidelity could aid in sustaining a small population at the range edge (Smith and Summers 2005). However, increased dispersal has been shown to result in accelerating rates of range expansion (Travis et al. 2009). Increased dispersal should be favoured during range expansion for example to get a colonisation advantage, to avoid inbreeding depressions or as a way of bet-hedging that some offspring survive in a patchy environment with different extinction probabilities (Hargreaves and Eckert 2014).

Although a species may be physically able to reach an uninhabited area, they may still be limited by various abiotic and biotic factors such as a lack of suitable habitat, differing climatic conditions, lack of prey items, or novel competitors and natural enemies (e.g., Sexton et al. 2009). Climate has long been regarded as having primary control over terrestrial species' ranges (Merriam 1894; Grinnell 1914), and distribution boundaries are often significantly associated with various climatic variables (Gaston 2003; Sexton et al. 2009). Indeed, this is why species are expected to shift their distributions under climate change.

Species expanding into new habitats will frequently encounter resident competitors. An experimental study using flour beetles *Tribolium castaneum* and *Tribolium confusum*, showed that interspecific competition between these two species significantly slowed down range expansion (Legault et al. 2020).

The expanding population has to maintain high fitness, both in terms of survival and reproduction, and the relative fitness of dispersing individuals will influence subsequent population growth. During range expansion, it is important for a species to maintain their adaptive potential, and species may be able to quickly adapt to novel environments either by adaptive evolution (Bonnet et al. 2022), or by plastic responses (Hoffmann and Sgrò 2011). However, if adaptive potential is not maintained, this may limit range expansion (Colautti et al. 2010; González-Martínez et al. 2017; Szűcs et al. 2017).

It can be difficult to determine what limits a species' range and what drives range expansion, as many complex processes are at work simultaneously. Individual traits such as dispersal ability and aggressive behaviour may be important, but population-level factors such as population density, intra- and interspecific competition or genetic diversity may offer more explanatory power (Crawford and Whitney 2010), and may in turn influence site fidelity and dispersal.

1.3   Possible consequences of range expansion

The genomic consequences of range expansions can be substantial. At the expanding range edge, populations undergo serial founder events, where each new colonisation creates a population bottleneck. When effective population sizes are small and rates of genetic drift are high, genetic diversity such as heterozygosity is expected to be reduced at the expansion front (Austerlitz et al. 1997; Wegmann et al. 2006; Excoffier et al. 2009). In the historic human range expansion out of Africa, it is thought that a very large number of small bottlenecks and limited subsequent gene flow have shaped the observed pattern of decreased genetic diversity along colonisation routes (Prugnolle et al. 2005). A number of empirical studies of contemporary range expansions in various species have also reported a reduction of genetic diversity towards the edge of the expansion front (Hill et al. 2005; Eckert et al. 2008; Colautti et al. 2010; White et al. 2013).

With small population sizes and limited gene flow, the risk of inbreeding and possible inbreeding depressions increases (Frankham et al. 2002). Population fitness may be negatively affected by the expression of recessive deleterious alleles, as well as the increased homozygosity and decrease in genetic diversity. Low levels of genetic diversity is thought to be one of the biggest challenges for populations colonising new habitat (Godefroid et al. 2011), and increased genetic diversity has been found to greatly increase colonisation success

(Crawford and Whitney 2010). Loss of genetic variation may constrain a population from adapting to changes in the environment, increasing the risk of extinction (Keller and Waller 2002; Frankham 2005; Willi et al. 2006). If selection is less efficient, strongly deleterious alleles may persist and reach high frequencies at range edges (Peischl and Excoffier 2015), through the process known as allele surfing (Klopfstein et al. 2006). However, empirical studies of contemporary species introductions indicates that reductions in genetic diversity associated with small population sizes during colonisation is less common than theory would suggest (e.g., Kolbe et al. 2004). Multiple introductions into the same area, or gene flow between populations may reduce the loss of genetic diversity, and mitigate possible inbreeding depressions. In particular, long-distance dispersal may preserve genetic diversity at range edges (Ray and Excoffier 2010; Berthouly-Salazar et al. 2013).

Although gene flow increases genetic diversity, which is essential for local adaptation, theoretical studies suggest that gene flow between different populations may inhibit local adaptation or cause maladaptations at range edges (Kirkpatrick and Barton 1997). When the environment changes rapidly in space, gene flow from an environmentally diverse part of the range may hinder local adaptation. However, gene flow between similar environments may facilitate local adaptation during range expansion by introducing new alleles that are locally beneficial (Sexton et al. 2011). Partially restricted gene flow due to environmental differences has been found to explain patterns of genetic differentiation between populations in many species, and to have positive effects on fitness at range edges (Kottler et al. 2021). Immigrants from climatically different environments may be selected against or mating could be non-random (Sexton et al. 2014).

Another possible consequence of range expansion is an increase in genetic structuring in the metapopulation (Austerlitz et al. 1997; Wegmann et al. 2006; Excoffier et al. 2009). The same processes that decrease genetic diversity in edge populations (bottlenecks, drift, etc.) may also lead to an increase in population structure. Many empirical studies have found this pattern following range expansion (Vandepitte et al. 2012; Swaegers et al. 2015). However, range expansions are not always connected with an increase in population structure; some studies find loss in genetic diversity at range edges, but no populations structuring (e.g., Garroway et al. 2011), whereas others find no loss in genetic diversity at range edges despite clear population structuring (Rózsa et al. 2016; Dayton et al. 2017; Heppenheimer et al. 2018). This suggest that the relationship between range expansion and population structure is complex and can be hard to predict. For instance, population structure in migratory whiskered tern (*Chlidonias hybrida hybrida*) was found to be caused by differences in wintering distribution and seasonal migratory route (Dayton et al. 2017).

At the range edges of an expanding species, strong selection pressure may act on the colonising populations, for example in dispersal capacity (Hill et al. 2011) or related to novel environmental conditions (Hoffmann and Sgrò 2011). Rapid and significant phenotypic changes can take place after colonisation, and its effects are of interest for both scientific and conservation purposes. Adaptive evolution of traits that increase survival and reproduction in new environments will facilitate establishment and increase the number of colonists (García-

Ramos and Rodríguez 2002). Phenotypic plasticity may be important in determining colonisation success (Donohue 2005; Chevin and Lande 2011), as it may allow more colonising individuals to survive and establish a population that can then become locally adapted to the new environmental conditions. Adaptive evolution and phenotypic plasticity are not mutually exclusive, and may also interact (e.g., Kelly 2019). Plasticity can be positively selected during the initial stage of encountering a novel environment (Wang and Althoff 2019), and plastic responses themselves can evolve and contribute to climatic adaptation. For example, in populations of cabbage white butterfly (*Pieris rapae*), plastic responses were found to evolve rapidly depending on different temperature regimes (Kingsolver et al. 2007). One way in which plasticity may evolve is through a process called genetic accommodation (West-Eberhard 2003), where increased plasticity evolves if the most plastic genotype brings phenotypes closest to the new adaptive optimum (Bock et al. 2018; Kelly 2019).

It can be difficult to determine whether changes in traits have evolved (are genetic) or have occurred through plasticity. Some studies fail to prove directly that trait changes have evolved, but find evidence indicative of the change being adaptive. For example, Van Buskirk et al. (2010) found that several bird species have exhibited steadily decreasing body size, consistent with an adaptive response to warmer climate. Common garden experiments, reciprocal transplant experiments or studies using multigenerational pedigree data may provide more direct evidence of evolution (Duckworth and Kruuk 2009; Hereford 2009; Hoffmann and Sgrò 2011). Duckworth and Kruuk (2009) used a multigeneration pedigree and long-term data on lifetime fitness of western bluebirds (*Sialia mexicana*), and found a significant genetic correlation between aggression and dispersal, concordant with consistent selection for coexpression of these behaviours. An experimental study of red flour beetles demonstrated that rapid evolution drove both population growth and expansion speed (Szűcs et al. 2017), and a study of house finches (*Carpodacus mexicanus*) involving cross-foster exchanges of freshly laid eggs found rapid evolution of eggshell structure in response to novel environmental conditions after colonisation (Stein and Badyaev 2011). A recent study found that the amount of additive genetic variance in relative fitness is larger than previously thought, suggesting species should be able to adapt quickly to environmental change (Bonnet et al. 2022).

As natural selection acts on phenotypic traits, changes in allele frequencies of genetic loci underlying those traits can shift the population toward a local optimum. Selection for beneficial alleles may lead to selective sweeps (Storz et al. 2004; Melo et al. 2020). In a selective sweep, newly advantageous alleles as well as the linked neutral alleles become fixed in a population, and local genetic diversity decreases (Hohenlohe et al. 2010). These marks of selection in the genome can be analysed to aid our understanding of species' evolution.

1.4   The value of a genomic resource

Having access to a high-quality genome assembly for a focal species can contribute greatly to increasing our knowledge of many research question regarding its ecology and evolutionary past. A genome assembly can be used for comparative studies to for instance compare genomic structure between species. In the study of genomic consequences of range expansion, a genome assembly provides the opportunity for reference-based mapping of markers to investigate population genomics and genome-wide diversity. Genome-scale data also makes it possible to identify some of the loci responsible for adaptive differences among populations, through different approaches. For instance, differentiation outlier methods identify loci with unusually high genetic differentiation among populations, and genetic-environment association methods identify correlations between local population allele frequencies and local environments (Hoban et al. 2016). Analyses of extended haplotype homozygosity may detect recent selective sweeps in the genome (Sabeti et al. 2002). An annotated genome enables the investigation of the function of genes associated with loci or haplotypes putatively under selection.

From the first projects to sequence and assemble the genomes of model species, the number of published vertebrate genomes has increased greatly in the past decade (Rice and Green 2019). Sequencing costs have decreased drastically in recent times (e.g., Goodwin et al. 2016), making genome sequencing and assembly ever more accessible (Ekblom and Wolf 2014). Still, each sequencing method has its limitations. For instance, Illumina reads are relatively short, but contain few errors (Glenn 2011), whereas sequencing methods creating much longer reads such as PacBio contain more errors (Rhoads and Au 2015). Several sequencing methods may be needed to produce a high-quality genome assembly (Richards and Murali 2015). Combining short Illumina reads with long-read sequencing such as PacBio and chromosome conformation capture techniques, such as Hi-C, may help span and resolve repetitive regions and therefore improve genome assemblies (Elbers et al. 2019).

In this thesis, we sequence and assemble the genome of the reed warbler (*Acrocephalus scirpaceus*) as an aid to study the genomic consequences of range expansion.

1.5   The study system

The Eurasian reed warbler (*Acrocephalus scirpaceus*) is a long-distance migrant passerine nesting in wetland habitats, and wintering in sub-Saharan Africa. They breed mainly in common reed *Phragmites australis* (hereafter "reed") or bulrush (*Typha* sp.), building their nests in between the vegetation above the water. There are four subspecies of the reed warbler, of which the nominal subspecies *A. scirpaceus scirpaceus* is the focal subspecies in this study. This subspecies breeds in most of Europe, Northern Africa and western Asia Minor (Figure 1, Cramp and Brooks 1992; Arbabi et al. 2014). The reed warbler expanded throughout Europe after the last glacial maximum (Arbabi et al. 2014), as temperatures

increased and suitable habitat became available, and has successively expanded its distribution towards the north and north-east into Fennoscandia during the 20th century (Järvinen and Ulfstrand 1980; Røed 1994; Stolt 1999; Brommer et al. 2012). Reed warblers colonised Finland in the 1920s (Järvinen and Ulfstrand 1980), and had already started expanding in Sweden earlier than that (Stolt 1999). In Norway, the reed warbler was first seen in 1937, and first recorded breeding in 1947 (Røed 1994), and has particularly expanded from the 1970's (Shimmings and Øien 2015). This northwards expansion is likely driven by recent climate warming, since recent changes in temperature have increased productivity in the northern range margin (Eglington et al. 2015; Meller et al. 2018), and there is a strong positive relationship between the presence of reed warblers and temperature in the northern part of Europe (Virkkala et al. 2005; Davies 2019). An increase in the amount of reed beds, likely due to eutrophication and climate change, also facilitated the expansion (Altartouri et al. 2014).

Interestingly, reed warblers very recently colonised Malta, which is at the southern border of the reed warbler breeding distribution, with a warm and dry climate. There was a lack of suitable habitat there until the early 1990's, when a wetland area was restored and conserved, containing both reed beds and tamarisk groves (*Tamarix* sp.). From around 1995 and onwards, Malta has sustained a small breeding population of reed warblers.

To what extent the range expansions and new climatic selection pressures have influenced the population structure and the genomic composition of the reed warbler is not yet known. Most previous studies have focused on ecological questions, at most using a few genetic markers. In studies using microsatellites as genetic markers, genetic differentiation has been found to be generally low between reed warbler populations, but moderate levels of differentiation have been connected to for instance migratory behaviour (Procházka et al. 2011) and wing shape (Kralj et al. 2010). Genetic diversity has also found to be similar between different reed warbler populations (Procházka et al. 2011; Arbabi et al. 2014).

Figure 1. Approximate breeding range of the reed warbler in Europe.

2   Aims

The overall aim of this thesis is to study contemporary range expansions. A lot is known about historic range expansions, and many phylogeographical studies have aimed at reconstructing the history of for instance the human expansion out of Africa, but there is still a large knowledge gap of the phenotypic and genomic consequences of contemporary range expansions. The current speed of range expansions due to human-induced factors such as climate change and alterations of habitat may have major consequences for contemporary species.

The reed warbler is an ideal study system to increase our knowledge of contemporary range expansions, since they have expanded their range both northwards and southwards likely due to different human-induced factors, namely climate change and habitat restoration. These expansions differ in scale, and the expansion northwards is more progressive whereas the colonisation southwards is more abrupt. In both cases, I focus on how locally adapted the new populations are and if I detect signals of rapid adaptation at the phenotypic and genomic level, because to be successful, a range expansion needs to improve the overall population fitness.

The goal of **Paper I** was to investigate the colonisation of Malta in terms of phenotypic changes and population fitness. In **Paper II** we created a high-quality chromosome-length

genome assembly of the reed warbler as an aid to **Paper III**, but also to assess the synteny between the reed warbler genome and the genomes of closely related species. The goal of **Paper III** was to investigate the genomic consequences of range expansion and colonisation, looking at population differentiation and diversity and identifying signs of selection.

3   Paper summaries

3.1   Paper I

A) Rapid adaptive phenotypic change following colonization of a newly restored habitat
B) Reply to 'Inconclusive evidence for rapid adaptive evolution'

In **Paper I**, we wanted to investigate potential evolutionary changes through 19 years of study of the reed warblers after they colonised Malta, and to assess the success of the colonisation. In **Paper IA**, we used the average body mass of individuals caught in each year, along with corresponding sample sizes and standard deviations as input in a time series analysis. We compared the goodness of fit of a neutral (unbiased random walk) and an adaptive Ornstein-Uhlenbeck (OU) model to our data.

Body mass is a plastic trait, and though diurnal variation in body mass was limited through our sampling design, we did not account for the seasonal variation. Especially in the beginning and end of the breeding period (May and August), body mass estimates may be higher since the birds gain weight in preparation for migration, and there could be a higher probability of heavy migrant birds being included in the sampling. To correct for seasonal body mass variation and exclude the effect from potential migrants being present in the dataset, we repeated the same analyses, but we first corrected for seasonality in three different ways through ANCOVA models in **Paper IB**. We regressed capture date on the natural logarithm of body mass, having year as a factor. In the first ANCOVA model, capture date was measured as "days away from the centre of the breeding season (July 1st)". In the second model, capture date was implemented as "days after May 1st" as a quadratic term (with the linear term also included in the model). The third model was a linear mixed-effect model with log body mass as the dependent variable, year as a fixed effect and month of capture as a random factor. We used the predicted mean and variance from each model as input in the time-series analysis to compare the neutral and the adaptive model. In both **Paper IA** and **Paper IB**, we found that the adaptive OU model consistently fits our data significantly better that the neutral model, meaning the decrease in body mass is consistent with adaptive evolution. In **Paper IB**, we also ran the time-series analysis on juvenile birds, corrected for seasonality with the first ANCOVA model, and found they had also undergone a decrease in average body mass consistent with adaptive evolution. We further corroborate our results with analyses of ecological and molecular microsatellite data. We estimated individual survival and mean population fitness using field observations and mark-recapture data. In

Paper IA, we found that individual survival is correlated with body mass, and more than half of the variation in mean population fitness is explained by variation in body mass. In **Paper IB**, we investigated the possibility that variation in capture date biased these results in a variety of ways, but we did not find any signs of bias. In **Paper IA**, we used the microsatellite data to estimate effective population size (N_e) and F_{IS}. The molecular N_e estimate corroborated an independent N_e estimate from the OU model, lending support to the OU model's validity. According to the F_{IS} estimate and compared to published estimates of other populations, the Maltese population show signs of inbreeding, which could potentially reduce the adaptability of this population in the future. However, our results demonstrate the reed warbler's high potential for adaptation and to our knowledge, the study constitutes one of the most rapid cases of adaptive evolution documented in the context of habitat restoration.

3.2   Paper II

A chromosome-level genome assembly of the reed warbler (*Acrocephalus scirpaceus*)

The goal of **Paper II** was to create a high-quality chromosome-length genome assembly of the reed warbler, both to provide a tool for numerous future studies, and as an aid to **Paper III**, to utilise the genome assembly to investigate the genomic consequences of range expansion. We used a combination of three different sequencing methods to assemble the reed warbler genome; PacBio, 10x and Hi-C. We first assembled the long-read PacBio sequencing data and then scaffolded this assembly with 10x and Hi-C reads. We polished the assembly with PacBio reads and then with 10x reads for two rounds, and the assembly was then decontaminated and manually curated using the gEVAL browser (Chow et al. 2016; Howe et al. 2021). Curation identified and confirmed 29 autosomes and the Z and W chromosomes, to which 98.6% of the assembly sequences were assigned. The completeness of the assembled genome was found to be high; BUSCO analysis showed that 95.7% of BUSCO genes were complete. We softmasked repeats and annotated the genome assembly using the Comparative Annotation Toolkit (CAT) (Fiddes et al. 2018). We annotated 14,645 protein-coding genes, and a BUSCO analysis of the protein sequences indicated 97.5% completeness. We aligned the reed warbler assembly against the great tit (*Parus major*) and the garden warbler (*Sylvia borin*) genome assemblies, and created circos plots to assess synteny. We found unequivocal evidence of two separate macrochromosomal fusions in the reed warbler genome, in addition to the previously identified fusion between chromosome Z and a part of chromosome 4A in the Sylvioidea superfamily. The genome assembly enabled us to utilise reference-based mapping of RAD sequences in **Paper III**, and to learn more about specific genes and their function. The interesting additional finding of the two macrochromosomal fusions in the reed warbler genome which are not present in the great tit or garden warbler genomes warrant future research.

