Standing genetic variation and *de novo* mutations underlie parallel evolution of island bird phenotypes

Andrea Estandía[1,✉], Ashley T. Sendell-Price[1], Bruce C. Robertson[2], and Sonya M. Clegg[1]

[1]Edward Grey Institute of Field Ornithology, Department of Biology, University of Oxford, United Kingdom
[2]Department of Zoology, University of Otago, Dunedin, New Zealand

Parallel evolution occurs when the same trait evolves in closely related lineages in response to similar ecological contexts and provides some of the best examples of determinism in evolutionary biology. However, from a genetic standpoint, this process can be driven by either new mutations that appear independently in each diverging population or by selection on existing genetic variation common to both lineages. Small-bodied birds, for example, tend to increase in size after they colonise a new island, following what is known as the 'island rule'. Such is the case of the silvereye (*Zosterops lateralis*), a prolific natural coloniser of southwest Pacific islands. Island forms of this bird species increase in body size after they establish, with the pattern and pace of change consistent with directional natural selection and evident even in the most recent colonisations within the last 200 years. The system provides an exceptional opportunity to explore the genomic basis of repeated body size evolution. We sequenced 377 whole genomes from 31 different silvereye populations, which revealed that both mechanisms are at play: in some lineages, new mutations are highly associated with body and bill size, but there are also highly associated polymorphisms present across all populations. Our research sheds light on the genomic basis of repeated body size evolution and emphasises that multiple molecular mechanisms can underlie similar evolutionary trajectories even within a single taxon.

parallel evolution | adaptation | island rule | GWAS | inversion | structural rearrangements | mutations | body size | standing genetic variation | birds
Correspondence: ***andrea.estandia@biology.ox.ac.uk***

Introduction

Parallel evolution refers to the independent emergence of similar traits in different lineages that share a common ancestor (Cerca 2023), and illustrates deterministic evolutionary outcomes that are widespread across the Tree of Life (Blount et al. 2018). Natural selection is typically responsible for parallel evolution, as genetic drift is unlikely to produce such similar changes (Arendt and Reznick 2008; Stayton 2008; Losos 2011; McGhee 2011). Selection can act on standing genetic variation to produce the same phenotype (Lai et al. 2019; Therkildsen et al. 2019; Konečná et al. 2021), on different *de novo* mutations in each population (Fulgione et al. 2022; Cerca et al. 2023), or a combination of the two (Ament-Velásquez et al. 2022). However, a faster rate of adaptive change is more often facilitated from standing variation than from mutations as the latter are mostly neutral or mildly deleterious (Hedrick 1983; Barrett and Schluter 2008). For example, the rapid repeated evolution of stickleback low-plated phenotypes is driven by reassembling ancestral variation (Colosimo et al. 2005; Archambeault et al. 2020). Additionally, standing genetic variation can be supplemented with genetic variants that contribute to parallel evolution shared via admixture among recently diverged lineages (Taylor et al. 2020; Montejo-Kovacevich et al. 2022; Cerca et al. 2023; Wang et al. 2023). The same processes can affect single polymorphisms or full blocks in the genome: chromosomal rearrangements, like inversions, can play a role in parallel evolution by keeping sets of adaptive alleles together in strong linkage disequilibrium (Faria et al. 2019; Harringmeyer and Hoekstra 2022). Understanding the relative importance of alternative genetic mechanisms behind parallel evolution has received considerable interest over the last few decades (MacPherson and Nuismer 2017; Westram et al. 2022). Genomic methods now allow us to investigate examples spanning diverse systems, providing empirical data that is necessary to draw generalities.

Islands provide environments where comparable evolutionary pressures can result in the rapid emergence of similar phenotypes (Baeckens and Van Damme 2020). The 'island syndrome' refers to the phenomenon where sets of morphological, physiological, and behavioural traits evolve in a predictable way when organisms colonise islands. The syndrome ultimately results from the exposure to a common suite of biotic conditions on islands; a combination of reduced predation and parasitism, and a shift in the balance of intra- and interspecific competition, features that fundamentally change the selective pressures in predictable and repeatable ways (Clegg and Owens 2002; Meiri et al. 2004; Lomolino 2005). Arguably the best-characterised component of this syndrome is the 'island rule' that describes evolution towards medium body sizes in insular settings; small colonising species evolve larger body sizes and vice versa (Foster 1964; Clegg and Owens 2002; Benítez-López et al. 2021). There is evidence for the island rule across vertebrate taxa, including birds (Benítez-López et al. 2021). Bill size also follows this rule whereby short-billed species show an increase in size (Clegg and Owens 2002). Body and bill size are complex multivariate quantitative traits predicted to be highly polygenic (Yang et al. 2011b; Bosse et al. 2017; Duntsch et al. 2020), but some studies in natural populations show that a few genes of major effect can also impact these traits (Lamichhaney et al. 2015, 2016; vonHoldt et al. 2018; Duntsch et al. 2020).

Members of the silvereye subspecies complex (*Zosterops lateralis*, Aves) have repeatedly colonised islands across the South Pacific from the Australian mainland and Tasmania, with up to 17 morphological subspecies currently recognised (Clements et al. 2021). Insular silvereyes are larger than

mainland ones, and the degree of this shift increases with time since colonisation (Clegg et al. 2002a), a pattern that has been shown to be a consequence of directional selection rather than drift (Clegg and Owens 2002; Clegg et al. 2008). The island populations represent at least four independent colonisations from the Australian mainland: to Southern Melanesia (∼150,000 years ago), Lord Howe Island (∼100,000 years ago), southern Great Barrier Reef Islands, including Heron Island (∼4000 years ago) and a recent historically recorded sequential colonisation sequence from Tasmania to New Zealand and outlying islands (<200 years ago) (Black 2010) (Fig. 1). The populations arising from these independent colonisations are genetically isolated from the original mainland source and other colonisations, though gene flow occurs among Southern Melanesian islands and among recently colonised populations (Estandía et al. 2023). The combination of independent and known colonisation sequences, along with broadly repeated shifts in body and bill size in similar ecological contexts provides an excellent system to explore the genetic mechanisms of parallel evolution. Here, we ask whether *de novo* mutations or standing genetic variation lead to the same phenotype: increased body and bill size in small-bodied species after island colonisation. We find that both are associated with body size variation in Southern Melanesia, whereas only standing genetic variation appears important for repeated patterns in all other populations. This demonstrates empirically that two genetic mechanisms hypothesised to underlie repeated evolution can operate simultaneously in a single system.

Results

The genetic basis of repeated body and bill size evolution.

To describe and quantify the bill and body size shifts in the populations studied here, we summarised univariate traits as principal component scores, resulting in a PC1 for body, and a PC1 for bill that represented size. In both cases 75% of the variation was explained by PC1 and roughly equal trait loadings of univariate traits (Fig. 1; Table S1). We statistically confirmed increases in body and bill size (Fig. S1; Table S2). To examine the genetic basis of repeated evolution, we generated low-to-intermediate coverage (5X) whole-genome sequencing data of 377 silvereyes from 24 islands and sampled across seven Australian States and Territories representative of most of the silvereye range (Table S3). After alignment of the sequenced individuals, we analysed genetic variation in a probabilistic framework accounting for low coverage (ANGSD (Korneliussen et al. 2014)) and obtained 14.3 million SNPs with minor allelic frequencies higher than 5% for differentiation analyses.

Population structure was described by a genomic PCA based on whole-genome data that revealed three major groups (Fig. S2 A). PC1 explained 20.18% of the total variance and separated two clusters: one containing individuals from Southern Melanesia (SM cluster henceforth) and the other containing the remaining individuals (ANZO cluster). PC2 explained 6.8% of the variance and separated the two main archipelagos within SM: New Caledonia and Vanuatu. This population structure aligns with previous analyses using fewer individuals (Estandía et al. 2023). We visualised further substructure within the SM and ANZO clusters (Fig. S2 B-C).

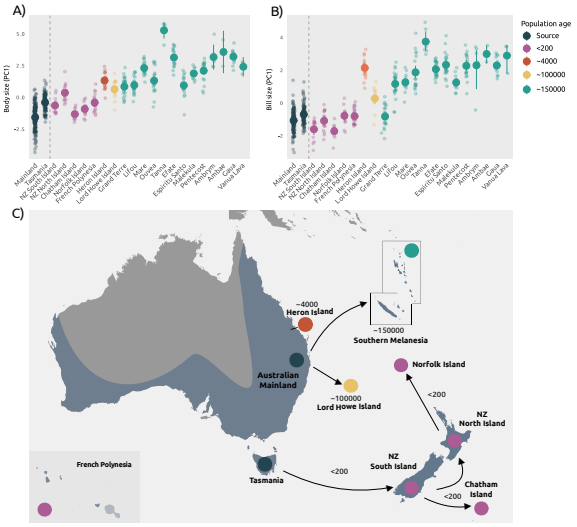

Fig. 1. Morphological divergence in A) body and B) bill size (PC1) among silvereye forms. The source populations are placed left to the vertical dashed line. Colonisations are arranged chronologically: the most recent ones on the left and the most ancient on the right. C) The map shows the silvereye distribution highlighted in blue and the time of colonisation of each island or archipelago. French Polynesia is shown as inset.

