Received: 11 August 2019 | Revised: 27 December 2019 | Accepted: 2 January 2020

DOI: 10.1111/eva.12916

ORIGINAL ARTICLE

Evolutionary Applications Open Access WILEY

A comparison of pedigree, genetic and genomic estimates of relatedness for informing pairing decisions in two critically endangered birds: Implications for conservation breeding programmes worldwide

Stephanie J. Galla[1] | Roger Moraga[2] | Liz Brown[3] | Simone Cleland[3] | Marc P. Hoeppner[4] | Richard F. Maloney[5] | Anne Richardson[6] | Lyndon Slater[7] | Anna W. Santure[8] | Tammy E. Steeves[1]

[1]School of Biological Sciences, University of Canterbury, Christchurch, New Zealand

[2]Tea Break Bioinformatics, Ltd, Palmerston North, New Zealand

[3]New Zealand Department of Conservation, Twizel, New Zealand

[4]Institute for Clinical Molecular Biology, Christian-Albrechts-University Kiel, Kiel, Germany

[5]New Zealand Department of Conservation, Christchurch, New Zealand

[6]The Isaac Conservation and Wildlife Trust, Christchurch, New Zealand

[7]New Zealand Department of Conservation, Rangiora, New Zealand

[8]School of Biological Sciences, University of Auckland, Auckland, New Zealand

Correspondence
Stephanie J. Galla, School of Biological Sciences, University of Canterbury, Christchurch, New Zealand.
Email: sgalla32@gmail.com

Funding information
Mohua Charitable Trust; Brian Mason Scientific and Technical Trust; Ministry for Business Innovation and Employment, Grant/Award Number: UOCX1602

Abstract

Conservation management strategies for many highly threatened species include conservation breeding to prevent extinction and enhance recovery. Pairing decisions for these conservation breeding programmes can be informed by pedigree data to minimize relatedness between individuals in an effort to avoid inbreeding, maximize diversity and maintain evolutionary potential. However, conservation breeding programmes struggle to use this approach when pedigrees are shallow or incomplete. While genetic data (i.e., microsatellites) can be used to estimate relatedness to inform pairing decisions, emerging evidence indicates this approach may lack precision in genetically depauperate species, and more effective estimates will likely be obtained from genomic data (i.e., thousands of genome-wide single nucleotide polymorphisms, or SNPs). Here, we compare relatedness estimates and subsequent pairing decisions using pedigrees, microsatellites and SNPs from whole-genome resequencing approaches in two critically endangered birds endemic to New Zealand: kakī/black stilt (*Himantopus novaezelandiae*) and kākāriki karaka/orange-fronted parakeet (*Cyanoramphus malherbi*). Our findings indicate that SNPs provide more precise estimates of relatedness than microsatellites when assessing empirical parent–offspring and full sibling relationships. Further, our results show that relatedness estimates and subsequent pairing recommendations using *PMx* are most similar between pedigree- and SNP-based approaches. These combined results indicate that in lieu of robust pedigrees, SNPs are an effective tool for informing pairing decisions, which has important implications for many poorly pedigreed conservation breeding programmes worldwide.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2020 The Authors. *Evolutionary Applications* published by John Wiley & Sons Ltd.

KEYWORDS
conservation breeding, conservation genetics, conservation genomics, pairing recommendations, PMx, relatedness

1 | INTRODUCTION

In order to recover the world's rarest species, a multifaceted approach is needed to address the factors that cause species decline and those that promote species resilience (Grueber et al., 2019; Jamieson 2015; Soulé, 1985). A critical facet of threatened species recovery is genetic management (Frankham, 2005; O'Grady et al., 2006; Spielman, Brook, & Frankham, 2004), including conservation breeding, where individuals in intensively managed populations are paired to minimize inbreeding and maximize genetic diversity in an effort to maintain evolutionary potential (Ballou & Lacy, 1995; Ballou et al., 2010; de Villemereuil et al., 2019; Giglio, Ivy, Jones, & Latch, 2016; Ivy & Lacy, 2012). In these conservation breeding programmes, offspring may remain in captivity as an insurance population (e.g., the Tasmanian devil, *Sarcophilius harissii*, Hogg et al. 2015) or they may be translocated to the wild (e.g., California condor, *Gymnogyps californianus*; Walters et al., 2010) to enhance recovery efforts. In addition to demographic considerations (e.g. Moore, Converse, Folk, Runge, & Nesbitt, 2012; Slotta-Bachmayr, Boegel, Kaczensky, Stauffer, & Walzer, 2004; Tenhumberg, Tyre, Shea, & Possingham, 2004), current best practice for making pairing decisions in conservation breeding programmes is to use available ancestry data from multigenerational pedigrees to estimate kinship—a metric of pairwise coancestry or relatedness—between all living individuals in a population (Ballou & Lacy, 1995; Ballou et al., 2010; Ivy & Lacy, 2012; Lacy, 1995). Individuals are paired to minimize mean kinship (i.e., average pairwise relatedness among all others in the population, including oneself), which has been shown to maximize founder representation and minimize inbreeding over time (Ballou & Lacy, 1995; Lacy, 2012; Willoughby et al., 2015).

Pedigrees are the tool of choice for estimating relatedness in conservation breeding programmes, including hundreds managed by the worldwide zoo and aquarium community (e.g., the Association of Zoos and Aquariums, or AZA; Hammerly, de la Cerda, Bailey, & Johnson, 2016; Jiménez-Mena, Schad, Hanna, & Lacy, 2016; Long, Dorsey, & Boyle, 2011). Still, there are inherent assumptions that, when violated, hinder pedigree accuracy. For example, pedigrees assume no variance in founder relationships (i.e., all founders are equally unrelated; Ballou, 1983), which is unlikely for many highly threatened wild populations, as most have experienced one or more historical population bottlenecks and founders sourced from these remnant wild populations will have variance in relatedness values (Bergner, Jamieson, & Robertson, 2014; Hogg et al., 2019). Simulation studies suggest that when kinship-based approaches are used for pairing, complete pedigrees with substantial depth (>5 generations recorded) are robust enough to reflect true relatedness and inbreeding estimates despite violating this assumption (Balloux, Amos, & Coulson, 2004; Pemberton, 2004; Rudnick & Lacy, 2008). However, in many conservation breeding programmes, wild founders are routinely sourced to supplement captive populations (e.g., kākāriki karaka, *Cyanoramphus malherbi*, this manuscript) and to reduce the risk associated with adaptation to captivity (Frankham, 2008). Under these circumstances, the assumption of no variance in founder relationships can be repeatedly violated, leading to significant underestimation of relatedness and inbreeding (Hogg et al., 2019). In addition to these caveats, many intensively managed populations are poorly pedigreed, meaning these pedigrees contain missing information (i.e., unknown parentage due to matings that include unidentified individuals or extra-pair parentage; Bérénos, Ellis, Pilkington, & Pemberton, 2014; Lacy, 2009; Pemberton, 2008; Putnam & Ivy, 2014) or record-keeping errors (e.g., Hammerly et al., 2016).

Even when pedigrees are of high depth, have no missing information, and contain no errors, expected relatedness between individuals can differ from realized relatedness, as pedigrees rely on probabilities as opposed to empirical estimates of genome sharing (Hill & Weir, 2012; Kardos, Luikart, & Allendorf, 2015; Speed & Balding, 2015; Willoughby et al., 2015). Based on Mendelian inheritance, pedigrees estimate the probability that two alleles, one chosen at random from each of two individuals, are identical by descent (IBD) from a common ancestor (Ballou, 1983; Lacy, 1995). When using a pedigree, the relatedness coefficient (R) for parents and offspring, as well as for full siblings, is 0.5 when inbreeding is not present, indicating each pair shares 50% of their genomic information. While parents do contribute 50% of their autosomal genomic information to their gametes, the combined effects of recombination, independent assortment and random fertilization can lead to a larger range of realized relatedness between full siblings (Hill & Weir, 2011, 2012; Speed & Balding, 2015). For example, a simulation study in humans revealed that realized relatedness between full siblings could range anywhere from 0.37 to 0.61 (Visscher et al., 2006); however, this range can vary depending on the genome architecture of the species in question (e.g., number and size of chromosomes and the frequency and location of recombination events; Hill & Weir, 2011; Kardos et al., 2015; Knief, Kempenaers, & Forstmeier, 2017).