3.3   Paper III

Genomic consequences of range expansion and colonisation in the reed warbler (*Acrocephalus scirpaceus*)

The results in **Paper I** demonstrated the reed warbler's high adaptive potential following colonisation. To set this southwards colonisation into context with the range expansion northwards, and to increase our knowledge of the reed warbler's adaptability and population dynamics, we wanted to investigate genomic data across its breeding range. The goal of **Paper III** was to investigate the genomic consequences of range expansion and colonisation. To achieve this, we used RAD sequencing of 10 populations in Europe from north to south, including both the northern range edge (Finland and Norway) and the southern range edge, using a subset of the Maltese individuals from **Paper I**. The reads were aligned to the reed warbler genome assembly from **Paper II** and subsequently analysed. We investigated population structure looking at both SNPs and haplotypes, and compared genome-wide diversity between populations. We also assessed the role of selection in divergence between populations across the species range with three different genome scan analyses as well as analyses of extended haplotype homozygosity. We found low levels of population structuring, and no apparent loss in genome-wide diversity at the range edges. We found that isolation by environment is much stronger than isolation by distance amongst the reed warbler populations, and we identified disparity in precipitation variability as the main barrier of gene flow. We found no evidence that the loci involved in population divergence and adaptation in the core populations are the same that are involved in adaptation at the range edges, suggesting partly novel variation at range edges. We found several SNPs putatively under selection, most of which are correlated with differences in temperature and precipitation variables. Some genome scan outliers also show signs of being part of nascent selective sweeps. These results suggest that in the reed warbler, contemporary range expansion has had little effect on molecular diversity and has been rapidly followed by local adaptation to climatic conditions, which could further corroborate the rapid pace at which colonisation of novel environments has occurred both northwards and southwards.

4   Main findings and discussion

4.1   Rapid adaptive phenotypic change following colonisation

Over a period of 19 years (1996–2014) after colonising a restored wetland habitat in Malta, we found that the population of reed warblers had undergone a decrease in body mass, following a trajectory consistent with a population ascending an adaptive peak, a so-called Ornstein–Uhlenbeck (OU) process (**Paper IA** and **Paper IB**). An OU process models how a trait evolves towards a new optimum; the trait shows directional change in the beginning of the time series as the population ascends the adaptive peak, followed by a 'stationary phase' where the trait is subjected to stabilizing selection (Lande 1976). The observed evolutionary trajectory of body mass in our dataset in **Paper IA** followed the expected pattern of an OU process, and the OU model outperformed the neutral model of an unbiased random walk by several bias-corrected Akaike Information Criterion (AICc) units. The relative support for the neutral model with genetic drift as an evolutionary driver was correspondingly weak (<1%). In **Paper IA**, however, we did not account for capture date, and body mass fluctuates throughout the season as the birds breed, moult and accumulate fat to prepare for migration. Also, in the beginning and end of the breeding season (May and August), there is a higher probability that the dataset may include some migrant birds, which are heavier than resident breeding birds. To investigate the possibility that variations in capture date and possible migrants in the dataset may have biased our results, we controlled for capture date in three different ways in **Paper IB**. With all three approaches, the OU model always fitted the data significantly better than the neutral model, further strengthening our original conclusion of adaptive evolution. Interestingly, even juveniles, born and ringed on site, exhibited a decrease in mean annual body mass corrected for capture date over the study period, and these data also showed a much better fit to an OU model than to a neutral model (**Paper IB**). We further corroborated our results with analyses of field observations and mark-recapture data.

The yearly estimates of mean population fitness (proportion of breeding adults) increased as the population approached the adaptive optimum, and more than 58% of the variation in mean population fitness was explained by variation in body mass (**Paper IA**). The mean annual variation in body mass decreased as mean population fitness increased, suggesting that body mass had been under strong selection. Body mass also correlated with our estimate of survival. Recaptured individuals (after at least 21 days) had a significantly lower body mass than individuals never recaptured, and their average body mass was extremely close to the adaptive optimum from the OU model (**Paper IA**). The proportion of birds recaptured and not recaptured were not significantly different in June/July than they were in May/August, suggesting that our survival estimates were not merely artifacts of including migrants (**Paper IB**). Furthermore, juvenile birds recaptured as adults were significantly lighter than the ones not recaptured (**Paper IB**). Most importantly, we also corroborate the validity of the OU model by comparing the estimate of effective population size (N_e) from

the OU model with an independent estimate of N_e based on molecular data, and we found the two estimates were very similar (**Paper IA**).

These results suggest that the reed warbler's colonisation of Malta represent one of the most rapid cases of adaptive evolution ever documented in the context of habitat restoration. Body mass is a plastic trait, but has been found to be significantly heritable in several passerine bird species (Charmantier et al. 2004; Rønning et al. 2007), and to be a highly evolvable trait (Teplitsky et al. 2014). Seasonal variation in body mass alone is an insufficient explanation for the decrease in body mass we observe in the Maltese population (**Paper IB**), and if the trait change were solely due to plasticity, the new population would immediately have reached the new optimal trait value, as the reaction norm for the trait in the founding population would have had to cover the optimal trait value. A recent long-term study of several species of birds and mammals found that most populations harbour much higher levels of additive genetic variance in fitness than previously estimated, indicating that adaptive evolution often operates at a substantial pace from generation to generation (Bonnet et al. 2022). The fact that we found a similar reduction in body mass consistent with adaptive evolution for juveniles as well may seem counter-intuitive, if one assumes that heavier juveniles should have higher survival than lighter juveniles because they are better fed and have greater body reserves. There is on average a positive effect of juvenile body mass on survival in bird species, though less pronounced than for mammal species (Ronget et al. 2018). However, these positive effects may be more pronounced for sedentary birds in colder regions, where higher body reserves may be beneficial during winter (Ringsby et al. 1998), or when food availability is low (Lee et al. 1991). Also, being too heavy may negatively affect flight performance and increase predation risk (Bonter et al. 2013), and perhaps a leaner body aids thermoregulation in the warm and dry Maltese environment (discussed more below).

4.2   Climate-driven northwards range expansion

In **Paper III**, we confirmed the range expansion northwards into Fennoscandia (Järvinen and Ulfstrand 1980; Røed 1994; Stolt 1999; Brommer et al. 2012) with a range expansion statistical analysis. The strongest signals from the analysis identified Norway and Finland as furthest away from the estimated point of origin of the range expansion. Reed warblers colonised Finland in the 1920s (Järvinen & Ulfstrand 1980), and was first recorded breeding in Norway in 1947 (Røed 1994), expanding particularly from the 1970's (Shimmings and Øien 2015). This northwards range expansion is likely climate-driven. Since climate has warmed in Northern Europe (IPCC 2022), and the presence of Phragmites reed beds have increased in the northern range (Altartouri et al. 2015), reed warblers may have more easily been able to track their existing niche. There is a strong positive relationship between the presence of reed warblers and temperature in the northern part of Europe (Virkkala et al. 2005; Davies 2019), and the recent increase in temperature have increased productivity in the northern range margin (Eglington et al. 2015; Meller et al. 2018). Climate change has been found to drive polewards or upwards range shifts across taxa (Parmesan and Yohe 2003; Thomas 2010). Our results also support this hypothesis, since climatic variables of

temperature and precipitation correlated with several loci putatively under selection, and served as the strongest explanatory variables for genomic differentiation among the reed warbler populations (**Paper III**).

Although population structure is weak between the reed warbler populations, Norway and Finland clearly cluster together, and are very little differentiated from each other (**Paper III**). Long-distance dispersal events probably aided the expansion northwards, but climatic differences act as a barrier of gene flow in the reed warbler populations. However, gene flow between similar environments may have facilitated local adaptation during range expansion by introducing new alleles that were locally beneficial (Sexton et al. 2011).

4.3   Climatic barriers of gene flow

Compared to the other populations sampled, Malta is the most genetically different, although levels of population structure are still low (**Paper III**). In **Paper III**, we found that isolation by environment (IBE) best explained genomic differentiation between the populations, to a much higher extent than isolation by distance (IBD). Disparity in precipitation seasonality, meaning how much precipitation fluctuates throughout the year, had by far the highest effect, explaining by itself 34% of the variance in F_{ST}. In comparison, geographic distance explained only 4% of the variance in genome-wide differentiation. We found several loci under selection based on three genome scan analyses, and most of these were correlated with climatic predictors, both temperature variables and in particular precipitation variables (**Paper III**).

Malta clearly has the highest precipitation seasonality of all populations, and it gets very dry in the warmest, driest quarter. In the driest month, there is practically no precipitation. It is probable that climatic variables are strong drivers of local adaptation, which in turn acts as a barrier of gene flow. Isolation by environment has been found to better explain the patterns of genetic differentiation between populations than isolation by distance, and to have positive effects on fitness at range edges (Kottler et al. 2021). Immigrants from climatically different environments may be selected against or mating could be non-random (Sexton et al. 2013). Prolonged periods of drought have been shown to negatively affect productivity in reed warblers in Spain, likely due to decreased availability of insects (Jiménez et al. 2018). However, egg laying and nestling mortality may be negatively affected by heavy rain in the breeding season (Halupka et al. 2008; Vafidis et al. 2016). Both precipitation and temperature have been shown to drive selection in reed warblers, for example in wing length (Nowakowski 2000), migration strategy (Chamorro et al. 2019) and egg coloration (Avilés et al. 2007).

It is likely that climatic conditions were the main drivers of the decrease in body mass found in **Paper I**. Body mass should decrease with increasing temperatures (Lima 1986; Yom-Tov 2001), in accordance with Bergmann's rule. Lower body mass has been interpreted as an

adaptation to warmer climates, as the relatively larger body surface areas of smaller individuals serve as efficient heat dissipators (Yom-Tov 2001). Smaller-bodied individuals are able to retain water at higher temperatures than large-bodied individuals (Gardner et al. 2011). Body mass also correlates strongly with metabolic rate, and in hot climates a lower resting metabolic rate could be advantageous, also in poorer environments. Other ecological factors linked to climatic variation, such as resource availability, may also affect body mass (Eglington et al. 2015). However, rising temperatures have previously been shown to correlate with a decrease in body mass (Salewski et al. 2010; Van Buskirk et al. 2010), which strongly point towards a thermoregulation hypothesis.

4.4   Novel genomic variation at range edges

We found that divergence between the core populations explain essentially nothing of the variation in divergence between edge populations, suggesting it is not the same loci involved in divergence between edge populations as between core populations (**Paper III**). Looking at the first principal component of a PCA of only the genome scan outlier loci, Malta is on the opposite end from Finland and Norway, meaning it is likely that conditions on range edges drives differentiation through selection. One of the outlier loci found by two of the three genome scan analyses and correlating with all climatic variables tested had relatively balanced allele frequencies only in Finland and Norway, whereas the other populations were fixed for one of the alleles. The private allele present in Finland and Norway may be a new mutation, which has arisen either independently in the two populations, or in one and transferred via gene flow. Another possibility is that this variant is present in the range core, but at a very low frequency which has not been captured by our sampling, and that the variant has been picked up by selection during founder events. According to the iHS analysis, the outlier is involved in a recent selective sweep. However, this should be interpreted with caution (discussed in **Paper III**). As an experiment, we used the Hi-C data used for the reed warbler genome assembly from **Paper II** to determine the significance of interactions between SNPs and transcription start sites of genes, and to see if the fraction of significant interactions for the genome scan outlier loci was significantly higher than random sets of loci. However, we did not find such an effect for our outliers.

4.5   Genomic diversity and inbreeding

In **Paper III**, we found very similar levels of genomic diversity amongst our population, meaning the Maltese population has had no apparent loss in genomic diversity after colonisation. However, in **Paper I** we found that Malta had significantly higher levels of inbreeding (F_{IS}) compared to other reed warbler populations. Malta had 12% fewer heterozygotes than expected under Hardy–Weinberg equilibrium, and the average value of 31 sites from Procházka et al. (2011) was <5%. These estimates were based on microsatellite data, and the F_{IS} estimates based on RAD-seq in **Paper III** are all higher than equivalent sites

in Procházka et al. (2011). Microsatellites have a much higher mutation rate and typically produce higher heterozygosity estimates than RAD loci (Sunde et al. 2020). Depending on the ratio of observed and expected number of heterozygotes, F_{IS} estimates may be lower for microsatellites than for RADs, but may also be higher (Hodel et al. 2017). It is difficult to know how much differences in study design affects the results, but it is possible that the high mutation rate in microsatellites allows for earlier detection of possible inbreeding occurring in Malta. We found in **Paper III** that several of the individuals from Malta had high relatedness between each other, which showed up in various analyses of population structure. However, values of allelic richness and nucleotide diversity were consistently high in all populations.

Range expansion should in theory decrease genetic diversity at the range edge due to repeated founder events and strong genetic drift at the expansion front, unless gene flow remains high (Ray et al. 2003; Excoffier 2004; Excoffier et al. 2009). Despite high site fidelity and recruitment rates, reed warblers are also capable of long-distance dispersal events (Paradis et al. 1998), and especially juveniles may disperse far away from their natal site (Bulyuk et al 2000; Mukhin 2004). Long-distance dispersal is thought to be an important factor in preserving genetic diversity at range edges during range expansion (Ray and Excoffier 2010, Berthouly-Salazar et al. 2013, Engler et al. 2015), and our results support this hypothesis. High genetic diversity may increase the adaptive potential in an expanding species, or increase the chance that individuals already are pre-adapted to the new area, thus increasing the success of colonisation (Crawford and Whitney 2010). Plasticity may also boost adaptive potential during range expansion. Reed warblers have shown behavioural plasticity for example in response to rapid changes in brood parasitism risk (e.g., Thorogood and Davies 2013).

4.6 Success of the southern colonisation and northern range expansion

It is likely that Malta was colonised by a large, genetically variable source population in the southern part of the range. Malta was on average identified as a source population by the range expansion statistic analysis in **Paper III**, meaning it was equally close or closer to the estimated point of origin of the range expansion in pairwise comparisons. This probably facilitated the colonisation and their ability to adapt to the novel environment (**Paper I**). The relatively weak population structure and the similar levels of genomic diversity between populations found in **Paper III** suggests high gene flow, or that there has not been enough time for divergence. Site fidelity in reed warblers is high (Paradis et al. 1998; Ceresa et al. 2016), and we found high recruitment rate in the Maltese population in **Paper I**, and considering the restricted size of the wetland in Malta, inbreeding may potentially become an increasing problem in the future.

We found no apparent loss in genome-wide diversity in the northern range edge (**Paper III**). Although none of the populations differed significantly from the others, Norway had the lowest number of observed heterozygotes, and amongst the highest F_{IS} values. The

heterozygosity values for Finland could be slightly inflated by a minor batch effect, but accounting for that, they are most likely not lower than the values for Norway.

4.7   Genome assembly of the reed warbler - additional findings

In addition to utilising the reed warbler genome assembly for **Paper III**, the assembly allowed us to discover unequivocal evidence of two novel macrochromosomal fusions in the reed warbler genome (**Paper II**). The reed warbler chromosome 6 is a fusion of great tit chromosomes 7 and 8, and reed warbler chromosome 8 is a fusion of great tit chromosomes 6 and 9 (Figure 2). Interestingly, these chromosomes are not fused in the garden warbler genome, but correspond to the great tit chromosomes. This suggests that the fusions evolved relatively recently, perhaps at the base of the Acrocephalidae branch within Sylvioidea, but further research is needed to determine this. We furthermore confirm the previously identified neo-sex chromosome (Pala et al. 2012; Sigeman et al. 2020), a fusion between the ancestral chromosome Z and a part of chromosome 4A (according to chromosome naming from the zebra finch). This fusion is thought to have occurred at the base of the Sylvioidea branch (Pala et al. 2012) and is shared with all species of Sylvioidea studied so far (Sigeman et al. 2020). In **Paper III**, we found several outliers with the genome scan analyses that were located on chromosome 6 and 8 (8 outliers out of 49 were located on either of these). We know for instance that the role of inversions in local adaptation may be very important (Barth et al. 2017), especially in the face of gene flow (Schaal et al. 2022).
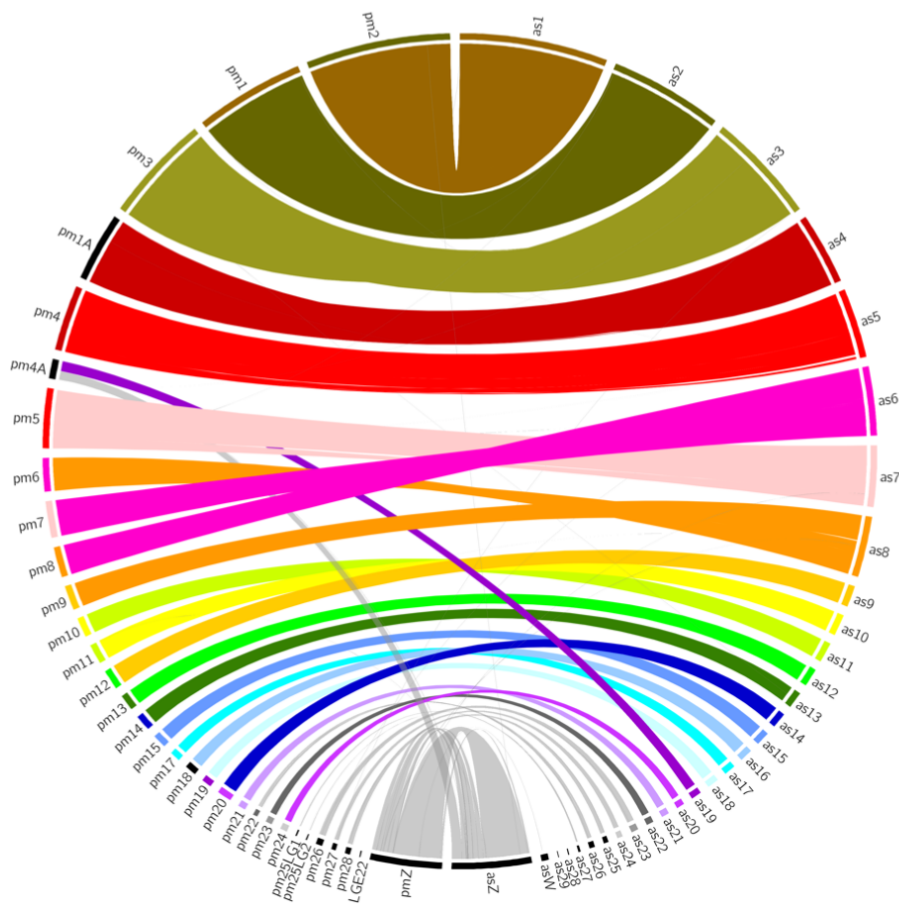
Figure 2. Circos plot showing the synteny between the reed warbler (on the right side, denoted with the prefix as [Acrocephalus scirpaceus]) and the great tit (left side, prefix pm [Parus major]) genome assemblies. The reed warbler chromosome 6 is a fusion of great tit chromosomes 7 and 8, whereas reed warbler chromosome 8 is a fusion of great tit chromosomes 6 and 9 (see **Paper II** for details). The reed warbler chromosome Z corresponds to great tit chromosome Z, and a part of great tit chromosome 4A.

