Received: 28 August 2019 | Revised: 9 November 2019 | Accepted: 2 December 2019

DOI: 10.1111/1755-0998.13124

RESOURCE ARTICLE

MOLECULAR ECOLOGY RESOURCES WILEY

Genome of an iconic Australian bird: High-quality assembly and linkage map of the superb fairy-wren (*Malurus cyaneus*)

Joshua V. Peñalba[1] | **Yuan Deng**[2] | **Qi Fang**[2] | **Leo Joseph**[3] | **Craig Moritz**[4,5] | **Andrew Cockburn**[5]

[1]Division of Evolutionary Biology, Ludwig Maximilian University of Munich, Munich, Germany

[2]BGI-Shenzhen, Shenzhen, China

[3]Australian National Wildlife Collection, CSIRO National Research Collections, Australia, Canberra, ACT, Australia

[4]Centre for Biodiversity Analysis, Acton, ACT, Australia

[5]Division of Ecology and Evolution, Australian National University, Acton, ACT, Australia

Correspondence
Joshua V. Peñalba, Division of Evolutionary Biology, Ludwig Maximilian University of Munich, Munich, Germany.
Email: josh.penalba@gmail.com

Funding information
Australian Research Council, Grant/Award Number: DP150100298; Australian National University

Abstract

The superb fairy-wren, *Malurus cyaneus*, is one of the most iconic Australian passerine species. This species belongs to an endemic Australasian clade, Meliphagides, which diversified early in the evolution of the oscine passerines. Today, the oscine passerines comprise almost half of all avian species diversity. Despite the rapid increase of available bird genome assemblies, this part of the avian tree has not yet been represented by a high-quality reference. To rectify that, we present the first high-quality genome assembly of a Meliphagides representative: the superb fairy-wren. We combined Illumina shotgun and mate-pair sequences, PacBio long-reads, and a genetic linkage map from an intensively sampled pedigree of a wild population to generate this genome assembly. Of the final assembled 1.07-Gb genome, 975 Mb (90.4%) was anchored onto 25 pseudochromosomes resulting in a final superscaffold N50 of 68.11 Mb. This high-quality bird genome assembly is one of only a handful which is also accompanied by a genetic map and recombination landscape. In comparison to other pedigree-based bird genetic maps, we find that the fairy-wren genetic map more closely resembles those of *Taeniopygia guttata* and *Parus major* maps, unlike the *Ficedula albicollis* map which more closely resembles that of *Gallus gallus*. Lastly, we also provide a predictive gene and repeat annotation of the genome assembly. This new high-quality, annotated genome assembly will be an invaluable resource not only regarding the superb fairy-wren species and relatives but also broadly across the avian tree by providing a novel reference point for comparative genomic analyses.

KEYWORDS

genetic map, *Malurus cyaneus*, Meliphagides, recombination landscape, reference genome

1 | INTRODUCTION

We present a high-quality annotated assembly of the genome of the superb fairy-wren, *Malurus cyaneus* (Maluridae). Although genome assembly resources for bird species are accumulating quickly, the phylogenetic coverage of high-quality genomes, particularly those anchored and oriented into chromosomes, is currently very biased (Figure 1). Among the oscine passerines, by far the largest

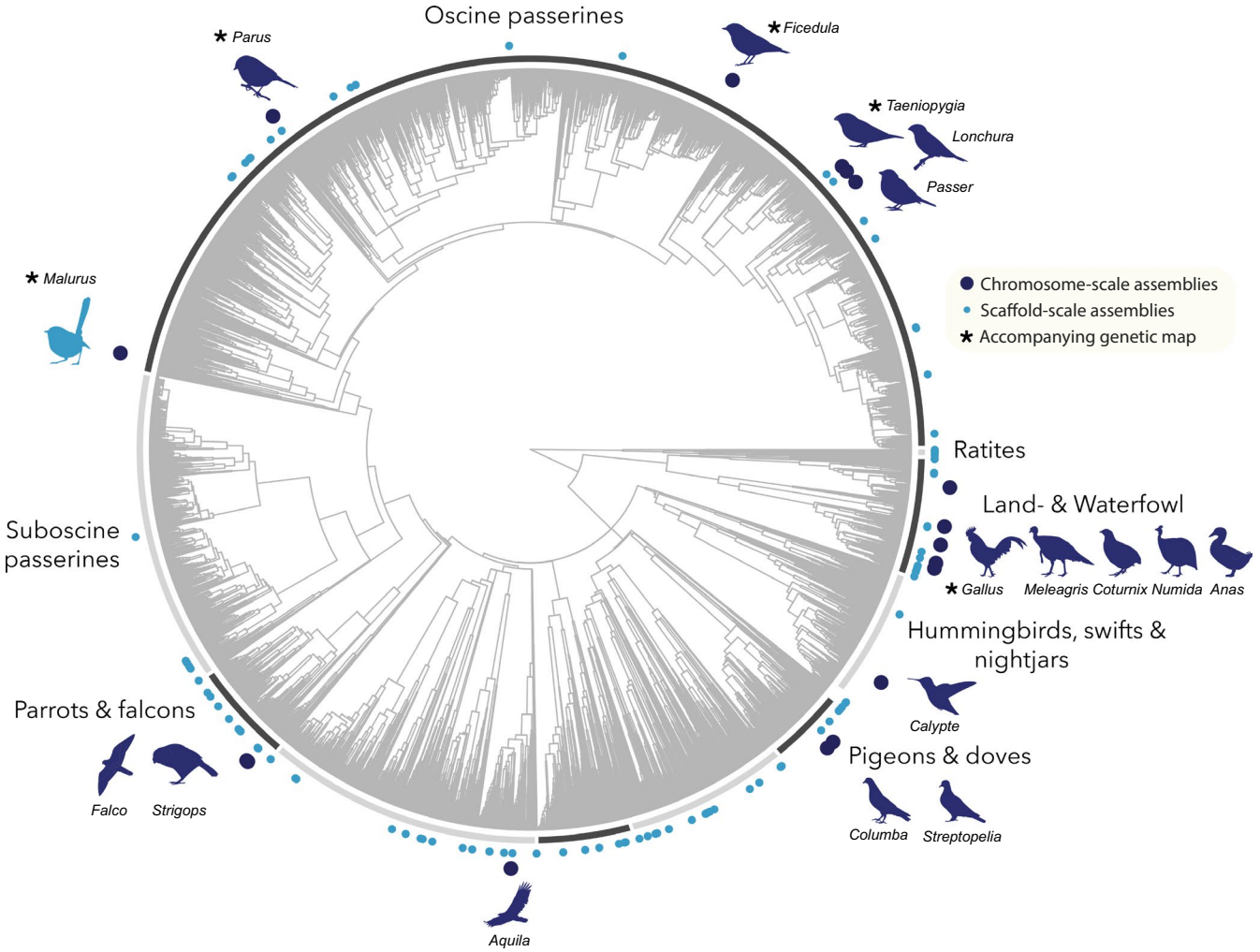

FIGURE 1 Phylogenetic coverage of available bird genome assemblies. The placement of bird genome assemblies available in NCBI GenBank (as of August 19, 2019) on the Jetz, Thomas, Joy, Hartmann, and Mooers (2012) avian phylogeny. The dark blue circles and associated silhouettes represent the chromosome-scale genomes and light blue circles represent the scaffold-scale genome assemblies [Colour figure can be viewed at wileyonlinelibrary.com]

radiation of birds, these resources are limited to five species from the Passerides clade. Yet the early evolution of the Oscines involved multiple branching in the Australo-Papuan region in the Oligocene before the emergence of the Corvides and Passerides, the two groups that ultimately gave the oscines their numerical and ecological dominance (Marki et al., 2017; Oliveros et al., 2019). The superb fairy-wren is an exemplar of the largest clade of that early radiation, the Meliphagides, which has almost 300 species. Furthermore, we complement this high-level assembly with a detailed genetic map from which we can infer the variation in recombination across the genome. Although genome assembly and genetic map resources have been developed separately for various species, there are currently only four other bird species for which we have both resources readily available.

Superb fairy-wrens are small, insectivorous birds found throughout south-eastern Australia. The males boast a familiar, striking bright blue breeding plumage contrasted against dark black patches. They have become a model species in behavioural and evolutionary ecology because they combine two conceptually interesting behaviours in an unpredicted way. First, they were the first bird anywhere in the world shown to exhibit cooperative breeding, and the single female on each territory can be assisted by as many as five males in provisioning offspring (Boland & Cockburn, 2002; Rowley & Russell, 1997). Experimental studies provided a textbook example of how limitations to mating opportunities preclude dispersal by young males, which instead defer dispersal and help the breeding pair raise offspring (Pruett-Jones & Lewis, 1990). Second, this cooperative breeding occurs in a novel context. Paternity in this species is dominated by extra-group fertilizations, which are sought by the female during predawn forays to the territories of a small number of attractive males (Cockburn, Brouwer, Double, Margraf, & Pol, 2013; Cockburn et al., 2009; Double & Cockburn, 2000; Mulder, Dunn, Cockburn, Cohen, & Howell, 1994). As a consequence, males often provision offspring to which they are completely unrelated (Dunn, Cockburn, & Mulder, 1995). This paradox has provoked more than three decades of demographic, behavioural and genetic

study of cooperative breeding in this species (Cockburn et al., 2016; Cockburn, Sims, et al., 2008). Further issues investigated include mate choice and sexual selection (Cockburn et al., 2016; Cockburn, Sims, et al., 2008; Peters, 2000), inbreeding (Hajduk et al., 2018), dispersal (Cockburn, Osmond, Mulder, Green, & Double, 2003), life history evolution and parental investment (Russell, Langmore, Cockburn, Astheimer, & Kilner, 2007), parent–offspring communication (Colombelli-Négrel et al., 2012), evolution of brood parasitism (Langmore, Hunt, & Kilner, 2003) and behavioural responses to predator risk (Magrath & Bennett, 2012; Magrath, Haff, McLachlan, & Igic, 2015; Potvin, Ratnayake, Radford, & Magrath, 2018), as well as population responses to climate change (Kruuk, Osmond, & Cockburn, 2015; van de Pol, Osmond, & Cockburn, 2012).

Parallel studies of mating systems and extragroup paternity in other *Malurus* species have revealed considerable diversity, supporting the idea that evolutionary pathways can be traced through phylogenetically-based comparative analysis (Brouwer et al., 2017; Buchanan & Cockburn, 2013). Other *Malurus* species have also been the focus of studies on breeding biology (Karubian, 2008; Leitão, Hall, Venables, & Mulder, 2019; Varian-Ramos & Webster, 2012), song and vocalizations (Dowling & Webster, 2016; Greig & Pruett-Jones, 2008; Yandell, Hochachka, Pruett-Jones, Webster, & Greig, 2018), plumage and ornamentation (Karubian, 2013; Lindsay, Webster, & Schwabl, 2011), ecology and conservation (Driskell et al., 2011; Murphy, Legge, Heathcote, & Mulder, 2010; Skroblin, Lanfear, Cockburn, & Legge, 2012), and phylogeography (Baldassarre, White, Karubian, & Webster, 2014; Kearns, Joseph, Edwards, & Double, 2009; McLean, Toon, Schmidt, Joseph, & Hughes, 2012). Providing a high-quality, annotated genome assembly sets up the foundation to understand the genetic mechanisms underlying some of the behaviours and natural history traits of the superb fairy-wren and allies.

Upgrading genome assemblies from scaffold-level to chromosome-level superscaffolds provides additional genomic context by orienting genes relative to each other and other genomic features such as centromeres, telomeres, various repeat elements and regulatory regions. Knowledge of this organization aids in understanding how genome architecture can influence variation in evolution within the genome but also provide insight into genome evolution between populations and species (Joseph et al., 2018; Sävilammi et al., 2019; Thind et al., 2018). To date, there are only 16 bird species with genome assemblies classified as 'chromosome' level in GenBank (as of August 19, 2019). Twelve of these assemblies were assembled de novo and four of these were assembled using other reference assemblies as a guide. These assemblies are distinguished from 'scaffold' level by higher contiguity where scaffolds are anchored onto chromosomes. This is performed using molecular cytogenetics (BAC and FISH; Damas et al., 2017; O'Connor et al., 2018), physical mapping (HiC, BioNano, Burton et al., 2013) or genetic mapping (Fierst, 2015). In the near future, we expect to see an exponential accumulation of bird genome assemblies to be released, particularly from the B10K consortium which is currently sequencing at least one bird species per family (~300 assemblies; Zhang et al., 2015). Despite this, only a handful of species are in the pipeline to be sequenced to chromosome level, all using HiC chromatin interaction maps (Stiller & Zhang, 2019). Genome assembly aided by a genetic map will remain valuable despite the onslaught of new assemblies being, as yet, limited to a few exemplar species.

The utility of a genetic map extends far beyond upgrading a genome assembly. The maps themselves serve as an invaluable resource by enabling us to associate a particular phenotypic trait to specific loci through quantitative trait loci (QTL) mapping (Su et al., 2017), associate different traits through tight linkage of the genes that code for them (Schwander, Libbrecht, & Keller, 2014), and provide an understanding of the variation in genetic diversity within the genome (Burri et al., 2015). It is also of interest to understand how the recombination landscape itself varies between individuals, populations and species, and how it may be influenced by selection, demographic history, genomic features and chromosomal rearrangements (Barton, 1995; Dapper & Payseur, 2017; Ortiz-Barrientos, Engelstädter, & Rieseberg, 2016; Stapley, Feulner, Johnston, Santure, & Smadja, 2017). From the existing chromosome-level genome assemblies, only four other species have a high-density, single nucleotide polymorphism (SNP)-based genetic map to complement the assembly (Groenen et al., 2009; Kawakami et al., 2014; Stapley, Birkhead, Burke, & Slate, 2008; van Oers et al., 2014). Providing a genetic map and recombination landscape to accompany a high-quality genome assembly greatly expands the array of possible research questions and avenues.