We controlled for the population structure described above by using the eigenvectors that explained most of the variation, and conducted a GWAS that identified multiple loci linked to body size variation suggesting a polygenic basis for the trait. However, given the complicated population structure of our system, we retained those associated variants within the top 50 for which we had evidence of intra-population variation associated with body size, a total of 36 variants. We found that eight were classified as putative *de novo* mutations exclusive to islands in Southern Melanesia (Fig. 2A; Table S4), five of which were part of an association peak on Chromosome 2, which also contained one annotated gene, *GRHPR*, important in metabolic processes (Rumsby and Cregeen 1999). For the top outliers, those individuals with the genotype BB showed a five times increase in median body size relative to those with the AA genotype (Table S5), and the effect was still strong when we only explored Central Vanuatu, where the putative new mutation seems to be restricted to (Fig. 2A). This associated peak emerged again when we conducted GWAS on Southern Melanesia alone (Fig. S3). The effect size of the SNPs across this region was greater than any of the covariates included in the model to capture any population structure, suggesting no signs of conflation (Fig. S4). Other size-associated loci showed a similar effect in direction and magnitude on body size. Among these, were genes involved in calmodulin binding, cytoskeletal motor activity, organisation of the extracellular matrix and other

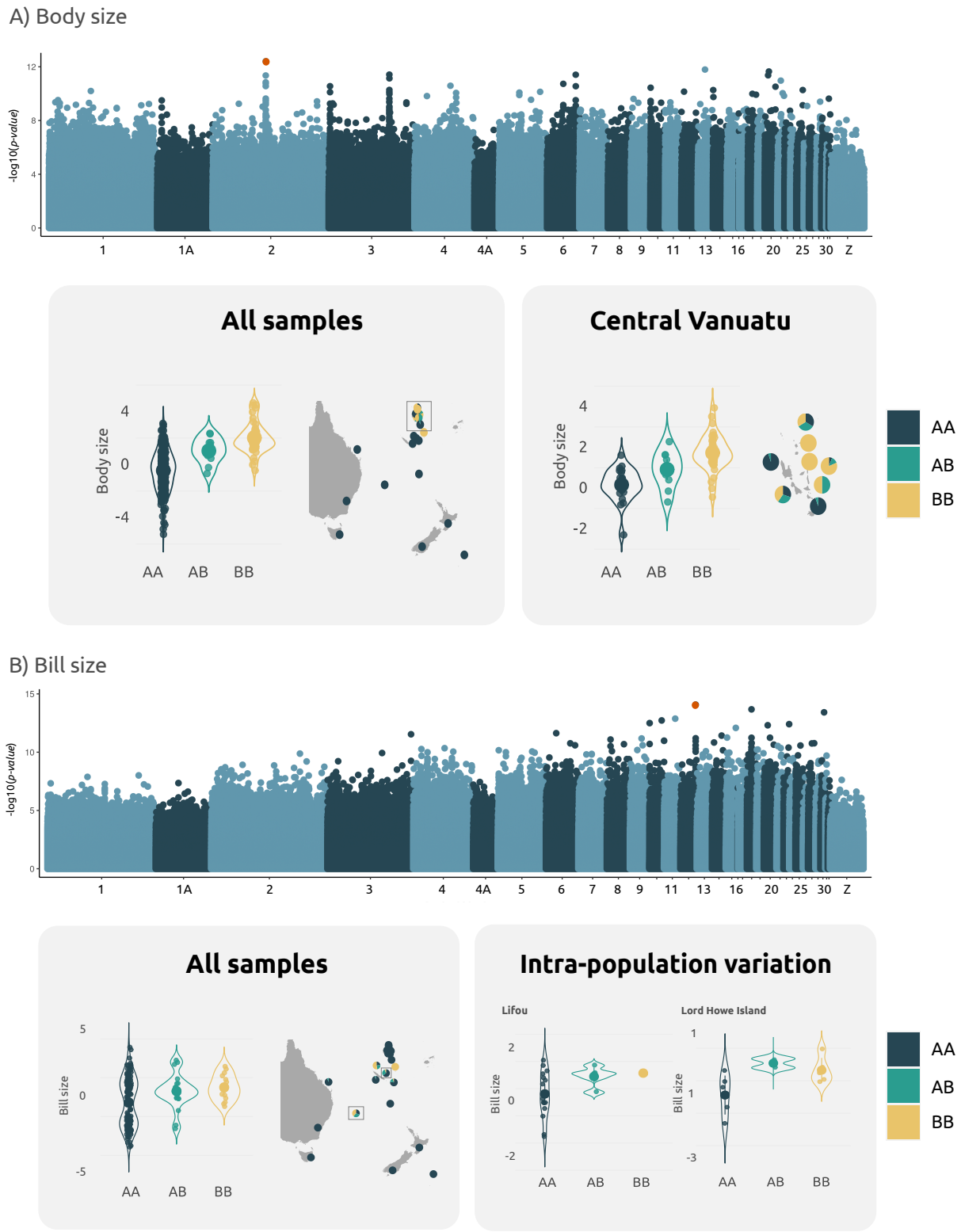

Fig. 2. A) Manhattan plot for body size (PC1) and correlation between allele frequency for the top-ranked SNP and its geographic distribution across all samples (left) and restricted to Central and North Vanuatu, partly where the putative *de novo* mutations occur (right). It is important to note that this set of islands show minimal population structure. B) Manhattan plot for bill size (PC1) and correlation between allele frequency for the top-ranked SNP and its geographic distribution across all samples (left) and across two of the populations where there is intra-population variation associated with body size (right). These two populations are part of separate colonisations.

metabolic processes (Table S6).

All of the top associated loci with bill size that were variable within populations represented standing genetic variation (Fig. 2B; Table S7). A number of adjacent variants on Chromosomes 12 and 18 were found in several of the large bill-size populations: islands in Southern Melanesia, Heron Island and Lord Howe Island. Among the annotated genes for bill size, we found gene *LGR6* (Table S8).

We quantified the proportion of phenotypic variance explained by all genome-wide SNPs with GCTA (Yang et al. 2011a). These narrow-sense heritability estimates were high but showed wide confidence intervals (CI) (Table S9). Tarsus length is a good proxy for body size and it is a highly heritable trait based on pedigree data, with a heritability of 0.7 for Heron Island silvereyes (Clegg et al. 2008), a higher value than our SNP-based estimate (h^2=0.5 (95% CI =0.3-0.8)).

Structural variation: a neo-sex chromosome and multiple putative inversions.

To examine the structural variation along the genome, we identified clusters of PCA windows (local PCAs) based on Multidimensional Scaling (MDS) (Fig. 3-4). This method uses common patterns to identify structural variation such as non-recombining haplotype blocks (Li and Ralph 2019; Mérot et al. 2020). We found a large outlier region on Chromosome 4A (Fig. 3A-3B). MDS1 separated two distinct groups that aligned with sex (Fig. 3C), where sex was independently confirmed by PCR-based sexing of individuals.

To specifically search for putative inversions, we filtered those sections of the genome with at least five adjacent outlier windows, pooled the windows and repeated MDS analysis for each separate region. Several chromosomes showed large outlier regions that formed three discrete equidistant clusters along MDS1. This type of pattern is highly suggestive of a potential inversion with each alternative homokaryotype separated by the intermediate heterokaryotype (Westram et al. 2022) (Fig. 4B). Consistent with the occurrence of inversions, LD was high inside these regions contrasting to the LD along the rest of the chromosome (Fig. S5), indicating limited recombination in heterokaryotypes. For the potential inversion on Chromosome 27, the AA karyotype (where A indicates the most common version of the arrangement) was present at high frequencies in all intermediate and old colonisations (Southern Melanesia, Lord Howe Island and Heron Island), and at very low frequency on the mainland. The potential inversion on Chromosome 28 showed a different pattern with the AA karyotype exclusively present in Southern Melanesia. On Chromosome 29, BB was only recovered from individuals in Southern Melanesia, but some individuals from New Caledonia displayed the AA karyotype (Fig. 4). Body size was not associated with karyotype for Chromosomes 27 and 29 after controlling for population structure. Body size was associated with the AA karyotype in Southern Melanesia, but we could not control for population structure as karyotype and neutral variation structure were entirely coincident. Using our custom annotation, we found genes involved in various metabolic processes on Chromosomes 27 and 28, and lipogenesis, adipocyte differentiation and muscle tissue development on Chromosome 29 (Table S10).