An alternative approach for populations lacking robust pedigrees is to use genetic-based estimates of pairwise relatedness to inform pairing decisions (Attard et al., 2016; Pemberton, 2004, 2008; Premachandra, Nguyen, & Knibb, 2019; Slate et al., 2004). This approach typically uses 8–30 microsatellite markers and empirical allele frequencies to estimate the probability that two alleles chosen at random from two individuals are IBD from a common ancestor (Speed & Balding, 2015). To date, numerous conservation breeding programmes have used a genetic approach to inform pairing recommendations, repair studbooks and resolve unknown parentage assignments, including programmes for the near-threatened Parma wallaby (*Macropus parma*;

Ivy, Miller, Lacy, & DeWoody, 2009), the vulnerable Jamaican yellow boa (*Epicrates subflavus*; Tzika, Remy, Gibson, & Milinkovitch, 2009), the critically endangered Anegada iguana (*Cyclura pinguis*; Mitchell et al., 2011) and the critically endangered Attwater's prairie-chicken (*Tympanuchus cupido attwateri*; Hammerly et al., 2016; Hammerly, Morrow, & Johnson, 2013). While some empirical research indicates that a large and diverse panel of microsatellites produces diversity estimates that are representative of genome-wide diversity and can be more useful than shallow pedigrees (e.g., Forstmeier, Schielzeth, Mueller, Ellegren, & Kempenaers, 2012), more recent simulation studies indicate that microsatellites provide less precision for relatedness and inbreeding, particularly in genetically depauperate endangered species where allelic diversity is low (i.e., <4 alleles per locus in the founding population; Robinson, Simmons, & Kennington, 2013; Taylor, 2015; Taylor, Kardos, Ramstad, & Allendorf, 2015). While the use of larger panels of diverse microsatellites may circumvent this issue for some species (e.g., Bergner et al., 2014; Gooley, Hogg, Belov, & Grueber, 2017), one simulation study for genetically depauperate endangered species shows that little precision is gained beyond 40 microsatellites, leading to inaccurate estimates of relatedness (Taylor et al., 2015). Recent studies argue that a better indication of genome-wide diversity can be obtained from genomic-based estimates of relatedness based on large numbers of genome-wide single nucleotide polymorphisms (i.e., SNPs; Knief et al., 2015; Taylor, 2015; Taylor et al., 2015).

Given the decreasing cost of high-throughput sequencing (Hayden, 2014) and the increasing amount of genomic resources readily available for nonmodel species (Galla et al., 2019), producing thousands of SNPs is now possible for many highly threatened species and provides an exciting opportunity for use in conservation breeding programmes (Galla et al., 2016; He, Johansson, & Heath, 2016). Indeed, there are several recent examples of genome-wide SNPs being used for relatedness in conservation, ecology and evolution (e.g., De Fraga, Lima, Magnusson, Ferrão, & Stow, 2017; Escoda, González-Esteban, Gómez, & Castresana, 2017), with some studies indicating that genome-wide SNPs provide greater precision in estimating relatedness and inbreeding compared to robust pedigrees (Kardos et al., 2015; Santure et al., 2010; Wang, 2016) or microsatellites (Attard, Beheregaray, & Möller, 2018; Bérénos et al., 2014; Hellmann et al., 2016; Keller, Visscher, & Goddard, 2011; Lemopoulos et al., 2019; Li, Strandén, Tiirikka, Sevón-Aimonen, & Kantanen, 2011; Thrasher, Butcher, Campagna, Webster, & Lovette, 2018).

To our knowledge, no study has compared pedigree-, genetic- and genomic-based approaches for estimating relatedness to inform pairing decisions for conservation breeding programmes, despite there being over 350 vertebrates worldwide that are captive bred for release to the wild (Smith et al., 2011). Here, we evaluate these three approaches using two critically endangered birds endemic birds to Aotearoa New Zealand—the kakī/black stilt (*Himantopus novaezelandiae*) and kākāriki karaka/orange-fronted parakeet (*C. malherbi*)—as proof of concept. Kakī and kākāriki karaka are excellent candidates for this research as both have active conservation breeding programmes, as well as multigenerational pedigrees (this study), microsatellite panels (Andrews, Hale, & Steeves, 2013; Steeves, Hale, & Gemmell, 2008) and genomic resources including species-specific reference genomes and whole-genome resequencing data (Galla et al., 2019; this study). In addition, because captive breeding pairs for both species are housed in separate enclosures and all offspring are intensively managed, kakī and kākāriki karaka present an excellent opportunity to examine relatedness in known family groups.

Once found on both main islands of Aotearoa, kakī experienced significant population declines throughout the 20th century due to introduced mammalian predators (e.g., feral cats, *Felis catus*; stoats, *Mustela erminea*; and hedgehogs, *Erinaceus europaeus*) along with braided river habitat loss and degradation (Sanders & Maloney, 2002). Today, an estimated 129 kakī are largely restricted to braided rivers of Te Manahuna/The Mackenzie Basin (Department of Conservation, *personal comm.*; Figure 1a) and recovery efforts include a conservation breeding programme that was initiated in the early 1980s (Reed, 1998). In addition to breeding birds in captivity, the kakī recovery programme also collects eggs from intensively monitored wild nests and rears them in captivity before wild release. Similar to kakī, kākāriki karaka were also once found on both main islands of Aotearoa and experienced population declines in the 19th and 20th centuries due to introduced mammalian predators (e.g., brushtail possums, *Trichosurus vulpecula*; rats, *Rattus rattus* and *R. norvegicus*; and stoats) and habitat loss (Kearvell & Legault, 2017). Today, breeding populations of an estimated 100–300 kākāriki karaka are restricted to beech (*Nothofagus* spp.) forests in three North Canterbury Valleys (the Hawdon, Hurunui, and Poulter) and to Oruawairua/Blumine Island in the Marlborough Sounds (Department of Conservation, *unpublished data*; Figure 1b). Recovery efforts include a conservation breeding programme initiated in 2003, with founders sourced from the Poulter, Hawdon and Hurunui Valleys. In most instances, offspring from pairings are released to the Hurunui Valley for wild supplementation. More recently, offspring are also released into the Poulter Valley to encourage pairing with the remaining birds from an extremely small remnant wild population (Department of Conservation, *personal comm.*). Eggs from these pairs are harvested, brought into captivity and fostered under surrogate birds, with hatchlings incorporated into the conservation breeding programme. Kakī belong to the Order Charadriiformes and are relatively long-lived braided river specialists that breed predictably within the bounds of a spring and summer season (Pierce, 2013). In contrast, kākāriki karaka belong to the Order Psittaciformes and are relatively short-lived beech forest specialists capable of breeding year-round, with prolific breeding periods associated with food abundance (e.g., beech forest masting events; Kearvell & Legault, 2017).

Here, we compare relatedness estimates from pedigree, microsatellites and genome-wide SNPs using known parent–offspring and full sibling relationships. We then compare pairing recommendations among these three approaches to assess how each translates to effective conservation management. Given that kakī and kākāriki karaka represent two taxonomically distinct bird species