Here we combine both conventional short-read (Illumina shotgun and mate-pair) and long-read sequencing (PacBio) with an extensive pedigree of our long-term study population to achieve a highly contiguous assembly and recombination map. The fairy-wren genome assembly has 975 Mb (out of the 1.07 Gb total assembled) of sequences anchored onto 25 pseudochromosomes (out of $n = 36$; L. Christidis, pers. comm.) with a contig and superscaffold N50 of 465 kb and 68.11 Mb, respectively. We also provide comparisons of recombination rate and other genomic features to help understand the sources of variation of these features within the genome. Lastly, we add prediction-based repeat and gene annotations using existing libraries from other bird species. This assembly will greatly facilitate ecological and evolutionary studies of fairy-wrens, and provide a resource for understanding the evolution and diversification of the Meliphagides as a whole.

2 | MATERIALS AND METHODS

2.1 | Reference genome sequencing

We chose a female individual (ANWC:B45704) from the Flinders Island subspecies (*Malurus cyaneus samueli*) for reference genome sequencing (Figure 2); a previous microsatellite survey of *M. cyaneus* suggested that this population has the lowest heterozygosity across the species, (D. Etemadmoghadam, et al., undated thesis, unpublished data, Dept. of Genetics, University of Melbourne) making it an ideal candidate for genome assembly. We extracted the DNA from a

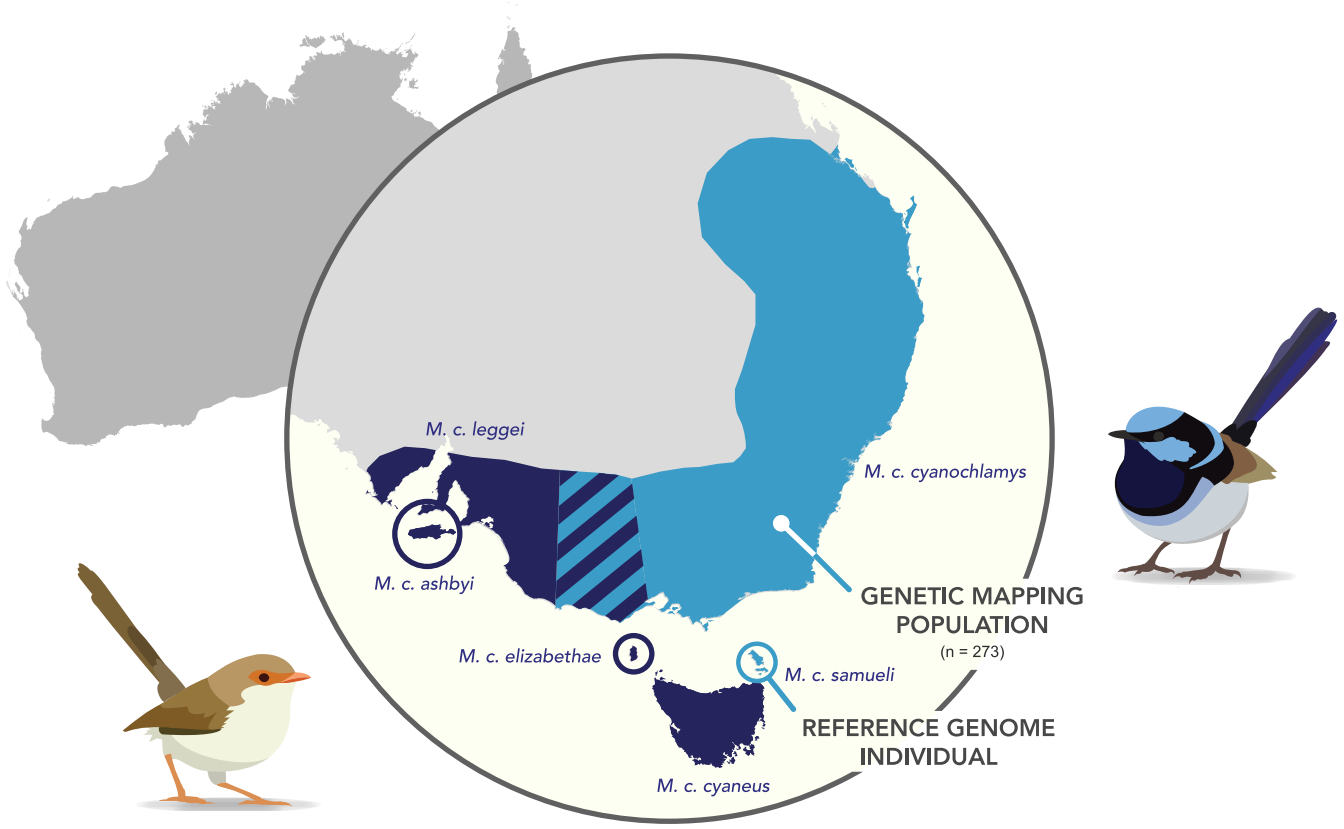

FIGURE 2 Superb fairy-wren range map and sampling. The entire range of the superb fairy-wren is shown here with the six subspecies designations. The individual chosen for the reference genome (ANWC:B45704) comes from the *samueli* subspecies from Flinders Island off Tasmania, while the linkage mapping population comes from the mainland *cyanochlamys* subspecies [Colour figure can be viewed at wileyonlinelibrary.com]

liver tissue sample using a standard salting out procedure. For small insert sizes, we prepared two size ranges centred on 250 and 500 bp using the Meyer and Kircher (2010) protocol. The DNA was sheared using a BIORUPTOR (Diagenode), and a double-sided bead size selection was then used to obtain the correct insert size. The 250-bp library was sequenced using half a lane of an Illumina HiSeq 2500 (100 bp, paired-end) and the 500-bp library was sequenced using a lane of Illumina MiSeq (300 bp, paired-end). Three Illumina mate-pair libraries were prepared and sequenced for insert sizes centred around 3.5, 5.5 and 7.5 kb by the ACRF Biomolecular Resource Facility (ANU). Each mate-pair library was sequenced using 1/6th of an Illumina HiSeq 2500 (100 bp, paired-end). The DNA was extracted for the long-read libraries using the same method as for the Illumina libraries. For the single-cell, long-read libraries, libraries were prepared and sequenced on 27 SMRT cells on the Pacific Biosciences RSII platform by the DNA Sequencing Facility (UC Davis).

2.2 | Linkage map sampling and sequencing

We used a subset of the extensively sampled, wild pedigree of a population of the subspecies *M. cyaneus cyanochlamys* in the Australian National Botanical Garden for linkage mapping (Figure 2). We collected a small sample of blood from all captured birds. Male philopatry and high skew in extragroup mating success resulted in a complex pedigree (Figure S1; Cockburn, Osmond, & Double, 2008; Dunn & Cockburn, 1999). We chose 273 individuals spanning three to nine generations. DNA was extracted from blood samples with the ammonium acetate precipitation protocol (Burke, Bruford, Hanotte, & Brookfield, 1998). Paternity was inferred a priori using microsatellite data and these paternity assignments were used to select the individuals for mapping. The paternity analysis through microsatellites relied on an exclusion approach described most fully by Hajduk et al. (2018). We then sequenced tens of thousands of SNPs distributed randomly across the genome using a method provided by Diversity Arrays Technology Pty Ltd (DaRTseq; Kilian et al., 2012). All prior paternity assignments through microsatellite data were later supported by the SNP genotyping data.

2.3 | De novo genome assembly

The nuclear genome assembly pipeline can be divided into five main stages which improved contiguity, reduced gaps and/or improved quality. Each stage is detailed in Figure S2. The first stage was the Illumina assembly. The raw Illumina reads for both the short insert sizes and mate-pair libraries were trimmed for adapter sequences and low-quality bases using the LIBNGS toolkit (https://github.com/sylvainfor

et/libngs). The genome size was initially estimated using SGA PREQC (Simpson, 2014). This size was used to roughly guide the assembly and assessment of its resultant contiguity. We used the ALLPATHS-LG assembler for the initial assembly and scaffolding (Butler et al., 2008). The second stage was PacBio gap filling: spanning the gaps within and between scaffolds using error-corrected PacBio subreads. Coverage of the PacBio reads was not sufficient to be included in the de novo assembly but was sufficient for filling gaps. These PacBio reads were first error-corrected using the higher quality Illumina reads using the tool LORDEC (Salmela & Rivals, 2014). We then used the error-corrected PacBio reads to upgrade the Illumina scaffolds by filling in the gaps and extending the assembly using PBJELLY (English et al., 2012). The third stage was superscaffolding: anchoring and orienting the scaffolds into pseudochromosome superscaffolds. This assumes that the genetic order, as inferred from the pedigree-based linkage map, corresponds to the physical order along the chromosome. Details on how the genetic map was built for scaffolding is described in more detail below and in the supporting methods. We used the software LEPMAP3 to build our genetic map using the DaRTseq data of our pedigree (Rastas, 2017). Before assembling the superscaffolds, we first identified the large misassemblies in the scaffolds and split them accordingly. We then used ALLMAPS to anchor and orient the scaffolds into 25 superscaffolds using the information from the genetic map (Tang et al., 2015). This stage resulted in new gaps and new associations within superscaffolds, which led to the fourth stage: superscaffold gap filling. Similar to the second stage, the fourth stage used the error-corrected PacBio reads and PBJELLY to fill in any new gaps. The fifth and final stage was to polish the genome assembly and convert it to a pseudohaploid reference. This stage involved mapping all of the Illumina reads onto the stage four assembly using BWA. We then used PILON (version 1.22) to correct incorrectly called bases and small indels (Walker et al., 2014).

The remaining steps were curating and naming of the superscaffolds. All superscaffolds were aligned onto the *Taeniopygia guttata* assembly (version 3.2.4, Zhang, Jarvis, & Gilbert, 2014) using LASTZ and the 25 superscaffolds were assigned to pseudochromosomes based on the match to the *T. guttata* genome, which in turn had been initially named for homology to the *Gallus gallus* genome. The main nomenclature adopted from the *T. guttata* genome was the fissions relative to the *G. gallus* genome such as chromosomes 1A and 4A. W-linked scaffolds were identified using genotyping of the DArTseq SNPs in the mapping population. Any variants that were found only to be genotyped in females and missing in males were considered putative W-linked SNPs. The scaffolds which these SNPs map to were labelled as unmapped W chromosome scaffolds. The remaining small scaffolds were also aligned to the other passerine chromosome assemblies: *Parus major* (version 1), *Ficedula albicollis* (version 1.5) and *Passer domesticus* (version 1). If these unmapped scaffolds were associated with the same chromosome in at least two out of the four genomes, they were assigned as 'unplaced' to that chromosome. Any remaining scaffolds were labelled as 'chromosome unknown'.

The mitochondrial genome was assembled independently from the nuclear genome. To assemble the mitochondrial genome, we used MITOBIM, which used a reference mitochondrial genome of a closely related species to seed the assembly (Hahn, Bachmann, & Chevreux, 2013). We used a reference assembly from a congeneric species, *Malurus melanocephalus* (GenBank NC024873), to bait mitochondrial sequences. We used the Illumina HiSeq (~250-bp insert size) library as it yielded the most consistent results through multiple trials. After choosing the best assembly, we mapped the reads again and corrected any misincorporated sequencing errors using the majority call for each base. Finally, we used MITOS Webserver to identify the location of the genes and tRNAs in the genome (Bernt et al., 2013).

2.4 | Genetic linkage mapping for assembly

The third stage of the genome assembly involved linkage mapping to inform the placement and orientation of each scaffold along each chromosome. First, each DaRTseq library was trimmed to remove Illumina adapters and barcode sequences using TRIMMOMATIC (version 0.32, Bolger, Lohse, & Usadel, 2014). The trimmed reads were then mapped onto the PacBio gap-filled scaffolds (stage 2) using BWA MEM (Li & Durbin, 2009). The remaining steps follow the pipeline provided by LEPMAP3 (version 0.2). This used genotype likelihoods for mapping (Rastas, 2017). The pedigree was split into 37 full-sibling families with a total of 273 individuals. Individuals were included more than once if they belonged to multiple families and grandparents were included to aid phasing.

We first built a framework map by following the LEPMAP3 (LM3) pipeline. We highlight the relevant parameters here but a detailed description of the mapping can be found in the supporting methods and Figure S3. The *SeparateChromosomes2* module split the loci into linkage groups that should be associated with a chromosome. We decided on log of odds (LOD) = 13 as our cut-off for the linkage groups after testing various cut-offs for their generation groups (Figure S4). We then added remaining markers using the *JoinSingles2All* module with a LOD threshold of 10. The *OrderingMarkers2* module finds the best order for markers within a linkage group. This module was run repeatedly, and spuriously mapped markers were manually removed. We continued this process until no more spuriously mapped markers remained. The final marker order constituted our 'framework map'.

Next, we built a 'forced map' using the framework map and forcibly grouping all SNPs within a scaffold into the linkage group to which that scaffold belonged. We performed another round of *JoinSingles2All* to map any additional scaffolds that were not mapped in the framework map. Within each linkage group, we then filtered out all SNPs that were below a LOD threshold of 3. Finally, we performed multiple rounds of the *OrderingMarkers2* and manual curation to build the map we used to anchor and orient the scaffolds.

2.5 | Genome assembly assessment

We summarized the contiguity of the genome assembly using standard metrics provided by the ASSEMBLATHON 2 STATS perl script

(https://github.com/ucdavis-bioinformatics/assemblathon2-analysis). The contiguity of the genome assembly was assessed between each major assembly step and the final assembly was compared to other chromosome-level bird assemblies. The completeness and quality of the genome assembly was assessed using BUSCO (3.0.2) analyses, which searches the assembly for 4,915 universal single-copy orthologues from the AVES (odb9) database (Simão, Waterhouse, Ioannidis, Kriventseva, & Zdobnov, 2015). As for assembly contiguity, this was measured between each assembly step and between the final assembly and existing chromosome-level assemblies.