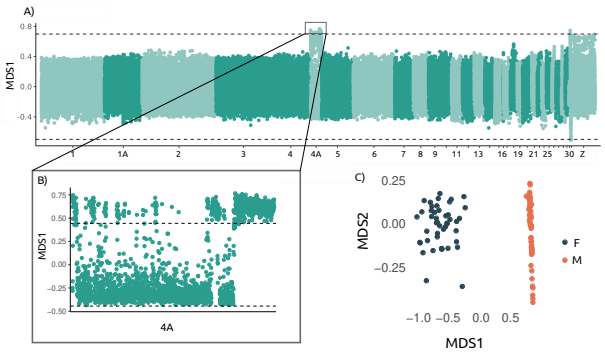

Fig. 3. A) Local PCA results (MDS1) plotted across the genome B) where chromosome 4A shows outlier windows each representing 100 SNPs. C) MDS1 and MDS2 separate by sex in this region.

Discussion

Selection on standing genetic variation is a common route to repeated evolution of larger body and bill sizes in island-colonising silvereyes. Size-associated variants were shared across colonisation histories spanning hundreds to over a hundred thousand years. However, putative *de novo* mutations present only in Southern Melanesia silvereye populations were also linked to increased body sizes. This empirical system suggests the joint role of standing genetic variation and *de novo* mutations in producing a repeated island phenotype.

We identified several candidate genes common across all populations. This pattern suggests that variation present in ancestral source populations was represented in multiple, independent colonising flocks, and maintained across sequential island colonisations. Given that population founding can result in loss of genetic variation and the fate of rare alleles can be highly stochastic especially when populations remain small (Whitlock 2000; Sendell-Price et al. 2021), such an outcome might seem unlikely. However, silvereyes colonise in reasonably large flocks (>100 individuals), and they recover rapidly from small population sizes, minimising loss of variation during founding and population establishment (Clegg et al. 2002a; Sendell-Price et al. 2021). Historically colonised silvereye populations and Heron Island silvereyes have experienced rapid trait evolution driven by natural selection (Clegg et al. 2002b, 2008; Sendell-Price et al. 2020), and such rapid evolution is more likely to stem from standing genetic variation than *de novo* mutations (Marques et al. 2019).

However, an association peak harbouring putative *de novo* mutations was present exclusively in Southern Melanesia, the first region to be colonised by silvereyes. These mutations were not fixed across all populations: some of them had fully fixed BB genotypes associated with very large body sizes, while others lacked the minor allele entirely, and many displayed all genotypes. This pattern held even within popula-

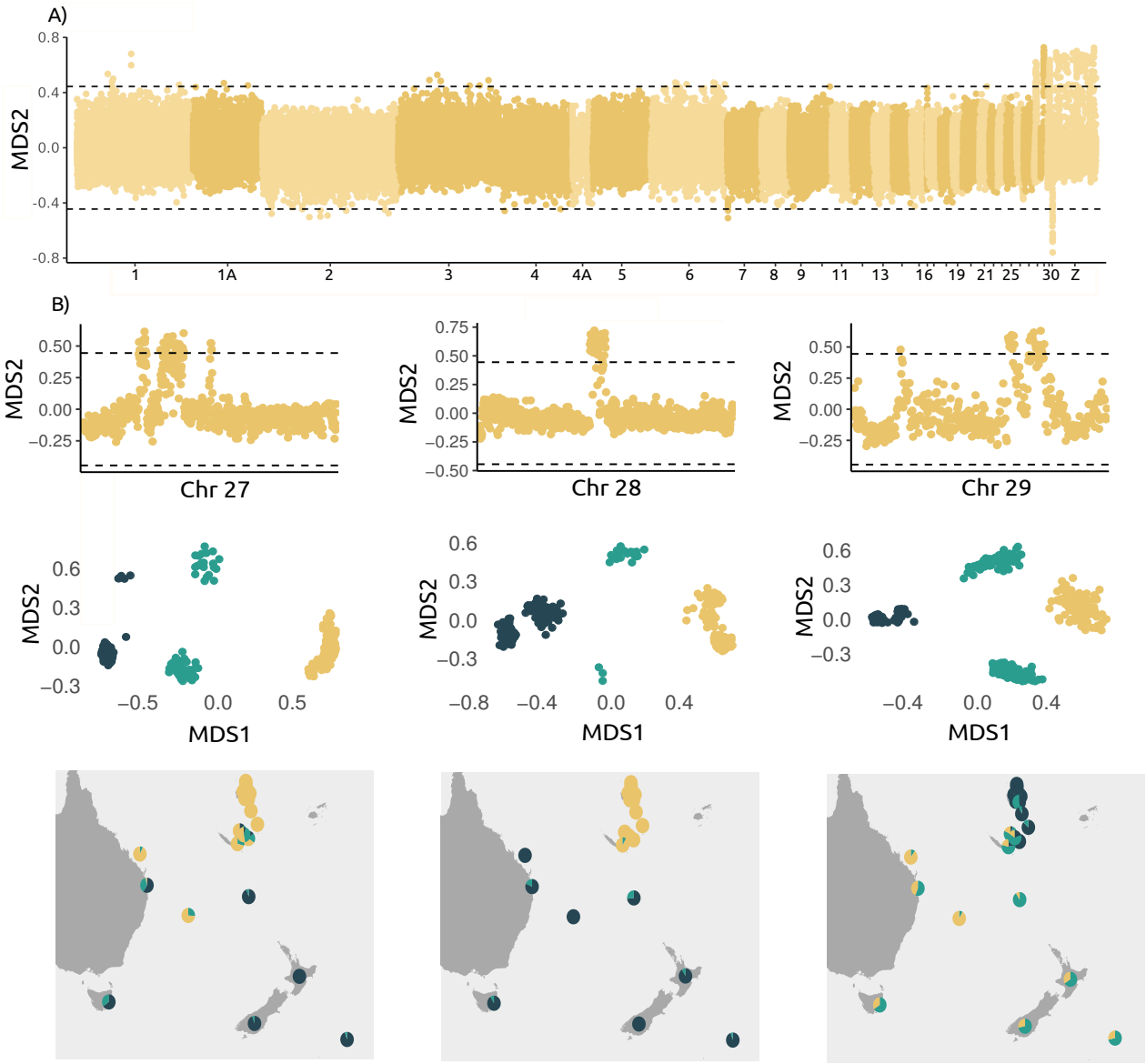

Fig. 4. A) Local PCA results (MDS2) are displayed across the entire genome. B) Outlier regions in chromosomes 27, 28, and 29. Each dot corresponds to 100 SNPs. MDS1 identifies three distinct clusters within the outlier regions, suggesting potential inversions. Chartpies provide information about the relative frequencies of each cluster within different populations.

tions, indicating no bias due to population structure. Variation could be maintained across the archipelagos if different degrees of size are favoured in different populations connected by gene flow as seen across Southern Melanesian silvereye populations (Clegg and Phillimore 2010; Estandía et al. 2023). While larger size of island silvereyes appears generally to be favoured, other species, and indeed other *Zosterops* white-eye species that co-occur, could impose local constraints on size. For example, on Lifou, an island in New Caledonia, three Zosterops species occur in sympatry: the relatively small-sized endemic *Z. minutus*, the large endemic *Z. inornatus*, with the silvereye being intermediate in body size between the two (Mees 1969). An alternative but untested possibility is that variants arise from introgression with sympatric *Zosterops* species (Gill 1970). Other mechanisms for maintaining variation, such as negative frequency-dependent selection, are more difficult to reconcile in systems like islands where the usual drivers, e.g. predation risk, are limited (Beauchamp 2021).

Our study was unable to determine the prevalence of *de novo* mutations versus standing genetic variation. Even though we were able to detect new mutations in the oldest colonisation, our sample size was denser for this region than for any of the other of the old-to-intermediate colonisations which might have affected detection of association loci private to these populations. For these reasons, we can only say that in our study system, both routes seem to underlie body size variation, but their relative importance remains an important

question to tackle in future studies.

Although linking annotated gene function to phenotypes in non-model organisms is highly speculative, any associations can provide the basis for future investigation. The putative *de novo* association peak on Chromosome 2 associated with body size has a hit for *GRHPR*. This gene encodes the enzyme glyoxylate reductase/hydroxypyruvate reductase, and is involved in key metabolic processes, including serine metabolism for gluconeogenesis and oxalate metabolism (Rumsby and Cregeen 1999). No other previous study has linked body size variation with this gene, yet, avian comparative genomic studies suggest that enzymes involved in glyoxylate metabolic routes might be responsible for diet-related adaptations in birds (Zhang et al. 2014). The function of other genes associated with body size is variable, from bone and muscle development to metabolic processing, and even though their role in body size is feasible, exact mechanisms remain to be found.

LGR6 was one of the genes associated with bill size variation. This gene acts as a mediator for the Wnt/β-catenin signalling pathway, known to be involved in the formation of the avian beak (Bhullar et al. 2015). Interestingly, we did not find significant associations with *ALX1* or *HMGA2*, two large-effect genes that have been repeatedly demonstrated to be implicated in bill morphology variation, including in Darwin's finches (Lamichhaney et al. 2015, 2016). However, different genetic architectures are not entirely surprising; even other species of finches show different pathways underlying bill size variation (vonHoldt et al. 2018).