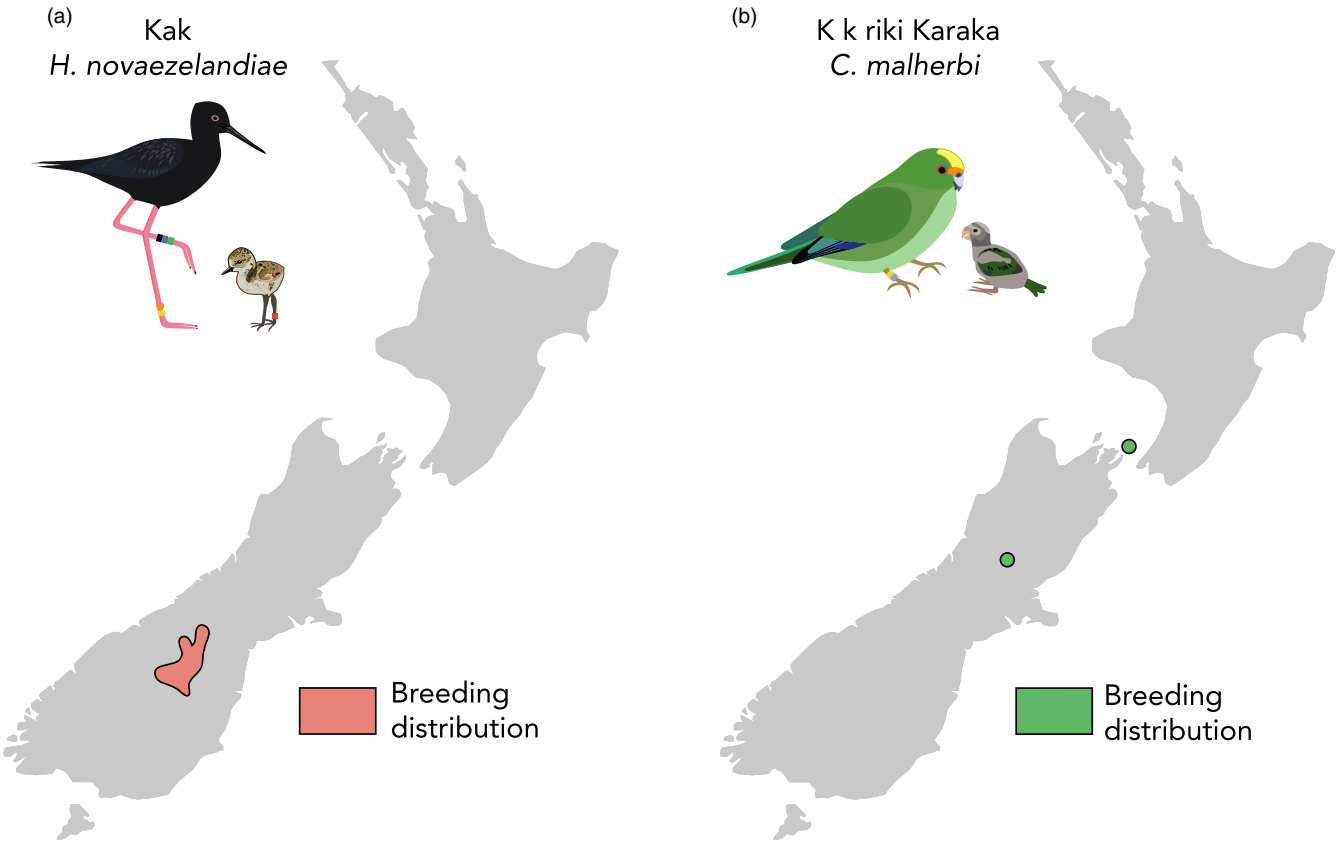

FIGURE 1 Current breeding distributions of wild kakī (a) and kākāriki karaka (b) in Aotearoa

with different life history strategies, we anticipate the results of our research may be applicable to the wider conservation breeding community.

2 | MATERIALS AND METHODS

2.1 | Sample collection and DNA extraction

Animal ethics approval for this project has been granted by the New Zealand Department of Conservation (i.e., DOC; permit number AEC 283). Captive kakī and kākāriki karaka are managed by DOC at two facilities in Aotearoa: the DOC Kakī Management Centre in Twizel and the Isaac Conservation and Wildlife Trust in Christchurch. Kakī used in this study are 36 individuals sampled between 2014 and 2017, including 24 individuals from six captive family groups and 12 individuals from wild parents that represent diverse lineages based on the pedigree. Kākāriki karaka sampled in this study are 36 individuals sampled between 2015 and 2019, including individuals from eight captive family groups and one wild individual from the Poulter Valley of North Canterbury (Table 1).

Blood, feather or tissue samples were sampled from each bird during routine health checks by DOC and Isaac Conservation and Wildlife Trust staff and immediately transferred into 95% molecular grade ethanol and stored at −80°C. High quantity and quality DNA was extracted using a lithium chloride chloroform extraction method (Galla et al., 2019) at the University of Canterbury School of Biological Sciences. Extractions were assessed for quality by running 2 µl of DNA on a 2% agarose gel. A Qubit® 2.0 Fluorometer (Fisher Scientific) was used for DNA quantification.

Familial relationships are known for all samples collected, as they were sampled from birds of known provenance in captive conditions. However, to verify that no sample was mislabelled during genetic and genomic processing, parent–offspring relationships were verified through an allele mismatch exclusion analysis (Jones & Ardren, 2003) using microsatellite panels previously developed for kakī (Steeves et al., 2008) and kākāriki karaka (Andrews, 2013), with a maximum allowed mismatch of one allele at one locus (see *Microsatellite data*

TABLE 1 Family group sampling strategy used in this study, based on pedigree data

Species	No. of Sampled Individuals	No. of family groups	No. of parent–offspring relationships	No. of sibling relationships
Kakī	36	6	24	7
Kākāriki Karaka	36	8	52	48

below). Family groups were further verified by clustering genome-wide SNP relatedness values calculated using the KGD method (Dodds et al., 2015, see below) using principal component analysis and visualization of family groups using the TensorFlow Embedding Projector (Smilkov et al., 2016; *data not shown*).

2.2 | Pedigree-based relatedness

Multigenerational pedigrees were constructed for kakī and kākāriki karaka by entering studbook information (i.e., hatch date, sex, parentage and status) into the programme PopLink v. 2.5.1 (Faust, Bergström, Thompson, & Bier, 2018). Sex for all individuals was verified using molecular markers 2550F/2718R (Fridolfsson & Ellegren, 1999) for kakī and P2/P8 (Griffiths, Double, Orr, & Dawson, 1998) for kākāriki karaka, with PCR products run on a 3% agarose gel for visual characterization, with positive controls included. Due to the short distance between P2/P8 alleles on the Z and W chromosomes in kākāriki karaka (Robertson & Gemmell, 2006), 2 µl of PCR products using a tagged forward primer was combined with 11.7 µl formamide and 0.3 µl of GeneScan™ LIZ® 500 size standard (Applied Biosystems) and genotyped on an ABI 3739xl (Applied Biosystems), with alleles manually scored using GeneMarker v. 2.2 (State College, PA, USA). Inconsistencies in pedigrees were identified using the validation tool in PopLink and corrected using observations by DOC and the Isaac Conservation and Wildlife Trust. Pairwise estimates of kinship and inbreeding were produced using the programme PMx v. 1.6.20190628 (Lacy, Ballou, & Pollak, 2012), selecting only the individuals used in this study (n = 36 in kakī and kākāriki karaka) and treating all unknown individuals in the pedigree as wild founders. In order to produce direct comparisons of pairwise relatedness coefficients (R) between pedigree, genetic and genomic data, R was calculated from pedigree kinship data using $R(xy) = 2 * f(xy)/\sqrt{(1 + Fx)(1 + Fy)}$. In this formula, $f(xy)$ is the kinship between two individuals (x and y) and Fx and Fy are the inbreeding coefficients of individuals x and y (Crow & Kimura, 1970).

2.3 | Genetic- and genomic-based relatedness

2.3.1 | Microsatellite data

Microsatellite loci (n = 8) previously described for kakī were amplified using PCR protocols by Steeves et al. (2008). Microsatellite loci (n = 17) designed for kākāriki karaka and one locus (*Cfor0809*) for Forbe's parakeet (*C. forbesi*) that cross-amplified in kākāriki karaka were amplified using PCR protocols by Andrews et al. (2013) and Chan, Ballantyne, Lambert, and Chambers (2005), respectively. Samples were prepared for genotyping by adding 0.5 µl of PCR product to 11.7 µl formamide and 0.3 µl GeneScan™ LIZ® 500 size standard (Applied Biosystems) and were genotyped on either a 3130xl or 3730xl Genetic Analyser (Applied Biosystems). Chromatograms were visualized using GeneMarker v. 2.2. To avoid bias by potential dye shifts (Sutton, Robertson, & Jamieson, 2011), peaks were scored manually. The number of alleles and standard estimates of per locus diversity—including expected heterozygosity (H_O) and observed heterozygosity (H_E)—were produced using GenAlEx v. 6.5 (Peakall & Smouse, 2006; Smouse & Peakall, 2012). Tests for deviations from Hardy–Weinberg and linkage disequilibrium for these loci using samples that are representative of larger kakī and kākāriki karaka populations can be found in Steeves et al. (2008), Steeves, Maloney, Hale, Tylianakis, and Gemmell (2010) and Andrews (2013), respectively. For kākāriki karaka, only eight of the 18 microsatellite markers previously described were polymorphic in this study and these eight loci were used in all downstream analyses.