2.6 | Predictive gene and repeat annotation

The predictive gene annotation followed the same annotation pipeline that was run for other avian assemblies in the B10K project for comparability (Zhang et al., 2015). Briefly, the primary gene set was derived from ENSEMBL85 (from the *G. gallus* galGal4 and *T. guttata* taeGut2 genome assemblies) totalling 20,194 genes. A supplementary gene set was also compiled using 20,169 human genes, and genes from 71 transcriptomes for a second set. The two rounds of annotation using the two different gene sets involved a rough alignment using TBLASTN (version 2.2.2) and generating a gene model using GENEWISE (wise2.4.1). We then filtered out short proteins (<30 amino acids), pseudogenes, retrogenes, highly duplicated genes with 70% repeats or with a single exon, and redundant or overlapping genes. For de novo repeat discovery, we ran REPEATMODELER (version 1.0.8) on the assembly to create a *Malurus*-specific repeat library. To characterize the final repeat content, we ran REPEATMASKER (version 4.0.7) using the database containing the *Malurus*-specific repeat library and those from other avian databases: *T. guttata*, *F. albicollis* and *Corvus cornix* (Vijay et al., 2016).

2.7 | Synteny comparisons and intrachromosomal rearrangement (inversion) inference

Large-scale synteny between the final fairy-wren genome assembly and existing chromosome-scale assemblies were compared using LASTZ (version 1.04.00) alignments. The synteny was represented in CIRCOS plots to qualitatively compare interchromosomal rearrangements such as chromosomal fissions, fusions and translocations. We performed a preliminary exploration of potential inversions between *M. cyaneus* and other species. For this inference we included only the autosomes which were homologous to the 10 largest autosomes in the *G. gallus* genome (i.e. Chromosomes 1–10, 1A and 4A) and the Z chromosome. We also omitted *Meleagris gallopavo*, *Aquila chrysaetos*, *Falco peregrinus* and *Strigops habroptila* from the analysis due to the pronounced levels of interchromosomal rearrangements that made collinear comparisons intractable. First, the chromosomes were masked using REPEATMASKER (version 4.0.7). We then performed pairwise alignments between *M. cyaneus* and comparison species using PROGRESSIVEMAUVE (version 2.4.0, Darling, Mau, & Perna, 2010) for each chromosome separately. We then used GRIMM-SYNTENY (version 2.02, Tesler, 2002) to identify syntenic blocks between alignments using two block sizes of 50 and 500 kb with gap sizes of 10 and 50 kb, respectively. We used two sizes to see if we observe similar trends, although we are mostly interested in larger (megabase-scale) inversions and will mostly focus on the 500-kb block size results. Lastly, we used MGR (version 2.01; Bourque & Pevzner, 2002) to find the optimal number of intrachromosomal rearrangements necessary to obtain the current order and orientation of syntenic blocks.

2.8 | Recombination landscape and comparisons

The process to generate the recombination landscape was identical to the linkage mapping but used the information from the physical position on the final genome assembly as the physical order of the markers. First, we remapped the DaRTseq reads on the final polished genome assembly using BWA. The first stages of the LM3 pipeline were run as previously. As with the forced map, within each chromosome we filtered out markers which fell below the LOD score limit of 3. We then ran the *OrderMarkers2* module using the physical order and obtained the sex-specific and sex-averaged genetic maps for the 24 autosomes and the male-specific map for the Z chromosome.

We obtained recombination rate estimates from the Marey map representation: genetic distance plotted against physical distance. First, we smoothed the map using a LOESS (local polynomial regression) smoothing with a span of 0.2. This regression smooths over windows of a fixed number of SNPs instead of physical length to reduce bias in regions where more SNPs were recovered. Using this regression, we estimated recombination rates in 200-kb non-overlapping windows and correlated these estimates with other genomic features. GC content was calculated as the percentage of G or C bases within the given window. Gene density was the percentage of coding sequence within the 200-kb window. Relative distance from chromosome end was standardized from 0 to 1, 0 corresponding to the chromosome end and 1 corresponding to chromosome centre. Before analyses, recombination rate was $\log_{10}$-transformed to reduce skewness. Gene density and GC content were both square root-transformed. We explicitly tested the correlation of recombination rate to both gene density and distance to chromosome end using Pearson's *R*. GC content was not explicitly correlated with recombination rate as this feature was not an explanatory variable but rather probably a byproduct of the variation in recombination (Bolívar, Mugal, Nater, & Ellegren, 2016; Clément & Arndt, 2013). To minimize the effects of spatial autocorrelation within the genome, we performed permutation tests to obtain the distribution of Pearson's *R* statistic under a null model (shuffled recombination rate and explanatory variables among windows) and our observed data. Furthermore, we performed bootstrapping by subsampling 20% of the windows

with replacement. We performed 2,000 iterations of bootstrapping to obtain the null distribution of Pearson's *R* and another 2,000 for the observed distribution. Because chromosome sizes ranged over an order of magnitude, standardizing distance from a chromosome end from 0 to 1 across all chromosomes may not always be entirely comparable. We therefore included a separate representation of the effect of the distance to the chromosome end by comparing the $\log_{10}$-transformed recombination rate to the $\log_{10}$-transformed exact distance to each chromosome end in megabases. We then performed a LOESS smoothing of this correlation to illustrate the shift in recombination rates from the ends of the chromosomes to the centre. Between chromosomes, we also compared recombination rate (genetic map distance/chromosome length) to chromosome length and gene density (number of genes/chromosome length). Additionally, we qualitatively compared the fairy-wren genetic map (25,841 SNPs) with the pedigree-derived genetic maps of *G. gallus* (9,176 SNPs; Groenen et al., 2009), *T. guttata* (700 SNPs; Backström et al., 2010), *F. albicollis* (4,161 SNPs; Kawakami et al., 2014) and *P. major* (1,213 and 2,478 SNPs; van Oers et al., 2014). Both mapping populations of *P. major* were included in the comparison. The markers were initially filtered for those available in both maps and were concordant with the physical order. Note that the *G. gallus* and *T. guttata* maps were derived from a captive population and may not be reflective of the genetic maps of their corresponding wild populations.

3 | RESULTS

3.1 | Genome sequencing effort

The raw data comprised 128,683,241 read pairs from the 250-bp insert library, 12,456,651 read pairs from the 500-bp insert library, 47,009,583 from the 3.5-kb mate-pair library, 56,672,592 from the 5.5-kb mate-pair library, and 39,069,768 from the 7.5-kb mate-pair library. The SGA PREQC results estimated the genome size to be 1.07 and 1.04 Gb from the 250- and 500-bp insert size libraries, respectively. Using the 1.07-Gb genome size estimate, sequencing coverage corresponded to 24.0× from the 250 bp insert library, 7.0× from the 500-bp insert library, 8.7× from the 3.5-kb mate-pair library, 10.5× from the 5.5-kb mate-pair library and 7.3× from the 7.5-kb mate-pair library for a total of 55.7× coverage from Illumina sequencing. After filtering and cleaning of the raw reads, we retained 91,417,772 read pairs and 21,503,225 merged or unpaired reads from the 250-bp insert size Illumina library and 6,303,745 read pairs and 3,496,087 merged or unpaired reads from the 500-bp insert size Illumina library. For the mate-pair libraries, we retained 12,287,646 read pairs for the 3.5-kb insert size, 15,656,984 read pairs for the 5.5-kb insert size and 9,326,889 read pairs for the 7.5-kb insert size libraries (Table S1).

From the 27 PacBio RSII SMRT cells, we sequenced 2,676,850 subreads. The mean length of the subreads was 8.7 kb with the longest subread being 53.6 kb. Total coverage of the PacBio filtered subreads was 22× using the estimated 1.07-Gb size genome. The distribution of subread lengths is given in Figure S5. Of the total number of reads, 1,271,158 (47%) were error-corrected by LORDEC.

3.2 | Genome assembly

Here we focus on the stages of genome assembly, but the specific sections that require more attention (such as linkage mapping and misassembly detection) will be discussed in detail below. The initial Illumina assembly yielded a scaffold N50 of 6.0 Mb and a contig N50 of 15 kb. The size of the scaffold assembly was 1.01 Gb (94.4% of the predicted 1.07-Gb genome size) with a total gap length of 113 Mb. PacBio gap filling yielded a marginally increased scaffold N50 but a substantial 31-fold increase in contig N50 to 465 kb. Additionally, gap length was greatly reduced to 27 Mb. The size of the scaffold assembly was 1.05 Gb (98.1%) after stage 2. The largest contribution to the PacBio reads here was the increase in contig N50 and reduction in gap length. The superscaffolding stage used the genetic map to place and orient the scaffolds into pseudochromosomes. Because our subsampling of the pedigree consisted only of 273 individuals, we were only able to assemble 25 chromosomes with confidence and not all are as large as their homologues in other bird species. This step anchored 298 scaffolds (975 Mb = 90.4% of the genome) into pseudochromosomes and oriented 247 scaffolds (894 Mb = 82.9% of the genome). This left 104 Mb (9.6%) of the genome unplaced. The superscaffold N50 increased to 67.7 Mb (~8.5-fold increase) but the contig N50 stayed approximately the same (465 kb). Total scaffold assembly length was also 1.05 Gb. Gap length remained at 27 Mb. Stage 4 (superscaffold gap filling) yielded only a marginal increase in superscaffold N50 to 68.1 Mb with an increase in the contig N50 to 540 kb. This stage elevated the assembly length to the predicted length of 1.07 Gb. This further reduced the gap length to 16 Mb. Final assembly polishing had a superscaffold N50 of 68.11 Mb and contig N50 of 560 kb. The total assembly length of the superscaffold was 1.07 Gb and the contig was 1.06 Gb, both of which are around the estimated genome length. In total there were 25 superscaffolds, 4,314 scaffolds and 15,027 contigs. The final gap length remained at 16 Mb. A more detailed breakdown of the assembly statistics (via the ASSEMBLATHON2 STATS script) can be found in Table S2. The progression of contiguity and the BUSCO assessment of the stages of genome assembly are presented in Figure 3. The assembled chromosome sizes are given in Table 1. The final MITOBIM mitochondrial genome assembly size was 17,031 bp. The annotation resulted in the expected 13 genes (*COI–III*, *ND1–6*, *cytB*, *ATP6 and ATP8*), two ribosomal RNAs (12S and 16S) and 22 tRNAs (Figure S6).

3.3 | Genetic linkage mapping for superscaffolding (Stage 2 to Stage 3)

We recovered 35,276 SNPs after mapping the DArTseq reads onto the stage 2 assembly. The *SeparateChromosomes2* analysis with LOD limit of 13 yielded 150 linkage groups with 7,331 markers

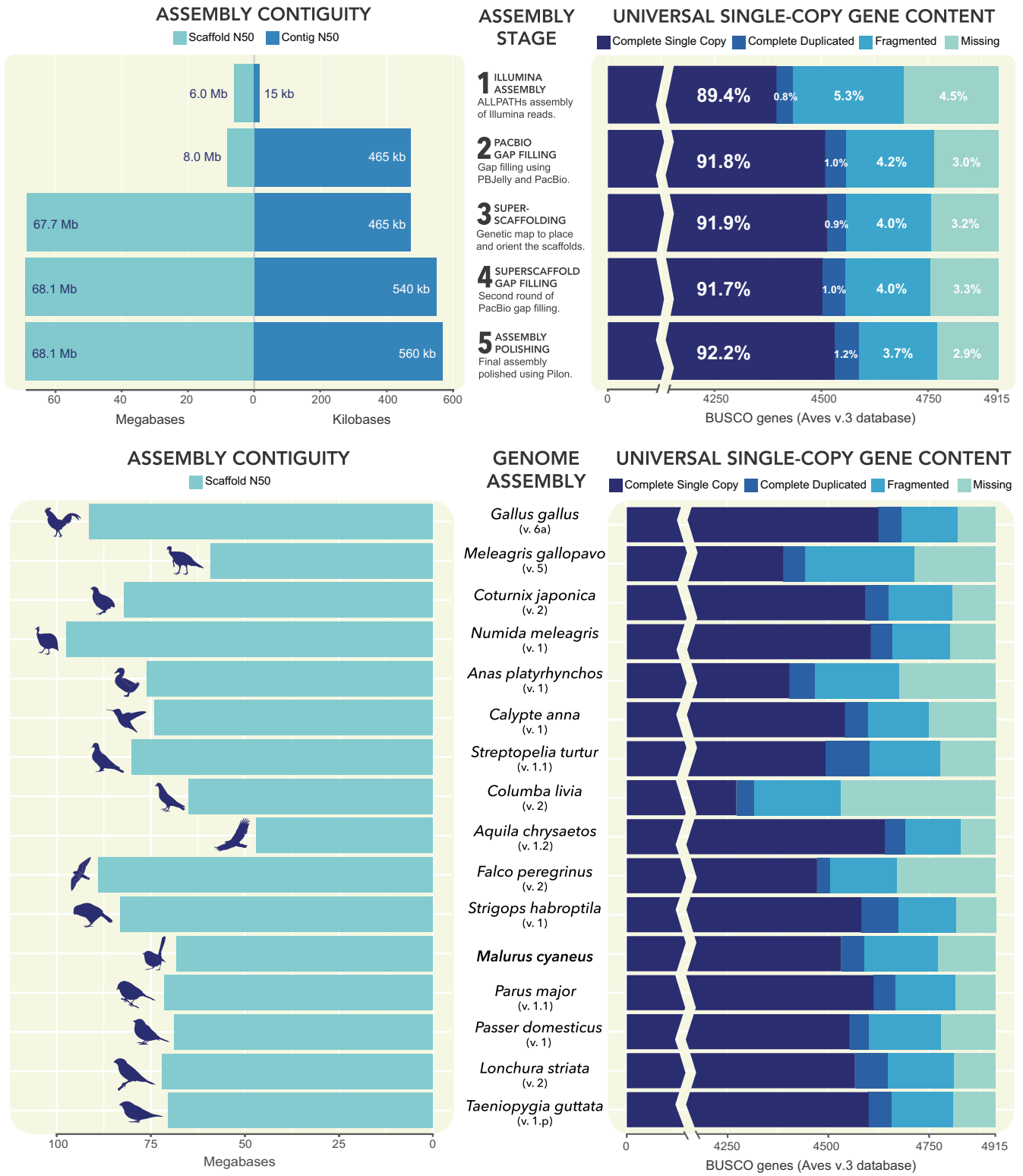

FIGURE 3 Genome assembly quality assessments. The assembly contiguity metrics on the left plots show scaffold and contig N50 estimated and the assembly completeness metrics on the right plots show the BUSCO gene search using the aves database. The plots on the top row show the progression of increasing assembly quality between the five stages of genome assembly. The plots on the bottom row compare assembly quality between bird genomes, showing the *Malurus cyaneus* assembly to be comparable to other chromosome assemblies [Colour figure can be viewed at wileyonlinelibrary.com]