We identified signatures of structural variants, one of which could potentially contribute to body size variation on chromosome 28, but others where such a relationship was clearly absent. Rearrangements on chromosomes 27, 28 and 29 each showed a pattern consistent with an inversion: three different clusters separated by MDS1 and high LD within the putative inversion compared with the rest of each chromosome (Mérot et al. 2020; Westram et al. 2022). Our conservative size estimates of the putative inversions observed here (approximately 200kb for those on Chromosomes 28 and 29 and 650 kb for Chromosome 27) place them in the range of inversion sizes known to be most frequently associated with adaptive phenotypes (100kb to several Mb) (Faria et al. 2019). Reduced recombination in the heterokaryotypes preserves linkage among alleles despite mating between alternative homokaryotypes, maintaining the adaptive benefit of the haplotype blocks (Faria et al. 2019; Westram et al. 2022). For example, marine snails (*Littorina saxatilis*) and seaweed flies (*Coelopa frigida*) show inverted regions linked to adaptive phenotypes (Morales et al. 2019; Koch et al. 2021); the latter has an inversion under balancing selection that works as a supergene for body size. The karyotype distribution of the putative inversion identified on Chromosome 28 was associated with body size but was conflated with geographic population structure. The function of genes inside this region varied widely, some of them are involved in immunity, others in cholesterol metabolism and some are even directly linked to fitness proxies in birds (Krist 2011), like egg quality (Gao et al. 2021). However, it is unclear whether they contributed to the evolution of large body sizes in Southern Melanesia. To further study the role of structural variants we could generate long-read sequencing and recombination maps that help to confirm that they are indeed inversions and to determine breakpoint locations (Mérot et al. 2020; Akopyan et al. 2022).

The three putative inversions displayed karyotype distribution patterns that suggested varying evolutionary histories. The inversion on Chromosome 28 could have emerged after colonisation from the Australian mainland and could have been subsequently fixed in the majority of Southern Melanesian populations. Aside from sampling considerations, alternative scenarios are more complex, requiring different combinations of karyotype loss and maintenance that seem less likely. For putative inversions on both Chromosomes 27 and 29, the maintenance of the heterokaryotype and alternative homokaryotype in many populations suggest they could represent ancient inversions under balancing selection. Such polymorphism can be maintained when the inversion provides a fitness advantage against different selective pressures: if there are epistatic interactions among alleles at different loci (Charlesworth and Charlesworth 1973), if there is associative overdominance (Ohta 1971), or if a locus that is under balancing selection is located within the inversion (Faria et al. 2019). Similar to alternative processes that could explain SNP variation, karyotypes could appear to be balanced if there is introgression resulting from gene flow among different sympatric Zosterops species occurring across Southern Melanesia, on Norfolk Island and prior to 20th-century extinction, also on Lord Howe Island, but not Heron Island or New Zealand.

A sizeable structural variant detected via local PCA and visualised by MDS was on chromosome 4A, coinciding with a known neo-sex chromosome in the Zosteropidae (Leroy et al. 2021), formed via fusion between an autosome and a sex chromosome approximately 24 Ma in the Sylvioidea superfamily to which the Zosteropidae belong (Pala et al. 2012; Sigeman et al. 2020). Neo-sex chromosomes can play a role in speciation as they can contain loci for traits that contribute to reproductive isolation (e.g. male courtship display behavioural traits (Kitano et al. 2009)). However, the role of this neo-sex chromosome in speciation within the Sylvioidea, including in the Zosteropidae, is unclear (Pala et al. 2012).

The study of non-model systems in the wild has inherent practical and bioinformatic limitations. Like any empirical study, we had to balance sample size and genomic coverage. GWAS usually relies on a high number of whole genomes with attached individual data. Capturing and taking samples from the birds used in this study covering a vast geographical distribution was a tremendous effort that spanned decades, and sequencing hundreds of genomes remains relatively expensive. Additionally, body and bill sizes are complex multivariate phenotypes that are highly polygenic (Duntsch et al. 2020) and correlate with several other traits (e.g. Estandía et al. 2021) that we did not investigate. Lastly, identifying the causal genes is a major challenge and often needs extra

steps following GWAS to confirm the causality of the candidate loci, for example by doing knock-outs (Bengston et al. 2018), which is not possible in wild birds.

Despite these caveats and limitations, work on non-model systems is essential to gain a broad perspective on the different genetic routes that lead to similar phenotypes. The application of genomic methods to the silvereye system, which encompasses a wide range of island colonisation times and repeated evolution of body and bill size, provides a rare example of the role of standing genetic variation on both short and longer timescales of colonisation, with putative *de novo* mutations contributing only to the oldest island populations.

Methods

Sample collection.

We sampled 377 silvereyes (*Z. lateralis*) from 31 islands or sites on the Australian mainland between 1997 and 2017 (Table S3). Birds were caught in the wild using mist nets or traps and we collected 20–40 µl of blood from the brachial wing vein and stored it in 1 ml of Queen's lysis buffer (Seutin et al. 1991). We took a suite of morphological measurements, including the wing length using a metal ruler to determine the maximum flattened chord of the longest primary feather, and the length of the central tail feathers from the base to the tip using dividers. We used dial calipers (accuracy ±0.1 mm) to measure metatarsal length, head length from the rear of the skull to the bill's tip. Additionally, we took various bill measurements using dial calipers, including mandible length, width, and depth at the posterior nostril opening, and mandible length, width, and depth at the anterior nostril opening. Finally, we measured the body weight. Sex cannot be determined in silvereyes with morphological measurements. We selected a random subset of individuals and sexed them using molecular methods described in the section below. All sampling was non-destructive, and we released the birds at point of capture. Complete sampling information is provided in Table S3.

Construction of a pseudochromosome assembly.

We used the Chromosembler tool available in Satsuma2 (Grabherr et al. 2010) to order and orient scaffolds of the Z. l. melanops genome assembly (GCA_001281735.1) (Cornetti et al. 2015) into 33 pseudo-chromosomes according to synteny with the Vertebrate Genome Project's Zebra finch (*Taeniopygia guttata*) genome assembly (GCA_003957565.1) (Rhie et al. 2021).

To reconstruct a transcriptome that we could use to generate hints for gene prediction, we used available RNA-seq data for *Zosterops borbonicus* (Leroy et al. 2021). We filtered out low-quality sequence reads (phred quality score < 20) and adaptor sequences using TrimGalore (Martin 2011) and mapped them onto our pseudo-chromosome assembly using HISAT2 (Kim et al. 2015). We used Trinity (Grabherr et al. 2011) to assemble the transcriptome.

Assembly annotation.

We annotated the resulting pseudo-chromosome assembly using a multifaceted approach of repeat identification and *ab initio* gene prediction using AUGUSTUS (Stanke et al. 2006). We incorporated hints containing extrinsic evidence about the location and structure of three types of genomic features: repetitive regions, short-read transcriptome sequences, and single-copy orthologs. To obtain our hints, we first identified and masked repetitive and low-complexity DNA sequences using RepeatMasker (Smit et al. 2013) with the chicken (*Gallus gallus*) database of repeats. Second, we used BLAT (Kent 2002) to incorporate RNA-seq information as hints. Third, we used single-copy genes in the Aves lineage group odb10 to detect conserved orthologous genes with BUSCO (Manni et al. 2021). We then used BLASTx (Camacho et al. 2009) in the AUGUSTUS output to obtain functional information. More information about specific parameters used, including cut-offs, can be found at https://github.com/andreaestandia/1.0_island_rule_GWAS/tree/main/pipeline/9_annotation

DNA extraction, molecular sexing, library preparation, and sequencing.

We extracted DNA from blood samples using a standard Phenol-chloroform protocol (Seutin et al. 1991). We added 100 uL of blood stored in Queen's buffer to a microcentrifuge tube containing 250 uL of DIGSOL extraction buffer (0.02 M EDTA, 0.05 M Tris-HCl (pH 8.0), 0.4 M NaCl, 0.5% sodium dodecyl sulphate (SDS)) and 10uL of Proteinase K (20mg/mL). Samples were incubated at 55C overnight. Following incubation, we added 250 uL of phenol:chloroform:isoamyl alcohol (25:24:1) and gently mixed the samples for 10 minutes. We then centrifuged the samples at 10,000 rpm for 10 minutes. We transferred the aqueous layer to a new microcentrifuge tube and repeated the previous step by adding the phenol mixture, mixing and centrifuging it. We recovered a new aqueous phase, transferred it to a new tube, added 250uL of chloroform:isolamyl alcohol (24:1), mixed and centrifuged the samples as in the previous steps. We precipitated the DNA by adding 2 volumes of ice-cold 100% ethanol, 1 volume of 2.5M ammonium acetate and 2 uL of glycogen, leaving it overnight at -20C. The next morning we centrifuged the tubes for 10 min at 15,000 rpm at 4C and removed the supernatant. We rinsed the precipitate with 500 uL of ice-cold 70% ethanol and centrifuged the tubes as in the previous step. We removed the supernantant and the left the precipitated DNA to dry at room temperature. Once dried, the DNA was resuspended in 50 uL of TE (TrisEDTA) buffer (0.01 M Tris-HCL (pH 8.0), 0.0001 M EDTA). We quantified the DNA concentration with a Qubit 2.0. DNA extracts were sent to Novogene UK for library preparation and whole genome sequencing at 5X on the Illumina Novaseq 6000 platform (Illumina, San Diego) using paired-end 150bp sequencing reads.