Genetic-based R estimates were produced in the programme COANCESTRY v. 1.0.1.9 (Wang, 2011). COANCESTRY offers seven different estimators of relatedness, and to choose the most appropriate estimator for the kakī and kākāriki karaka microsatellite data sets, we employed the simulation module within COANCESTRY using allele frequencies, missing data and error rates from our microsatellite data sets. To produce dyads that represent the relationships and degree of inbreeding found within kakī and kākāriki karaka, we used R package "identity" (Li, 2010) to generate 10,879 dyads for kakī and 1,484 dyads for kākāriki karaka based on the pedigrees of both species. The frequency of each unique dyad in the kakī and kākāriki karaka data sets was scaled to create 1,000 dyads for each set that are representative of relationships between individuals used in this study. The COANCESTRY simulations were conducted using allele frequencies, error rates and missing data rates from each microsatellite data set, with settings changed to account for inbreeding. The triadic likelihood approach (Wang, 2007) was selected given it had the highest Pearson's correlation with "true" relatedness for both data sets (see Supporting Information for details). This approach is also preferred, as it is one of the few estimators that accounts for instances of inbreeding (Wang, 2007).

To estimate R with our genetic data set, COANCESTRY programme parameters were set to account for inbreeding, with the number of reference individuals and bootstrapping samples set to 100.

2.3.2 | Genomic data

Reference genomes

A reference genome for kakī has already been assembled (Galla et al., 2019) and was used in this study. To assemble a de novo reference genome for kākāriki karaka, a paired-end library was prepared at the Institute of Clinical Molecular Biology (IKMB) at Kiel University using the Nextera™ DNA Flex Library Prep Kit according to the manufacturer's specifications and sequenced on an Illumina NovaSeq™ 6,000 with 2 × 150 bp reads at a depth of approximately 70×.

FastQC v. 0.11.8 (Andrews, 2010) was used to evaluate the quality of the raw Illumina data and assess potential sample contamination. Initial read trimming was performed using TrimGalore v. 0.6.2 (Krueger, 2019) and Cutadapt v. 2.1 (Martin, 2011) with an end trim

quality of 30, a minimum length of 54 and using the *--nextseq* two-colour chemistry option. A median Phred score of 20 was also used for initial read trimming to remove obvious data errors; it should be noted that the assembly programmes used here (i.e., Meraculous-2D v. 2.2.10 and MaSuRCA v. 3.2.9; see below) have their own error corrections embedded in their respective pipelines. Kmer analyses were performed using Jellyfish v. 2.2.10 (Marçais & Kingsford, 2011) prior to assembly to assess heterozygosity and contamination. Two genome assembly programmes were tested for assembly performance: Meraculous-2D v. 2.2.5.1 (Chapman et al., 2011) and MaSuRCA v. 3.2.9 (Zimin et al., 2013). Meraculous was run using trimmed reads in diploid mode 1, with all other assembly parameters set to default. MaSuRCA was run using untrimmed reads, as it incorporates its own error correction pipeline. MaSuRCA parameter adjustments include a grid batch size of 300,000,000, the longest read coverage of 30, a Jellyfish hash size of 14,000,000,000 and the inclusion of scaffold gap closing; all other parameters were set to default for nonbacterial Illumina assemblies. The final assembly using the Meraculous pipeline was more fragmented (i.e., an N50 of 28.5 kb with 67,046 scaffolds > 1 kb), while the MaSuRCA genome was less fragmented (i.e., an N50 of 107.4 kb with 66,212 scaffolds > 1 kb) but contained possible artefacts due to heterozygosity (i.e., tandem repeats flanking short stretches of "N"s). To correct for these issues, the Meraculous assembly was first aligned to the MaSuRCA assembly using Last v. 959 (Kielbasa, Wan, Sato, Horton, & Frith, 2011); then, alignments were filtered to find matches where the Meraculous assembly spans the entirety of the tandem repeat in the MaSuRCA scaffolds, but lacking the tandem repeat or stretch of "N"s (i.e., gaps). In those cases, the aligned sequence in the MaSuRCA scaffold was replaced with the Meraculous match. All compute requirements needed to assemble the kākāriki karaka genome are available in Supporting Information.

Whole-genome resequencing

Kakī resequencing libraries were prepared at IKMB using a TruSeq® Nano DNA Library Prep Kit following the manufacturer's protocol and were sequenced across 34 lanes of an Illumina HiSeq 4,000. Twenty-four individuals were sequenced at high coverage depth (approximately 50×) for an aligned study, and all others were sequenced at a lower coverage depth (approximately 10×). Kākāriki karaka libraries were prepared at IKMB using the Nextera™ DNA Flex Library Prep Kit according to the manufacturer's specifications and sequenced across one lane of an Illumina NovaSeq™ 6,000 at IKMB at a coverage depth of approximately 10×, with one individual sequenced at a depth of approximately 70×, which was additionally used for the reference genome (see above).

FastQC v. 0.11.4 and 0.11.8 (Andrews, 2010) were used to evaluate the quality of the raw Illumina data for kakī and kākāriki karaka, respectively. Kakī resequencing reads were subsequently trimmed for the Illumina barcode, a minimum Phred quality score of 20 and a minimum length of 50 bp using Trimmomatic v. 0.38 (Bolger, Lohse, & Usadel, 2014). Because kākāriki karaka libraries were produced using different library preparation protocols and nextera chemistry, reads were trimmed using TrimGalore v. 0.6.2 (Krueger, 2019) for nextera barcodes and two-colour chemistry, using a median Phred score of 20, end trim quality of 30 and a minimum length of 54. Prior to mapping, the kakī reference genome was concatenated to a single chromosome using the custom perl script "concatenate_genome.pl" (Moraga, 2018a) for use in an aligned project that used both resequencing and genotyping-by-sequencing reads (see Galla et al., 2019). The kakī and kākāriki karaka reference genomes were indexed, and resequencing reads were mapped using Bowtie2 v. 2.2.6 and v. 2.3.4.1 (Langmead & Salzberg, 2012), respectively, with the setting *--very-sensitive*. Resulting SAM files were converted to BAM and were sorted using Samtools v. 1.9 (Li et al., 2009). Read coverage and variant calling were performed using *mpileup* in BCFtools v. 1.9 (Li et al., 2009). The custom perl script "split_bamfile_tasks.pl" (Moraga, 2018b) was used to reduce the computational time needed for mpileup by increasing parallelization. SNPs were detected, filtered and reported using BCFtools. Filtering settings were set to retain biallelic SNPs with a minor allele frequency (MAF) greater than 0.05, a quality score greater than 20 and a maximum of 10% missing data per site. After a series of filtering trials for each species (see Supporting Information for details), depth for kakī was set to have an average mean depth greater than 10, while kākāriki karaka depth was set so that each site had a minimum depth of 5 and a maximum depth of 200. Resulting SNPs for both data sets were pruned for linkage disequilibrium using BCFtools with the r^2 set to 0.6 and a window size of 1,000 sites. Sites were not filtered for Hardy–Weinberg equilibrium, as the nature of these data sets (mostly family groups) violates the assumptions of random mating. Per site missingness, depth and diversity—including proportion of observed and expected heterozygous SNP sites per individual, nucleotide diversity and SNP density per kb—were evaluated in the final sets using VCFtools v. 1.9 (Danecek et al., 2011). Diversity statistics were calculated using polymorphic markers only.

SNP-based relatedness

To produce estimates of *R* using whole-genome SNPs, the programme KGD (Dodds et al., 2015) was used, as it was designed to estimate relatedness using reduced representation and resequencing data while taking into account read depth. Pairwise *R* values derived from KGD were scaled so that self-relatedness for all individuals was equal to 1 using the formula $MS = D \times MO \times D$ where *MS* is the scaled matrix, *MO* is the original matrix, and *D* is a diagonal matrix with elements $D = 1/\sqrt{(diag(MO))}$. This scaling was performed to simplify downstream Mantel tests by creating a standardized diagonal value. This scaling was maintained throughout all analyses, as the scaled approach better approximated parent–offspring relationships, while demonstrating minimal bias to downstream analyses (see Supporting Information for details).

We evaluated the scaled KGD approach with other maker-based relatedness estimators, including the triadic likelihood approach (Wang, 2007), the KING estimator (Waples, Albrechtsen, & Moltke, 2019) and the r_{xy} method (Hedrick & Lacy, 2015), using parent–offspring relatedness as a benchmark for precision. We found that the

scaled KGD approach estimates parent–offspring relatedness closer to 0.5 compared to other relatedness estimators, while still providing estimates that are significantly concordant with all other approaches in both kakī (Pearson's r = 0.80–0.96, p < .001) and kākāriki karaka (Pearson's r = 0.89–96, p < .001; see Supporting Information for details).