5   Concluding remarks and future perspectives

The reed warbler has provided a unique opportunity to investigate genomic and phenotypic consequences of range expansions of different scales and due to different anthropogenic drivers. This thesis has provided valuable insights for understanding human-mediated range expansion and colonisation in a highly mobile species, and we have studied contemporary evolution both at the phenotypic and genomic level, integrating ecology and evolution in an innovative way.

We found in **Paper I** that the reed warblers had undergone rapid adaptive phenotypic change following colonisation at the southern range edge. It is difficult to definitively prove evolution without an experimental study design or access to pedigree data, but we have shown that adaptive evolution is the most likely explanation for the colonising reed warblers' decrease in body mass observed in **Paper I**. Our results in **Paper III** support the hypothesis

that highly mobile species are likely to retain high levels of genomic diversity at range edges and maintain low population structure (Ray and Excoffier 2010; Berthouly-Salazar et al. 2013). This likely makes the range expansion process more successful, and may have contributed to the rapid adaptation found in **Paper I**.

Our results point to climatic variables as a key driving force of adaptation in the reed warbler, and supports the hypothesis that the northwards range expansion is climate-driven (**Paper III**). The success of range expansions in the reed warbler is likely aided by its high mobility and its retainment of high genome-wide diversity which may have boosted its adaptability. Although gene flow is partly restricted by differences in climate, gene flow from similar environments may have helped during range expansion. Future studies may investigate more directly genes linked to for instance thermoregulation, metabolism or expansion-facilitating traits. We found several candidate genes under selection (**Paper III**), but more research is needed to establish the phenotypic effect of these loci in the reed warbler, and studies of transcriptomics may be needed to see if the loci are expressed.

A high-quality genome assembly for non-model species is a very useful tool for in-depth evolutionary studies, and the reed warbler genome assembly from **Paper II** enabled us to gain important knowledge of the genomic consequences of range expansion in **Paper III**. The reed warbler genome assembly (**Paper II**) will serve as a tool for future studies in for instance population genomics and comparative genomics. Our finding of the fusions in the reed warbler genome (**Paper II**) is interesting and should be investigated at a deeper level. Further research is needed to determine the evolutionary age of these fusions, especially because they are not present in the garden warbler genome, suggesting they are relatively new. Studies could also focus on determining the effects of such fusions on the species.

Acknowledgements

References

Altartouri, A., Nurminen, L., & Jolma, A. (2014). Modeling the role of the close-range effect and environmental variables in the occurrence and spread of *Phragmites australis* in four sites on the Finnish coast of the Gulf of Finland and the Archipelago Sea. Ecology and Evolution, 4(7), 987-1005.

Arbabi, T., Gonzalez, J., Witt, H. H., Klein, R., & Wink, M. (2014). Mitochondrial phylogeography of the Eurasian Reed Warbler *Acrocephalus scirpaceus* and the first genetic record of *A. s. fuscus* in Central Europe. Ibis, 156(4), 799-811.

Austerlitz, F., Jung-Muller, B., Godelle, B., & Gouyon, P. H. (1997). Evolution of coalescence times, genetic diversity and structure during colonization. Theoretical Population Biology, 51(2), 148-164.

Avilés, J. M., Stokke, B. G., Moksnes, A., Røskaft, E., & Møller, A. P. (2007). Environmental conditions influence egg color of reed warblers *Acrocephalus scirpaceus* and their parasite, the common cuckoo *Cuculus canorus*. Behavioral Ecology and Sociobiology, 61(3), 475-485.

Barth, J. M., Berg, P. R., Jonsson, P. R., Bonanomi, S., Corell, H., Hemmer-Hansen, J., ... & Andre, C. (2017). Genome architecture enables local adaptation of Atlantic cod despite high connectivity. Molecular Ecology, 26(17), 4452-4466.

Beale, C. M., Lennon, J. J., & Gimona, A. (2008). Opening the climate envelope reveals no macroscale associations with climate in European birds. Proceedings of the National Academy of Sciences, 105(39), 14908-14912.

Berthouly-Salazar, C., Hui, C., Blackburn, T. M., Gaboriaud, C., Van Rensburg, B. J., Van Vuuren, B. J., & Le Roux, J. J. (2013). Long-distance dispersal maximizes evolutionary potential during rapid geographic range expansion. Molecular Ecology, 22(23), 5793-5804.

Bock, D. G., Kantar, M. B., Caseys, C., Matthey-Doret, R., & Rieseberg, L. H. (2018). Evolution of invasiveness by genetic accommodation. Nature Ecology and Evolution, 2(6), 991-999.

Bonnet, T., Morrissey, M. B., de Villemereuil, P., Alberts, S. C., Arcese, P., Bailey, L. D., ... & Kruuk, L. E. (2022). Genetic variance in fitness indicates rapid contemporary adaptive evolution in wild animals. Science, 376(6596), 1012-1016.

Bonter, D. N., Zuckerberg, B., Sedgwick, C. W., & Hochachka, W. M. (2013). Daily foraging patterns in free-living birds: exploring the predation–starvation trade-off. Proceedings of the Royal Society B: Biological Sciences, 280(1760), 20123087.

Brommer, J. E., Lehikoinen, A., & Valkama, J. (2012). The breeding ranges of Central European and Arctic bird species move poleward. PLoS ONE, 7(9), 1-7.

Bulyuk, V. N., Mukhin, A., Fedorov, V. A., Tsvey, A., & Kishkinev, D. (2000). Juvenile dispersal in Reed Warblers *Acrocephalus scirpaceus* at night. Avian Ecology and Behaviour, 5, 45-63.

Ceresa, F., Belda, E. J., & Monrós, J. S. (2016). Similar dispersal patterns between two closely related birds with contrasting migration strategies. Population Ecology, 58(3), 421-427.

Chamorro, D., Nieto, I., Real, R., & Muñoz, A. R. (2019). Wintering areas on the move in the face of warmer winters. Ornis Fennica, 96(2), 41-54.

Charmantier, A., Kruuk, L. E. B., Blondel, J., & Lambrechts, M. M. (2004). Testing for microevolution in body size in three blue tit populations. Journal of Evolutionary Biology, 17(4), 732-743.

Chen, I. C., Hill, J. K., Ohlemüller, R., Roy, D. B., & Thomas, C. D. (2011). Rapid range shifts of species associated with high levels of climate warming. Science, 333(6045), 1024-1026.

Chevin, L. M., & Lande, R. (2011). Adaptation to marginal habitats by evolution of increased phenotypic plasticity. Journal of Evolutionary Biology, 24(7), 1462-1476.

Chow, W., Brugger, K., Caccamo, M., Sealy, I., Torrance, J., & Howe, K. (2016). gEVAL— a web-based browser for evaluating genome assemblies. Bioinformatics, 32(16), 2508-2510.

Colautti, R. I., Eckert, C. G., & Barrett, S. C. (2010). Evolutionary constraints on adaptive evolution during range expansion in an invasive plant. Proceedings of the Royal Society B: Biological Sciences, 277(1689), 1799-1806.

Cramp, S., & Brooks, D. J. (1992). Handbook of the birds of Europe, the Middle East and North Africa. The birds of the western Palearctic, vol. VI. Warblers (pp. 396-405). Oxford University Press, Oxford, UK.

Crawford, K. M., & Whitney, K. D. (2010). Population genetic diversity influences colonization success. Molecular Ecology, 19(6), 1253–1263.

Davies, J. G. (2019). The role of dispersal in range change in birds. PhD Thesis. University of York, UK.

Dayton, J., Ledwoń, M., Paillisson, J. M., Atamas, N., & Szczys, P. (2017). Genetic diversity and population structure of the Eurasian whiskered tern (*Chlidonias hybrida hybrida*), a species exhibiting range expansion. Waterbirds, 40(2), 105-117.

Donohue, K. (2005). Niche construction through phenological plasticity: life history dynamics and ecological consequences. New Phytologist, 166(1), 83-92.

Duckworth, R. A., & Kruuk, L. E. (2009). Evolution of genetic integration between dispersal and colonization ability in a bird. Evolution: International Journal of Organic Evolution, 63(4), 968-977.

Eckert, C. G., Samis, K. E., & Lougheed, S. C. (2008). Genetic variation across species' geographical ranges: the central–marginal hypothesis and beyond. Molecular Ecology, 17(5), 1170-1188.

Eglington, S. M., Julliard, R., Gargallo, G., van der Jeugd, H. P., Pearce-Higgins, J. W., Baillie, S. R., & Robinson, R. A. (2015). Latitudinal gradients in the productivity of e uropean migrant warblers have not shifted northwards during a period of climate change. Global Ecology and Biogeography, 24(4), 427-436.

Ekblom, R., & Wolf, J. B. (2014). A field guide to whole-genome sequencing, assembly and annotation. Evolutionary Applications, 7(9), 1026-1042.

Elbers, J. P., Rogers, M. F., Perelman, P. L., Proskuryakova, A. A., Serdyukova, N. A., Johnson, W. E., ... & Burger, P. A. (2019). Improving Illumina assemblies with Hi-C and long reads: An example with the North African dromedary. Molecular Ecology Resources, 19(4), 1015-1026.

Emerson, B. C., & Hewitt, G. M. (2005). Phylogeography. Current Biology, 15(10), R367-R371.

Engler, J. O., Secondi, J., Dawson, D. A., Elle, O., & Hochkirch, A. (2016). Range expansion and retraction along a moving contact zone has no effect on the genetic diversity of two passerine birds. Ecography, 39(9), 884-893.

Excoffier, L. (2004). Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. Molecular Ecology, 13(4), 853-864.

Excoffier, L., Foll, M., & Petit, R. J. (2009). Genetic consequences of range expansions. Annual Review of Ecology, Evolution and Systematics, 40, 481-501.

Fiddes, I. T., Armstrong, J., Diekhans, M., Nachtweide, S., Kronenberg, Z. N., Underwood, J. G., ... & Paten, B. (2018). Comparative Annotation Toolkit (CAT)—simultaneous clade and personal genome annotation. Genome Research, 28(7), 1029-1038.

Frankham, R., Ballou, S. E. J. D., Briscoe, D. A., & Ballou, J. D. (2002). Introduction to conservation genetics. Cambridge University Press, Cambridge, UK.

Frankham, R. (2005). Genetics and extinction. Biological Conservation, 126, 131-140.

Fraser, E. J., Lambin, X., Travis, J. M., Harrington, L. A., Palmer, S. C., Bocedi, G., & Macdonald, D. W. (2015). Range expansion of an invasive species through a heterogeneous landscape–the case of American mink in Scotland. Diversity and Distributions, 21(8), 888-900.

Fristoe, T. S., Iwaniuk, A. N., & Botero, C. A. (2017). Big brains stabilize populations and facilitate colonization of variable habitats in birds. Nature Ecology and Evolution, 1(11), 1706-1715.

García-Ramos, G., & Rodríguez, D. (2002). Evolutionary speed of species invasions. Evolution, 56(4), 661-668.

Gardner, J. L., Peters, A., Kearney, M. R., Joseph, L., & Heinsohn, R. (2011). Declining body size: a third universal response to warming?. Trends in Ecology and Evolution, 26(6), 285-291.

Garroway, C. J., Bowman, J., Holloway, G. L., Malcolm, J. R., & Wilson, P. J. (2011). The genetic signature of rapid range expansion by flying squirrels in response to contemporary climate warming. Global Change Biology, 17(5), 1760-1769.

Gaston, K. J. (2003). The structure and dynamics of geographic ranges. Oxford University Press on Demand.

Glenn, T. C. (2011). Field guide to next-generation DNA sequencers. Molecular Ecology Resources, 11(5), 759-769.

Godefroid, S., Piazza, C., Rossi, G., Buord, S., Stevens, A. D., Aguraiuja, R., ... & Vanderborght, T. (2011). How successful are plant species reintroductions?. Biological Conservation, 144(2), 672-682.

González-Martínez, S. C., Ridout, K., & Pannell, J. R. (2017). Range expansion compromises adaptive evolution in an outcrossing plant. Current Biology, 27(16), 2544-2551.

Goodwin, S., McPherson, J. D., & McCombie, W. R. (2016). Coming of age: ten years of next-generation sequencing technologies. Nature Reviews Genetics, 17(6), 333-351.

Grinnell, J. (1914). Barriers to distribution as regards birds and mammals. The American Naturalist, 48(568), 248-254.

Halupka, L., Dyrcz, A., & Borowiec, M. (2008). Climate change affects breeding of reed warblers *Acrocephalus scirpaceus*. Journal of Avian Biology, 39(1), 95-100.

Hargreaves, A. L., & Eckert, C. G. (2014). Evolution of dispersal and mating systems along geographic gradients: implications for shifting ranges. Functional Ecology, 28(1), 5-21.

Henn, B. M., Cavalli-Sforza, L. L., & Feldman, M. W. (2012). The great human expansion. Proceedings of the National Academy of Sciences, 109(44), 17758-17764.

Heppenheimer, E., Brzeski, K. E., Hinton, J. W., Patterson, B. R., Rutledge, L. Y., DeCandia, A. L., ... & vonHoldt, B. M. (2018). High genomic diversity and candidate genes under selection associated with range expansion in eastern coyote (*Canis latrans*) populations. Ecology and Evolution, 8(24), 12641-12655.

Hereford, J. (2009). A quantitative survey of local adaptation and fitness trade-offs. The American Naturalist, 173(5), 579-588.

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. Nature, 405(6789), 907-913.

Hill, J. K., Hughes, C. L., Dytham, C., & Searle, J. B. (2006). Genetic diversity in butterflies: interactive effects of habitat fragmentation and climate-driven range expansion. Biology Letters, 2(1), 152-154.

Hill, J. K., Griffiths, H. M., & Thomas, C. D. (2011). Climate change and evolutionary adaptations at species' range margins. Annual Review of Entomology, 56, 143-159.

Hoban, S., Kelley, J. L., Lotterhos, K. E., Antolin, M. F., Bradburd, G., Lowry, D. B., ... & Whitlock, M. C. (2016). Finding the genomic basis of local adaptation: pitfalls, practical solutions, and future directions. The American Naturalist, 188(4), 379-397.

Hodel, R. G., Chen, S., Payton, A. C., McDaniel, S. F., Soltis, P., & Soltis, D. E. (2017). Adding loci improves phylogeographic resolution in red mangroves despite increased missing data: comparing microsatellites and RAD-Seq and investigating loci filtering. Scientific Reports, 7(1), 1-14.

Hoffmann, A. A., & Sgrò, C. M. (2011). Climate change and evolutionary adaptation. Nature, 470(7335), 479-485.

Hohenlohe, P. A., Phillips, P. C., & Cresko, W. A. (2010). Using population genomics to detect selection in natural populations: key concepts and methodological considerations. International Journal of Plant Sciences, 171(9), 1059-1071.

Holt, R. D. (2009). Bringing the Hutchinsonian niche into the 21st century: ecological and evolutionary perspectives. Proceedings of the National Academy of Sciences, 106(supplement_2), 19659-19665.

Howe, K., Chow, W., Collins, J., Pelan, S., Pointon, D. L., Sims, Y., ... & Wood, J. (2021). Significantly improving the quality of genome assemblies through curation. Gigascience, 10(1), giaa153.

Hudina, S., Hock, K., & Žganec, K. (2014). The role of aggression in range expansion and biological invasions. Current Zoology, 60(3), 401-409.

Hutchinson, G. E. (1978). An Introduction to Population Biology. Yale University Press, New Haven, Connecticut, USA.

Jezkova, T., Oláh-Hemmings, V., & Riddle, B. R. (2011). Niche shifting in response to warming climate after the last glacial maximum: inference from genetic data and niche assessments in the chisel-toothed kangaroo rat (*Dipodomys microps*). Global Change Biology, 17(11), 3486-3502.

Jezkova, T., Jaeger, J. R., Oláh-Hemmings, V., Jones, K. B., Lara-Resendiz, R. A., Mulcahy, D. G., & Riddle, B. R. (2016). Range and niche shifts in response to past climate change in the desert horned lizard *Phrynosoma platyrhinos*. Ecography, 39(5), 437-448.

Jiménez, J., Hernández, J. M., Feliú, J., Carrasco, M., & Moreno-Opo, R. (2018). Breeding in a dry wetland. Demographic response to drought in the common reed-warbler *Acrocephalus scirpaceus*. Ardeola, 65(2), 247-259.

Järvinen, O., & Ulfstrand, S. (1980). Species turnover of a continental bird fauna: Northern Europe, 1850–1970. Oecologia, 46(2), 186-195.

Keller, L. F., & Waller, D. M. (2002). Inbreeding effects in wild populations. Trends in Ecology and Evolution, 17, 230-241.

Kelly, M. (2019). Adaptation to climate change through genetic accommodation and assimilation of plastic phenotypes. Philosophical Transactions of the Royal Society B, 374(1768), 20180176.

Kingsolver, J. G., Massie, K. R., Ragland, G. J., & Smith, M. H. (2007). Rapid population divergence in thermal reaction norms for an invading species: breaking the temperature–size rule. Journal of Evolutionary Biology, 20(3), 892-900.

Kirkpatrick, M., & Barton, N. H. (1997). Evolution of a species' range. The American Naturalist, 150(1), 1-23.

Klopfstein, S., Currat, M., & Excoffier, L. (2006). The fate of mutations surfing on the wave of a range expansion. Molecular Biology and Evolution, 23(3), 482-490.

Kolbe, J. J., Glor, R. E., Rodríguez Schettino, L., Lara, A. C., Larson, A., & Losos, J. B. (2004). Genetic variation increases during biological invasion by a Cuban lizard. Nature, 431(7005), 177-181.

Kottler, E. J., Dickman, E. E., Sexton, J. P., Emery, N. C., & Franks, S. J. (2021). Draining the swamping hypothesis: little evidence that gene flow reduces fitness at range edges. Trends in Ecology and Evolution, 36(6), 533-544.

Kralj, J., Procházka, P., Fainová, D., Patzenhauerová, H., & Tutiš, V. (2010). Intraspecific variation in the wing shape and genetic differentiation of reed warblers *Acrocephalus scirpaceus* in Croatia. Acta Ornithologica, 45(1), 51-58.