Quality control and genotype calling.

We trimmed reads to remove adapter content and base calls of low quality using fastp (Chen et al. 2018) with default settings and 10 bp trimmed from the start of each read.

We aligned cleaned reads of each individual to our pseudo-chromosome assembly using the Burrows-Wheeler Aligner (BWA)-mem algorithm (Li and Durbin 2010). Genotype calling can result in bias when using low-coverage sequencing (Johnson and Slatkin 2007). Given the moderately low sequencing depth of our dataset (5X) downstream analyses were based on genotype probabilities rather than directly-called genotypes. We calculated genotype probabilities in ANGSD (Korneliussen et al. 2014) using the samtools model and keeping SNPs with a minimum allele frequency (MAF) of 0.05. All analyses described below were run considering only biallelic SNPs with a mapping quality score 20. We filtered the bam files by excluding SNPs that showed significant deviation from Hardy-Weinberg equilibrium, recovering a total of 14,330,920 SNPs.

Statistical tests of repeated increase in body and bill size.

Body size is most often represented by multivariate measures, such as Principal Component Analysis (PCA), that account for the covariation among characters and summarise most of the variation in the first two axes, the first representing body size and the second body shape (Rising and Somers 1989). We conducted a PCA that included all morphological measurements that represent body size: wing length, tail length, tarsus length, and head length; and separately bill size: bill length to posterior nostril opening, bill width, and depth at anterior nostril opening. We tested whether colonisations resulted in changes in body size and bill size relative to the mainland using *brms* (Bürkner 2017). Model 1 included PC1 of body size as the response variable and Model 2 PC1 included bill size. Colonisation time was set as a categorical population-level effect with a weak prior in the form of a normal distribution centered in 0 with a standard deviation of 10. This prior gives equal support to the idea that body size can remain unchanged, increase or decrease after colonisation. We controlled for population by setting it as a group-level effect. We used a Gaussian family and the default link function. We ran four chains each with 5000 iterations which included a warm-up of 1000.

The genetic basis of parallel evolution.

To test whether *de novo* mutations or ancestral standing genetic variation led to similar phenotypes, we did a Genome-wide Association Study (GWAS) under the idea that if the associated variants are common to all populations then they are likely due to standing genetic variation but if they are present in a specific population or specific lineage, they may represent *de novo* mutations that have accumulated since the lineage split. This framework assumes no potential gene flow among colonisations, which is supported by genome-wide data (Estandía et al. 2023).

Genome-wide Association Study.

We conducted GWAS on the first two PCs of body size (body size PC1 and body size PC2) and bill size (bill size PC1 and bill size PC2). We ran a generalised linear model implemented in PLINK (Purcell et al. 2007) and controlled for population structure by generating a covariance matrix with PCAngsd (Meisner and Albrechtsen 2018) and decomposing it by using a scaling transformation. We used the first 8 eigenvectors that accounted for most of the variation and adding more components accounted for less than 1% of the variation and did not split populations further. The two main clusters (ANZO and SM) have diverged for a long time, and many of the associated variants could have arisen due to population substructure not represented by the top eigenvectors and phenotypes that did not fully overlap between the ANZO and SM groups (but see Fig. S6). To check that our results held on each subset, we repeated the GWAS on each subgroup while accounting for their own population substructure (Fig S2 B-C). For the top 50 associated variants, we excluded samples with genotype likelihoods lower or equal to 0.33, which indicates that all genotypes have the same probability of being the true genotype. We classified putative *de novo* mutations as those only present in a single colonisation. For example, if an associated variant is only present on Heron Island, it would be classified as a new mutation. However, if the associated variant is present in various populations from separate colonisations, we classified them as arising due to standing genetic variation. This is because it is unlikely that the same new mutation would arise more than once in separate lineages stemming from the same source. Given that our system shows a complicated population structure, we decided to be conservative and further filtered associated variants by keeping those that showed at least one population where there was evidence of intra-population variability associated with body size. We did this by conducting an ANOVA between genotype and body size in each population separately. We also explored the frequency of genotypes and plotted it on maps to explore geographic structuring. We used our custom annotation to find the regions where the outliers were located by selecting all annotated coding regions within 5000 bp upstream and downstream from the associated locus. To confirm that our annotation was robust, we pulled the same number of bp from the reference genome and performed a BLAST search against all organisms. To estimate the proportion of phenotypic variance explained by all genome-wide SNPs while correcting for LD bias we used GCTA (Yang et al. 2011a). We created four quartiles of LD scores and used each individual SNP's LD score for binning as recommended by Evans et al. 2018. We also controlled for population structure using the principal components arising from the PCAngsd analysis. We computed the h^2-SNP 95% CI as +- 1.96 * SE.

Detection of structural variants.

To detect structural variants, we conducted a localised PCA of genetic variation based on the methodology outlined in Li and Ralph (2019). We used PCAngsd and the R package *lostruct* to identify non-recombining haploblocks across the genome. Structural variants like inversions are characterised by high levels of linkage disequilibrium because recombination is often to be negligible in heterokaryotes (Kirkpatrick 2010). For this reason, we used the unpruned SNP dataset containing 14 million SNPs. First, we split each chromosome into non-overlapping windows containing 100 SNPs

each, ran PCAngsd and obtained a covariance matrix for each window. We then descomposed each matrix by using scaling transformation and generated a pairwise distance matrix based on PC1 and PC2 for each window. Finally, we mapped similarity using MDS of up to 20 axes. We defined outlier regions as windows with a coefficient of correlation above four standard deviations from the mean. Adjacent outliers with less than 10 windows between them were pooled, and clusters with less than 5 windows were not considered following Mérot et al. (2021). A typical signature of a polymorphic inversion is three groups of individuals appearing on a PCA: the two homokaryotypes for the alternative arrangements and, as an intermediate group, the heterokaryotypes (Mérot et al. 2020). All clusters of outlier windows were examined by MDS as single blocks. We then used K-means clustering to identify putative groups of haplotypes. We explored LD patterns in heterokaryotype versus homokaryotype individuals for the chromosomes where we found outlier regions. We calculated pairwise r^2 values with ngsLD based on genotype likelihoods obtained by ANGSD and created a heatmap where we could visually explore LD levels. We built a model in *brms* to test whether karyotype identity was associated with body size after controlling for population structure. Karyotype was set as a categorical population-level effect and body size as the explanatory variable. Population was set as a group-level effect. We used a Gaussian family and the default link function. We ran four chains each with 5000 iterations which included a warm-up of 1000. We refer to karyotypes AA, AB, and BB where A represents the most frequent variant, whether inverted or not since we do not have information about this.

ACKNOWLEDGEMENTS

The samples used in this study were collected over two decades and we thank the many people who helped in numerous ways to facilitate the work. We are grateful to the chiefs and landholders of Vanuatu and New Caledonia for granting access to field sites, and field and logistic assistance from many people including the following: N. Clark, D. Treby, J. LeBreton, F. Cugny, W. Waheoneme, O. Hebert and A. Rouquie (New Caledonia); F. Robertson and C. Sendell-Price (Heron Island); S. Geiger, O. Boissier, R. Hills, S. Totterman, O. Drew, K. Ser, W. Ser, J. Saksak (Vanuatu), I.P.F. Owens, N. Clark (Australian mainland); J. Kikkawa (Norfolk Island); I.P.F. Owens (Lord Howe Island); P. Park, A. Fletcher, P. Gray (Tasmania); P. Schweigman, D. Onely (New Zealand); M. Bell (Chatham Island); N. Davies, O. Grant, C. Sendell-Price (Moorea, French Polynesia). For additional samples we thank A. Phillimore, R. Black, M. Massaro, D. Potvin and N. Clark. The work was conducted under permits from the Direction de l'Environnement Province Sud and Direction Du Développement Economique (New Caledonia and Loyalty Islands with thanks to G. Kakue); Vanuatu Environment Unit letters of permission and permits provided by E. Bani and we further thank D. Kalfatak and T. Tiwok for their assistance (Vanuatu); Délégations régionales à la recherche et à la technologie (French Polynesia); Lord Howe Island Board; Environment Australia (Norfolk Island); Queensland Department of Environment and Resource Management; Parks and Wildlife Service (Tasmania). New Zealand Department of Conservation Te Papa Atawhai; Australian Bird and Bat Banding Scheme project and individual permits to SMC. Ethics clearances were provided by University of Queensland ethics committee (ZOO/165/94/ARC, ZOO/520/96/ARC/PHD, ZOO/520/97/ARC/PHD) and Griffith University ethics committee (ENV/01/12/AEC, ENV/07/16/AEC, ENV/06/20/AEC, ENV/24/13/AEC) to SMC. We thank F. Robertson for conducting labwork. We thank the funders of this work: a Marsden Fund grant (UOO1410) from the New Zealand Royal Society Te Apārangi awarded to BCR and SMC to support fieldwork on mainland Australia and Heron Island portions of the molecular work; National Geographic Society Committee for Research and Exploration Grant (9383-13) to SMC to support fieldwork in New Caledonia; Natural Environment Research Council (NERC) postdoctoral fellowship to SMC to support fieldwork in Vanuatu and New Caledonia; Percy Sladen Memorial fund to SMC to support fieldwork in French Polynesia; the Department of Zoology, Oxford start-up to SMC to support sequencing. ASP was supported by a NERC studentship (NE/L002612/1); AE was supported by a NERC studentship (NE/S007474/1) and a St John's College Graduate Scholarship.