(minimum of three markers per linkage group). From the linkage groups, 36 had at least six markers or more. For the remaining stages we retained 36 linkage groups (6,589 loci) ensuring each group had multiple markers that mapped to two or more scaffolds. After joining additional markers to these 36 groups, we retained a total of 11,414 markers. Multiple iterations of the *OrderMarkers2*

TABLE 1 Summary chromosome metrics: assembly lengths, linkage map lengths, and annotation of each chromosome in the final superb fairy-wren genome assembly; *Taeniopygia guttata* chromosomes 19 and 26 are within chromosomes 2 and 5, respectively, in these metrics

			Genetic distance (cM)			
Chromosome	Size (Mb) (unplaced)	GC (%)	Average	Male	Female	Genes (unplaced)
Chrom 1	125.32 (1.70)	39.96	90.11	95.55	85.58	1,307 (7)
Chrom 1A	68.11 (1.56)	39.92	74.00	84.65	63.91	846 (8)
Chrom 2	163.99 (0.52)	40.06	104.54	115.68	94.83	1,788 (4)
Chrom 3	108.86 (0.67)	39.83	75.75	92.09	60.07	1,111 (15)
Chrom 4	72.11 (0.60)	39.48	94.59	99.70	90.28	777 (1)
Chrom 4A	18.69 (1.28)	44.01	48.56	52.50	45.16	303 (54)
Chrom 5	54.65 (5.10)	41.68	86.62	84.02	90.64	907 (137)
Chrom 6	33.73 (1.01)	41.96	66.35	73.08	60.36	555 (16)
Chrom 7	35.83 (1.65)	41.13	84.07	89.82	79.96	509 (51)
Chrom 8	30.90 (0.20)	42.47	84.70	90.56	80.01	532 (3)
Chrom 9	24.52 (0.04)	43.31	54.55	56.75	52.83	439 (0)
Chrom 10	20.32 (0.00)	43.56	78.61	82.08	75.93	416 (0)
Chrom 11	19.37 (0.22)	42.46	47.16	52.09	42.8	303 (14)
Chrom 12	21.60 (0.04)	44.12	60.19	62.79	58.47	364 (0)
Chrom 13	19.04 (0.02)	45.45	46.88	43.06	51.09	376 (0)
Chrom 14	19.34 (0.02)	45.38	79.74	75.23	86.83	490 (0)
Chrom 15	14.69 (0.13)	46.33	67.23	61.94	73.41	375 (6)
Chrom 17	12.35 (0.01)	48.20	50.47	45.97	55.39	294 (0)
Chrom 18	12.16 (1.30)	46.76	48.22	52.08	44.82	308 (36)
Chrom 20	14.96 (0.06)	46.77	47.28	44.82	50.17	341 (0)
Chrom 21	5.21 (2.14)	47.45	25.93	26.52	25.75	154 (77)
Chrom 23	3.93 (1.45)	48.84	26.90	30.21	24.00	96 (43)
Chrom 27	2.33 (2.18)	50.26	17.14	19.27	15.25	138 (120)
Chrom 28	3.94 (2.20)	50.74	22.16	22.92	21.6	173 (67)
Chrom W	– (0.39)	–	–	–	–	– (3)
Chrom Z	68.91 (16.80)	37.23	–	72.81	–	581 (112)
Unknown	– (62.72)	–	–	–	–	– (1,202)
Total	1,078.87	42.04	1,618.30	1,686.92	1,583.19	15,458

module and manual curation eventually resulted in 26 linkage groups that we can confidently assign to chromosomes. Some of the smaller linkage groups were associated with the same microchromosome and were merged when appropriate. Because average recombination rates were markedly higher in microchromosomes, they would probably have required a lower LOD score limit relative to the macrochromosomes. The first stage of the forced map resulted in 32,708 SNPs associated with an existing linkage group. Adding new scaffolds onto the existing map resulted in an additional 390 SNPs being mapped. The final *OrderingMarkers2* iterations and manual curation dropped one linkage group that was nested within a single scaffold. This was the final set of 25,078 SNPs in 25 linkage groups that was used for anchoring and ordering of the scaffolds.

Linkage mapping consistently associated markers between two pairs of chromosomes, suggesting potential fusions relative to the *Taeniopygia guttata* karyotype. Markers mapping to multiple scaffolds associated with *T. guttata* chromosomes 26 and 19 were consistently associated with chromosomes 5 and 2, respectively (Figure S4). Fusion of chromosomes 2 and 19 was further supported by a single, large scaffold which had segments associated with both chromosomes. Furthermore, the Marey map (Figure 4; Figure S8) and recombination landscapes (Figure S9) within the fused chromosomes were undisrupted by large jumps, signifying incorrectly mapped loci. The karyological evidence showed a diploid number of 72 for *Malurus cyaneus* and 80 for *T. guttata* (L. Christidis, pers. comm), suggesting fusions relative to the *T. guttata* karyotype. The two fusions were retained during the superscaffolding and the final pseudochromosomes labelled Chrom 2 and Chrom 5 contain Chrom 19 and Chrom 26, respectively. Because there was insufficient data to reconstruct all of the microchromosomes, we were unable to fully characterize the chromosomal rearrangements with confidence. A summary of the number of SNPs retained for each step can be found in Figure S3.

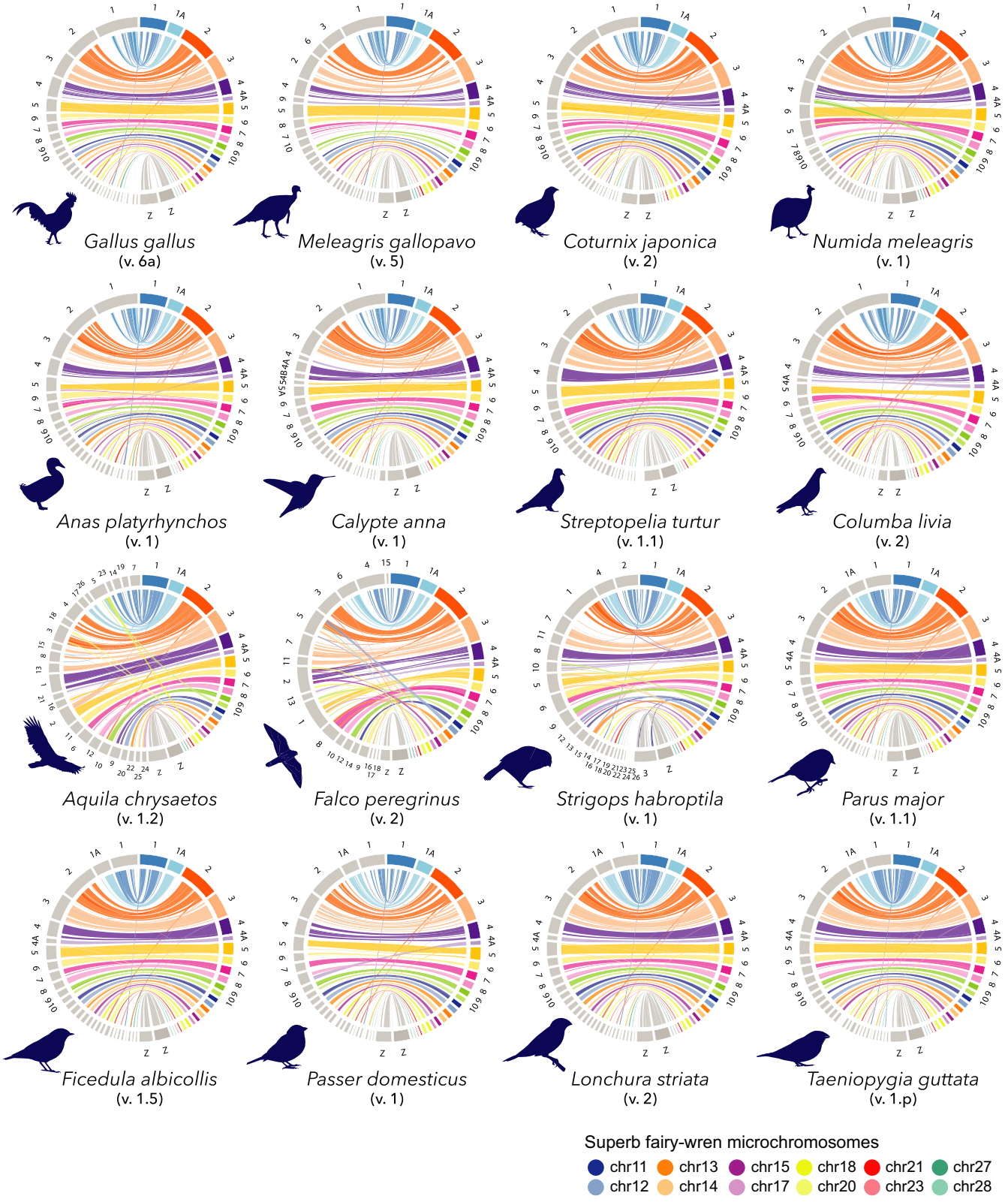

FIGURE 4 Superb fairy-wren synteny plots. These plots show pairwise synteny comparisons between the *Malurus cyaneus* chromosomes and the chromosomes of 16 other bird assemblies [Colour figure can be viewed at wileyonlinelibrary.com]

3.4 | Gene and repeat annotation

In total, the gene annotation pipeline predicted 15,458 genes across the genome. Of these, 13,483 (87.2%) were annotated in assembled pseudochromosome superscaffolds, 774 (5.0%) were annotated in unplaced scaffolds associated with chromosomes, and 1,202 (7.8%) were annotated in scaffolds of unknown location. Of the unplaced scaffolds, we annotated three genes associated with the W chromosome (Table 1). With the repeat annotation and masking, REPEATMASKER masked 7.90% of the total genome length. Of this fraction of the genome, 41.1% were LINEs (long interspersed elements), 28.5% were LTR (long terminal repeat) elements, 1.4% were SINEs (short interspersed elements), 1.0% were DNA elements and 8.2% were unclassified. Of the remaining, 15.4% were simple repeats, 3.8% were low complexity, 0.6% were small RNAs and 0.5% were satellite DNA (Table S3).

3.5 | Comparison with existing chromosome-level assemblies

We compared the final *M. cyaneus* genome assembly to 16 other high-quality, chromosome-level bird genome assemblies for contiguity, completeness and synteny. The scaffold N50 of the other genome assemblies ranged from 46.93 to 97.48 Mb (mean ± *SD* = 74.7 ± 12.61 Mb) while that of *M. cyaneus* was at 68.1 Mb, and hence falls within the contiguity range of other chromosome-level bird genome assemblies. The same BUSCO search was run for the other bird genomes. The *Ficedula albicollis* genome was omitted due to an undeterminable error during the gene search although the other genomes were sufficient for comparison. On average, 92.15% (±0.02%) of the 4,915 AVES (odb9) genes were recovered as complete and single-copy for the other genomes and 92.2% for the superb fairy-wren assembly (Figure 3).

3.6 | Chromosomal rearrangements

Synteny of the macrochromosomes is largely conserved between the fairy-wren and other bird genome assemblies (Figure 4). While most assemblies had chromosomes that were named after the homology to the *Gallus gallus* genome, the *Meleagris gallopavo*, *Numida meleagris* and *Strigops habroptila* chromosomes were named from decreasing size. The chromosomes were reordered in the CIRCOS plot to match the homologous chromosomes of the *M. cyaneus* genome for comparison. The *Falco peregrinus* genome, like other falconid genomes, has consistently been shown to have many rearrangements (Nishida et al., 2008; O'Connor et al., 2018). Interestingly the other raptor genome, *Aquila chrysaetos*, also exhibits high levels of rearrangements despite being distantly related and converging on raptorial lifestyles. *S. habroptila* is the closest relative of *F. peregrinus* here and also shows many interchromosomal rearrangements, although not as substantial (Figure 1). Comparisons with nonpasserine species show quite a few interchromosomal rearrangements, even among the macrochromosomes, relative to the oscine passerine species, which show fewer. The CIRCOS plots also show probable fusions or translocations of microchromosomes to macrochromosomes in the *M. cyaneus* genome relative to all other genomes. This is indicated by fewer chromosomes in the *M. cyaneus* karyotype but may also be due to misassemblies. Currently, there are not enough data to distinguish these alternatives.

Inference of intrachromosomal rearrangement yielded a sizable number of rearrangements for each pairwise comparison. Note that the software assumes that the discrepancy between the order of syntenic blocks was generated by inversions rather than intrachromosomal translocations. Chromosomes were concatenated if the synteny suggested fusions and fissions (Table S4). The number of rearrangements ranged from 129 (*Columbia livia*) to 223 (*Passer domesticus*) with an average of 167 (*SD* 28.68; Table S5). There was no notable difference between pairwise comparisons with other oscine passerines (160.2 ± 36.49) and non-passerines (171.0 ± 23.59). Among chromosomes, inversion number and chromosome size were positively correlated (Figure S7). In the 50-kb syntenic block sizes, chromosome Z showed a much higher number of inversions, as expected for a chromosome its size but that is lost in the more conservative 500-kb syntenic block size inference.