Lande, R. (1976). Natural selection and random genetic drift in phenotypic evolution. Evolution, 314-334.

Lee, P. C., Majluf, P., & Gordon, I. J. (1991). Growth, weaning and maternal investment from a comparative perspective. Journal of Zoology, 225(1), 99-114.

Legault, G., Bitters, M. E., Hastings, A., & Melbourne, B. A. (2020). Interspecific competition slows range expansion and shapes range boundaries. Proceedings of the National Academy of Sciences, 117(43), 26854-26860.

Lima, S. L. (1986). Predation risk and unpredictable feeding conditions: determinants of body mass in birds. Ecology, 67(2), 377-385.

Lustenhouwer, N., & Parker, I. M. (2022). Beyond tracking climate: Niche shifts during native range expansion and their implications for novel invasions. Journal of Biogeography.

Matthiopoulos, J., Harwood, J., & Thomas, L. E. N. (2005). Metapopulation consequences of site fidelity for colonially breeding mammals and birds. Journal of Animal Ecology, 74(4), 716-727.

Matthysen, E. (2005). Density-dependent dispersal in birds and mammals. Ecography, 28(3), 403-416.

McLeod, T., & MacDonald, G. (1997). Postglacial range expansion and population growth of *Picea mariana*, *Picea glauca* and *Pinus banksiana* in the western interior of Canada. Journal of Biogeography, 24(6), 865-881.

Meller, K., Piha, M., Vähätalo, A. V., & Lehikoinen, A. (2018). A positive relationship between spring temperature and productivity in 20 songbird species in the boreal zone. Oecologia, 186, 883-893.

Melo, W. A., Vieira, L. D., Novaes, E., Bacon, C. D., & Collevatti, R. G. (2020). Selective sweeps lead to evolutionary success in an Amazonian hyperdominant palm. Frontiers in Genetics, 11, 596662.

Merriam, C. H. (1894). Laws of temperature control of the geographic distribution of terrestrial animals and plants. National Geographic, 6, 229-38.

Mestre, A., Poulin, R., & Hortal, J. (2020). A niche perspective on the range expansion of symbionts. Biological Reviews, 95(2), 491-516.

M'Gonigle, L. K., Ponisio, L. C., Cutler, K., & Kremen, C. (2015). Habitat restoration promotes pollinator persistence and colonization in intensively managed agriculture. Ecological Applications, 25(6), 1557-1565.

Mukhin, A. (2004). Night movements of young reed warblers (*Acrocephalus scirpaceus*) in summer: is it postfledging dispersal?. The Auk, 121(1), 203-209.

Nowakowski, J. J. (2000). Long-term variability of wing length in a population of the Reed Warbler *Acrocephalus scirpaceus*. Acta Ornithologica, 35(2), 173-182.

Pala, I., Naurin, S., Stervander, M., Hasselquist, D., Bensch, S., & Hansson, B. (2012). Evidence of a neo-sex chromosome in birds. Heredity, 108(3), 264-272.

Paradis, E., Baillie, S. R., Sutherland, W. J., & Gregory, R. D. (1998). Patterns of natal and breeding dispersal in birds. Journal of Animal Ecology, 67(4), 518-536.

Parmesan, C., & Yohe, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. Nature, 421(6918), 37-42.

Peischl, S., & Excoffier, L. (2015). Expansion load: recessive mutations and the role of standing genetic variation. Molecular Ecology, 24(9), 2084-2094.

Phillips, B. L., Brown, G. P., Greenlees, M., Webb, J. K., & Shine, R. (2007). Rapid expansion of the cane toad (*Bufo marinus*) invasion front in tropical Australia. Austral Ecology, 32(2), 169-176.

Procházka, P., Stokke, B. G., Jensen, H., Fainová, D., Bellinvia, E., Fossøy, F., ... & Soler, M. (2011). Low genetic differentiation among reed warbler *Acrocephalus scirpaceus* populations across Europe. Journal of Avian Biology, 42(2), 103-113.

Prugnolle, F., Manica, A., & Balloux, F. (2005). Geography predicts neutral genetic diversity of human populations. Current Biology, 15(5), R159-R160.

Ray, N., Currat, M., & Excoffier, L. (2003). Intra-deme molecular diversity in spatially expanding populations. Molecular Biology and Evolution, 20(1), 76-86.

Ray, N., & Excoffier, L. (2010). A first step towards inferring levels of long-distance dispersal during past expansions. Molecular Ecology Resources, 10(5), 902-914.

Rhoads, A., & Au, K. F. (2015). PacBio sequencing and its applications. Genomics, Proteomics and Bioinformatics, 13(5), 278-289.

Rice, E. S., & Green, R. E. (2019). New approaches for genome assembly and scaffolding. Annual Review of Animal Biosciences, 7(1), 17-40.

Richards, S., & Murali, S. C. (2015). Best practices in insect genome sequencing: what works and what doesn't. Current Opinion in Insect Science, 7, 1-7.

Ringsby, T. H., Sæther, B. E., & Solberg, E. J. (1998). Factors affecting juvenile survival in house sparrow *Passer domesticus*. Journal of Avian Biology, 241-247.

Ronget, V., Gaillard, J. M., Coulson, T., Garratt, M., Gueyffier, F., Lega, J. C., & Lemaître, J. F. (2018). Causes and consequences of variation in offspring body mass: Meta-analyses in birds and mammals. Biological Reviews, 93(1), 1-27.

Rózsa, J., Strand, T. M., Montadert, M., Kozma, R., & Höglund, J. (2016). Effects of a range expansion on adaptive and neutral genetic diversity in dispersal limited Hazel grouse (*Bonasa bonasia*) in the French Alps. Conservation Genetics, 17(2), 401-412.

Røed, T. 1994. Rørsanger. In Norsk Fugleatlas, 382–83. Norsk Ornitologisk Forening.

Rønning, B., Jensen, H., Moe, B., & Bech, C. (2007). Basal metabolic rate: heritability and genetic correlations with morphological traits in the zebra finch. Journal of Evolutionary Biology, 20(5), 1815-1822.

Saastamoinen, M., Bocedi, G., Cote, J., Legrand, D., Guillaume, F., Wheat, C. W., ... & del Mar Delgado, M. (2018). Genetics of dispersal. Biological Reviews, 93(1), 574-599.

Sabeti, P. C., Reich, D. E., Higgins, J. M., Levine, H. Z., Richter, D. J., Schaffner, S. F., ... & Lander, E. S. (2002). Detecting recent positive selection in the human genome from haplotype structure. Nature, 419(6909), 832-837.

Sakai, A. K., Allendorf, F. W., Holt, J. S., Lodge, D. M., Molofsky, J., With, K. A., ... & Weller, S. G. (2001). The population biology of invasive species. Annual Review of Ecology and Systematics, 305-332.

Salewski, V., Hochachka, W. M., & Fiedler, W. (2010). Global warming and Bergmann's rule: do central European passerines adjust their body size to rising temperatures?. Oecologia, 162(1), 247-260.

Schaal, S. M., Haller, B. C., & Lotterhos, K. E. (2022). Inversion invasions: when the genetic basis of local adaptation is concentrated within inversions in the face of gene flow. Philosophical Transactions of the Royal Society B, 377(1856), 20210200.

Sexton, J. P., McIntyre, P. J., Angert, A. L., & Rice, K. J. (2009). Evolution and ecology of species range limits. Annual Review of Ecology, Evolution and Systematics, 40(1), 415-436.

Sexton, J. P., Strauss, S. Y., & Rice, K. J. (2011). Gene flow increases fitness at the warm edge of a species' range. Proceedings of the National Academy of Sciences, 108(28), 11704-11709.

Sexton, J. P., Hangartner, S. B., & Hoffmann, A. A. (2014). Genetic isolation by environment or distance: which pattern of gene flow is most common?. Evolution, 68(1), 1-15.

Shimmings, P., & I. J. Øien. 2015. Bestandsestimater for Norske Hekkefugler. Norsk Ornitologisk Forening.

Sigeman, H., Strandh, M., Proux-Wéra, E., Kutschera, V. E., Ponnikas, S., Zhang, H., ... & Hansson, B. (2020). Genomics of an avian neo-sex chromosome reveals the evolutionary dynamics of recombination suppression and sex-linked genes. bioRxiv.

Skellam, J. G. (1951). Random dispersal in theoretical populations. Biometrika, 38(1/2), 196-218.

Smith, J. G., & Phillips, B. L. (2006). Toxic tucker: the potential impact of cane toads on Australian reptiles. Pacific Conservation Biology, 12(1), 40-49.

Smith, R. D., & Summers, R. W. (2005). Population size, breeding biology and origins of Scottish purple sandpipers. British Birds, 98(11), 579.

Stein, L. R., & Badyaev, A. V. (2011). Evolution of eggshell structure during rapid range expansion in a passerine bird. Functional Ecology, 25(6), 1215-1222.

Stolt, B. O. 1999. The Swedish Reed Warbler Acrocephalus scirpaceus Population Estimated by a Capture-Recapture Technique. Ornis Svecica, 35–45.

Storz, J. F., Payseur, B. A., & Nachman, M. W. (2004). Genome scans of DNA variability in humans reveal evidence for selective sweeps outside of Africa. Molecular Biology and Evolution, 21(9), 1800-1811.

Swaegers, J., Mergeay, J., Van Geystelen, A., Therry, L., Larmuseau, M. H. D., & Stoks, R. (2015). Neutral and adaptive genomic signatures of rapid poleward range expansion. Molecular Ecology, 24(24), 6163-6176.

Szűcs, M., Vahsen, M. L., Melbourne, B. A., Hoover, C., Weiss-Lehman, C., & Hufbauer, R. A. (2017). Rapid adaptive evolution in novel environments acts as an architect of population range expansion. Proceedings of the National Academy of Sciences, 114(51), 13501-13506.

Teplitsky, C., & Millien, V. (2014). Climate warming and Bergmann's rule through time: is there any evidence?. Evolutionary Applications, 7(1), 156-168.

Thomas, C. D. (2010). Climate, climate change and range boundaries. Diversity and Distributions, 16(3), 488-495.

Thorogood, R., & Davies, N. B. (2013). Reed warbler hosts fine-tune their defenses to track three decades of cuckoo decline. Evolution, 67(12), 3545-3555.

Travis, J. M., Mustin, K., Benton, T. G., & Dytham, C. (2009). Accelerating invasion rates result from the evolution of density-dependent dispersal. Journal of Theoretical Biology, 259(1), 151-158.

Urban, M. C., Phillips, B. L., Skelly, D. K., & Shine, R. (2008). A toad more traveled: the heterogeneous invasion dynamics of cane toads in Australia. The American Naturalist, 171(3), E134-E148.

Vafidis, J. O., Vaughan, I. P., Jones, T. H., Facey, R. J., Parry, R., & Thomas, R. J. (2016). The effects of supplementary food on the breeding performance of Eurasian reed warblers *Acrocephalus scirpaceus*; implications for climate change impacts. PLoS One, 11(7), e0159933.

Vandepitte, K., Gristina, A. S., De Hert, K., Meekers, T., Roldán-Ruiz, I., & Honnay, O. (2012). Recolonization after habitat restoration leads to decreased genetic variation in populations of a terrestrial orchid. Molecular Ecology, 21(17), 4206-4215.

Van Buskirk, J., Mulvihill, R. S., & Leberman, R. C. (2010). Declining body sizes in North American birds associated with climate change. Oikos, 119(6), 1047-1055.

Van Kleunen, M., Dawson, W., & Maurel, N. (2016). Characteristics of successful alien plants. Invasion Genetics: The Baker and Stebbins Legacy, 40-56.

Van Petegem, K., Moerman, F., Dahirel, M., Fronhofer, E. A., Vandegehuchte, M. L., Van Leeuwen, T., ... & Bonte, D. (2018). Kin competition accelerates experimental range expansion in an arthropod herbivore. Ecology Letters, 21(2), 225-234.

Virkkala, R., Luoto, M., Heikkinen, R. K., & Leikola, N. (2005). Distribution patterns of boreal marshland birds: modelling the relationships to land cover and climate. Journal of Biogeography, 32, 1957-70.

WallisDeVries, M. F., & Ens, S. H. (2010). Effects of habitat quality and isolation on the colonization of restored heathlands by butterflies. Restoration Ecology, 18(3), 390-398.

Wang, S. P., & Althoff, D. M. (2019). Phenotypic plasticity facilitates initial colonization of a novel environment. Evolution, 73(2), 303-316.

Wegmann, D., Currat, M., & Excoffier, L. (2006). Molecular diversity after a range expansion in heterogeneous environments. Genetics, 174(4), 2009-2020.

West-Eberhard, M. J. (2003). Developmental plasticity and evolution. Oxford University Press, Oxford, UK.

White, T. A., Perkins, S. E., Heckel, G., & Searle, J. B. (2013). Adaptive evolution during an ongoing range expansion: the invasive bank vole (*Myodes glareolus*) in Ireland. Molecular Ecology, 22(11), 2971-2985.

Willi, Y., Van Buskirk, J., & Hoffmann, A. A. (2006). Limits to the adaptive potential of small populations. Annual Review of Ecology, Evolution, and Systematics, 37, 433-458.

Williams, D. A., Lawton, M. F., & Lawton, R. O. (1994). Population growth, range expansion, and competition in the cooperatively breeding brown jay, *Cyanocorax morio*. Animal Behaviour, 48(2), 309-322.

Winkler, D. E., Chapin, K. J., François, O., Garmon, J. D., Gaut, B. S., & Huxman, T. E. (2019). Multiple introductions and population structure during the rapid expansion of the invasive Sahara mustard (*Brassica tournefortii*). Ecology and Evolution, 9(14), 7928-7941.

Yom-Tov, Y. (2001). Global warming and body mass decline in Israeli passerine birds. Proceedings of the Royal Society of London. Series B: Biological Sciences, 268(1470), 947-952.

Zera, A. J. & Brisson, J. A. (2012). Quantitative, physiological, and molecular genetics of dispersal/migration. In Dispersal Ecology and Evolution (eds J. Clobert, M. Baguette, T. G. Benton and J. M. Bullock), pp. 63–82. Oxford University Press, Oxford, UK.

Paper IA

ARTICLE

Rapid adaptive phenotypic change following colonization of a newly restored habitat

Camilla Lo Cascio Sætre[1], Charles Coleiro[2], Martin Austad[2], Mark Gauci[2], Glenn-Peter Sætre[1], Kjetil Lysne Voje[1] & Fabrice Eroukhmanoff[1]

Real-time observation of adaptive evolution in the wild is rare and limited to cases of marked, often anthropogenic, environmental change. Here we present the case of a small population of reed warblers (*Acrocephalus scirpaceus*) over a period of 19 years (1996–2014) after colonizing a restored wetland habitat in Malta. Our data show a population decrease in body mass, following a trajectory consistent with a population ascending an adaptive peak, a so-called Ornstein–Uhlenbeck process. We corroborate these findings with genetic and ecological data, revealing that individual survival is correlated with body mass, and more than half of the variation in mean population fitness is explained by variation in body mass. Despite a small effective population size, an adaptive response has taken place within a decade. A founder event from a large, genetically variable source population to the southern range margin of the reed warbler distribution likely facilitated this process.

---

[1] Centre for Ecological and Evolutionary Synthesis, Department of Biosciences, University of Oslo, P. O. Box 1066 Blindern, N-0316 Oslo, Norway. [2] BirdLife Malta, Xemxija Waterfront Apartments, Flat ½, Triq Is-Simar, Xemxija SPB 9025, Malta. Correspondence and requests for materials should be addressed to F.E. (email: fabrice.eroukhmanoff@ibv.uio.no).

Contemporary evolution, or evolutionary changes observable over less than a few hundred generations, has been documented for a variety of species[1,2]. The threespine stickleback (*Gasterosteus aculeatus*), for example, has shown extraordinary abilities to undergo adaptive evolution within a few decades following natural colonization of novel habitats[3,4]. Evolutionary changes are often associated with shifts in the environment, as changes in adaptive optima cause new selective pressures to operate, resulting in phenotypic evolution[5,6]. These shifts in the environment are often brought on by climate change and various human activities, which have detrimental effects on many species[7,8]. Contemporary evolution is therefore often connected with conservation biology[1,9]. Many different conservation strategies are put to use, such as habitat restoration and assisted colonization, though they may not always work as well as intended[10]. When a population relocates to a new habitat, which may constitute a shift from its geographic range, new selection pressures may cause unforeseen evolutionary changes. One of the biggest challenges for populations colonizing new or restored habitat is their relatively low level of genetic diversity[11]. When effective population sizes are small, genetic variation is rapidly lost due to high rates of genetic drift, and the risk of inbreeding increases[12]. Loss of genetic variation may in return constrain a population from adapting to changes in the environment, increasing the risk of extinction[13–15]. The extent to which natural populations are able to colonize and rapidly adapt to novel habitats in connection with conservation projects has rarely been studied[16].

In this study, we follow a population of reed warblers (*Acrocephalus scirpaceus*) in Is-Simar nature reserve (Fig. 1a) in Malta from 1996 to 2014. Is-Simar was originally a waterlogged marsh, but was drained for agricultural purposes, and later used as a dump. In 1992, it became a special protection area belonging to the NATURA 2000 network and Birdlife Malta began transforming it into a nature reserve of $0.58 \, km^2$ by restoring this lost wetland habitat. Is-Simar currently contains the highest number of *Phragmites* reed beds of the Maltese islands, in addition to tamarisk groves (*Tamarix* sp.), which are both suitable breeding habitats for reed warblers (Fig. 1b). Indeed, it was rapidly visited by numerous reed warbler migrants, and within 2 years, they had established a small breeding population. We show that the population has adapted to this restored habitat very rapidly, despite a small population size. There has been a decrease in body mass in the population, which correlates with an increase in mean population fitness and higher individual survival. Our results have important implications for conservation biology and evolution, especially regarding the potential success of habitat restoration in relation to a species' ability to rapidly adapt to a new environment.

Results

Time-series analysis. Through the course of 19 years, the population has gone through a decrease in body mass consistent with a model of adaptive evolution (Fig. 1c). The adaptive nature of this trend is strongly supported, as it fits an Ornstein–Uhlenbeck (OU) process[17] that outperforms the neutral model of an unbiased random walk by several Akaike Information Criterion (AICc) units (Table 1). The relative support for the neutral model with genetic drift as an evolutionary driver is correspondingly weak (<1%). An OU process models how a trait evolves towards a new optimum; the trait shows directional change in the beginning of the time series as the population ascends the adaptive peak, followed by a 'stationary phase' where the trait is subjected to stabilizing selection. The trait fluctuates around the optimum due to genetic drift, plasticity and

unmeasured direct and indirect selective forces acting on the trait. The observed evolutionary trajectory of body mass follows the expected pattern of a population ascending an adaptive peak (Fig. 1c). Initially, the decrease in mean body mass is substantial, but the changes become progressively smaller and more non-directional when the population is close to the new optimum.