Author contributions

SMC and AE formulated the idea and SMC supervised the project. SMC, BCR, and ATSP conducted fieldwork. AE, ATSP and BCR conducted laboratory work. AE conducted the genetic and statistical analyses with input from ATSP. AE drafted the manuscript with input from SMC and all authors provided comments.

Data and code availability

Code to reproduce all analyses is available at https://github.com/andreaestandia/1.0_island_rule_GWAS

References

- Akopyan, M., Tigano, A., Jacobs, A., Wilder, A. P., Baumann, H., & Therkildsen, N. O. (2022). Comparative linkage mapping uncovers recombination suppression across massive chromosomal inversions associated with local adaptation in Atlantic silversides. Molecular Ecology, 31(12), 3323–3341. https://doi.org/10.1111/mec.16472
- Ament-Velásquez, S. L., Gilchrist, C., Rêgo, A., Bendixsen, D. P., Brice, C., Grosse-Sommer, J. M., Rafati, N., & Stelkens, R. (2022). The Dynamics of Adaptation to Stress from Standing Genetic Variation and de novo Mutations. Molecular Biology and Evolution, 39(11). https://doi.org/10.1093/molbev/msac242
- Archambeault, S. L., Bärtschi, L. R., Merminod, A. D., & Peichel, C. L. (2020). Adaptation via pleiotropy and linkage: Association mapping reveals a complex genetic architecture within the stickleback Eda locus. Evolution Letters, 4(4), 282–301. https://doi.org/10.1002/evl3.175
- Arendt, J., & Reznick, D. (2008). Convergence and parallelism reconsidered: What have we learned about the genetics of adaptation? Trends in Ecology & Evolution, 23(1), 26–32. https://doi.org/10.1016/j.tree.2007.09.011
- Baeckens, S., & Van Damme, R. (2020). The island syndrome. Current Biology, 30(8), R338–R339. https://doi.org/10.1016/j.cub.2020.03.029
- Barrett, R. D. H., & Schluter, D. (2008). Adaptation from standing genetic variation. Trends in Ecology & Evolution, 23(1), 38–44. https://doi.org/10.1016/j.tree.2007.09.008
- Beauchamp, G. (2021). Do avian species survive better on islands? Biology Letters, 17(1), 20200643. https://doi.org/10.1098/rsbl.2020.0643
- Bengston, S. E., Dahan, R. A., Donaldson, Z., Phelps, S. M., van Oers, K., Sih, A., & Bell, A. M. (2018). Genomic tools for behavioural ecologists to understand repeatable individual differences

in behaviour. Nature Ecology & Evolution, 2(6). https://doi.org/10.1038/s41559-017-0411-4

- Benítez-López, A., Santini, L., Gallego-Zamorano, J., Milá, B., Walkden, P., Huijbregts, M. A. J., & Tobias, J. A. (2021). The island rule explains consistent patterns of body size evolution in terrestrial vertebrates. Nature Ecology & Evolution, 5(6). https://doi.org/10.1038/s41559-021-01426-y

- Bhullar, B.-A. S., Morris, Z. S., Sefton, E. M., Tok, A., Tokita, M., Namkoong, B., Camacho, J., Burnham, D. A., & Abzhanov, A. (2015). A molecular mechanism for the origin of a key evolutionary innovation, the bird beak and palate, revealed by an integrative approach to major transitions in vertebrate history. Evolution, 69(7), 1665–1677. https://doi.org/10.1111/evo.12684

- Black, R. A. (2010). Phylogenetic and phenotypic divergence of an insular radiation of birds [PhD]. Imperial College London.

- Blount, Z. D., Lenski, R. E., & Losos, J. B. (2018). Contingency and determinism in evolution: Replaying life's tape. Science, 362(6415). https://doi.org/10.1126/science.aam5979

- Bosse, M., Spurgin, L. G., Laine, V. N., Cole, E. F., Firth, J. A., Gienapp, P., Gosler, A. G., McMahon, K., Poissant, J., Verhagen, I., Groenen, M. A. M., Van Oers, K., Sheldon, B. C., Visser, M. E., & Slate, J. (2017). Recent natural selection causes adaptive evolution of an avian polygenic trait. Science, 358(6361), 365–368. https://doi.org/10.1126/science.aal3298

- Bürkner, P.-C. (2017). *brms*: An R Package for Bayesian Multilevel Models Using Stan. Journal of Statistical Software, 80, 1–28. https://doi.org/10.18637/jss.v080.i01

- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: Architecture and applications. BMC Bioinformatics, 10(1), 421. https://doi.org/10.1186/1471-2105-10-421

- Cerca, J. (2023). Understanding natural selection and similarity: Convergent, parallel and repeated evolution. Molecular Ecology, 32(20), 5451–5462. https://doi.org/10.1111/mec.17132

- Cerca, J., Cotoras, D. D., Santander, C. G., Bieker, V. C., Hutchins, L., Morin-Lagos, J., Prada, C. F., Kennedy, S., Krehenwinkel, H., Rominger, A. J., Meier, J., Dimitrov, D., Struck, T. H., & Gillespie, R. G. (2023). Multiple paths toward repeated phenotypic evolution in the spiny-leg adaptive radiation (Tetragnatha; Hawai'i). Molecular Ecology, 32(18), 4971–4985. https://doi.org/10.1111/mec.17082

- Charlesworth, B., & Charlesworth, D. (1973). Selection of new inversions in multi-locus genetic systems. Genetical Research, 21(2), 167–183. https://doi.org/10.1017/S0016672300013343

- Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: An ultra-fast all-in-one FASTQ preprocessor. Bioinformatics, 34(17), i884–i890. https://doi.org/10.1093/bioinformatics/bty560

- Clegg, S. M., Degnan, S. M., Kikkawa, J., Moritz, C., Estoup, A., & Owens, I. P. F. (2002). Genetic consequences of sequential founder events by an island-colonizing bird. Proceedings of the National Academy of Sciences, 99(12), 8127–8132.

- Clegg, S. M., Degnan, S. M., Moritz, C., Estoup, A., Kikkawa, J., & Owens, I. P. F. (2002). Microevolution in Island Forms: The Roles of Drift and Directional Selection in Morphological Divergence of a Passerine Bird. Evolution, 56(10), 2090–2099. https://doi.org/10.1111/j.0014-3820.2002.tb00134.x

- Clegg, S. M., Frentiu, F. D., Kikkawa, J., Tavecchia, G., & Owens, I. P. F. (2008). 4000 Years of Phenotypic Change in an Island Bird: Heterogeneity of Selection Over Three Microevolutionary Timescales. Evolution, 62(9), 2393–2410. https://doi.org/10.1111/j.1558-5646.2008.00437.x

- Clegg, S. M., & Owens, P. F. (2002). The 'island rule' in birds: Medium body size and its ecological explanation. Proceedings of the Royal Society of London. Series B: Biological Sciences, 269(1498), 1359–1365. https://doi.org/10.1098/rspb.2002.2024

- Clegg, S. M., & Phillimore, A. B. (2010). The influence of gene flow and drift on genetic and phenotypic divergence in two species of Zosterops in Vanuatu. Philosophical Transactions of the Royal Society B: Biological Sciences, 365(1543), 1077–1092. https://doi.org/10.1098/rstb.2009.0281

- Clements, J. F., Schulenberg, T. S., Iliff, M. J., Fredericks, T. A., Gerbracht, J. A., Lepage, D., Billerman, S. M., Sullivan, B. L., & Wood, C. L. (2021). The eBird/Clements checklist of Birds of the World: V2021. https://www.birds.cornell.edu/clementschecklist/

- Colosimo, P. F., Hosemann, K. E., Balabhadra, S., Villarreal, G., Dickson, M., Grimwood, J., Schmutz, J., Myers, R. M., Schluter, D., & Kingsley, D. M. (2005). Widespread Parallel Evolution in Sticklebacks by Repeated Fixation of Ectodysplasin Alleles. Science, 307(5717), 1928–1933. https://doi.org/10.1126/science.1107239

- Cornetti, L., Valente, L. M., Dunning, L. T., Quan, X., Black, R. A., Hébert, O., & Savolainen,

V. (2015). The Genome of the "Great Speciator" Provides Insights into Bird Diversification. Genome Biology and Evolution, 7(9), 2680–2691. https://doi.org/10.1093/gbe/evv168