3.7 | Genetic map and recombination landscape comparisons

On average, the male-specific, autosomal genetic map of the fairy-wren was longer than that of the female-specific map (Figure S8). This resulted in the total male-specific map being 1.08 times longer than that of the females. Other passerine systems also show deviation from sex-equal recombination rates but differ regarding which sex has the longer map. In the *F. albicollis* system, the male genetic map is 1.13× longer than that of the females, and in two *P. major* populations the female maps are longer by a factor of 1.04 and 1.05 (Kawakami et al., 2014; van Oers et al., 2014). The difference between male and female map lengths also varied with respect to chromosome (Table 1). When map length was converted to recombination rate within the chromosomes, we did not always consistently find that male-specific local recombination rate was higher than that of the female-specific rate.

We compared the sex-averaged genetic distance (cM) along the length of the chromosome between five different bird species: the *M. cyaneus*, *G. gallus*, *T. guttata*, *F. albicollis* and *P. major* genetic maps (Figure 5). For the *P. major* map, both mapping populations from van Oers et al. (2014) were included in the comparison. Generally, the two *P. major* maps were near identical or were more similar to each other rather than the other species. While the detailed comparison of the *F. albicollis* map has emphasized that their map is more similar to that of *G. gallus* than to that of *T. guttata* (Kawakami et al., 2014), the *M. cyaneus* map is more concordant with the *T. guttata* and *P. major* maps, despite *F. albicollis* also being a passerine species. The *M. cyaneus*, *T. guttata* and *P. major* maps were shorter than the

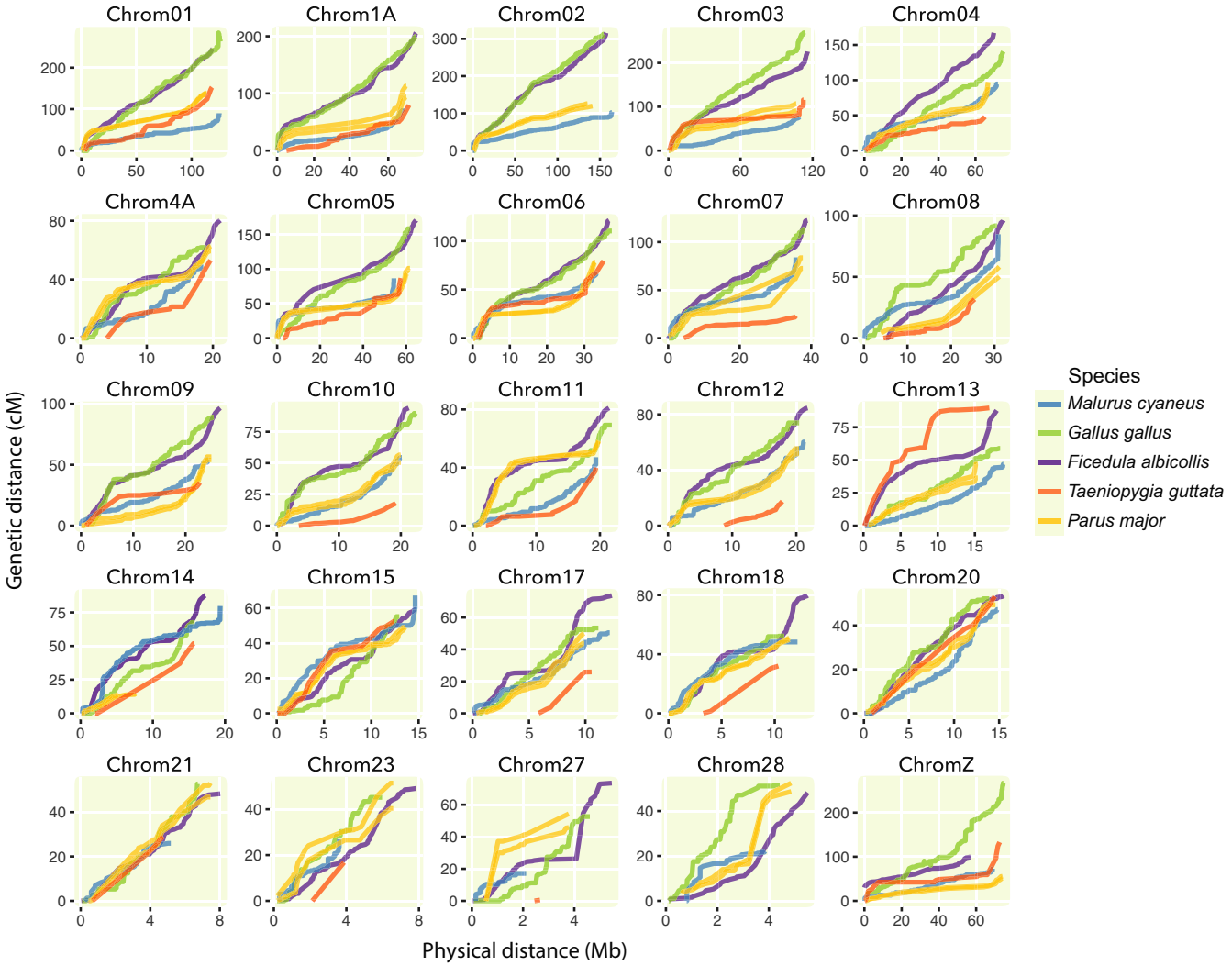

FIGURE 5 Comparative Marey maps. Sex-averaged genetic distances (cM) plotted against physical distances (Mb) across 25 autosomes and the Z chromosome. The pedigree-based linkage maps of the chicken (*Gallus gallus*), collared flycatcher (*Ficedula albicollis*), zebrafinch (*Taeniopygia guttata*) and great tit (*Parus major*) are plotted with the *Malurus cyaneus* map for comparison. The remaining *M. cyaneus* autosomes were omitted due to short size and lack of information [Colour figure can be viewed at wileyonlinelibrary.com]

other two maps. Even in chromosome 2, where the *T. guttata* genetic map from Backström et al. (2010) has usually been omitted due to inconsistencies with the physical map, the *M. cyaneus* map was shorter than that of the other three species, suggesting generally lower recombination rates. The *T. guttata* map generated by Stapley et al. (2008) estimated the total map length of chromosome 2 to be 34.3 cM. Although this short map length may be confounded by the number of markers, it is almost an order of magnitude lower than that of either *F. albicollis* or *G. gallus*. This pattern was most apparent in the largest chromosomes, and within the smaller chromosomes the patterns were less consistent, although the *M. cyaneus* map was generally within the range of the other four species' maps. In having recombination 'deserts' across most macrochromosomes, the *M. cyaneus* map is also more similar to the *T. guttata* and *P. major* maps. There is no trend between the maps derived from the captive populations (*G. gallus* and *T. guttata*) and wild populations (*M. cyaneus*, *F. albicollis* and *P. major*).

We compared the recombination landscape (cM/Mb) across the genome estimated from the genetic map with various genomic features (Figure 6). All comparisons were made using measures in 200-kb nonoverlapping sliding windows. Gene density had the lower strength of association (Pearson's R = 0.12 ± 0.03) but was still significantly different from zero and the null distribution. This positive correlation between gene density and recombination rate was consistent within and between chromosome comparisons. The macrochromosomes were the least gene-dense and had the lowest average recombination, and the microchromosomes were both gene-dense and had high average recombination rates. A higher correlation from our comparisons was for the relative distance from chromosome ends (Pearson's R = −0.59 ± 0.02). The negative relationship showed a decreasing recombination rate from the end to the centre of the chromosome. This pattern was also found in the LOESS-smoothed relationship between recombination rate and physical distance from chromosome ends, which also showed a rapid

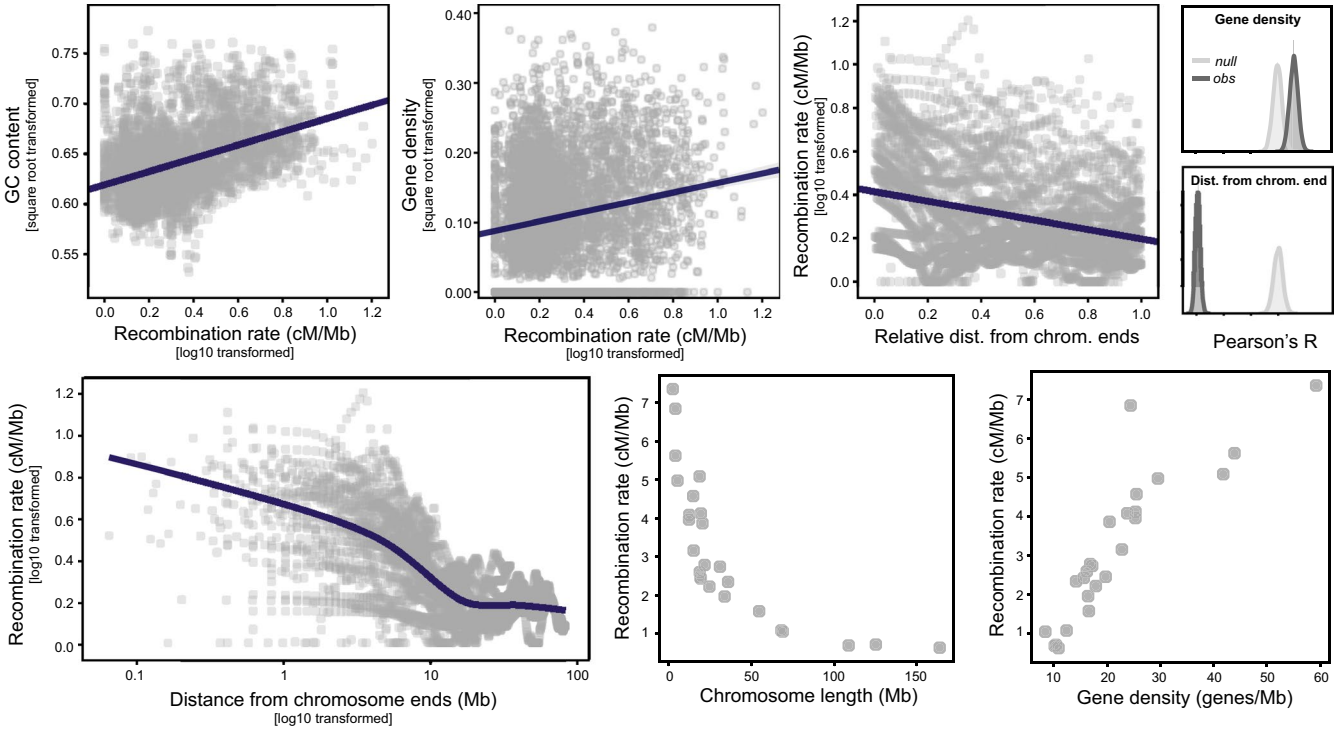

FIGURE 6 Recombination rate correlations. Plots showing relationship between recombination rate and different genome features. From left to right the top panel shows the relationship between recombination rate and GC content, gene density, and relative distance from chromosome ends. The smaller plots on the right show the distribution of the Pearson's *R* values from 2,000 bootstrap permutations. From left to right, the bottom panel shows the decrease in recombination rate from the end to the centre of the chromosome, between-chromosome comparisons of recombination rate and chromosome length, and between-chromosome comparisons of recombination rate and gene density [Colour figure can be viewed at wileyonlinelibrary.com]

decrease in recombination rate ~10 Mb (Figure 6). This can also be observed in the recombination rate plotted across the length of the chromosomes (Figure S9).

4 | DISCUSSION

The superb fairy-wren genome is particularly important in that it fills a gap in the passerine tree as it is the first member of the Meliphagides infraorder to have a reference genome. Not only does it provide a useful resource for species within that clade (Peñalba, Joseph, & Moritz, 2019) and the nonpasserine part of the tree but also a useful comparison point in the broader passerine and avian tree for comparative evolution. As part of the early Australasian radiation within oscine passerines, the *Malurus cyaneus* reference genome anchors the clade which comprises almost half of all avian diversity. This filled gap would allow us to search for shared genomic features that evolved uniquely within the oscine clade. On the other side of the coin, we will also be able to compare how rapidly large-scale genomic-scale variation has accumulated in oscine passerines, a relatively young radiation, in comparison to the remaining nonpasserines.

One initial example of potential for comparative genomics is the evolution of chromosomal rearrangements. Avian genomes have generally been shown to have high synteny (Ellegren, 2010). Previous karyotypic and genome-scale studies, however, have shown that species from certain clades have genomes that have undergone more chromosomal rearrangements than others. Further across birds certain types of rearrangements are more common than others (Ellegren, 2010; O'Connor et al., 2018; Skinner & Griffin, 2012). Although gene synteny along a chromosome might remain high, chromosome number between bird species has a wide range, suggesting that chromosome fusions and fissions might be fairly common (Kapusta & Suh, 2017). The *M. cyaneus* genetic map provides evidence of fusion events between macrochromosome 2–19 and macrochromosome 5–26 (nomenclature based on the *Gallus gallus* reference). There are other notable interchromosomal rearrangements in other genome assemblies even in some of the macrochromosomes. More points of comparison would help to show whether certain chromosomes or particular motifs found in chromosomes might be more prone to rearrangements than others. Furthermore, the dynamic fission and fusion of the gene-rich microchromosomes may have implications regarding species or clade-specific adaptation (Guerrero & Kirkpatrick, 2014; Hansmann et al., 2009; Wellband et al., 2019). Physically creating or breaking linkages between sets of genes will change the local recombination rates. In turn, this would affect the efficiency of selection acting on certain combinations of alleles.

Our preliminary inference of intrachromosomal rearrangements yielded many potential inversions among the macrochromosomes.