2.4 | Comparison of relatedness

Mantel tests using the R package *ape* (Paridis & Schliep, 2018) were performed with 1,000 iterations to determine whether pedigree-, microsatellite- and SNP-based R were significantly correlated compared to a null distribution. Pearson's correlation coefficient (r) was additionally calculated to provide an additional measure of concordance between approaches. While our relatedness data sets are nonparametric, Pearson's was used over nonparametric tests, such as rank correlations, as our pedigree and microsatellite data sets have an excess of tied values.

2.5 | Pairing recommendations

We used two complementary methods in PMx v. 1.6.20190628 (Lacy et al., 2012) to determine whether pairing recommendations change using pedigree-, microsatellite- and SNP-based approaches for estimating R. First, we used mate suitability index (MSI), which scores how valuable offspring of a potential pair would be by taking into account four factors: deltaGD (i.e., the net positive or negative genetic diversity provided to the population), the difference of mean kinship values of the pair, the inbreeding coefficients of resulting offspring and the extent of unknown ancestry (Ballou, Earnhardt, & Thompson, 2001; Lacy et al. 2012). MSI scores scale from 1 to 6, with 1 being "very beneficial," and 6 being "very detrimental." An additional category denoted with a "-" indicates "very highly detrimental" pairings. Here, we assign this category with a numerical MSI score of 7. MSI scores provide a standardized approach for comparing pairing recommendations within and among species, including those based on the three approaches used in this study. However, Ballou et al. (2001) recommend caution when using automated pairing recommendations such as MSI in small and inbred populations. Thus, we also used mean kinship (MK) rank, which is an approach known to perform well in small and inbred populations (Ballou & Lacy, 1995; Rudnick & Lacy, 2008). This approach ranks individuals from lowest to highest MK among males and females, thereby creating a list of individuals for pairing prioritization.

For MSI score and MK rank analyses, only the individuals used in this study (n = 36 for both kakī and kākāriki karaka) were selected for analysis. PMx settings were set to default, with the exception of treating all unknown individuals in the pedigree as wild (i.e., 100% analytics known in the pedigree) to minimize bias from unknown pedigree assignments. Pedigree-based MSI scores and MK ranks were produced using pedigree-based kinship, while pairing recommendations using microsatellites and SNPs were produced using coefficients of relatedness. These genetic and genomic estimates of relatedness were uploaded to PMx, which divides these values by two to create an empirical metric of kinship. These empirical values were weighted to 1 to produce MSI scores and MK ranks that relied only on empirical data.

Pearson's correlation (r) was used to evaluate whether pairwise MSI scores and MK ranks between approaches were concordant. To test whether the distribution of MSI scores was statistically different

TABLE 2 Descriptive statistics based on pedigree data, as produced by PMx, including number of individuals, sex ratio (% males), maximum age, gene diversity, number of founders, number of founder genome equivalents, average inbreeding, average mean kinship, average generation time, % ancestry and analytic known, and effective population size

	All pedigreed individuals		Individuals in study	
Pedigree statistic	**Kakī**	**Kākāriki Karaka**	**Kakī**	**Kākāriki Karaka**
No. of Individuals	2,481	618	36	36
Sex ratio	0.27[a]	0.5	0.44	0.5
Max. age (years)	24	16.4	19.3	19.6
Gene diversity	0.96	0.915	0.9112	0.886
No. of founders	94	16	29	12
Founder genome equivalents	12.9	12	5.6	4.4
Average inbreeding	0.027	0.03	0.034	0.016
Average mean Kinship	0.039	0.085	0.089	0.114
Average generation time	4.82	1.31	5.25	3.79
% Ancestry known	55	100	58	100
% Analytic known	100	100	100	100
Ne/N	0.103	0.072	0.353	0.541

[a]The sex ratio for all pedigreed individuals for kakī is biased by a large number of individuals with unknown sex.

from one another, we used a nonparametric Kruskal–Wallis test with Bonferroni correction and a Tukey honest significant difference test.

3 | RESULTS

3.1 | Pedigree-based relatedness

This study has produced the first multigenerational pedigrees for two critically endangered endemic birds from Aotearoa. The kakī pedigree includes 2,481 wild and captive individuals recorded from 1977 to present, with a pedigree depth ranging from 1 to 8 generations (3.35 average). The number of founders and founder genome equivalents in the kakī pedigree (94 and 12.9, respectively) is high relative to the kākāriki karaka pedigree (16 and 12, respectively), and the % known ancestry is lower (55% and 100%, respectively; Table 2). Pedigree-based *R* between the 36 focal kakī ranged from 0 to 0.56, with an average *R* of 0.13 ± *SD* 0.13. The average coefficient of relatedness between all known kakī parent–offspring was higher than the expected 0.5 contribution from each parent (0.52 ± *SD*

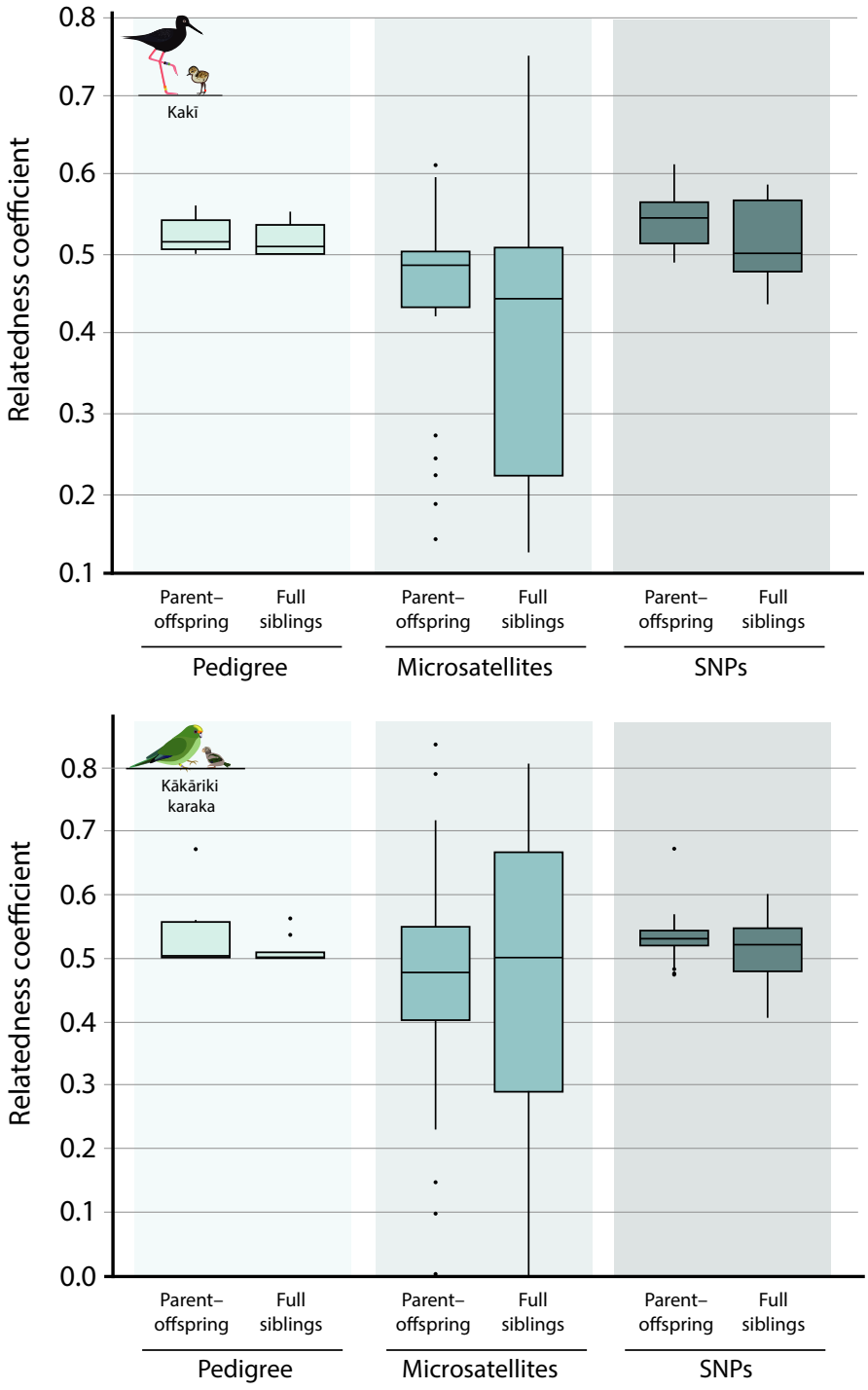

FIGURE 2 Parent–offspring and full sibling relatedness values derived from pedigree- (pale blue), microsatellite- (medium blue) and SNP-based (dark blue) methods in kakī (top graph) and kākāriki karaka (bottom graph) for the 36 focal individuals in this study

0.02), with averaged full sibling *R* of 0.52 ± *SD* 0.02 (Figure 2). The kākāriki karaka pedigree includes 624 captive individuals from 2003 to present, with an a pedigree depth ranging from 1 to 5 generations (2.48 average). Pedigree-based *R* for the 36 focal kākāriki karaka ranged from 0 to 0.67, with an average *R* of 0.19 ± *SD* 0.18. Average *R* between all parent–offspring was 0.52 ± *SD* 0.03, with averaged full sibling *R* being 0.51 ± *SD* 0.02 (Figure 2).