The rate of evolutionary change that occurred before the population reached the adaptive optimum is substantial (1.478 haldanes over the first 6 years). This corresponds to a very rapid change, but it is within the normal range of evolutionary rates measured in populations affected by human-induced environmental changes[9]. The alpha parameter (α) from the OU model (Table 1) represents the strength of the restraining force around the optimum[6]. This parameter can be used to express the phylogenetic half-life (ln $(2)/\alpha$), which is the expected amount of time it takes for the population to evolve halfway from the ancestral state to the optimal phenotype[18]. In our case, the phylogenetic half-life is 1.76, meaning that the population is estimated to have evolved halfway to the optimal body mass in >2 years. This is, as far as we know, the fastest rate of adaptation ever recorded using time series data.

The estimated adaptive optimum was reached after ~7 years (Fig. 1c). This gradual change in the initial phase of directional selection followed by stabilization around the new optimum suggests that the population has undergone adaptive evolution, as this pattern would not be expected by adaptive plasticity alone. If the trait change were solely due to plasticity, the new population would immediately have reached the new optimal trait value, as the reaction norm for the trait in the founding population would have had to cover the optimal trait value. However, without direct evidence of heritability, the heritable and environmental components of phenotypic change cannot be partitioned definitively. While some studies on contemporary evolution have found evidence for adaptive evolution[19,20], others have pointed towards an important role of plasticity[21]. However, we are not aware of any studies that have found evidence for phenotypic changes of a plastic nature that would fit an OU model as is the case here.

Ecological and molecular data analysis. We further corroborate our results with analyses of field observations and mark-recapture data. The yearly estimates of mean population fitness (proportion of breeding adults) increased as the population approached the adaptive optimum (Fig. 2a). More than 58% of the variation in mean population fitness was explained by variation in body mass (Fig. 2a). The mean annual variation in body mass decreased as mean population fitness increased, suggesting that body mass has been under strong selection in this population (Supplementary Fig. 1). Also, individuals with a lower body mass survive better than heavier individuals, as can be seen from mark-recapture data (Fig. 2b). Recaptured individuals were much closer to the adaptive optimum of body mass than individuals not recaptured, the latter group being well above the optimal value. These results confirm that the decrease in body mass is an adaptation to the Maltese environment, and that this evolutionary change has resulted in a more successful breeding population. Body mass remained significantly correlated with recapture rate throughout the years (analysis of variance (ANOVA): $P = 0.007$), and there was no significant year effect (ANOVA: $P = 0.588$). Hence, selection remains fairly constant over time, despite slowing evolutionary rates.

We estimated an average mean-standardized selection gradient across years for body mass to be equal to -0.39 (linear regression: $P = 0.006$), which is consistent with a significant but moderate amount of directional selection on body mass[22]. Body mass has been found to be significantly heritable in the great reed warbler (*Acrocephalus arundinaceus*) as in many other passerine
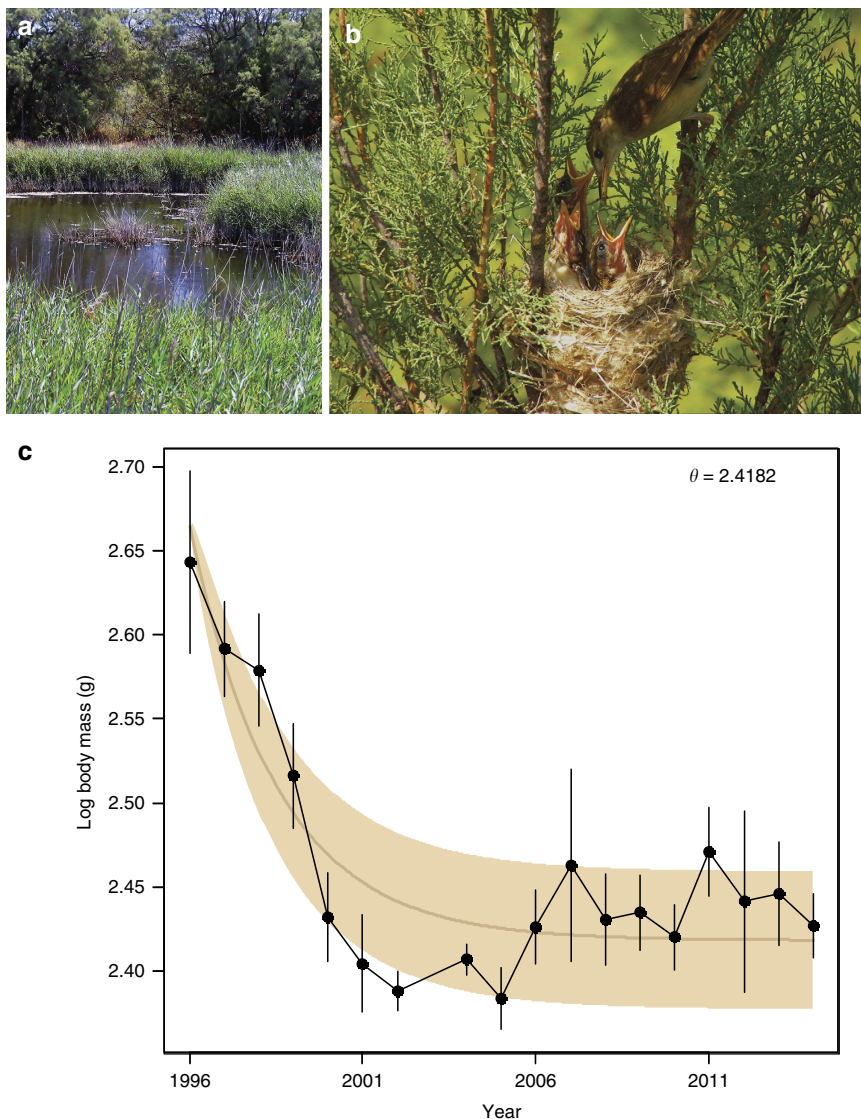
Figure 1 | Evolution of body mass towards an adaptive optimum. (**a**) The Is-Simar nature reserve is situated on the island of Malta, in the Mediterranean Sea. Restoration of the wetland began in 1992, where a network of pools, canals and islands were created and vegetation replanted. Shortly after, in 1994, it was colonized by reed warblers (*Acrocephalus scirpaceus*). (**b**) Reed warblers nest in reed beds (*Phragmites*) or *Tamarix* trees (almost exclusively in the latter at Is-Simar) and usually lay three to five eggs, which are incubated by both parents. (**c**) The evolution of log body mass over time (years; $N = 392$). Vertical error bars signify one standard error. The expected evolutionary trajectory of the best-fit adaptive model (OU) is shown as a line, with a 95% probability interval around in brown. The adaptive optimum (θ) for log body mass is 2.42. No samples available for year 2003.

Table 1 | Estimates of model fit for a neutral and an adaptive model of evolution for mean body mass.

Trait	Model	logL	K	AICc	Akaike weights	LRT
Body mass	Neutral	30.45	2	− 56.10	0.002	
	Adaptive	36.73	4	− 62.38	0.998	12.55, $P = 0.002$

Neutral evolution was modelled as an unbiased random walk, and adaptive evolution was modelled as an Ornstein–Uhlenbeck (OU) process. For the OU model, the adaptive optimum (θ) for log body mass is 2.42, the step variance (σ^2_{step}) is 0.0004 and the alpha (α), the strength of the restraining force around the optimum, is 0.39. The log-likelihood (logL), number of parameters (K), bias-corrected Akaike Information Criterion (AICc) and Akaike weights suggest that the adaptive model is the more likely model. A likelihood ratio test (LRT), which tests the significance of the improved fit of the adaptive over the neutral model, with the latter treated as the null model, confirmed that indeed the observed changes in body mass are of an adaptive nature. The LRT statistic is distributed as a χ^2, with two degrees of freedom.

bird species[23], and to be a highly evolvable trait[23]. In our data set, annual mid-parent and offspring body mass is significantly correlated irrespective of annual fluctuations in body mass, which supports the assumption that body mass has a significant heritable component in our population (Supplementary Fig. 2). Migrants presumably coming from Italy[24] have an average body mass that is very close to the initial body mass of the Maltese

population in 1996 (Supplementary Fig. 3). This suggests that evolution of body mass has occurred *in situ* in Malta and is not the result of biased immigration.

Assuming that the fluctuations in body mass during the stationary phase are caused by genetic drift, we can estimate effective population size (N_e) from the OU model. This N_e estimate represents an important investigation of the reliability of
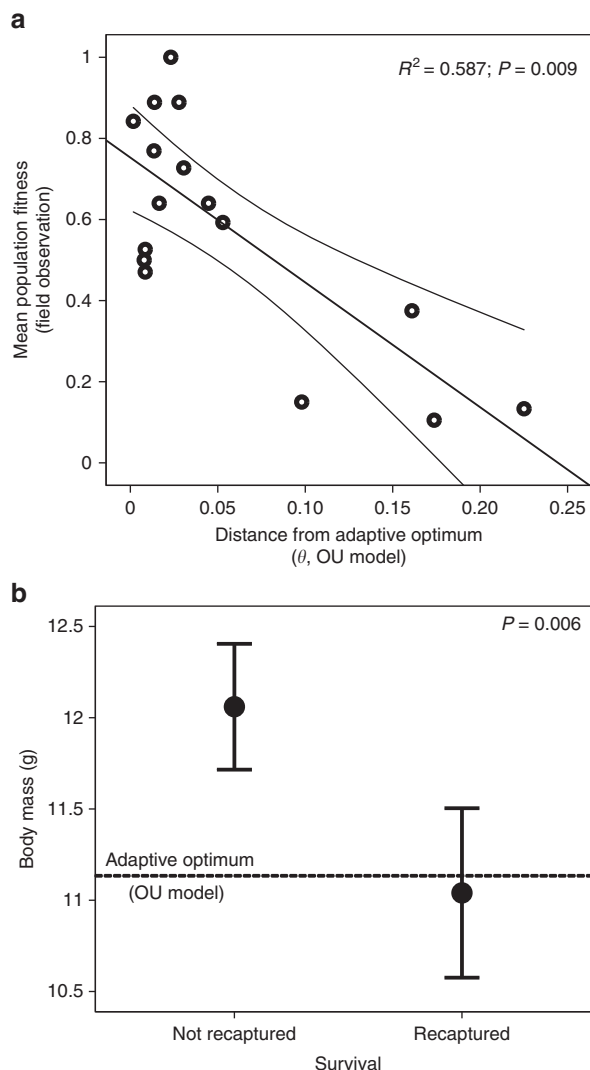
a

b

Figure 2 | Fitness and body mass in the Maltese reed warbler population.
(**a**) Linear regression of mean population fitness (proportion of breeding adults each year (N years $= 16$)) against the yearly distance from the adaptive optimum for body mass (θ) estimated from the OU model. As the population evolved towards the optimum, the mean population fitness increased significantly (linear regression: $P = 0.009$) and overall, 58.7% of the variation in mean population fitness can be explained by variation in the distance from the adaptive optimum for body mass. (**b**) The effect of body mass on survival. Individuals marked and recaptured have a significantly lower body mass than individuals that are marked but never recaptured (total $N = 198$). The dashed line corresponds to the adaptive optimum (θ) for body mass estimated from the OU model. The mean-standardized selection gradient is equal to -0.39 (linear regression: $P = 0.006$).

the model as it can be compared with an independent molecular estimate of N_e (Table 2). Both estimates of N_e are very similar, and also close to an ecological estimate of N_e (based on the number of breeding pairs observed in the field). Such a consistent match suggests that the OU model portrays the adaptive evolution of body mass in an accurate way.

Discussion
We suggest that reed warblers with a lower body mass could be better suited for the hot and dry Maltese environment. Body mass should decrease with increasing temperature[25–29], in accordance

Table 2 | Estimates of effective populations size (N_e) from three independent methods.

N_e (field)	N_e (OU)	N_e (molecular)	F_{IS}
7.75 (2–16)	23.39 (4.68–32.75)	23.60 (14.10–51)	0.123 ($P = 0.01$)

Field, harmonic mean of number of breeding pairs with the range shown in parentheses. OU, $Ne1/4h^2s^2p/s^2$ step, where h^2 is the trait heritability (set to 0.5, with 0.1–0.7 shown in parentheses), s^2p is the phenotypic variance of the sample and s^2 step is the step variance. molecular, linkage disequilibrium method with N_e estimator with 95% confidence intervals shown in parentheses.

with Bergmann's rule. Lower body mass has been interpreted as an adaptation to warmer climates, as the relatively larger body surface areas of smaller individuals serve as efficient heat dissipators[27]. Smaller-bodied individuals are able to retain water at higher temperatures than large-bodied individuals[30]. Other ecological factors linked to climatic variation, such as resource availability, may also affect body size[31]. However, rising temperatures have previously been shown to correlate with a decrease in body mass[19,32], which strongly point towards a thermoregulation hypothesis. The reed warbler has been singled out as a species that is very sensitive to global warming, which has resulted in a recent range shift northwards[31]. Malta is the southernmost location within the European reed warbler distribution[33], and it has a dry subtropical Mediterranean climate, which is likely to have imposed a strong selective pressure on colonizers.

Is-Simar, the reed warbler's main breeding site in Malta, is small (5.8 hectares), meaning that the maximum number of breeding pairs is constrained (5–8 pairs). However, the probability of recruitment is high, with an estimated 10.6% of ringed nestlings recaptured as breeding adults over the 19 years of our study. This high recruitment rate has probably played an important role in the adaptive process, especially in the early stages of the colonization event. The founder population on Is-Simar most likely consisted of individuals from a large, genetically variable source population in Southern Europe.

With small effective population sizes, the possibilities for adaptation should be constrained[15] and directional selection is expected to deplete genetic variation. However, this population has managed to reach its adaptive optimum very rapidly, despite its small size. The F_{IS}-value from Table 2 signifies that there are currently 12% fewer heterozygotes than expected under Hardy–Weinberg equilibrium, indicating that the population is inbred. This value is well above the average F_{IS} for other European populations (F_{IS} Malta $= 0.123$, average F_{IS} in Europe[33] (N populations $= 31$) $= 0.05$ (s.d. $= 0.04$), $t = -10.31$, $P < 0.001$). This may have other detrimental effects for the population in the future, as inbreeding, for the same set of microsatellite loci in other warbler species, has been shown to be correlated with fitness[34,35] (reproductive output).

Furthermore, after correcting for variation due to sampling error in trait means after the population has reached the optimum, only $\sim 13\%$ of the stationary variance in body mass is left unexplained (observed stationary variance $= 0.000692$, corrected stationary variance ($\pm$ S.E.) $= 0.000092 \pm 0.00014$). This may indicate very strong stabilizing selection for the optimal body size in the stationary phase with little effect of drift or plasticity, an interpretation supported by the very rapid rate of adaptation (half-life) estimated in this system. Another possibility is that the optimum is non-stationary, which means that part of the trait fluctuations during the stationary phase shows how the population is tracking the optimum's movement across years. However, to our knowledge, the environmental conditions in terms of habitat and resources have remained stable since the restoration was completed in 1994, with relatively few competitors or

predators[24] and a relatively stable climate over the years[24]. Since we do not possess any pedigree data that would have allowed for a quantitative genetic assessment of the changes observed here, we cannot exclude a possible role of plasticity. Although an adaptive plastic process alone is unlikely to generate a reaction norm consistent with an OU model, we acknowledge that at least part of the changes we report here may be of a plastic nature.

The population has so far shown high potential for adaptation, but increased inbreeding could reduce the adaptability of the population. Inbreeding may amplify tendencies to deviate from the optimum and lead to maladaptation, which would threaten the future of this minuscule but evolutionary successful population. We therefore recommend further restoration of this wetland habitat allowing for a larger population, thus decreasing the risk of inbreeding depression[36] and the impact of genetic drift.

Our study may represent one of the most rapid cases of adaptive evolution ever documented in the context of habitat restoration. It also demonstrates the importance of population monitoring in evolutionary and conservation biology, as the success of a conservation project may be difficult to predict and depend largely on the evolutionary potential of the focal population or species.

Methods

Sampling. From 1996 to 2014 (except for 2003 where data is not available), we sampled a total of 392 adult reed warblers during the breeding season (May–August), as part of the BirdLife Malta project. All birds included in the main analyses (unless otherwise stated) were resident individuals captured during the breeding season when no migration occurs. Birds were ringed with unique ID rings and body mass recorded to nearest 0.1 g using a digital scale. All measurements took place during the morning hours between 06:00 and 10:00 to minimize daily variation in body mass. We estimated the minimum number of breeding pairs through the intensive monitoring of nests and other frequent field observations during the entire breeding season. We also captured migrants stopping over for several days in the population during autumn migration (September–October) to investigate differences in body mass with local residents. All sampling and handling of birds was in compliance with ethical regulations, and permits for sampling were obtained from the local authorities (BirdLife Malta).

DNA extraction. In 2014, we sampled blood from 18 individuals. DNA was extracted from the blood samples using DNeasy Blood & Tissue kit (Qiagen), and subsequently genotyped. We amplified eight microsatellites previously used for reed warblers[33]; Aar4, Aar5, Aar8, Ase34, Ase58, Pca3, Pdoµ1 and POCC2 (Supplementary Table 1).

Statistical analyses. To investigate evolutionary changes through our time series of 19 years ($N = 392$), we used the average body mass of individuals caught in each year, along with corresponding sample sizes and standard deviations. If individuals were captured and measured more than once, we used the average across measurements. The average difference between two measurements at two different time points was non-significant (ANOVA: -0.0268 g; df $= 204$; $P = 0.765$). We compared the goodness of fit of a neutral (unbiased random walk) and an adaptive (OU) model to our data using the PaleoTS package[6] in R. We used bias-corrected AICc as a measure of model fit, and to show the relative support for the two models we used Akaike weights (transformations of the AICc scores to make them sum to one). We also conducted a log-likelihood ratio test using the log-likelihood estimates from the models.

To investigate the relationship between mean population fitness (proportion of breeding adults each year) and body mass, we conducted a linear regression analysis. We plotted mean population fitness to the population distance from the adaptive optimum of body mass (estimated from the OU model). We also estimated the correlation between s.d. in body mass and mean population fitness. Both these estimates should be expected to be negative if natural selection is acting on body mass. From recapture data over the entire study period ($N = 198$), we distinguished between individuals recaptured after a minimum of 21 days (but also taking in account individuals recaptured the following seasons) and individuals never recaptured, and used this as a proxy for individual survival, which we acknowledge could also be partially affected by other factors such as emigration, although our estimates on recruitment rate suggest very high philopatry. We subsequently estimated the mean-standardized selection gradient for body mass. To investigate whether there is any temporal variation in selection during the study period, we conducted a linear model of recapture probability where both body mass and year of capture were included as covariates.