Interestingly, there is no obvious difference in inversion counts between the *M. cyaneus* versus other oscine passerines comparison and the *M. cyaneus* versus nonpasserines comparison. This may be an artefact of the ability to detect syntenic blocks or variation in the rate of accumulating inversions along the different lineages. The Z chromosome was shown to have a higher number of inversions similar to other studies (Hooper & Price, 2015, 2017) but only in the comparison using the 50-kb syntenic block size. This difference between Z chromosome and autosomes is negligible with a conservative block size of 500 kb. It is possible that the high number of inversions may make it difficult to establish larger syntenic block sizes. Although theoretical studies support the idea that inversions in the sex chromosome are expected to fix at a higher rate due to their role in sex chromosome evolution, local adaptation and reproductive isolation, it is also possible that this is more relevant over shallower evolutionary timescales and may slow down through time (Charlesworth, Coyne, & Barton, 1987; Connallon et al., 2018; Kirkpatrick, 2010). Although we use the terms inversion and intrachromosomal rearrangement interchangeably here, we cannot rule out intrachromosomal translocations. Furthermore, variation in assembly quality, particularly in properly orienting scaffolds within chromosomes, will affect the inference. The reported numbers are probably an overestimate but provide a good starting point for comparison. As more highly contiguous bird genome assemblies accumulate, it will be possible to explore both inter- and intrachromosomal rearrangements in greater detail using more robust inference methods and in a phylogenetic context.

A fine-scale genetic map and recombination landscape is the first step in understanding the causes and consequences of variation in local recombination rates. In particular, a pedigree-based recombination landscape is not as biased by effective population size or selection as a population-based recombination landscape might be. From the existing chromosome-scale genomes, this is only the fifth that is accompanied by a genetic map. In comparison with the four available genetic maps, we find that the *M. cyaneus* map bears more similarity to the *Taeniopygia guttata* map and *Parus major* maps. By contrast, the map for *Ficedula albicollis*, which is also a passerine, is more similar to the distantly related *Gallus gallus* (Figure 5). Within the *M. cyaneus* map, we also find variation between male and female recombination rate. Genetic maps from a larger variety of populations and species would be required to start characterizing the variation in this trait and testing hypotheses as to what may be driving this variation between sexes and between species.

Delving into the fine-scale variation within the genome, we found a positive correlation between local recombination rate and GC content, and gene density as well as a stronger and negative relationship with distance from chromosome ends. More broadly we also see a relationship between average recombination rate and chromosome size. These correlations are consistent with previous studies of recombination (Kawakami et al., 2014; Paape et al., 2012). The correlation with GC content is probably due to G-biased gene conversion. Errors during repair of double strand breaks during meiosis tend to be biased towards guanine, resulting in a higher GC content in regions which have high rates of recombination (Bolívar et al., 2016; Clément & Arndt, 2013; Fullerton, Bernardo Carvalho, & Clark, 2001). Our sliding window analysis shows only a slight positive correlation between gene density and local recombination rate. Selection pressure on local recombination rate would depend probably not only on the density of genes but also the type of genes in a given location and how they are regulated (*cis*- vs. *trans*-). In the between-chromosome comparison, we found that smaller chromosomes tend to have much higher recombination rate and also tend to have much higher gene density (Axelsson, Webster, Smith, Burt, & Ellegren, 2005). Furthermore, recombination tends to be higher closer to the ends of chromosomes (Haenel, Laurentino, Roesti, & Berner, 2018). While we can only scratch the surface of these comparisons here, understanding the causes and consequences of recombination rate variation is critical in understanding how genome organization influences variation in the efficiency of selection across the genome. The recombination rate itself is a unique trait in that it can shift under the influence of selection and yet simultaneously influence the efficacy of selection on various other traits (Comeron, 2017; Schumer et al., 2018; Wang, Street, Scofield, & Ingvarsson, 2016). Providing these resources from different species would lead us closer to disentangling the complex relationship between the organization of the genome and the evolutionary forces that act on it.

This initial draft of the *M. cyaneus* genome is a well-developed foundation that can be further improved through rapidly developing sequencing technologies and scaffolding methods. As with many of the existing bird assemblies with assembled chromosomes, not all of the microchromosomes are fully resolved and may still benefit from physical mapping methods such as HiC or BioNano optical mapping (Dudchenko et al., 2017; Jiao et al., 2017; Lehmann et al., 2019). A larger sampling of the pedigree would also provide more informative meioses. That would help in orienting scaffolds already placed in chromosomes. It would also better scaffold the smaller microchromosomes, which tend to have higher recombination rate and lower linkage (Fierst, 2015). The capability of long-read sequencing technology has also advanced such that longer fragments can be sequenced. This would help close even more gaps and potentially extend through the telomeres (Kingan et al., 2019; Michael et al., 2018). Highly contiguous, chromosome-scale reference genomes are becoming easily accessible, further expanding our ability to test various hypotheses. More work can still be done, however, to further refine these assemblies and get closer to the true structure of the genome.

Here we provide the first high-quality reference assembly of an intensively studied Australian bird species, the superb fairy-wren, and the exceptionally diverse family and infraorder to which it belongs. The species has been a focal organism in studying cooperative breeding and sexual selection. This new resource will complement the wealth of behavioural and ecological data, particularly the long-term, multidecade study of the population in the Australian National Botanical Gardens, Canberra. This reference assembly accompanied by a genetic map and annotation will open up the system to the field of behavioural genomics and strengthen the system as a key

player in evolutionary and ecological studies. This reference genome provides a good foundation in future genomic studies in the superb fairy-wren and more broadly across other bird species.

ACKNOWLEDGEMENTS

We would like to acknowledge the late Sylvain Forêt who guided the design for the sequencing strategy and assembly pipeline of the fairy-wren genome. We thank Kaiman Peng and Angela Higgins from the ACRF Biomolecular Resource Facility at the Australian National University who prepared and sequenced the Illumina mate-pairs and also consulted for the sequencing design; Lutz Froenicke from the DNA Sequencing Facility at UC Davis who guided us through the PacBio sequencing; Andrzej Kilian from the Diversity Arrays Technology who designed and guided us through the sequencing of the pedigree for the linkage map; Takeshi Kawakami and Homa Papoli for advice on linkage mapping and upgrading genome assemblies to the chromosome-scale; Hans Ellegren for access to the collared flycatcher genetic map and Anna Santure, Kees van Oers, and Jon Slate for access to the great tit genetic map; and Ana Catalán for advice on polishing the genome assembly and the BUSCO assessments. Funding was provided by the Australian Research Council (DP150100298), and the office of the DVC (Research) at the Australian National University. Ethics for sampling the pedigree population was provided by the Animal Experimentation Ethics Committee of the Australian National University. Lastly, we thank Mozes Blom, Ana Silva, Sergio Tusso and Claire Peart for discussions and comments on the manuscript and various members of the Moritz and Wolf laboratories for helpful discussions.

AUTHOR CONTRIBUTIONS

J.V.P., L.J., C.M. and A.C. designed and conceived ideas for this project. Y.D. and Q.F. performed the annotation. J.V.P. performed the research, analysed the data, and wrote the paper.

DATA AVAILABILITY STATEMENT

The data that support these data are openly available in NCBI: Accession VKON00000000, NCBI BioProject & SRA Accession PRJNA553115, and DRYAD repository DRYAD https://doi.org/10.5061/dryad.3m7875s.

This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession VKON00000000. The version described in this paper is version VKON01000000.

NCBI BioSample: SAMN12217597.

NCBI BioProject & SRA accession (raw data): PRJNA553115.

Genome annotation (GFF file): DRYAD https://doi.org/10.5061/dryad.3m7875s.

Repeat annotation (REPEATMASKER Output): DRYAD https://doi.org/10.5061/dryad.3m7875s.

Genetic map & recombination landscape data (Sliding window recombination rates, LEPMAP3 pedigree genotype input, *Malurus cyaneus* genetic map): DRYAD https://doi.org/10.5061/dryad.3m7875s. Other chromosome-scale reference genomes on GenBank: (Dalloul et al., 2014; Damas et al., 2016; Damas et al., 2016; Elgvin et al., 2016; Ellegren et al., 2013; International Chicken Genome Sequencing Consortium, 2018; Jarvis, Howard, et al., 2019; Jarvis, Rhei, et al., 2019; Jarvis, Rhie, et al., 2019; Laine et al., 2019; Vignal & Warren, 22017; Warren et al., 2016, 2013; Wellcome Sanger Institute Data Sharing, 2019a; Wellcome Sanger Institute Data Sharing, 2019b; Zhou, Jiang, Liu, Cheng, & Hua, 2018).

ORCID

Joshua V. Peñalba https://orcid.org/0000-0001-6549-6885
Leo Joseph https://orcid.org/0000-0001-7564-1978

REFERENCES

Axelsson, E., Webster, M. T., Smith, N. G. C., Burt, D. W., & Ellegren, H. (2005). Comparison of the chicken and turkey genomes reveals a higher rate of nucleotide divergence on microchromosomes than macrochromosomes. *Genome Research*, *15*(1), 120–125. https://doi.org/10.1101/gr.3021305

Backström, N., Forstmeier, W., Schielzeth, H., Mellenius, H., Nam, K., Bolund, E., … Ellegren, H. (2010). The recombination landscape of the zebra finch *Taeniopygia guttata* genome. *Genome Research*, *20*(4), 485–495. https://doi.org/10.1101/gr.101410.109

Baldassarre, D. T., White, T. A., Karubian, J., & Webster, M. S. (2014). Genomic and morphological analysis of a semipermeable avian hybrid zone suggests asymmetrical introgression of a sexual signal. *Evolution; International Journal of Organic Evolution*, *68*(9), 2644–2657. https://doi.org/10.1111/evo.12457

Barton, N. H. (1995). A general model for the evolution of recombination. *Genetical Research*, *65*(2), 123–145. https://doi.org/10.1017/S0016672300033140

Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., … Stadler, P. F. (2013). MITOS: Improved de novo metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution*, *69*(2), 313–319. https://doi.org/10.1016/j.ympev.2012.08.023

Boland, C. R. J., & Cockburn, A. (2002). Short sketches from the long history of cooperative breeding in Australian birds. *The Emu*, *102*(1), 9–17. https://doi.org/10.1071/MU01039

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bolívar, P., Mugal, C. F., Nater, A., & Ellegren, H. (2016). Recombination rate variation modulates gene sequence evolution mainly via GC-biased gene conversion, not Hill-Robertson interference, in an avian system. *Molecular Biology and Evolution*, *33*(1), 216–227. https://doi.org/10.1093/molbev/msv214

Bourque, G., & Pevzner, P. A. (2002). Genome-scale evolution: Reconstructing gene orders in the ancestral species. *Genome Research*, *12*(1), 26–36.

Brouwer, L., van de Pol, M., Aranzamendi, N. H., Bain, G., Baldassarre, D. T., Brooker, L. C., … Cockburn, A. (2017). Multiple hypotheses explain variation in extra-pair paternity at different levels in a single bird family. *Molecular Ecology*, *26*(23), 6717–6729. https://doi.org/10.1111/mec.14385

Buchanan, K. L., & Cockburn, A. (2013). Fairy-wrens and their relatives (Maluridae) as model organisms in evolutionary ecology: The scientific legacy of Ian Rowley and Eleanor Russell. *Emu - Austral Ornithology*, *113*(3), i–vii. https://doi.org/10.1071/MUv113n3_ED

Burke, T. A., Bruford, M. W., Hanotte, O., & Brookfield, J. F. Y. (1998). Multilocus and single-locus DNA fingerprinting. In A. R. Hoelzel (Ed.), *Molecular genetic analysis of populations* (pp. 287–336). University of Scheffield, Scheffield, UK: IRL Press.

Burri, R., Nater, A., Kawakami, T., Mugal, C. F., Olason, P. I., Smeds, L., … Ellegren, H. (2015). Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of *Ficedula* flycatchers. *Genome Research*, *25*(11), 1656–1665.

Burton, J. N., Adey, A., Patwardhan, R. P., Qiu, R., Kitzman, J. O., & Shendure, J. (2013). Chromosome-scale scaffolding of de novo genome assemblies based on chromatin interactions. *Nature Biotechnology*, *31*(12), 1119–1125. https://doi.org/10.1038/nbt.2727

Butler, J., MacCallum, I., Kleber, M., Shlyakhter, I. A., Belmonte, M. K., Lander, E. S., … Jaffe, D. B. (2008). ALLPATHS: De novo assembly of whole-genome shotgun microreads. *Genome Research*, *18*(5), 810–820. https://doi.org/10.1101/gr.7337908

Charlesworth, B., Coyne, J. A., & Barton, N. H. (1987). The relative rates of evolution of sex chromosomes and autosomes. *The American Naturalist*, *130*(1), 113–146. https://doi.org/10.1086/284701

Clément, Y., & Arndt, P. F. (2013). Meiotic recombination strongly influences GC-content evolution in short regions in the mouse genome. *Molecular Biology and Evolution*, *30*(12), 2612–2618. https://doi.org/10.1093/molbev/mst154

Cockburn, A., Brouwer, L., Double, M. C., Margraf, N., & van de Pol, M. (2013). Evolutionary origins and persistence of infidelity in *Malurus*: The least faithful birds. *The Emu*, *113*(3), 208–217.

Cockburn, A., Brouwer, L., Margraf, N., Osmond, H. L., Van de Pol, M., Koening, W. D., & Dickinson, J. L. (2016). *Superb fairy-wrens: Making the worst of a good job*. Cambridge, UK: Cambridge University Press.

Cockburn, A., Dalziell, A. H., Blackmore, C. J., Double, M. C., Kokko, H., Osmond, H. L., … Wells, K. (2009). Superb fairy-wren males aggregate into hidden leks to solicit extragroup fertilizations before dawn. *Behavioral Ecology: Official Journal of the International Society for Behavioral Ecology*, *20*(3), 501–510. https://doi.org/10.1093/beheco/arp024

Cockburn, A., Osmond, H. L., & Double, M. C. (2008). Swingin'in the rain: Condition dependence and sexual selection in a capricious world. *Proceedings of the Royal Society B: Biological Sciences*, *275*(1635), 605–612.