3.2 | Microsatellite diversity and relatedness

All eight microsatellite markers for kakī successfully amplified in all individuals used in this study. The number of alleles present across kakī loci ranged from 2 to 4 (average 3.13 ± *SD* 0.64; Table 3), with overall fewer alleles found here than reported in previous studies utilizing these loci with more individuals (Hagen, Hale, Maloney, & Steeves, 2011; Steeves et al., 2010). While 18 microsatellite markers were amplified in kākāriki karaka, one was removed from this study for not successfully amplifying in more than 50% of individuals (locus *OFK56*) and nine were removed for being monomorphic (Table 3). The number of alleles among polymorphic loci ranged from 2 to 4 (average 3.0 ± *SD* 0.93), with overall fewer alleles found here than reported in previous studies (Andrews, 2013; Andrews et al., 2013). Observed (H_O) heterozygosity and expected (H_E) heterozygosity for kakī (average H_O = 0.57 ± *SD* 0.17, average H_E = 0.54 ± *SD* 0.14) were higher than kākāriki karaka (average H_O = 0.43 ± *SD* 0.23, average H_E = 0.43 ± *SD* 0.20; Table 3).

Microsatellite-based *R* between all kakī used in this study ranged from 0 to 0.85, with an average *R* of 0.16 ± *SD* 0.19. Average *R* between all known kakī parent–offspring (0.44 ± *SD* 0.13) was below the minimum expected relatedness value of 0.5, with a larger standard deviation of *R* values compared to pedigree-based estimates. Averaged microsatellite-based full sibling *R* (0.41 ± *SD* 0.20) also had a larger deviation around the mean compared to the parent–offspring estimates (Figure 2).

Microsatellite-based *R* between all kākāriki karaka used in this study ranged from 0 to 0.84, with an average *R* of 0.18 ± *SD* 0.22. Similar to kakī, average *R* between all known kākāriki karaka parent–offspring relationships (0.47 ± *SD* 0.19) was below the minimum expected *R* value of 0.5, with a larger standard deviation of *R* values compared to pedigree-based estimates. Averaged full sibling *R* (0.49 ± *SD* 0.21) also had a larger deviation around the mean compared to microsatellite-based parent–offspring estimates (Figure 2).

3.3 | Reference genome assembly, SNP discovery, diversity and relatedness estimates

3.3.1 | Kākāriki karaka reference genome assembly

Reference genome library preparation and Illumina NovaSeq™ sequencing resulted in 584.47 million total reads for the kākāriki karaka genome. The final kākāriki karaka genome assembly was 1.15 GB in length, which is within the range of most assembled avian genomes (e.g., Zhang et al., 2014). The final assembly had 66,212 scaffolds with a scaffold N50 of 107.4 kb. See Data Availability section for access information.

TABLE 3 Descriptive statistics, including number of alleles, observed heterozygosity (H_O) and expected heterozygosity (H_E) for microsatellite loci used in this study. Loci from kākāriki karaka that were monomorphic (OFK12, OFK 19, OFK21, OFK26, OFK31, OFK33, OFK52, OFK56, OFK58, OFK61) are not included

Species	Locus	No. of alleles	H_O	H_E
Kakī	BS2	3	0.667	0.652
	BS9	3	0.611	0.59
	BS12	3	0.278	0.245
	BS13	2	0.528	0.5
	BS21	4	0.833	0.703
	BS27	4	0.667	0.551
	BS40	3	0.444	0.448
	BSdi7	3	0.556	0.596
Kākāriki Karaka	OFK9	4	0.472	0.477
	OFK41	4	0.75	0.702
	OFK50	3	0.444	0.513
	OFK54	4	0.722	0.574
	OFK55	2	0.278	0.346
	OFK60	2	0.222	0.239
	OFK62	2	0.083	0.08
	C for 809	3	0.44	0.52

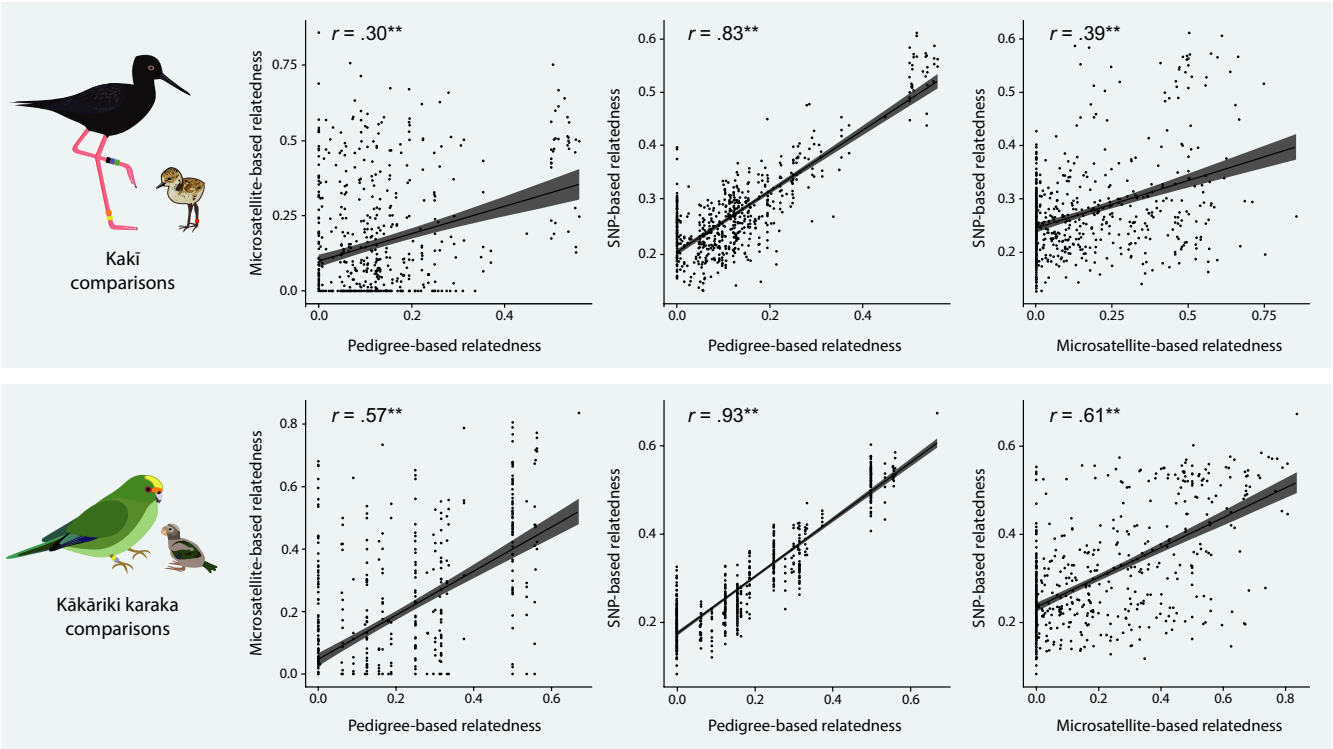

FIGURE 3 Scatterplots showing relationships between pedigree-, microsatellite- and SNP-based relatedness estimates in known family groups for kakī and kākāriki karaka. A trend line (black) and 95% confidence intervals (grey) are shown in each comparison

3.3.2 | SNP discovery and diversity

Library preparation and Illumina sequencing resulted in 6.07 billion total reads for kakī (168.69 ± *SD* 65.32 million reads). In addition to the individual used for the reference assembly, 3.64 billion total reads (average = 103.92 ± *SD* 29.76 million reads) were produced for the additional 35 kākāriki karaka in this study. More SNPs were discovered during initial SNP discovery using kākāriki karaka than kakī, and more remained postfiltering (Table 4). These filtered SNPs were used for all downstream analyses. Average missingness was low for both data sets (Table 4), but lower for kākāriki karaka than kakī, as kākāriki karaka had a hard minimum cut-off for depth during filtering that resulted in no missing data. Average depth for both data sets was relatively high (Table 4), with kakī having slightly higher average depth. Average diversity statistics (nucleotide diversity, and the average observed and expected SNP heterozygosity per individual postfiltering) were similar in both species, with diversity in kakī being slightly higher. SNP density using the kakī data set was higher than the kākāriki karaka data set (Table 4).