We calculated different estimates of N_e. From field data, we estimated N_e using the harmonic mean of the number of breeding pairs observed in the field across years. We further followed the procedure as described by Hunt et al.[6] to estimate N_e using the parameter estimates from the OU model. $N_e = h^2 \sigma^2_p / \sigma^2_{step}$, where h^2 is the trait heritability, σ^2_p is the phenotypic variance of the samples and σ^2_{step} is the step variance, which is estimated from the model fit. We solved the equation with three different values of h^2 (0.1, 0.5 and 0.7). Since we do not have pedigree-based information to calculate heritability, we estimated the correlation between annual mid-parent and offspring body mass to assess the plausibility of a significant additive component of genetic variation for body mass, irrespective of environmental variation across years. Finally, we used NeEstimator (v2.01)[37] to estimate N_e from our molecular data, using the linkage disequilibrium method[38].

Data availability. Data are available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.hj30r.

References

1. Stockwell, C. A., Hendry, A. P. & Kinnison, M. T. Contemporary evolution meets conservation biology. *Trends Ecol. Evol.* **18**, 94–101 (2003).
2. Huey, R. B., Gilchrist, G. W., Carlson, M. L., Berrigan, D. & Serra, L. Rapid evolution of a geographic cline in size in an introduced fly. *Science* **287**, 308–309 (2000).
3. Bell, M. A., Aguirre, W. E. & Buck, N. J. Twelve years of contemporary armor evolution in a threespine stickleback population. *Evolution.* **58**, 814–824 (2004).
4. Lescak, E. A. et al. Evolution of stickleback in 50 years on earthquake-uplifted islands. *Proc. Natl Acad. Sci. USA* **112**, E7204–E7212 (2015).
5. Grant, P. R. & Grant, B. R. Unpredictable evolution in a 30- year study of Darwin's finches. *Science* **296**, 707–711 (2002).
6. Hunt, G., Bell, M. A. & Travis, M. P. Evolution toward a new adaptive optimum: phenotypic evolution in a fossil stickleback lineage. *Evolution.* **62**, 700–710 (2008).
7. Tilman, D., May, R. M., Lehman, C. L. & Nowak, M. A. Habitat destruction and the extinction debt. *Nature* **371**, 65–66 (1994).
8. Thomas, C. D. et al. Extinction risk from climate change. *Nature* **427**, 145–148 (2004).
9. Hendry, A. P., Farrugia, T. J. & Kinnison, M. T. Human influences on rates of phenotypic change in wild animal populations. *Mol. Ecol.* **17**, 20–29 (2008).
10. Dobson, A. P., Bradshaw, A. D. & Baker, A. J. M. Hopes for the future: restoration ecology and conservation biology. *Science* **277**, 515–522 (1997).
11. Godefroid, S. et al. How successful are plant species reintroductions? *Biol. Cons.* **144**, 672–682 (2011).
12. Frankham, R., Ballou, J. D. & Briscoe, D. A. *Introduction to Conservation Genetics* (Cambridge University Press, 2002).
13. Keller, L. F. & Waller, D. M. Inbreeding effects in wild populations. *Trends Ecol. Evol.* **17**, 230–241 (2002).
14. Frankham, R. Genetics and extinction. *Biol. Cons.* **126**, 131–140 (2005).
15. Willi, Y., Van Buskirk, J. & Hoffmann, A. A. Limits to the adaptive potential of small populations. *Annu. Rev. Ecol. Evol. Syst.* **37**, 433–458 (2006).
16. Collyer, M. L., Heilveil, J. S. & Stockwell, C. A. Contemporary evolutionary divergence for a protected species following assisted colonization. *PLoS ONE* **6**, e22310 (2011).
17. Lande, R. Natural selection and random genetic drift in phenotypic evolution. *Evolution* **30**, 314–334 (1976).
18. Hansen, T. Stabilizing selection and the comparative analysis of adaptation. *Evolution* **51**, 1341–1351 (1997).
19. Salewski, V., Hochachka, W. M. & Fiedler, W. Global warming and Bergmann's rule: do central European passerines adjust their body size to rising temperatures? *Oecologica* **162**, 247–260 (2010).
20. Karell, P., Ahola, K., Karstinen, T., Valkama, J. & Brommer, J. E. Climate change drives microevolution in a wild bird. *Nat. Commun.* **2**, 208 (2011).
21. Merilä, J. & Hendry, A. P. Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. *Evol. Appl.* **7**, 1–14 (2014).
22. Hereford, J., Hansen, T. F. & Houle, D. Comparing strengths of directional selection: how strong is strong? *Evolution* **58**, 2133–2143 (2004).
23. Teplitsky, C. et al. Assessing multivariate constraints to evolution across ten long-term avian studies. *PLoS ONE* **9**, e90444 (2014).
24. Sultana, J. et al. *The Breeding Birds of Malta* (Birdlife Malta, 2011).
25. Lima, S. L. Predation risk and unpredictable feeding conditions: determinants of body mass in birds. *Ecology* **67**, 377–385 (1986).
26. Rogers, C. M. Predation risk and fasting capacity: do wintering birds maintain optimal body mass? *Ecology* **68**, 1051–1061 (1987).
27. Yom-Tov, Y. Global warming and body mass decline in Israeli passerine birds. *Proc. R. Soc. Lond. B* **268**, 947–952 (2001).
28. Ramirez, L., Diniz-Filho, J. A. F. & Hawkins, B. A. Partitioning phylogenetic and adaptive components of the geographical body-size pattern of New World birds. *Global Ecol. Biogeogr.* **17**, 100–110 (2008).
29. Olson, V. A. et al. Global biogeography and ecology of body size in birds. *Ecol. Lett.* **12**, 249–259 (2009).

30. Gardner, J. L., Peters, A., Kearney, M. R., Joseph, L. & Heinsohn, R. Declining body size: a third universal response to warming? *Trends Ecol. Evol.* **26,** 285–291 (2011).

31. Eglington, S. M. *et al.* Latitudinal gradients in the productivity of European migrant warblers have not shifted northwards during a period of climate change. *Global Ecol. Biogeogr.* **24,** 427–436 (2015).

32. Van Buskirk, J., Mulvihill, R. S. & Leberman, R. C. Declining body sizes in North American birds associated with climate change. *Oikos* **119,** 1047–1055 (2010).

33. Procházka, P. *et al.* Low genetic differentiation among reed warbler *Acrocephalus scirpaceus* populations across Europe. *J. Avian Biol.* **42,** 103–113 (2011).

34. Hansson, B., Bensch, S., Hasselquist, D. & Åkesson, M. Microsatellite diversity predicts recruitment of sibling great reed warblers. *Proc. R. Soc. Lond. B* **268,** 1287–1291 (2001).

35. Bensch, S., Hasselquist, D. & Von Schantz, T. Genetic similarity between parents predicts hatching failure: nonincestuous inbreeding in the great reed warbler? *Evolution* **48,** 317–326 (1994).

36. Hansson, B. *et al.* Increase of genetic variation over time in a recently founded population of great reed warblers (*Acrocephalus arundinaceus*) revealed by microsatellites and DNA fingerprinting. *Mol. Ecol.* **9,** 1529–1538 (2000).

37. Do, C. *et al.* NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. *Mol. Ecol. Resour.* **14,** 209–214 (2014).

38. Waples, R. S. & Do, C. LDNE: a program for estimating effective population size from data on linkage disequilibrium. *Mol. Ecol. Resour.* **8,** 753–756 (2008).

Author contributions

C.C., M.A. and M.G. monitored the population and collected the morphological data. G.-P.S. and F.E. designed the study and collected the genetic data. C.L.C.S. performed the DNA extractions and the analysis of the microsatellite data. C.L.C.S., K.L.V., G.-P.S. and F.E. analysed the morphological and mark-recapture data. C.L.C.S. wrote the paper with contributions and comments from all authors.

Additional information

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

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

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Sætre, C. L. C. *et al.* Rapid adaptive phenotypic change following colonization of a newly restored habitat. *Nat. Commun.* **8,** 14159 doi: 10.1038/ncomms14159 (2017).

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

Paper IB

CORRESPONDENCE

Reply to 'Inconclusive evidence for rapid adaptive evolution'

Camilla Lo Cascio Sætre[1], Charles Coleiro[2], Martin Austad[2], Mark Gauci[2], Glenn-Peter Sætre[1], Kjetil Lysne Voje[1] & Fabrice Eroukhmanoff[1]

In our study[1], we showed that a newly founded population of reed warblers in Malta had undergone a decrease in body mass through 19 years, following a trajectory consistent with a population ascending an adaptive peak, an Ornstein–Uhlenbeck process (OU)[2]. Neto et al.[3] claim that our result is an artifact of including migrants in the dataset, which inflated the average body mass in the initial years. Controlling for possible seasonal effects is important, which we thank Neto et al.[3] for pointing out. We now control for season in three different ways and the OU-model always fits better than the neutral model, further strengthening our original conclusion of adaptive evolution.

In Malta, the autumn migration mainly takes place in September, and spring migrants arrive in April or early May[4]. We think limiting our data to mid-June to mid-July, as Neto et al.[3] suggest, is unreasonably restricted for capturing local birds. We chose to include the entire breeding season (May–August)[4], as it has been described in several other studies[5–7]. We cannot exclude the possibility that there are some migrants in our dataset, but we investigate possible biases in our sampling design. The ratio of birds captured in the center of the breeding season (June and July) to birds captured in May and August, where the possibility of migrants is greater, shows no apparent trend throughout the years (Supplementary Figure 1). However, body mass is significantly correlated with capture date (linear regression: $R^2 = 0.23$, $P < 2e{-}16$; Supplementary Figure 2). Thus, we agree there is a need to correct for capture date in our models.

We controlled for capture date in three different ways, and in all three cases, the OU-model outcompeted the Random walk (neutral) model (see Supplementary Table 1, Fig. 1). In fact, some of the new results show an even larger difference in relative model fit than in our initial study.

Interestingly, mean annual body mass of juveniles corrected for seasonality also exhibits a negative trend over the study period, and these data also show a much better fit to an OU-model than to a neutral model (Supplementary Table 2, Supplementary Figure 3). Juveniles were born and ringed on site and the observed trend can therefore not be explained by the potential inclusion of migrants. Furthermore, juveniles recaptured as adults were

significantly lighter than the ones not recaptured (mean ± SD of those not recaptured: 12.23 ± 2.23 g, mean ± SD of those recaptured: 10.93 ± 1.87 g, linear model: Estimate ± SE = -1.30 ± 0.41, $P = 0.0016$.).

Also, when running a linear regression of body mass throughout the years within each month for adult birds, there is a significantly negative trend within June, July and August (Supplementary Table 3). This demonstrates that body mass decreased significantly from 1996 to 2014, without the possibility for migrants to influence the pattern, as the probability of capturing migrants in June or July is negligible. We acknowledge that without further experimental data, we cannot know the exact mechanisms behind the decrease in body mass, nor why it is apparent both in adults and juveniles.

In order to investigate whether our estimates of survival in the population could have been affected by the inclusion of migrants, we compared birds captured in June/July to birds captured in May/August in terms of proportion recaptured and proportion not recaptured. The proportions were not significantly different from each other (Supplementary Table 4; two-tailed Fisher's exact test, $P = 0.18$), suggesting that our survival estimates are not merely artifacts of including migrants.

Neto et al.[3] also seem skeptical towards our conclusion since our results would "constitute an example of exceptionally rapid adaptive evolution in the wild". Although we agree evolution was fast, we do not consider it to be exceptional: the haldanes calculated from our model are within the normal range of evolutionary rates measured in populations affected by human-induced environmental changes[8]. The half-life we report is as far as we know the shortest estimated from an OU-model, but this is due to the short time interval covered by our data. Also, the estimated selection gradient is within the normal range[9].

Neto et al.[3] include a boxplot (Fig. 1 in ref. [3]) depicting variation in body mass in adult and first year reed warblers in Portugal, showing that body mass is highest in April and from August and onwards. The largest difference between any sample median in Fig. 1 of ref. [3] represents a difference of 0.105 natural log units. In comparison, the difference between the initial

[1] Department of Biology, Centre for Ecological and Evolutionary Synthesis, University of Oslo, P. O. Box 1066 Blindern, 0316 Oslo, Norway. [2] BirdLife Malta, Xemxija Waterfront Apartments, Flat ½, Triq Is-Simar, Xemxija SPB 9025, Malta. Correspondence and requests for materials should be addressed to F.E. (email: fabrice.eroukhmanoff@ibv.uio.no)
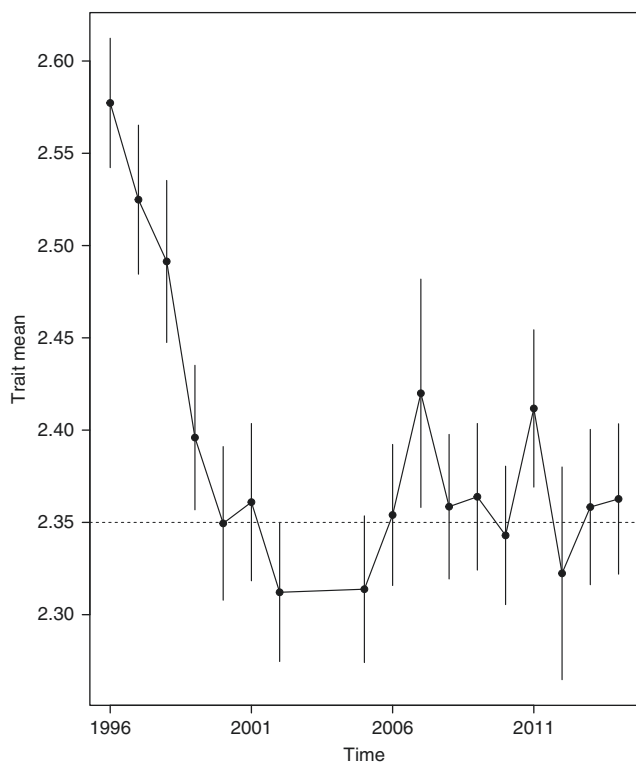
Fig. 1 The evolution of log body mass over time, corrected for seasonal variation linked to capture date. We regressed capture date (days away from the center of the breeding season (July 1st)) on log body mass (the dependent variable), having year a factor. Vertical error bars signify one standard error. The data reveal a negative trend in body mass consistent with an OU-model. The dotted line represents the estimated adaptive optimum (θ) for log body mass (2.35)
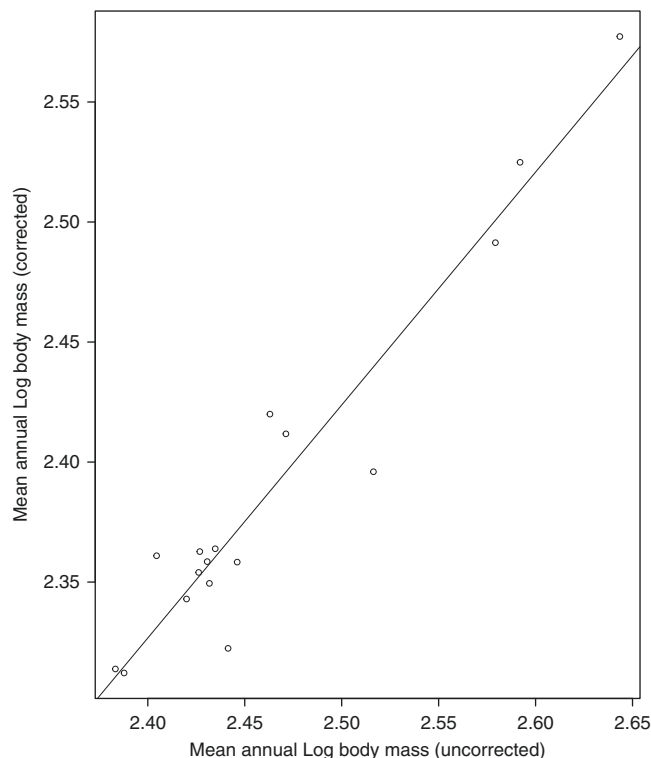

Fig. 2 Regression between uncorrected mean annual estimates of Log body mass (from Sætre et al.[1]) and corrected mean annual estimates of log body mass (the values predicted by the first ANCOVA model, where we regressed capture date (days away from the center of the breeding season (July 1st)) on log body mass (the dependent variable), having year a factor). The data show a strong concordance between both estimates ($R^2 = 0.92$; $P = 8.29e{-}10$)

population in our data and the estimated optimum (from the first ANCOVA model) is 0.23 (0.22 in the original analysis) natural log units, and the difference between the most extreme sample means in our data set is even larger. This suggests that seasonal variation in body mass alone is an insufficient explanation for the decrease in body mass we observe in the Maltese population.

Neto et al.[3] present the body mass distribution of reed warblers from May to August from their study site in Sweden and point out that the average body mass in the first years of the population in Malta exceeds the mass of the heaviest individual in their Swedish population. It is not clear to us why a single Swedish population should accurately reflect the whole phenotypic range of body mass in this species. The data from Portugal presented by Neto et al.[3] are indeed containing birds of similar size to what we observe in the Maltese population. Yet, the average body masses of the birds in the initial years were arguably high relative to other populations. We suggest that this may be a result of biased colonization, if the founding population consisted of relatively heavy individuals. Another possibility is that food availability was particularly high in the initial years, and density-dependent effects may be a confounding factor. We acknowledge that we cannot be certain of the origin of the population, and we welcome research to pinpoint the origin of the Maltese population.

Neto et al.[3] claim the pattern of body size reduction is coincidental and speculates that it is an artifact of the inclusion of migrants in the dataset. We have shown that this is unlikely to be the case. We fail to see how Neto et al.[3] alternative explanation predicts a trend in body size that is well described by an OU-process. We also note that Neto et al.[3] do not comment on the fact that the N_e estimated from the OU model parameters

corresponds accurately with our independent molecular estimate of N_e. Our molecular data stems from individuals that were definitely locals; either juveniles or nesting adults captured in June or July. If the data we used in our model selection had been heavily biased by migrants, we would not expect such a close correspondence with the molecular estimate.

We acknowledge, as we did in the original publication, that body mass is a plastic trait. We thank Neto et al.[3] for pointing out that seasonal variation should be corrected for in analyses of body mass evolution in birds. Doing so puts further strength to our claim that adaptive evolution is likely an important part of the observed trend in body mass in the Maltese reed warbler population.