Cockburn, A., Osmond, H. L., Mulder, R. A., Green, D. J., & Double, M. C. (2003). Divorce, dispersal and incest avoidance in the cooperatively breeding superb fairy-wren *Malurus cyaneus*. *The Journal of Animal Ecology*, *72*(2), 189–202. https://doi.org/10.1046/j.1365-2656.2003.00694.x

Cockburn, A., Sims, R. A., Osmond, H. L., Green, D. J., Double, M. C., & Mulder, R. A. (2008). Can we measure the benefits of help in cooperatively breeding birds: The case of superb fairy-wrens *Malurus cyaneus*? *The Journal of Animal Ecology*, *77*(3), 430–438. https://doi.org/10.1111/j.1365-2656.2007.01351.x

Colombelli-Négrel, D., Hauber, M. E., Robertson, J., Sulloway, F. J., Hoi, H., Griggio, M., & Kleindorfer, S. (2012). Embryonic learning of vocal passwords in superb fairy-wrens reveals intruder cuckoo nestlings. *Current Biology*, *22*(22), 2155–2160. https://doi.org/10.1016/j.cub.2012.09.025

Comeron, J. M. (2017). Background selection as null hypothesis in population genomics: Insights and challenges from Drosophila studies. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *372*(1736), 20160471.

Connallon, T., Olito, C., Dutoit, L., Papoli, H., Ruzicka, F., & Yong, L. (2018). Local adaptation and the evolution of inversions on sex chromosomes and autosomes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1757), 20170423. https://doi.org/10.1098/rstb.2017.0423

Dalloul, R. A., Long, J. A., Zimin, A. V., Aslam, L., Beal, K., Le Blomberg, A., …Reed, K. M. (2014). Meleagris gallopavo GenBank Assemby; version 5.1 (Turkey_5.1); GenBank Accession GCA_000146605.4 [data set].

Damas, J., O'Connor, R., Farre, M., Lenis, V. P., Martell, H., Mandawala, A., … Larkin, D. M. (2016). Falco peregrinus GenBank Assembly; version 2 (falPer2); GenBank Accession GCA_001887755.1. [data set].

Damas, J., O'Connor, R., Farre, M., Lenis, V. P., Martell, H., Mandawala, A., …Larkin, D. M. (2016). Columba livia GenBank Accession; version 2 (colLiv2); GenBank Accession GCA_001887795.1 [data set].

Damas, J., O'Connor, R., Farré, M., Lenis, V. P. E., Martell, H. J., Mandawala, A., … Larkin, D. M. (2017). Upgrading short-read animal genome assemblies to chromosome level using comparative genomics and a universal probe set. *Genome Research*, *27*(5), 875–884. https://doi.org/10.1101/gr.213660.116

Dapper, A. L., & Payseur, B. A. (2017). Connecting theory and data to understand recombination rate evolution. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *372*(1736). https://doi.org/10.1098/rstb.2016.0469

Darling, A. E., Mau, B., & Perna, N. T. (2010). progressiveMauve: Multiple genome alignment with gene gain, loss and rearrangement. *PLoS ONE*, *5*(6), e11147. https://doi.org/10.1371/journal.pone.0011147

Double, M., & Cockburn, A. (2000). Pre-dawn infidelity: Females control extra-pair mating in superb fairy-wrens. *Proceedings. Biological Sciences/The Royal Society*, *267*(1442), 465–470.

Dowling, J., & Webster, M. S. (2016). An experimental test of duet function in a fairy-wren (*Malurus*) with moderate cuckoldry rates. *Behavioral Ecology: Official Journal of the International Society for Behavioral Ecology*, *27*(1), 228–236.

Driskell, A. C., Norman, J. A., Pruett-Jones, S., Mangall, E., Sonsthagen, S., & Christidis, L. (2011). A multigene phylogeny examining evolutionary and ecological relationships in the Australo-papuan wrens of the subfamily Malurinae (Aves). *Molecular Phylogenetics and Evolution*, *60*(3), 480–485. https://doi.org/10.1016/j.ympev.2011.03.030

Dudchenko, O., Batra, S. S., Omer, A. D., Nyquist, S. K., Hoeger, M., Durand, N. C., … Aiden, E. L. (2017). De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. *Science*, *356*(6333), 92–95.

Dunn, P. O., & Cockburn, A. (1999). Extrapair mate choice and honest signaling in cooperatively breeding superb fairy-wrens. *Evolution; International Journal of Organic Evolution*, *53*(3), 938–946. https://doi.org/10.1111/j.1558-5646.1999.tb05387.x

Dunn, P. O., Cockburn, A., & Mulder, R. A. (1995). Fairy-wren helpers often care for young to which they are unrelated. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *259*, 339–343.

Elgvin, T. O., Trier, C. N., Toerresen, O. K., Hagen, I. J., Lien, S., Nederbragt, A. J., …Saetre, G.-P.(2016). Passer domesticus GenBank Assembly; version 1.0 (Passer_domesticus-1.0); GenBank Accession GCA_001700915.1 [data set].

Ellegren, H. (2010). Evolutionary stasis: The stable chromosomes of birds. *Trends in Ecology & Evolution*, *25*(5), 283–291. https://doi.org/10.1016/j.tree.2009.12.004

Ellegren, H., Smeds, L., Burri, R., Olason, P. I., Backstrom, N., Kawakami, T., … Wolf, J. B. (2013). Ficedula albicollis GenBank Assembly; version 1.5 (FicAlb1.5); GenBank Accession GCA_000247815.2 [data set].

English, A. C., Richards, S., Han, Y. I., Wang, M., Vee, V., Qu, J., … Gibbs, R. A. (2012). Mind the gap: Upgrading genomes with Pacific Biosciences RS long-read sequencing technology. *PLoS ONE*, *7*(11), e47768. https://doi.org/10.1371/journal.pone.0047768

Fierst, J. L. (2015). Using linkage maps to correct and scaffold de novo genome assemblies: Methods, challenges, and computational tools. *Frontiers in Genetics*, *6*, 220. https://doi.org/10.3389/fgene.2015.00220

Fullerton, S. M., Bernardo Carvalho, A., & Clark, A. G. (2001). Local rates of recombination are positively correlated with GC content in the human genome. *Molecular Biology and Evolution*, *18*(6), 1139–1142. https://doi.org/10.1093/oxfordjournals.molbev.a003886

Greig, E., & Pruett-Jones, S. (2008). Splendid songs: The vocal behaviour of Splendid Fairy-wrens (*Malurus splendens melanotus*). *Emu - Austral Ornithology*, *108*(2), 103–114.

Groenen, M. A., Wahlberg, P., Foglio, M., Cheng, H. H., Megens, H.-J., Crooijmans, R. P., … Andersson, L. (2009). A high-density SNP-based linkage map of the chicken genome reveals sequence features correlated with recombination rate. *Genome Research*, *19*(3), 510–519. https://doi.org/10.1101/gr.086538.108

Guerrero, R. F., & Kirkpatrick, M. (2014). Local adaptation and the evolution of chromosome fusions. *Evolution; International Journal of Organic Evolution*, *68*(10), 2747–2756. https://doi.org/10.1111/evo.12481

Haenel, Q., Laurentino, T. G., Roesti, M., & Berner, D. (2018). Meta-analysis of chromosome-scale crossover rate variation in eukaryotes and its significance to evolutionary genomics. *Molecular Ecology*, *27*(11), 2477–2497. https://doi.org/10.1111/mec.14699

Hahn, C., Bachmann, L., & Chevreux, B. (2013). Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads–a baiting and iterative mapping approach. *Nucleic Acids Research*, *41*(13), e129. https://doi.org/10.1093/nar/gkt371

Hajduk, G. K., Cockburn, A., Margraf, N., Osmond, H. L., Walling, C. A., & Kruuk, L. E. B. (2018). Inbreeding, inbreeding depression, and infidelity in a cooperatively breeding bird. *Evolution; International Journal of Organic Evolution*, *72*(7), 1500–1514. https://doi.org/10.1111/evo.13496

Hansmann, T., Nanda, I., Volobouev, V., Yang, F., Schartl, M., Haaf, T., & Schmid, M. (2009). Cross-species chromosome painting corroborates microchromosome fusion during karyotype evolution of birds. *Cytogenetic and Genome Research*, *126*(3), 281–304. https://doi.org/10.1159/000251965

Hooper, D. M., & Price, T. D. (2015). Rates of karyotypic evolution in Estrildid finches differ between island and continental clades. *Evolution*, *69*(4), 890–903. https://doi.org/10.1111/evo.12633

Hooper, D. M., & Price, T. D. (2017). Chromosomal inversion differences correlate with range overlap in passerine birds. *Nature Ecology & Evolution*, *1*(10), 1526–1534. https://doi.org/10.1038/s41559-017-0284-6

International Chicken Genome Sequencing Consortium. (2018). Gallus gallus GenBank assembly; NCBI; version 6a (GRCg6a); GenkBank Accession GCA_000002315.5 [data set].

Jarvis, E. D., Howard, J., Rhie, A., Phillippy, A., Korlach, J., Digby, A., …Fedrigo, O. (2019). Strigops habroptila GenBank assembly; version 1.p (bStrHab1_v1.p); GenBank Accession GCA_004027225.1 [data set].

Jarvis, E. D., Rhei, A., Korlach, J., Fedrigo, O., Koren, S., Howe, K., …Formenti, G. (2019). Calype anna GenBank assembly; version 1.p (bCalAnn1_v1.p); GenBank Accession GCA_003957555.2 [data set].

Jarvis, E. D., Rhie, A., Korlach, J., Fedrigo, O., Mountcastle, J., Koren, S., …Mello, C. V. (2019). Taeniopygia guttata GenBank assembly; version 1.p (bTaeGut1_v1.p); GenBank Accession GCA_003957565.2 [data set].

Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K., & Mooers, A. O. (2012). The global diversity of birds in space and time. *Nature*, *491*(7424), 444–448.

Jiao, W.-B., Accinelli, G. G., Hartwig, B., Kiefer, C., Baker, D., Severing, E., … Schneeberger, K. (2017). Improving and correcting the contiguity of long-read genome assemblies of three plant species using optical mapping and chromosome conformation capture data. *Genome Research*, *27*(5), 778–786. https://doi.org/10.1101/gr.213652.116

Joseph, S., O'Connor, R. E., Al Mutery, A. F., Watson, M., Larkin, D. M., & Griffin, D. K. (2018). Chromosome level genome assembly and comparative genomics between three falcon species reveals an unusual pattern of genome organisation. *Diversity*, *10*(4), 113. https://doi.org/10.3390/d10040113

Kapusta, A., & Suh, A. (2017). Evolution of bird genomes—a transposon's-eye view. *Annals of the New York Academy of Sciences*, *1389*(1), 164–185.

Karubian, J. (2008). Changes in breeding status are associated with rapid bill darkening in male red-backed fairy-wrens *Malurus melanocephalus*. *Journal of Avian Biology*, *39*(1), 81–86. https://doi.org/10.1111/j.0908-8857.2008.04161.x

Karubian, J. (2013). Female ornamentation in *Malurus* fairy-wrens: A hidden evolutionary gem for understanding female perspectives on social and sexual selection. *Emu - Austral Ornithology*, *113*(3), 248–258.

Kawakami, T., Smeds, L., Backström, N., Husby, A., Qvarnström, A., Mugal, C. F., … Ellegren, H. (2014). A high-density linkage map enables a second-generation collared flycatcher genome assembly and reveals the patterns of avian recombination rate variation and chromosomal evolution. *Molecular Ecology*, *23*(16), 4035–4058. https://doi.org/10.1111/mec.12810

Kearns, A. M., Joseph, L., Edwards, S. V., & Double, M. C. (2009). Inferring the phylogeography and evolutionary history of the splendid fairy-wren *Malurus splendens* from mitochondrial DNA and spectrophotometry. *Journal of Avian Biology*, *40*(1), 7–17.

Kilian, A., Wenzl, P., Huttner, E., Carling, J., Xia, L., Blois, H., … Uszynski, G. (2012). Diversity arrays technology: A generic genome profiling technology on open platforms. *Methods in Molecular Biology*, *888*, 67–89.

Kingan, S., Heaton, H., Cudini, J., Lambert, C., Baybayan, P., Galvin, B., … Lawniczak, M. (2019). A high-quality de novo genome assembly from a single mosquito using PacBio sequencing. *Genes*, *10*, 62. https://doi.org/10.3390/genes10010062

Kirkpatrick, M. (2010). How and why chromosome inversions evolve. *PLoS Biology*, *8*(9), e1000501. https://doi.org/10.1371/journal.pbio.1000501

Kruuk, L. E. B., Osmond, H. L., & Cockburn, A. (2015). Contrasting effects of climate on juvenile body size in a Southern Hemisphere passerine bird. *Global Change Biology*, *21*(8), 2929–2941. https://doi.org/10.1111/gcb.12926

Laine, V. N., Gossmann, T. I., Schachtschneider, K. M., Garroway, C. J., Madsen, O., Verhoeven, K. J., …Groenen, M. A. (2019). Parus major GenBank Assembly; version 1.1 (Parus_major1.1); GenBank Accession GCA_001522545.3 [data set].

Langmore, N. E., Hunt, S., & Kilner, R. M. (2003). Escalation of a coevolutionary arms race through host rejection of brood parasitic young. *Nature*, *422*(6928), 157–160.

Lehmann, R., Lightfoot, D. J., Schunter, C., Michell, C. T., Ohyanagi, H., Mineta, K., … Ravasi, T. (2019). Finding Nemo's Genes: A chromosome-scale reference assembly of the genome of the orange clownfish *Amphiprion percula*. *Molecular Ecology Resources*, *19*(3), 570–585.