3.3.3 | SNP-based relatedness

SNP-based *R* between all kakī used in this study ranged from 0.13 to 0.61, with an average *R* of 0.27 ± *SD* 0.09. Similar to pedigree-based estimates, average *R* between all known kakī parent–offspring was slightly higher than the expected relatedness value of 0.5 with a small standard deviation relative to microsatellite-based estimates (0.54 ± *SD* 0.03). Averaged full sibling *R* also had a larger deviation around the mean (0.52 ± *SD* 0.05) than parent–offspring relationships (Figure 2).

SNP-based *R* between all kākāriki karaka used in this study ranged from 0.08 to 0.67, with an average *R* of 0.30 ± *SD* 0.12. Similar to pedigree-based estimates, average *R* between all known kākāriki karaka parent–offspring was slightly above the expected *R* value of 0.5 with a small standard deviation relative to genetic-based

TABLE 4 Descriptive statistics, including number of SNPs pre- and postfiltering, average depth ± *SD*, average missingness ± *SD*, average nucleotide diversity (π) ± *SD*, average proportion of observed heterozygous SNP sites (H_O) ± *SD*, average proportion of expected heterozygous SNP sites (H_E) ± *SD* and average SNP density (number of SNPs per kilobase) ± *SD*

Species	No. of SNPs prefiltering	No. of SNPs postfiltering	Average depth	Average missingness	Average π	Average H_O of SNP sites	Average H_E of SNP sites	SNP density
Kakī	4,246,100	68,144	28.73 ± 10.29	0.002 ± 0.004	0.35 ± 0.14	0.40 ± 0.02	0.35 ± 0.00	0.58 ± 3.18
Kākāriki Karaka	22,435,128	90,949	25.1 ± 14.87	0.00 ± 0.00	0.33 ± 0.14	0.37 ± 0.02	0.33 ± 0.00	0.17 ± 0.42

estimates (0.53 ± *SD* 0.03). Averaged full sibling relatedness also had a larger deviation around the mean (0.52 ± *SD* 0.05) compared to the pedigree-based estimates (Figure 2).

3.4 | Comparison of relatedness estimates and pairing recommendations

All kakī and kākāriki karaka *R* estimates using pedigree-, microsatellite- and SNP-based approaches correlated with one another with high statistical significance (Mantel test, $p < .001$; Pearson's correlation, $p < .001$; Figure 3). Of all the approaches, the correlation coefficient between pedigree- and SNP-based approaches was markedly higher than between other approaches, indicating that they are the most concordant (Figure 3).

Mate suitability index scores and MK ranks were calculated as an approximation for pairing recommendations derived from *R* estimates using the different approaches. Average pedigree-based MSI scores for kakī (4.46 ± *SD* 1.59) were lower on average than microsatellite-based scores (4.73 ± *SD* 1.63), but not significantly different from each other (Kruskal–Wallis test with Bonferroni correction, $p = .2$). SNP-based MSI scores for kakī (average = 5.67 ± *SD* 1.39) were significantly higher than pedigree- and microsatellite-based scores (Kruskal–Wallis test with Bonferroni correction, $p < .001$), with SNP-based scores providing the highest frequency of category 7 (i.e., very highly detrimental) pairings (Figure 4). While the distributions of MSI scores between each approach were different, each approach produced scores that correlated significantly with one another (Pearson's correlation, $p < .01$–0.001). Similar to correlations between *R* estimates, correlation coefficients between pedigree- and SNP-based MSI scores were highest (Pearson's $r = 0.5$, see Figure S1 for details). Of the 320 possible kakī pairings, 38% did not experience a change in MSI score value between pedigree- and SNP-based approaches; however, 20% of pairings experienced an MSI score change that was 3+ categories different. In 2% of pairings, pedigree-based MSI scores were categorized as a 1 (i.e., preferred pairing), while SNP-based MSI scores were categorized as a 7 (i.e., very highly detrimental). Correlations between MK ranks provided by the three approaches were significant between pedigree- and SNP-based approaches only (Pearson's $r = 0.75$, $p < .001$; see Figure S2 for details). Among pedigree- and SNP-based MK ranks, 64% of individuals experienced a minimal rank shift of 0–3 categories, 22% experienced a moderate rank shift of 4–7 categories, and 3% experienced a high rank shift of ≥8 categories.

Similar to kakī, average kākāriki karaka SNP-based MSI scores (5.64 ± *SD* 1.47) were significantly higher than pedigree (5.20 ± *SD* 1.71) and microsatellite (5.04 ± *SD* 1.61) scores (Kruskal–Wallis test with Bonferroni correction, $p < .001$), while pedigree- and microsatellite-based scores did not significantly differ (Kruskal–Wallis test with Bonferroni correction, $p = .67$). SNP-based scores provided the highest frequency of category 7 (i.e., very highly detrimental) pairings (Figure 4). Each approach also produced scores that correlated significantly with one another (Pearson's correlation, $p < .001$), with

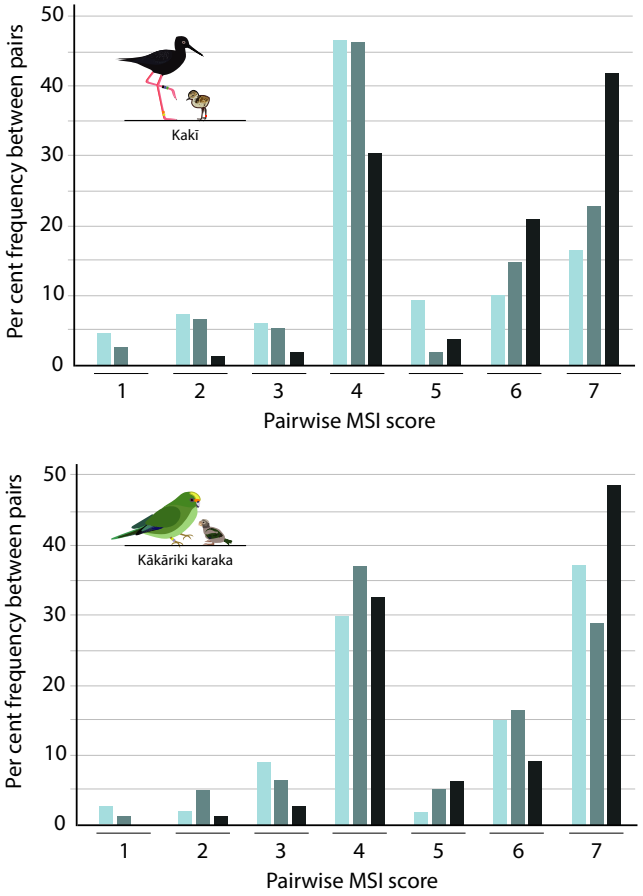

FIGURE 4 Frequency of MSI scores using pedigree- (pale blue), microsatellite- (medium blue) and SNP-based (dark blue) kinship/relatedness values in kakī and kākāriki karaka. MSI, mate suitability index; SNP, single nucleotide polymorphism

the highest correlation coefficients seen between pedigree- and SNP-based MSI scores (Pearson's $r = 0.65$, see Figure S1 for details). Of the 324 possible pairings for kākāriki karaka, 59% did not experience a change in MSI score value between pedigree- and SNP-based approaches; however, 9% of pairings experienced an MSI score change that was 3+ categories different. In 2% of pairings, pedigree-based MSI scores were categorized as a 1 (i.e., very beneficial), while SNP-based MSI scores were categorized as a 7 (i.e., very highly detrimental). Correlations between MK rank were significant between pedigree- and SNP-based approaches (Pearson's $r = 0.64$, $p < .001$) and microsatellite- and SNP-based approaches (Pearson's $r = 0.51$, $p = .002$, see Figure S2 for details). Among pedigree- and SNP-based MK ranks, 53% of individuals experienced a minimal rank shift of 0–3 categories, 31% experienced a moderate rank shift of 4–7 categories, and 8% experienced a high rank shift of ≤8 categories.