- Duntsch, L., Tomotani, B. M., de Villemereuil, P., Brekke, P., Lee, K. D., Ewen, J. G., & Santure, A. W. (2020). Polygenic basis for adaptive morphological variation in a threatened Aotearoa | New Zealand bird, the hihi (Notiomystis cincta). Proceedings of the Royal Society B: Biological Sciences, 287(1933), 20200948. https://doi.org/10.1098/rspb.2020.0948

- Estandía, A., Chesser, R. T., James, H. F., Levy, M. A., Obiol, J. F., Bretagnolle, V., González-Solís, J., & Welch, A. J. (2021). Substitution Rate Variation in a Robust Procellariiform Seabird Phylogeny is not Solely Explained by Body Mass, Flight Efficiency, Population Size or Life History Traits(p. 2021.07.27.453752). bioRxiv. https://doi.org/10.1101/2021.07.27.453752

- Estandía, A., Sendell-Price, A. T., Oatley, G., Robertson, F., Potvin, D., Massaro, M., Robertson, B. C., & Clegg, S. M. (2023). Candidate gene polymorphisms are linked to dispersive and migratory behaviour: Searching for a mechanism behind the "paradox of the great speciators." Journal of Evolutionary Biology, 36, 1503–1516. https://doi.org/10.1111/jeb.14222

- Evans, L. M., Tahmasbi, R., Vrieze, S. I., Abecasis, G. R., Das, S., Gazal, S., Bjelland, D. W., de Candia, T. R., Goddard, M. E., Neale, B. M., Yang, J., Visscher, P. M., & Keller, M. C. (2018). Comparison of methods that use whole genome data to estimate the heritability and genetic architecture of complex traits. Nature Genetics, 50(5). https://doi.org/10.1038/s41588-018-0108-x

- Faria, R., Johannesson, K., Butlin, R. K., & Westram, A. M. (2019). Evolving Inversions. Trends in Ecology & Evolution, 34(3), 239–248. https://doi.org/10.1016/j.tree.2018.12.005

- Foster, J. B. (1964). Evolution of Mammals on Islands. Nature, 202(4929), 234–235. https://doi.org/10.1038/202234a0

- Fulgione, A., Neto, C., Elfarargi, A. F., Tergemina, E., Ansari, S., Göktay, M., Dinis, H., Döring, N., Flood, P. J., Rodriguez-Pacheco, S., Walden, N., Koch, M. A., Roux, F., Hermisson, J., & Hancock, A. M. (2022). Parallel reduction in flowering time from de novo mutations enable evolutionary rescue in colonizing lineages. Nature Communications, 13(1). https://doi.org/10.1038/s41467-022-28800-z

- Gao, G., Gao, D., Zhao, X., Xu, S., Zhang, K., Wu, R., Yin, C., Li, J., Xie, Y., Hu, S., & Wang, Q. (2021). Genome-Wide Association Study-Based Identification of SNPs and Haplotypes Associated With Goose Reproductive Performance and Egg Quality. Frontiers in Genetics, 12. https://www.frontiersin.org/articles/10.3389/fgene.2021.602583

- Gill, F. B. (1970). Hybridization in Norfolk Island White-Eyes (Zosterops). The Condor, 72(4), 481–482. https://doi.org/10.2307/1366406

- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., Di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., … Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature Biotechnology, 29(7), 644–652. https://doi.org/10.1038/nbt.1883

- Grabherr, M. G., Russell, P., Meyer, M., Mauceli, E., Alföldi, J., Di Palma, F., & Lindblad-Toh, K. (2010). Genome-wide synteny through highly sensitive sequence alignment: Satsuma. Bioinformatics, 26(9), 1145–1151. https://doi.org/10.1093/bioinformatics/btq102

- Griffiths, R., Double, M. C., Orr, K., & Dawson, R. J. G. (1998). A DNA test to sex most birds. Molecular Ecology, 7(8), 1071–1075. https://doi.org/10.1046/j.1365-294x.1998.00389.x

- Harringmeyer, O. S., & Hoekstra, H. E. (2022). Chromosomal inversion polymorphisms shape the genomic landscape of deer mice. Nature Ecology & Evolution, 6(12). https://doi.org/10.1038/s41559-022-01890-0

- Hedrick, P. W. (1983). Genetics of populations. Science Books International.

- Johnson, P. L. F., & Slatkin, M. (2007). Accounting for Bias from Sequencing Error in Population Genetic Estimates. Molecular Biology and Evolution, 25(1), 199–206. https://doi.org/10.1093/molbev/msm239

- Kent, W. J. (2002). BLAT—The BLAST-Like Alignment Tool. Genome Research, 12, 656–664.

- Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: A fast spliced aligner with low memory requirements. Nature Methods, 12(4). https://doi.org/10.1038/nmeth.3317

- Kirkpatrick, M. (2010). How and Why Chromosome Inversions Evolve. PLOS Biology, 8(9), e1000501. https://doi.org/10.1371/journal.pbio.1000501

- Kitano, J., Ross, J. A., Mori, S., Kume, M., Jones, F. C., Chan, Y. F., Absher, D. M., Grimwood, J., Schmutz, J., Myers, R. M., Kingsley, D. M., & Peichel, C. L. (2009). A role for a neo-sex chromosome in stickleback speciation. Nature, 461(7267), 1079–1083. https://doi.org/10.1038/nature08441

- Koch, E. L., Morales, H. E., Larsson, J., Westram, A. M., Faria, R., Lemmon, A. R., Lemmon, E. M., Johannesson, K., & Butlin, R. K. (2021). Genetic variation for adaptive traits is associated with polymorphic inversions in Littorina saxatilis. Evolution Letters, 5(3), 196–213. https://doi.org/10.1002/evl3.227
- Konečná, V., Bray, S., Vlček, J., Bohutínská, M., Požárová, D., Choudhury, R. R., Bollmann-Giolai, A., Flis, P., Salt, D. E., Parisod, C., Yant, L., & Kolář, F. (2021). Parallel adaptation in autopolyploid Arabidopsis arenosa is dominated by repeated recruitment of shared alleles. Nature Communications, 12(1). https://doi.org/10.1038/s41467-021-25256-5
- Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. BMC Bioinformatics, 15(1), 356. https://doi.org/10.1186/s12859-014-0356-4
- Krist, M. (2011). Egg size and offspring quality: A meta-analysis in birds. Biological Reviews, 86(3), 692–716. https://doi.org/10.1111/j.1469-185X.2010.00166.x
- Lai, Y.-T., Yeung, C. K. L., Omland, K. E., Pang, E.-L., Hao, Y., Liao, B.-Y., Cao, H.-F., Zhang, B.-W., Yeh, C.-F., Hung, C.-M., Hung, H.-Y., Yang, M.-Y., Liang, W., Hsu, Y.-C., Yao, C.-T., Dong, L., Lin, K., & Li, S.-H. (2019). Standing genetic variation as the predominant source for adaptation of a songbird. Proceedings of the National Academy of Sciences, 116(6), 2152–2157. https://doi.org/10.1073/pnas.1813597116
- Lamichhaney, S., Berglund, J., Almén, M. S., Maqbool, K., Grabherr, M., Martinez-Barrio, A., Promerová, M., Rubin, C.-J., Wang, C., Zamani, N., Grant, B. R., Grant, P. R., Webster, M. T., & Andersson, L. (2015). Evolution of Darwin's finches and their beaks revealed by genome sequencing. Nature, 518(7539), 371–375. https://doi.org/10.1038/nature14181
- Lamichhaney, S., Han, F., Berglund, J., Wang, C., Almén, M. S., Webster, M. T., Grant, B. R., Grant, P. R., & Andersson, L. (2016). A beak size locus in Darwin's finches facilitated character displacement during a drought. Science, 352(6284), 470–474. https://doi.org/10.1126/science.aad8786
- Leroy, T., Anselmetti, Y., Tilak, M.-K., Bérard, S., Csukonyi, L., Gabrielli, M., Scornavacca, C., Milá, B., Thébaud, C., & Nabholz, B. (2021). A bird's white-eye view on avian sex chromosome evolution. Peer Community Journal, 1, e63. https://doi.org/10.24072/pcjournal.70
- Li, H., & Durbin, R. (2010). Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics, 26(5), 589–595. https://doi.org/10.1093/bioinformatics/btp698
- Li, H., & Ralph, P. L. (2019). Local PCA Shows How the Effect of Population Structure Differs Along the Genome. Genetics, 211, 289–304. https://doi.org/10.25386/GENETICS.7324526
- Lomolino, M. V. (2005). Body size evolution in insular vertebrates: Generality of the island rule. Journal of Biogeography, 32(10), 1683–1699. https://doi.org/10.1111/j.1365-2699.2005.01314.x
- Losos, J. B. (2011). Convergence, Adaptation, and Constraint. Evolution, 65(7), 1827–1840. https://doi.org/10.1111/j.1558-5646.2011.01289.x
- MacPherson, A., & Nuismer, S. L. (2017). The probability of parallel genetic evolution from standing genetic variation. Journal of Evolutionary Biology, 30(2), 326–337. https://doi.org/10.1111/jeb.13006
- Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A., & Zdobnov, E. M. (2021). BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. Molecular Biology and Evolution, 38(10), 4647–4654. https://doi.org/10.1093/molbev/msab199
- Marques, D. A., Meier, J. I., & Seehausen, O. (2019). A Combinatorial View on Speciation and Adaptive Radiation. Trends in Ecology & Evolution, 34(6), 531–544. https://doi.org/10.1016/j.tree.2019.02.008
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.Journal, 17(1). https://doi.org/10.14806/ej.17.1.200
- McGhee, G. R. (2011). Convergent Evolution: Limited Forms Most Beautiful. MIT Press.
- Mees, G. F. (1969). A systematic review of the Indo-Australian Zosteropidae, three parts.
- Meiri, S., Dayan, T., & Simberloff, D. (2004). Body Size of Insular Carnivores: Little Support for the Island Rule. The American Naturalist, 163(3), 469–479. https://doi.org/10.1086/382229
- Meisner, J., & Albrechtsen, A. (2018). Inferring Population Structure and Admixture Proportions in Low-Depth NGS Data. Genetics, 210(2), 719–731. https://doi.org/10.1534/genetics.118.301336
- Mérot, C., Berdan, E. L., Cayuela, H., Djambazian, H., Ferchaud, A.-L., Laporte, M., Normandeau, E., Ragoussis, J., Wellenreuther, M., & Bernatchez, L. (2021). Locally Adaptive Inversions Modulate Genetic Variation at Different Geographic Scales in a Seaweed Fly. Molecular Biology and Evolution, 38(9), 3953–3971. https://doi.org/10.1093/molbev/msab143