Leitão, A. V., Hall, M. L., Venables, B., & Mulder, R. A. (2019). Ecology and breeding biology of a tropical bird, the Lovely Fairy-Wren (*Malurus amabilis*). *Emu - Austral Ornithology*, *119*(1), 1–13.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Lindsay, W. R., Webster, M. S., & Schwabl, H. (2011). Sexually selected male plumage color is testosterone dependent in a tropical passerine bird, the red-backed fairy-wren (*Malurus melanocephalus*). *PLoS ONE*, *6*(10), e26067. https://doi.org/10.1371/journal.pone.0026067

Magrath, R. D., & Bennett, T. H. (2012). A micro-geography of fear: Learning to eavesdrop on alarm calls of neighbouring heterospecifics. *Proceedings. Biological Sciences/The Royal Society*, *279*(1730), 902–909.

Magrath, R. D., Haff, T. M., McLachlan, J. R., & Igic, B. (2015). Wild birds learn to eavesdrop on heterospecific alarm calls. *Current Biology*, *25*(15), 2047–2050. https://doi.org/10.1016/j.cub.2015.06.028

Marki, P. Z., Jønsson, K. A., Irestedt, M., Nguyen, J. M. T., Rahbek, C., & Fjeldså, J. (2017). Supermatrix phylogeny and biogeography of the

Australasian Meliphagides radiation (Aves: Passeriformes). *Molecular Phylogenetics and Evolution*, *107*, 516–529. https://doi.org/10.1016/j.ympev.2016.12.021

McLean, A. J., Toon, A., Schmidt, D. J., Joseph, L., & Hughes, J. M. (2012). Speciation in chestnut-shouldered fairy-wrens (*Malurus* spp.) and rapid phenotypic divergence in variegated fairy-wrens (*Malurus lamberti*): A multilocus approach. *Molecular Phylogenetics and Evolution*, *63*(3), 668–678.

Meyer, M., & Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols*, *2010*(6), db.prot5448. https://doi.org/10.1101/pdb.prot5448

Michael, T. P., Jupe, F., Bemm, F., Motley, S. T., Sandoval, J. P., Lanz, C., … Ecker, J. R. (2018). High contiguity *Arabidopsis thaliana* genome assembly with a single nanopore flow cell. *Nature Communications*, *9*(1), 541. https://doi.org/10.1038/s41467-018-03016-2

Mulder, R. A., Dunn, P. O., Cockburn, A., Cohen, K. A. L., & Howell, M. J. (1994). Helpers liberate female fairy-wrens from constraints on extra-pair mate choice. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *255*(1344), 223–229.

Murphy, S. A., Legge, S. M., Heathcote, J., & Mulder, E. (2010). The effects of early and late-season fires on mortality, dispersal, physiology and breeding of red-backed fairy-wrens (*Malurus melanocephalus*). *Wildlife Research*, *37*(2), 145–155. https://doi.org/10.1071/WR09007

Nishida, C., Ishijima, J., Kosaka, A., Tanabe, H., Habermann, F. A., Griffin, D. K., & Matsuda, Y. (2008). Characterization of chromosome structures of Falconinae (Falconidae, Falconiformes, Aves) by chromosome painting and delineation of chromosome rearrangements during their differentiation. *Chromosome Research*, *16*(1), 171–181. https://doi.org/10.1007/s10577-007-1210-6

O'Connor, R. E., Farré, M., Joseph, S., Damas, J., Kiazim, L., Jennings, R., … Griffin, D. K. (2018). Chromosome-level assembly reveals extensive rearrangement in saker falcon and budgerigar, but not ostrich, genomes. *Genome Biology*, *19*(1), 171. https://doi.org/10.1186/s13059-018-1550-x

Oliveros, C. H., Field, D. J., Ksepka, D. T., Barker, F. K., Aleixo, A., Andersen, M. J., … Faircloth, B. C. (2019). Earth history and the passerine superradiation. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(16), 7916–7925. https://doi.org/10.1073/pnas.1813206116

Ortiz-Barrientos, D., Engelstädter, J., & Rieseberg, L. H. (2016). Recombination rate evolution and the origin of species. *Trends in Ecology & Evolution*, *31*(3), 226–236. https://doi.org/10.1016/j.tree.2015.12.016

Paape, T., Zhou, P., Branca, A., Briskine, R., Young, N., & Tiffin, P. (2012). Fine-scale population recombination rates, hotspots, and correlates of recombination in the *Medicago truncatula* genome. *Genome Biology and Evolution*, *4*(5), 726–737. https://doi.org/10.1093/gbe/evs046

Peñalba, J. V., Joseph, L., & Moritz, C. (2019). Current geography masks dynamic history of gene flow during speciation in northern Australian birds. *Molecular Ecology*, *28*(3), 630–643. https://doi.org/10.1111/mec.14978

Peters, A. (2000). Testosterone treatment is immunosuppressive in superb fairy-wrens, yet free-living males with high testosterone are more immunocompetent. *Proceedings. Biological Sciences/The Royal Society*, *267*(1446), 883–889.

Potvin, D. A., Ratnayake, C. P., Radford, A. N., & Magrath, R. D. (2018). Birds learn socially to recognize heterospecific alarm calls by acoustic association. *Current Biology*, *28*(16), 2632–2637.e4. https://doi.org/10.1016/j.cub.2018.06.013

Pruett-Jones, S. G., & Lewis, M. J. (1990). Sex ratio and habitat limitation promote delayed dispersal in superb fairy-wrens. *Nature*, *348*(6301), 541–542.

Rastas, P. (2017). Lep-MAP3: Robust linkage mapping even for low-coverage whole genome sequencing data. *Bioinformatics*, *33*(23), 3726–3732. https://doi.org/10.1093/bioinformatics/btx494

Rowley, I., & Russell, E. M. (1997). *Fairy-wrens and grasswrens: Maluridae*. Oxford, UK: Oxford University Press.

Russell, A. F., Langmore, N. E., Cockburn, A., Astheimer, L. B., & Kilner, R. M. (2007). Reduced egg investment can conceal helper effects in cooperatively breeding birds. *Science*, *317*(5840), 941–944.

Salmela, L., & Rivals, E. (2014). LoRDEC: Accurate and efficient long read error correction. *Bioinformatics*, *30*(24), 3506–3514. https://doi.org/10.1093/bioinformatics/btu538

Sävilammi, T., Primmer, C. R., Varadharajan, S., Guyomard, R., Guiguen, Y., Sandve, S. R., … Lien, S. (2019). The chromosome-level Genome Assembly of European Grayling Reveals Aspects of a Unique Genome Evolution Process within Salmonids. *G3*, *9*(5), 1283–1294. https://doi.org/10.1534/g3.118.200919

Schumer, M., Xu, C., Powell, D. L., Durvasula, A., Skov, L., Holland, C., … Przeworski, M. (2018). Natural selection interacts with recombination to shape the evolution of hybrid genomes. *Science*, *360*(6389), 656–660.

Schwander, T., Libbrecht, R., & Keller, L. (2014). Supergenes and complex phenotypes. *Current Biology*, *24*(7), R288–R294.

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, *31*(19), 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Simpson, J. T. (2014). Exploring genome characteristics and sequence quality without a reference. *Bioinformatics*, *30*(9), 1228–1235. https://doi.org/10.1093/bioinformatics/btu023

Skinner, B. M., & Griffin, D. K. (2012). Intrachromosomal rearrangements in avian genome evolution: Evidence for regions prone to breakpoints. *Heredity*, *108*(1), 37–41. https://doi.org/10.1038/hdy.2011.99

Skroblin, A., Lanfear, R., Cockburn, A., & Legge, S. (2012). Inferring population connectivity across the range of the purple-crowned fairy-wren (*Malurus coronatus*) from mitochondrial DNA and morphology: Implications for conservation management. *Australian Journal of Zoology*, *60*(3), 199–209. https://doi.org/10.1071/ZO12093

Stapley, J., Birkhead, T. R., Burke, T., & Slate, J. (2008). A linkage map of the zebra finch *Taeniopygia guttata* provides new insights into avian genome evolution. *Genetics*, *179*(1), 651–667.

Stapley, J., Feulner, P. G. D., Johnston, S. E., Santure, A. W., & Smadja, C. M. (2017). Variation in recombination frequency and distribution across eukaryotes: Patterns and processes. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *372*(1736), 20160455. https://doi.org/10.1098/rstb.2016.0455

Stiller, J., & Zhang, G. (2019). Comparative phylogenomics, a stepping stone for bird biodiversity studies. *Diversity*, *11*(7), 115. https://doi.org/10.3390/d11070115

Su, C., Wang, W., Gong, S., Zuo, J., Li, S., & Xu, S. (2017). High density linkage map construction and mapping of yield trait QTLs in maize (*Zea mays*) using the genotyping-by-sequencing (gbs) technology. *Frontiers in Plant Science*, *8*, 706. https://doi.org/10.3389/fpls.2017.00706

Tang, H., Zhang, X., Miao, C., Zhang, J., Ming, R., Schnable, J. C., … Lu, J. (2015). ALLMAPS: Robust scaffold ordering based on multiple maps. *Genome Biology*, *16*(1), 3. https://doi.org/10.1186/s13059-014-0573-1

Tesler, G. (2002). GRIMM: Genome rearrangements web server. *Bioinformatics*, *18*(3), 492–493. https://doi.org/10.1093/bioinformatics/18.3.492

Thind, A. K., Wicker, T., Müller, T., Ackermann, P. M., Steuernagel, B., Wulff, B. B. H., … Krattinger, S. G. (2018). Chromosome-scale comparative sequence analysis unravels molecular mechanisms of genome dynamics between two wheat cultivars. *Genome Biology*, *19*(1), 104. https://doi.org/10.1186/s13059-018-1477-2

van de Pol, M., Osmond, H. L., & Cockburn, A. (2012). Fluctuations in population composition dampen the impact of phenotypic plasticity on trait dynamics in superb fairy-wrens. *The Journal of Animal Ecology*, *81*(2), 411–422. https://doi.org/10.1111/j.1365-2656.2011.01919.x

van Oers, K., Santure, A. W., De Cauwer, I., van Bers, N. E., Crooijmans, R. P., Sheldon, B. C., … Groenen, M. A. (2014). Replicated high-density genetic maps of two great tit populations reveal fine-scale genomic departures from sex-equal recombination rates. *Heredity*, *112*(3), 307–316. https://doi.org/10.1038/hdy.2013.107

Varian-Ramos, C. W., & Webster, M. S. (2012). Extrapair copulations reduce inbreeding for female red-backed fairy-wrens, *Malurus melanocephalus*. *Animal Behaviour*, *83*(3), 857–864. https://doi.org/10.1016/j.anbehav.2012.01.010

Vignal, A., & Warren, W. (2017).Numida Meleagris GenBank Assembly; version 1.0 (NumMel1.0); GenBank Accession GCA_002078875.2 [data set].

Vijay, N., Bossu, C. M., Poelstra, J. W., Weissensteiner, M. H., Suh, A., Kryukov, A. P., & Wolf, J. B. W. (2016). Evolution of heterogeneous genome differentiation across multiple contact zones in a crow species complex. *Nature Communications*, *7*, 13195. https://doi.org/10.1038/ncomms13195

Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., … Earl, A. M. (2014). Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE*, *9*(11), e112963. https://doi.org/10.1371/journal.pone.0112963

Wang, J., Street, N. R., Scofield, D. G., & Ingvarsson, P. K. (2016). Natural selection and recombination rate variation shape nucleotide polymorphism across the genomes of three related populus species. *Genetics*, *202*(3), 1185–1200. https://doi.org/10.1534/genetics.115.183152

Warren, W., Burt, D. W., Antin, P. B., Lanford, R., Gros, J., & Wilson, R. K. (2016). Coturnix japonica GenBank assembly; version 2.0 (Coturnix japonics 2.0); GenBank Accession GCA_001577835.1.

Warren, W. C., Clayton, D. F., Ellegren, H., Arnold, A. P., Hillier, L. W., Kunstner, A., …Wilson, R. K. (2013). Taeniopygia guttata GenBank Assembly; version 3.2.4 (Taeniopygia_guttata-3.2.4); GenBank Accession GCA_000151805.2 [data set].

Wellband, K., Mérot, C., Linnansaari, T., Elliott, J. A. K., Curry, R. A., & Bernatchez, L. (2019). Chromosomal fusion and life history-associated genomic variation contribute to within-river local adaptation of Atlantic salmon. *Molecular Ecology*, *28*(6), 1439–1459. https://doi.org/10.1111/mec.14965

Wellcome Sanger Institute Data Sharing.(2019a). Aquila chrysaetos GenBank assembly; NCBI; version 1.2 (bAquChr1.2); GenBank Accession GCA_900496995.2 [data set].

Wellcome Sanger Institute Data Sharing.(2019b). Streptopelia turtur GenBank assembly; version 1.1 (bStrTur1.1); GenBank Accession GCA_901699155.1 [data set].

Yandell, D. D., Hochachka, W. M., Pruett-Jones, S., Webster, M. S., & Greig, E. I. (2018). Geographic patterns of song variation in four species of *Malurus* fairy-wrens. *Journal of Avian Biology*, *49*(2), Jav – 01446.

Zhang, G., Jarvis, E. D., & Gilbert, M. T. P. (2014). A flock of genomes. *Science*, *346*(6215), 1308–1309.

Zhang, G., Rahbek, C., Graves, G. R., Lei, F., Jarvis, E. D., & Gilbert, M. T. P. (2015). Genomics: Bird sequencing project takes off. *Nature*, *522*(7554), 34. https://doi.org/10.1038/522034d

Zhou, Z., Jiang, Y., Liu, Z., Cheng, H., & Hua, P. (2018).Anas platyrhynchos GenBank Assembly; version 1.5 (IASCAAS_PekingDuck_PBH1.5); GenBank Accession GCA_003850225.1 [data set].

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Peñalba JV, Deng Y, Fang Q, Joseph L, Moritz C, Cockburn A. Genome of an iconic Australian bird: High-quality assembly and linkage map of the superb fairy-wren (*Malurus cyaneus*). *Mol Ecol Resour.* 2020;20:560–578. https://doi.org/10.1111/1755-0998.13124