Methods

Statistical analyses. We analyzed ANCOVA models where we regressed capture date on log body mass (the dependent variable), having year as a factor. We assumed a common coefficient for how body mass changes as a function of capture date due to the modest sample sizes of measured birds per year. In the first ANCOVA model, capture date was measured as "days away from the center of the breeding season (July 1st)". We tested the correlation between the annual body mass data used in our original publication and the annual body mass data corrected for seasonal variation with this model, and there is a strong concordance between both estimates ($R^2 = 0.92$; $P = 8.29e{-}10$, Fig. 2). In the second ANCOVA model, capture date was implemented as "days after May 1st" as a quadratic term (with the linear term also included in the model). The third model is a linear mixed-effect model implemented using the lme4 package[10] where log body mass was the dependent variable, year was a fixed effect and month (the month birds where captured) was implemented as a random factor. We note that there are several other non-linear models we could have used, which may have fitted the data better. However, we believe it is best to avoid complex, parameter-rich models given our limited sample size.

For each model, we used the predicted mean and variance to compare the goodness of fit of a neutral (unbiased random walk) and an adaptive (OU) model using the PaleoTS package[11] in R. We used bias-corrected AICc as a measure of model fit, and to show the relative support for the two models we used Akaike weights (transformations of the AICc scores to make them sum to one). We also conducted a log-likelihood ratio test using the log-likelihood estimates from the models.

References

1. Sætre, C. L. et al. Rapid adaptive phenotypic change following colonization of a newly restored habitat. *Nat. Commun.* **8**, 14159 (2017).
2. Lande, R. Natural selection and random genetic drift in phenotypic evolution. *Evolution* **30**, 314–334 (1976).
3. Neto, J. M., Bensch, S., Råberg, L. & Hansson, B. Inconclusive evidence for rapid adaptive evolution. *Nat. Commun.* https://doi.org/10.1038/s41467-018-05119-2 (2018).
4. Sultana, J. et al. *The Breeding Birds of Malta* (Birdlife Malta, 2011).
5. Schulze-Hagen, K., Leisler, B. & Winkler, H. Breeding success and reproductive strategies of two Acrocephalus warblers. *J. Ornithol.* **137**, 181–192 (1996).
6. Halupka, L., Dyrcz, A. & Borowiec, M. Climate change affects breeding of reed warblers Acrocephalus scirpaceus. *J. Avian Biol.* **39**, 95–100 (2008).
7. Vafidis, J. O. et al. The effects of supplementary food on the breeding performance of Eurasian reed warblers Acrocephalus scirpaceus; Implications for climate change impacts. *PLoS One* **11**, e0159933 (2016).
8. Hendry, A. P., Farrugia, T. J. & Kinnison, M. T. Human influences on rates of phenotypic change in wild animal populations. *Mol. Ecol.* **17**, 20–29 (2008).
9. Hereford, J., Hansen, T. F. & Houle, D. Comparing strengths of directional selection: how strong is strong? *Evolution* **58**, 2133–2143 (2004).
10. Bates, D., Maechler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–lme48 (2015).
11. Hunt, G., Bell, M. A. & Travis, M. P. Evolution toward a new adaptive optimum: phenotypic evolution in a fossil stickleback lineage. *Evolution* **62**, 700–710 (2008).

Author contributions

C.L.C.S performed the analyses with help from F.E. and K.L.V; F.E., K.L.V., and C.L.C.S wrote the paper with contributions and comments from G.-P.S; C.C., M.A., and M.G. collected the data.

Additional information

Paper II

A Chromosome-Level Genome Assembly of the Reed Warbler (*Acrocephalus scirpaceus*)

Camilla Lo Cascio Sætre[1],*, Fabrice Eroukhmanoff[1], Katja Rönkä[2,3], Edward Kluen[2,3], Rose Thorogood[2,3], James Torrance[4], Alan Tracey[4], William Chow[4], Sarah Pelan[4], Kerstin Howe[4], Kjetill S. Jakobsen[1], and Ole K. Tørresen[1]

[1]Centre for Ecological and Evolutionary Synthesis, University of Oslo, Norway

[2]HiLIFE Helsinki Institute of Life Sciences, University of Helsinki, Finland

[3]Research Programme in Organismal and Evolutionary Biology, Faculty of Biological and Environmental Sciences, University of Helsinki, Finland

[4]Tree of Life, Wellcome Sanger Institute, Cambridge, United Kingdom

*Corresponding author: E-mail: c.l.c.satre@ibv.uio.no.

Accepted: 6 September 2021

Abstract

The reed warbler (*Acrocephalus scirpaceus*) is a long-distance migrant passerine with a wide distribution across Eurasia. This species has fascinated researchers for decades, especially its role as host of a brood parasite, and its capacity for rapid phenotypic change in the face of climate change. Currently, it is expanding its range northwards in Europe, and is altering its migratory behavior in certain areas. Thus, there is great potential to discover signs of recent evolution and its impact on the genomic composition of the reed warbler. Here, we present a high-quality reference genome for the reed warbler, based on PacBio, 10×, and Hi-C sequencing. The genome has an assembly size of 1,075,083,815 bp with a scaffold N50 of 74,438,198 bp and a contig N50 of 12,742,779 bp. BUSCO analysis using aves_odb10 as a model showed that 95.7% of BUSCO genes were complete. We found unequivocal evidence of two separate macrochromosomal fusions in the reed warbler genome, in addition to the previously identified fusion between chromosome Z and a part of chromosome 4A in the Sylvioidea superfamily. We annotated 14,645 protein-coding genes, and a BUSCO analysis of the protein sequences indicated 97.5% completeness. This reference genome will serve as an important resource, and will provide new insights into the genomic effects of evolutionary drivers such as coevolution, range expansion, and adaptations to climate change, as well as chromosomal rearrangements in birds.

Key words: genome assembly, Hi-C sequencing, long reads, reference genome, *Acrocephalus scirpaceus*. .

Significance

The reed warbler (*Acrocephalus scirpaceus*) has been lacking a genomic resource, despite having been broadly researched in studies of coevolution, ecology, and adaptations to climate change. Here, we generated a chromosome-length genome assembly of the reed warbler, and found evidence of macrochromosomal fusions in its genome, which are likely of recent origin. This genome will provide the opportunity for a deeper understanding of the evolution of genomes in birds, as well as the evolutionary path and possible future of the reed warbler.

Introduction

The ecology and evolution of the reed warbler (*Acrocephalus scirpaceus*) has been of interest for over 40 years (Thorogood et al. 2019) as it is one of the favorite host species of the brood-

parasitic common cuckoo (*Cuculus canorus*) (Davies and Brooke 1989; Stokke et al. 2018). Decades of field experiments have demonstrated behavioral coevolution and spatial and temporal variation in species interactions (e.g., Thorogood

and Davies 2013). However, the reed warbler's response to climate change has begun to attract increasing attention. Reed warblers are experiencing far less severe declines in population size than is typical for long-distance migrants (Both et al. 2010; Vickery et al. 2014). In fact, they are expanding their breeding range northwards into Fennoscandia (Järvinen and Ulfstrand 1980; Røed 1994; Stolt 1999; Brommer et al. 2012), and have generally increased their productivity following the rise in temperature (Schaefer et al. 2006; Eglington et al. 2015; Meller et al. 2018). They are also showing rapid changes in phenology (Halupka et al. 2008), and migratory behavior; instead of crossing the Sahara, monitoring suggests that some reed warblers now remain on the Iberian Peninsula over winter (Chamorro et al. 2019). Morphological traits such as body mass and wing shape have been shown to change rapidly in reed warbler populations, indicating possible local adaptation (Kralj et al. 2010; Salewski et al. 2010; Sætre et al. 2017). Genetic differentiation is generally low between reed warbler populations, but moderate levels of differentiation have been connected to both migratory behavior (Procházka et al. 2011) and wing shape (Kralj et al. 2010). Reed warblers thus provide a promising system to study population, phenotypic, and genetic responses to climate change.

Although there has been an increasing number of avian genome assemblies in recent years (e.g., Feng et al. 2020), many nonmodel species, including the reed warbler, are still lacking a genome resource. To date, the closest relative to the reed warbler with a published reference genome is the great tit (*Parus major*) (GCA_001522545.3, deposited in NCBI; Laine et al. 2016), but the unpublished genome of the garden warbler (*Sylvia borin*) is available in public databases (GCA_014839755.1, deposited in NCBI). There is also a genome in preprint from the *Acrocephalus* genus, the great reed warbler (*A. arundinaceus*) (Sigeman, Strandh, et al. 2020), but the scaffolds are not chromosome length.

Here, we present the first genome assembly of the reed warbler, based on PacBio, 10×, and Hi-C sequencing, with descriptions of the assembly, manual curation, and annotation. This genome will be a valuable resource for a number of studies, including studies of coevolution, population genomics, adaptive evolution, and comparative genomics. For reduced-representation sequencing (e.g., RAD-seq) studies, it will help produce a more robust SNP set than with a de novo approach (Shafer et al. 2017). It will facilitate the detection of selective sweeps, and provide the physical localization of variants (Manel et al. 2016), thus giving insight into the potential genes involved in adaptation. Furthermore, the genome will be an important resource in the study of chromosomal rearrangements in birds.

Results and Discussion

Genome Assembly and Genome Quality Evaluation

We generated 3,810,665 reads with PacBio, with an average read length of 16 kb at 61× coverage. We further obtained 277,617,608 paired-end reads (2×150) with 10× Genomics, and 185,974,525 paired-end reads (2×150) with Hi-C, at 83× and 56× coverage, respectively. The final genome assembly was 1.08 Gb in length, and contains 1,081 contigs (contig N50 of 13 Mb) and 200 scaffolds (scaffold N50 of 74 Mb) (table 1). The completeness of the assembled genome is high: of the 8,338 universal avian single-copy orthologs, we identified 7,978 complete BUSCOs (95.7%), including 7,920 single-copy (95.0%), and 58 duplicated BUSCOs (0.7%). Fifty-nine BUSCOs (0.7%) were fragmented, and 301 BUSCOs (3.6%) were missing.

Genome Annotation

The GC content of the reed warbler genome assembly was 41.9%. The total repeat content of the assembly was 10.94%, with LTR elements as the most common type of repeat (4.50%) followed by LINEs (4.11%) (table 1).

Using the Comparative Annotation Toolkit, based on a whole-genome multiple alignments from Cactus, we predicted 14,645 protein coding genes, with an average Coding DNA Sequence (CDS) length of 1,782 bp, and an average intron length of 2,918 bp (table 1). The annotated genes had 97.5% completeness (based on predicted proteins).

Synteny Analysis

The reed warbler genome showed high synteny with the great tit genome, though with some notable differences (fig. 1). The reed warbler chromosome 6 is a fusion of great tit chromosomes 7 and 8, and reed warbler chromosome 8 is a fusion of great tit chromosomes 6 and 9. Interestingly, these chromosomes are not fused in the garden warbler genome (supplementary fig. 1, Supplementary Material online), but correspond to the great tit chromosomes. This suggests that the fusions evolved relatively recently, perhaps at the base of the Acrocephalidae branch within Sylvioidea, but further research is needed to determine this. Hi-C contact maps confirm that the chromosomes assembled in the reed warbler genome are unbroken (supplementary fig. 2, Supplementary Material online). Interchromosomal rearrangements are rare in avian evolution (Ellegren 2010; Skinner and Griffin 2012), with some exceptions, such as in the orders Falconiformes (Damas et al. 2017) and Psittaciformes (Furo et al. 2018). In fact, in all or most species of Psittaciformes, chicken chromosomes 6 and 7, and 8 and 9 are fused (Furo et al. 2018; Kretschmer et al. 2018)—the same chromosomes involved in the fusions discovered in the reed warbler genome. We can only speculate about the significance of this without more data. Passeriformes, the sister group of Psittaciformes, exhibit much lower rates of interchromosomal rearrangements, despite being a large, highly diverse order (Kretschmer et al. 2021). There is still a large knowledge gap in the cytogenetics of birds (Degrandi et al. 2020), and more

Table 1

Summary Statistics of the Reed Warbler Genome Assembly and Annotation

Genome Assembly			
	Estimated genome size	1.13 Gb	
	Guanine and cytosine content	41.91%	
	N50 length (contig)	13 Mb	
	Longest contig	48 Mb	
	Total length of contigs	1.07 Gb	
	N50 length (scaffold)	74.44 Mb	
	Longest scaffold	153.80 Mb	
	Total length of scaffolds	1.08 Gb	
	Complete BUSCOs	95.7%	
Transposable Elements		**Percent (%)**	**Total length**
	DNA	0.22	2.35 Mb
	LINE	4.11	44.2 Mb
	SINE	0.09	0.98 Mb
	LTR	4.50	48.4 Mb
	Unknown	0.55	5.9 Mb
	Other (satellites, simple repeats, low complexity)	1.49	16 Mb
	Total	10.94	117.6 Mb
Protein-Coding Genes			
	Predicted genes	14,645	
	Average coding sequence length (bp)	1,782	
	Average exon length (bp)	284	
	Average intron length (bp)	2918	
	Complete BUSCOs	97.5%	

research is needed to determine the rarity of the fusions we discovered in the reed warbler genome.

We furthermore confirm the previously identified neo-sex chromosome (Pala et al. 2012; Sigeman, Ponnikas, et al. 2020), a fusion between the ancestral chromosome Z and a part of chromosome 4A (according to chromosome naming from the zebra finch). This fusion is thought to have occurred at the base of the Sylvioidea branch (Pala et al. 2012) and is shared with all species of Sylvioidea studied so far (Sigeman, Ponnikas, et al. 2020). Figure 1 clearly shows that reed warbler chromosome Z corresponds to great tit chromosome Z, plus a part of great tit chromosome 4A, whereas reed warbler chromosome Z corresponds to garden warbler chromosome Z (supplementary fig. 1, Supplementary Material online).

Sylvioidea, we found unequivocal evidence of two novel macrochromosomal fusions in the reed warbler genome. Further research is needed to determine the evolutionary age of these fusions, especially because they are not present in the garden warbler genome, suggesting they are relatively new. This genome will serve as an important resource to increase our knowledge of chromosomal rearrangements in birds, both their prevalence and their significance for avian evolution. Furthermore, the genome will, through the identification of genetic variants and information of the function of associated genes, provide a deeper insight into the evolution of the reed warbler, a bird which will continue to fascinate researchers for years to come.

Conclusion

In this study, we present the first assembled and annotated genome for the reed warbler *A. scirpaceus*. We have accomplished this through utilizing long read PacBio sequencing, and scaffolding with paired-end 10× and Hi-C reads. In addition to the previously identified autosome-sex chromosome fusion shared by all members of

Materials and Methods

Sampling and Isolation of Genomic DNA

Blood was collected from a brachial vein of a female reed warbler (subspecies *A. scirpaceus scirpaceus*, NCBI Taxonomy ID: 126889) in Storminnet, Porvoo (60°19′24.9″N, 25°35′23.0″E), Finland, on May 22, 2019.
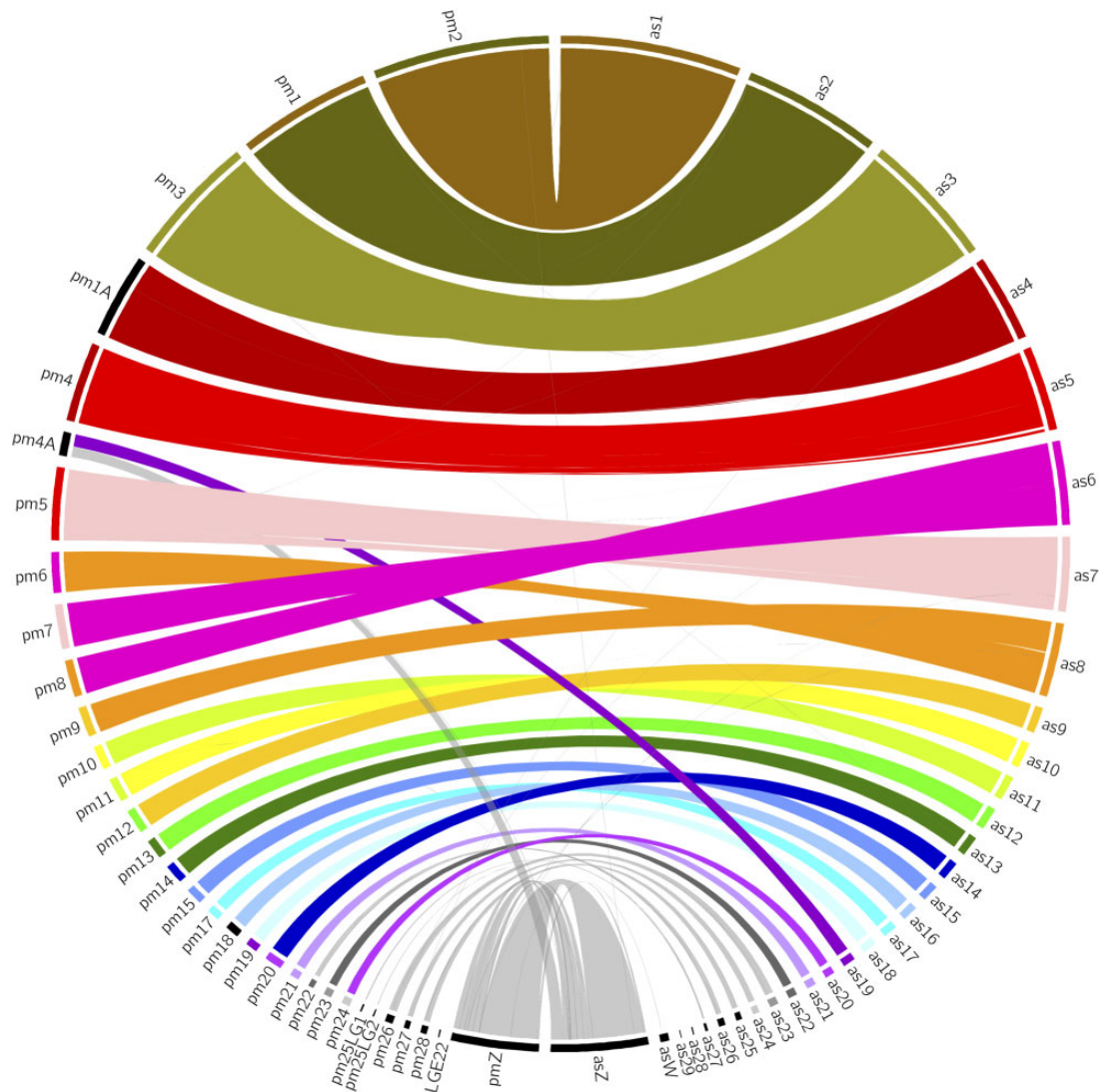
Fig. 1.—Circos plot showing the synteny between the reed warbler (on the right side, denoted with the prefix as [*Acrocephalus scirpaceus*]) and the great tit (left side, prefix pm [*Parus major*]) genome assemblies. The reed warbler chromosome 6 is a fusion of great tit chromosomes 7 and 8, whereas reed warbler chromosome 8 is a fusion of great tit chromosomes 6 and 9 (see Hi-C contact maps in supplementary fig. 2, Supplementary Material online). The reed warbler chromosome Z corresponds to great tit chromosome Z, and a part of great tit chromosome 4A.