- Mérot, C., Oomen, R. A., Tigano, A., & Wellenreuther, M. (2020). A Roadmap for Understanding the Evolutionary Significance of Structural Genomic Variation. Trends in Ecology & Evolution, 35(7), 561–572. https://doi.org/10.1016/j.tree.2020.03.002

- Montejo-Kovacevich, G., Meier, J. I., Bacquet, C. N., Warren, I. A., Chan, Y. F., Kucka, M., Salazar, C., Rueda-M, N., Montgomery, S. H., McMillan, W. O., Kozak, K. M., Nadeau, N. J., Martin, S. H., & Jiggins, C. D. (2022). Repeated genetic adaptation to altitude in two tropical butterflies. Nature Communications, 13(1). https://doi.org/10.1038/s41467-022-32316-x

- Morales, H. E., Faria, R., Johannesson, K., Larsson, T., Panova, M., Westram, A. M., & Butlin, R. K. (2019). Genomic architecture of parallel ecological divergence: Beyond a single environmental contrast. Science Advances, 5(12), eaav9963. https://doi.org/10.1126/sciadv.aav9963

- Ohta, T. (1971). Associative overdominance caused by linked detrimental mutations. Genetics Research, 18(3), 277–286. https://doi.org/10.1017/S0016672300012684

- Pala, I., Naurin, S., Stervander, M., Hasselquist, D., Bensch, S., & Hansson, B. (2012). Evidence of a neo-sex chromosome in birds. Heredity, 108(3). https://doi.org/10.1038/hdy.2011.70

- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. American Journal of Human Genetics, 81(3), 559–575.

- Rhie, A., McCarthy, S. A., Fedrigo, O., Damas, J., Formenti, G., Koren, S., Uliano-Silva, M., Chow, W., Fungtammasan, A., Kim, J., Lee, C., Ko, B. J., Chaisson, M., Gedman, G. L., Cantin, L. J., Thibaud-Nissen, F., Haggerty, L., Bista, I., Smith, M., . . . Jarvis, E. D. (2021). Towards complete and error-free genome assemblies of all vertebrate species. Nature, 592(7856), 737–746. https://doi.org/10.1038/s41586-021-03451-0

- Rising, J. D., & Somers, K. M. (1989). The Measurement of Overall Body Size in Birds. The Auk, 106, 666–674.

- Rumsby, G., & Cregeen, D. P. (1999). Identification and expression of a cDNA for human hydroxypyruvate/glyoxylate reductase. Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression, 1446(3), 383–388. https://doi.org/10.1016/S0167-4781(99)00105-0

- Sendell-Price, A. T., Ruegg, K. C., & Clegg, S. M. (2020). Rapid morphological divergence following a human-mediated introduction: The role of drift and directional selection. Heredity, 124(4). https://doi.org/10.1038/s41437-020-0298-8

- Sendell-Price, A. T., Ruegg, K. C., Robertson, B. C., & Clegg, S. M. (2021). An island-hopping bird reveals how founder events shape genome-wide divergence. Molecular Ecology, 30(11), 2495–2510. https://doi.org/10.1111/mec.15898

- Seutin, G., White, B. N., & Boag, P. T. (1991). Preservation of avian blood and tissue samples for DNA analyses. Canadian Journal of Zoology, 68. https://cdnsciencepub.com/doi/abs/10.1139/z91-013

- Sigeman, H., Ponnikas, S., & Hansson, B. (2020). Whole-genome analysis across 10 songbird families within Sylvioidea reveals a novel autosome–sex chromosome fusion. Biology Letters, 16(4), 20200082. https://doi.org/10.1098/rsbl.2020.0082

- Smit, A., Hubley, R., & Green, P. (2013). RepeatMasker Open(4.0) [Computer software]. http://www.repeatmasker.org

- Stanke, M., Keller, O., Gunduz, I., Hayes, A., Waack, S., & Morgenstern, B. (2006). AUGUSTUS: Ab initio prediction of alternative transcripts. Nucleic Acids Research, 34(Web Server), W435–W439. https://doi.org/10.1093/nar/gkl200

- Stayton, C. T. (2008). Is convergence surprising? An examination of the frequency of convergence in simulated datasets. Journal of Theoretical Biology, 252(1), 1–14. https://doi.org/10.1016/j.jtbi.2008.01.008

- Taylor, R. S., Manseau, M., Horn, R. L., Keobouasone, S., Golding, G. B., & Wilson, P. J. (2020). The role of introgression and ecotypic parallelism in delineating intraspecific conservation units. Molecular Ecology, 29(15), 2793–2809. https://doi.org/10.1111/mec.15522

- Therkildsen, N. O., Wilder, A. P., Conover, D. O., Munch, S. B., Baumann, H., & Palumbi, S. R. (2019). Contrasting genomic shifts underlie parallel phenotypic evolution in response to fishing. Science, 365(6452), 487–490. https://doi.org/10.1126/science.aaw7271

- vonHoldt, B. M., Kartzinel, R. Y., Huber, C. D., Le Underwood, V., Zhen, Y., Ruegg, K., Lohmueller, K. E., & Smith, T. B. (2018). Growth factor gene IGF1 is associated with bill size in the black-bellied seedcracker Pyrenestes ostrinus. Nature Communications, 9, 4855. https://doi.org/10.1038/s41467-018-07374-9

- Wang, Y., Wang, Y., Cheng, X., Ding, Y., Wang, C., Merilä, J., & Guo, B. (2023). Prevalent Introgression

Underlies Convergent Evolution in the Diversification of PungitiusSticklebacks. Molecular Biology and Evolution, 40(2). https://doi.org/10.1093/molbev/msad026

- Westram, A. M., Faria, R., Johannesson, K., Butlin, R., & Barton, N. (2022). Inversions and parallel evolution. Philosophical Transactions of the Royal Society B: Biological Sciences, 377(1856), 20210203. https://doi.org/10.1098/rstb.2021.0203

- Whitlock, M. C. (2000). Fixation of new alleles and the extinction of small populations: Drift load, beneficial alleles, and sexual selection. Evolution, 54(6), 1855–1861. https://doi.org/10.1111/j.0014-3820.2000.tb01232.x

- Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: A Tool for Genome-wide Complex Trait Analysis. American Journal of Human Genetics, 88(1), 76–82. https://doi.org/10.1016/j.ajhg.2010.11.011

- Yang, J., Weedon, M. N., Purcell, S., Lettre, G., Estrada, K., Willer, C. J., Smith, A. V., Ingelsson, E., O'Connell, J. R., Mangino, M., Mägi, R., Madden, P. A., Heath, A. C., Nyholt, D. R., Martin, N. G., Montgomery, G. W., Frayling, T. M., Hirschhorn, J. N., McCarthy, M. I., … Visscher, P. M. (2011). Genomic inflation factors under polygenic inheritance. European Journal of Human Genetics, 19(7). https://doi.org/10.1038/ejhg.2011.39

- Zhang, G., Li, C., Li, Q., Li, B., Larkin, D. M., Lee, C., Storz, J. F., Antunes, A., Greenwold, M. J., Meredith, R. W., Ödeen, A., Cui, J., Zhou, Q., Xu, L., Pan, H., Wang, Z., Jin, L., Zhang, P., Hu, H., Froman, D. P. (2014). Comparative genomics reveals insights into avian genome evolution and adaptation. Science, 346(6215), 1311–1320. https://doi.org/10.1126/science.1251385