Catching and sampling procedures complied with the Finnish law on animal experiments and permits were licensed by the National Animal Experiment Board (ESAVI/3920/2018) and Southwest Finland Regional Environment Centre (VARELY/758/2018). Reed warblers were trapped with a mist net, ringed and handled by E.K. under his ringing license.

The blood (∼80 ml) was divided and stored separately in 500 ml ethanol, and in 500 ml SET buffer (0.15M NaCl, 0.05M Tris, 0.001M EDTA, pH 8.0). The samples were immediately placed in liquid nitrogen, and kept at −80 °C when stored. We performed phenol−chloroform DNA isolation on the sample stored in SET buffer, following a modified protocol from Sambrook et al. (1989).

Library Preparation and Sequencing

DNA quality was checked using a combination of a fluorometric (Qubit, Invitrogen), UV absorbance (Nanodrop, Thermo Fisher) and DNA fragment length assays (HS-50 kb fragment kit from AATI, now part of Agilent Inc.). The PacBio library was prepared using the Pacific Biosciences Express library preparation protocol. DNA was fragmented to 35 kb. Size selection of the final library was performed using BluePippin with a 15 kb cut-off. Six single-molecule real-time (SMRT) cells were sequenced using Sequel Polymerase v3.0 and Sequencing chemistry v3.0 on a PacBio RS II instrument. The 10× Genomics Chromium linked-read protocol (10× Genomics Inc.) was used to prepare the 10× library,

and due to the reed warbler's smaller sized genome, only 0.7 ng/μl of high molecular weight DNA was used as input. A high-throughput chromosome conformation capture (Hi-C) library was constructed using 50 μl of blood, following step 10 and onwards in the Arima Hi-C (Arima Genomics) library protocol for whole blood. Adaptor ligation with Unique dual indexing (Illumina) were chosen to match the indexes from the 10× linked-read library for simultaneous paired-end sequencing (150 bp) on the same lane on an Illumina HiSeq X platform. Both libraries were quality controlled using a Fragment analyzer NGS kit (AATI) and qPCR with the Kapa library quantification kit (Roche) prior to sequencing.

The sequencing was provided by the Norwegian Sequencing Centre (https://www.sequencing.uio.no, last accessed September 17, 2021), a national technology platform hosted by the University of Oslo and supported by the "Functional Genomics" and "Infrastructure" programs of the Research Council of Norway and the South-Eastern Regional Health Authorities.

Genome Size Estimation and Genome Assembly

The genome size of the reed warbler was estimated by a k-mer analysis of 10× reads using Jellyfish v. 2.3.0 (Marçais and Kingsford 2011) and Genome Scope v. 1.0 (Vurture et al. 2017), with a k-mer size of 21. The estimated genome size was 1,130,626,830 bp.

We assembled the long-read PacBio sequencing data with FALCON and FALCON-Unzip (falcon-kit 1.5.2 and falcon-unzip 1.3.5) (Chin et al. 2016). Falcon was run with the following parameters: length_cutoff = −1; length_cutoff_pr = 1000; pa_HPCdaligner_option = −v −B128 −M24; pa_daligner_option = −e0.8 −l2000 −k18 −h480 −w8 −s100; ovlp_HPCdaligner_option = −v −B128 −M24; ovlp_daligner_option = −k24 −e.94 −l3000 −h1024 −s100; pa_DBsplit_option = −x500 −s200; ovlp_DBsplit_option = −x500 −s200; falcon_sense_option = −output-multi −min-idt 0.70 −min-cov 3 −max-n-read 200; overlap_filtering_setting = −max-diff 100 −max-cov 100 −min-cov 2. Falcon-unzip was run with default settings. The purge_haplotigs pipeline v. 1.1.0 (Roach et al. 2018) was used to curate the diploid assembly, with -l5, -m35, -h190 for the contig coverage, and -a60 for the purge pipeline. Next, we scaffolded the curated assembly with the 10× reads using Scaff10X v. 4.1 (https://github.com/wtsi-hpag/Scaff10X, last accessed September 17, 2021), and the Hi-C reads using SALSA v. 2.2 (Ghurye et al. 2017). Finally, we polished the assembly (combined with the alternative assembly from Falcon-Unzip), first with PacBio reads using pbmm2 v. 1.2.1, which uses minimap2 (Li 2018) internally (v. 2.17), and then with 10× reads for two rounds with Long Ranger v. 2.2.2 (Marks et al. 2019) and FreeBayes v. 1.3.1 (Garrison and Marth 2012).

Curation

The assembly was decontaminated and manually curated using the gEVAL browser (Chow et al. 2016; Howe et al. 2021),

resulting in 521 corrections (breaks, joins and removal of erroneously duplicated sequence). HiGlass (Kerpedjiev et al. 2018) and PretextView (https://github.com/wtsi-hpag/PretextView, last accessed September 17, 2021) were used to visualize and rearrange the genome using Hi-C data, and PretextSnapshot (https://github.com/wtsi-hpag/PretextSnapshot, last accessed September 17, 2021) was used to generate an image of the Hi-C contact map. The corrections made reduced the total length of scaffolds by 0.5% and the scaffold count by 44.6%, and increased the scaffold N50 by 20.2%. Curation identified and confirmed 29 autosomes and the Z and W chromosomes, to which 98.6% of the assembly sequences were assigned.

Genome Quality Evaluation

We assessed the quality of the assembly with the assemblathon_stats.pl script (Bradnam et al. 2013) and investigated the completeness of the genome with Benchmarking Universal Single-Copy Orthologs (BUSCO) v. 5.0.0 (Simão et al. 2015), searching for 8,338 universal avian single-copy orthologs (aves_odb10).

Genome Annotation

We used a repeat library provided by Alexander Suh called bird_library_25Oct2020 and described in Peona et al. (2020) to softmask repeats in the reed warbler genome assembly. Softmasked genome assemblies for golden eagle (*Aquila chrysaetos*), chicken (*Gallus gallus*), great tit (*Parus major*), Anna's hummingbird (*Calypte anna*), zebra finch (*Taeniopygia guttata*), great reed warbler (*Acrocephalus arundinaceus*), icterine warbler (*Hippolais icterina*), collared flycatcher (*Ficedula albicollis*), and New Caledonian crow (*Corvus moneduloides*) were downloaded from NCBI. The triangle subcommand from Mash v. 2.3 (Ondov et al. 2016) was used to estimate a lower-triangular distance matrix, and a Python script (https://github.com/marbl/Mash/issues/9#issue-comment-509837201, last accessed September 17, 2021) was used to convert the distance matrix into a full matrix. The full matrix was used as input to RapidNJ v. 2.3.2 (Simonsen et al. 2008) to create a guide tree based on the neighbor-joining method. Cactus v. 1.3.0 (Armstrong et al. 2020) was run with the guide tree and the softmasked genome assemblies as input.

We also downloaded the annotation for chicken, and used it as input to the Comparative Annotation Toolkit (CAT) v. 2.2.1-36-gfc1623d (Fiddes et al. 2018) together with the hierarchical alignment format file from Cactus. Chicken was used as reference genome, reed warbler as the target genome and the AUGUSTUS (Stanke et al. 2008) species parameter was set to "chicken." InterProScan v. 5.34-73 (Jones et al. 2014) was run on the predicted proteins to find functional annotations, and DIAMOND v. 2.0.7 (Buchfink et al. 2021) was used to compare the predicted proteins against

UniProtKB/Swiss-Prot release 2021_03 (The UniProt Consortium 2021). AGAT v. 0.5.1 (Dainat 2021) was used to generate statistics from the GFF3 file with annotations and to add functional annotations from InterProScan and gene names from UniProtKB/Swiss-Prot. BUSCO v. 5.0.0 was used to assess the completeness of the annotation.

Synteny Analysis

We aligned the assembly against the great tit (*Parus major*) and the garden warbler (*Sylvia borin*) genome assemblies with minimap2 v. 2.18-r1015 and extracted only alignments longer than 5,000 bp. The bundlelinks from circos-tools v. 0.23 was used to merge neighboring links using default options and a plot was created using circos v. 0.69-8.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Author Contributions

C.L.C.S., F.E., K.R., K.S.J., O.K.T., and R.T. designed the research. E.K., K.R., and R.T. collected the sample. K.R. extracted DNA. C.L.C.S. and O.K.T. performed the research and/or analyzed the data. A.T., J.T., K.H., S.P., and W.C. curated the assembly. C.L.C.S. drafted the manuscript. All authors read and approved the final manuscript.

Data Availability

The reference genome of *Acrocephalus scirpaceus* (bAcrSci1), and the raw sequence data, have been deposited in the European Nucleotide Archive under the BioProject number PRJEB45715. Genome annotations are available at Figshare and can be accessed at https://doi.org/10.6084/m9.figshare.16622302.v1.

Literature Cited

Armstrong J, et al. 2020. Progressive Cactus is a multiple-genome aligner for the thousand-genome era. Nature 587(7833):246–251.

Both C, et al. 2010. Avian population consequences of climate change are most severe for long-distance migrants in seasonal habitats. Proc Biol Sci. 277(1685):1259–1266.

Bradnam KR, et al. 2013. Assemblathon 2: evaluating de novo methods of genome assembly in three vertebrate species. Gigascience 2(1):10.

Brommer JE, Lehikoinen A, Valkama J. 2012. The breeding ranges of central European and arctic bird species move poleward. PLoS One 7(9):e43648.

Buchfink B, Reuter K, Drost HG. 2021. Sensitive protein alignments at tree-of-life scale using DIAMOND. Nat Methods 18(4):366–368.

Chamorro D, Nieto I, Real R, Muñoz AR. 2019. Wintering areas on the move in the face of warmer winters. Ornis Fennica 96:41–54.

Chin CS, et al. 2016. Phased diploid genome assembly with single molecule real-time sequencing. Nat Methods 13(12):1050–1054.

Chow W, et al. 2016. gEVAL—a web-based browser for evaluating genome assemblies. Bioinformatics 32(16):2508–2510.

Dainat J. 2021. AGAT: Another Gff Analysis Toolkit to Handle Annotations in Any GTF/GFF Format (Version v0.5.1). Zenodo. Available from: https://www.doi.org/10.5281/zenodo.3552717. Accessed September 17, 2021.

Damas J, et al. 2017. Upgrading short-read animal genome assemblies to chromosome level using comparative genomics and a universal probe set. Genome Res. 27(5):875–884.

Davies NB, Brooke ML. 1989. An experimental study of co-evolution between the cuckoo, *Cuculus canorus*, and its hosts. I. Host egg discrimination. J Anim Ecol. 58(1):207–224.

Degrandi TM, et al. 2020. Introducing the bird chromosome database: an overview of cytogenetic studies in birds. Cytogenet Genome Res. 160(4):199–205.

Eglington SM, et al. 2015. Latitudinal gradients in the productivity of European migrant warblers have not shifted northwards during a period of climate change. Global Ecol Biogeogr. 24(4):427–436.

Ellegren H. 2010. Evolutionary stasis: the stable chromosomes of birds. Trends Ecol Evol. 25(5):283–291.

Feng S, et al. 2020. Dense sampling of bird diversity increases power of comparative genomics. Nature 587(7833):252–257.

Fiddes IT, et al. 2018. Comparative Annotation Toolkit (CAT) – simultaneous clade and personal genome annotation. Genome Res. 28(7):1029–1038.

Furo IDO, et al. 2018. Chromosome painting in Neotropical long- and short-tailed parrots (Aves, Psittaciformes): phylogeny and proposal for a putative ancestral karyotype for tribe Arini. Genes 9: 491. Available from: 10.3390/genes9100491. Accessed September 17, 2021.

Garrison E, Marth G. 2012. Haplotype-based variant detection from short-read sequencing. arXiv Preprint. Available from: https://arxiv.org/abs/1207.3907. Accessed September 17, 2021.

Ghurye J, Pop M, Koren S, Bickhart D, Chin CS. 2017. Scaffolding of long read assemblies using long range contact information. BMC Genomics 18(1):527.

Halupka L, Dyrcz A, Borowiec M. 2008. Climate change affects breeding of reed warblers *Acrocephalus scirpaceus*. J Avian Biol. 39(1):95–100.

Howe K, et al. 2021. Significantly improving the quality of genome assemblies through curation. GigaScience 10(1):giaa153.

Järvinen O, Ulfstrand S. 1980. Species turnover of a continental bird fauna: northern Europe, 1850-1970. Oecologia 46(2):186–195.

Jones P, et al. 2014. InterProScan 5: genome-scale protein function classification. Bioinformatics 30(9):1236–1240.

Kerpedjiev P, et al. 2018. HiGlass: web-based visual exploration and analysis of genome interaction maps. Genome Biol. 19(1):125.

Kralj J, Procházka P, Fainová D, Patzenhauerová H, Tutiš V. 2010. Intraspecific variation in the wing shape and genetic differentiation of reed warblers Acrocephalus scirpaceus in Croatia. Acta Ornithol. 45(1):51–58.

Kretschmer R, Ferguson-Smith MA, de Oliveira EHC. 2018. Karyotype evolution in birds: from conventional staining to chromosome painting. Genes 9:181.

Kretschmer R, et al. 2021. Karyotype evolution and genomic organization of repetitive DNAs in the saffron finch, Sicalis flaveola (Passeriformes, Aves). Animals 11(5):1456. Available from: 10.3390/ani11051456. Accessed September 17, 2021.

Laine VN, Great Tit HapMap Consortium, et al. 2016. Evolutionary signals of selection on cognition from the great tit genome and methylome. Nat Commun. 7:10474.

Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34(18):3094–3100.

Manel S, et al. 2016. Genomic resources and their influence on the detection of the signal of positive selection in genome scans. Mol Ecol. 25(1):170–184.

Marçais G, Kingsford C. 2011. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. Bioinformatics 27(6):764–770.

Marks P, et al. 2019. Resolving the full spectrum of human genome variation using linked-reads. Genome Res. 29(4):635–645.

Meller K, Piha M, Vähätalo AV, Lehikoinen A. 2018. A positive relationship between spring temperature and productivity in 20 songbird species in the boreal zone. Oecologia 186(3):883–893.

Ondov BD, et al. 2016. Mash: fast genome and metagenome distance estimation using MinHash. Genome Biol. 17(1):132.

Pala I, et al. 2012. Evidence of a neo-sex chromosome in birds. Heredity (Edinb). 108(3):264–272.

Peona V, et al. 2020. The avian W chromosome is a refugium for endogenous retroviruses with likely effects on female-biased mutational load and genetic incompatibilities. Philos Trans R Soc B. 376:20200186.

Procházka P, et al. 2011. Low genetic differentiation among reed warbler Acrocephalus scirpaceus populations across Europe. J Avian Biol. 42(2):103–113.

Roach MJ, Schmidt SA, Borneman AR. 2018. Purge Haplotigs: allelic contig reassignment for third-gen diploid genome assemblies. BMC Bioinformatics 19(1):460.

Røed T. 1994. "Rørsanger." In Norsk Fugleatlas, 382–383. Norsk Ornitologisk Forening.

Salewski V, Hochachka WM, Fiedler W. 2010. Global warming and Bergmann's rule: do central European passerines adjust their body size to rising temperatures? Oecologia 162(1):247–260.

Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular cloning: a laboratory manual, 2nd ed. New York: Cold Spring, Harbor Laboratory Press.

Schaefer T, Ledebur G, Beier J, Leisler B. 2006. Reproductive responses of two related coexisting songbird species to environmental changes: global warming, competition, and population sizes. J Ornithol. 147(1):47–56.

Shafer ABA, et al. 2017. Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. Methods Ecol Evol. 8(8):907–917.

Sigeman H, Strandh, M, et al. 2020. Genomics of an avian neo-sex chromosome reveals the evolutionary dynamics of recombination suppression and sex-linked genes. bioRxiv Preprint, Available from: 10.1101/2020.09.25.314088. Accessed September 17, 2021.

Sigeman H, Ponnikas S, Hansson B. 2020. Whole-genome analysis across 10 songbird families within Sylvioidea reveals a novel autosome-sex chromosome fusion. Biol Lett. 16(4):20200082.

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics. 31(19):3210–3212.

Simonsen M, Mailund T, Pedersen CNS. 2008. Rapid neighbour-joining. In: Crandall KA, Lagergren J, editors. Algorithms in bioinformatics. WABI 2008. Lecture Notes in Computer Science. Vol 5251. Berlin, Heidelberg (Germany): Springer. p. 113–122. Available from: 10.1007/978-3-540-87361-7_10. Accessed September 17, 2021.

Skinner BM, Griffin DK. 2012. Intrachromosomal rearrangements in avian genome evolution: evidence for regions prone to breakpoints. Heredity (Edinb). 108(1):37–41.

Stanke M, Diekhans M, Baertsch R, Haussler D. 2008. Using native and syntenically mapped cDNA alignments to improve de novo gene finding. Bioinformatics 24(5):637–644.

Stokke B, et al. 2018. Characteristics determining host suitability for a generalist parasite. Sci Rep. 8(1):6285.

Stolt BO. 1999. The Swedish reed warbler Acrocephalus scirpaceus population estimated by a capture-recapture technique. Ornis Svecica 9:35–46.

Sætre CL, et al. 2017. Rapid adaptive phenotypic change following colonization of a newly restored habitat. Nat Commun. 8:14159.

The UniProt Consortium. 2021. UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Res. 49:D480–D489.

Thorogood R, Davies NB. 2013. Reed warbler hosts fine-tune their defenses to track three decades of cuckoo decline. Evolution 67(12):3545–3555.

Thorogood R, Spottiswoode CN, Portugal SJ, Gloag R. 2019. The coevolutionary biology of brood parasitism: a call for integration. Philos Trans R Soc B. 374:20180190.

Vickery JA, et al. 2014. The decline of Afro-Palaearctic migrants and an assessment of potential causes. Ibis 156(1):1–22.

Vurture GW, et al. 2017. GenomeScope: fast reference-free genome profiling from short reads. Bioinformatics 33(14):2202–2204.

Associate editor: Bonnie Fraser