4 | DISCUSSION

This study is the first to compare pedigree-, microsatellite- and SNP-based estimates of relatedness and subsequent pairing

recommendations for conservation breeding programmes. The results indicate that microsatellites provide the least precision when estimating relatedness in known family groups, with pedigree- and SNP-based estimates providing higher precision and a much closer approximation of parent–offspring and full sibling relatedness. Further, estimates of relatedness and downstream pairing recommendations using MSI scores and MK ranks are both more concordant when using pedigree- and SNP-based data sets compared to microsatellite-based data sets. Despite this, there were important differences in pairing recommendations between pedigree- and SNP-based approaches, with SNP-based MSI scores being statistically higher than pedigree-based scores, and some substantial disagreements existing between the two sets of MSI scores and MK ranks. Together, this study provides insight into the differences between pedigree-, microsatellite- and SNP-based approaches for making pairing recommendations and a pathway for estimating relatedness using genome-wide SNPs to inform pairing decisions in poorly pedigreed conservation breeding programmes worldwide.

4.1 | Relatedness comparisons

When producing empirical estimates of relatedness, simulations were performed to choose the most suitable estimator for microsatellites, and various relatedness estimators and filtering schemes for SNPs were evaluated to find the approach that best approximated known parent–offspring relationships as a biologically relevant benchmark. While different relatedness estimators and filtering schemes will result in different point estimates of relatedness, this study demonstrates an approach for producing relatedness estimates that are well suited for our particular data set and research question.

Pedigree-based estimates of parent–offspring and full sibling relatedness approximated 0.5 for both kakī and kākāriki karaka (Figure 2), with some measures being slightly higher, which likely reflects intergenerational inbreeding and/or higher pedigree depth from the baseline (reference) population. These results are consistent with expectations, as pedigrees are based on the probability of Mendelian inheritance, which postulates that first-order relationships (i.e., parents and offspring, and siblings) share 50% of their genomic information (Lacy, 1995; Wright, 1922). We expect realized (i.e., empirical) parent–offspring relationships to also approximate 0.5, but a broader range of realized relatedness estimates among full siblings, as they may receive different genetic material from each parent due to recombination and independent assortment during meiosis and random fertilization (Hill & Weir, 2011, 2012; Speed & Balding, 2015). Even when pedigrees are robust, this study highlights an unavoidable shortcoming as pedigrees do not adequately capture true relatedness between full siblings. We anticipate this uncaptured diversity may prove useful for maximizing existing diversity, especially in conservation breeding programmes with relatively few founders (Ballou & Lacy, 1995).

Compared to the pedigree-based approach, our empirical data sets (i.e., microsatellites and SNPs) capture more variation between siblings than parents and offspring (Figure 2). A broad range of microsatellite-based relatedness estimates were observed in both parent–offspring and sibling relationships, compared to the SNP-based approach. In some instances, even parent–offspring pairings appeared relatively unrelated using microsatellites (e.g., minimum parent–offspring $R = 0.14$ in kakī and $R = 0$ in kākāriki karaka), which underscores the lack of precision in this approach and how it could inadvertently lead to poorly informed pairing recommendations. These large ranges of relatedness values using microsatellites can be explained because genetic-based relatedness values between parent–offspring and full siblings are based on allele frequencies, and relatedness between individuals that share common alleles will be lower than individuals that share rare alleles (Speed & Balding, 2015; Wang, 2011). This bias in relatedness values can be exacerbated when samples sizes are small (Wang, 2017), which is typical of conservation breeding programmes. Furthermore, the lack of precision using microsatellites shown here is consistent with studies that suggest relatively few markers with low allelic diversity are insufficient for estimating relatedness and inbreeding, especially in genetically depauperate species (e.g., Attard et al., 2018; Escoda et al., 2017; Hellmann et al., 2016; Taylor, 2015; Taylor et al., 2015).

Compared to microsatellite-based relatedness, SNP-based relatedness showed a relatively small range with parent–offspring and full sibling relatedness estimates approximating 0.5, and full siblings showing a wider range of values than parent–offspring relationships (Figure 2). Not only is this pattern consistent with expectations given the behaviour of chromosomes during meiosis and random fertilization, but it also shows more precision than the microsatellite data sets. Other researchers have found similar results in a diverse range of wild taxa, indicating that thousands of genome-wide SNPs show more precision than microsatellites when measuring relatedness and inbreeding (e.g., Attard et al., 2018; Hellmann et al., 2016; Hoffman et al., 2014; Lemopoulos et al., 2019; Thrasher, Butcher, Campagna, Webster, & Lovette, 2018).

Beyond parent–offspring and full sibling relationships, pedigree- and SNP-based relatedness estimates showed the highest concordance with one another among the three approaches used (Figure 3). In kakī, the data sets used here include non-captive-bred individuals with intensively monitored wild parents. These results provide more credibility to the semi-wild kakī pedigree, where socially monogamous wild pairs of kakī are assumed to be the genetic parents of offspring at nests (but see also Overbeek et al., 2017). Still, it should be noted that many pairs with pedigree-based relatedness values of 0 had SNP-based relatedness values ranging upwards of 0.40 in kakī 0.33 and in kākāriki karaka, which approximates first- and second-order relationships in both species (Figure 2). This indicates that pedigree-based R between these individuals may be downwardly biased by the assumption that no variance in relatedness exists among founders, missing information and/or low pedigree depth (Balloux et al., 2004; Bérénos et al., 2014; Hammerly et al., 2016; Hogg et al., 2019;

Kardos et al., 2015; Lacy, 1995; Pemberton, 2008; Rudnick & Lacy, 2008; Tzika et al., 2009).

4.2 | Pairing recommendations

When these relatedness values are translated into pairing recommendations using MSI scores and MK rank, there is a high concordance between pedigree- and SNP-based approaches, with SNP-based MSI scores being significantly higher than pedigree- and microsatellite-based approaches. The latter result is somewhat expected, given that average relatedness estimates using SNPs was highest among the approaches used here, and empirical estimates of relatedness and inbreeding are usually higher than pedigrees as they more effectively capture relatedness between founders or misassigned individuals (Hammerly et al., 2016; Hogg et al., 2019). With that said, when making pairing recommendations using kinship-based pairing decisions (e.g., Ballou & Lacy, 1995), it is often the relative kinships between individuals that are more important than absolute values (Galla et al., 2019; McLennan, Wright, Belov, Hogg, & Grueber, 2019). This suggests that pedigree- and SNP-based approaches both yield similar results for pairing recommendations, with some important differences. For example, while correlation coefficients between these two sets of MSI scores are high relative to other comparisons, there are instances where pairings are considered "highly beneficial" (i.e., MSI category 1) when using the pedigree and "very highly detrimental" (i.e., MSI category 7) when using SNPs. When comparing MK ranks between pedigree- and SNP-based approaches, some kakī and kākāriki karaka experienced large shifts in rank (i.e., ≥8 positions difference) depending on the approach used. Although we expect some differences between pedigree- and SNP-based MSI scores and MK ranks, we attribute these very large differences to errors in the pedigree (e.g., Hammerly et al., 2016) or violations of the assumption that there is no variance in founder relationships (e.g., Hogg et al., 2019). Of all kakī and kākāriki karaka pairings that experienced a large shift between pedigree- and SNP-based MSI scores, most feature recurring individuals with wild parentage (i.e., founders), and in one recurring occasion, a wild individual (kakī) with high pedigree depth that likely represents an entry error in the pedigree.

4.3 | Management implications

Pedigree-, genetic- and genomic-based tools each have their advantages to inform conservation management. For example, pedigrees capture both genetic and demographic considerations dating back to the founding of the population, while empirical estimates of relatedness can circumvent pedigree errors and issues surrounding founder relationships by expressing realized relatedness between all sampled individuals. From the results shown here, we recommend that when conservation breeding programmes are poorly pedigreed (i.e., pedigrees of low depth or containing missing data), SNPs should be incorporated to provide a precise indicator of relatedness to genetically inform pairing decisions. The microsatellite panels used here have shown low precision in estimating relatedness, with demonstrated downstream effects for pairing recommendations compared to pedigree- and SNP-based approaches. More microsatellites could be developed to mitigate this shortcoming; however, other studies indicate that a larger number microsatellites (e.g., 20–40 markers) may not equate to higher precision for relatedness estimates and inbreeding coefficients, especially in threatened species with low allelic diversity (Nietlisbach et al., 2017; Robinson et al., 2013; Taylor, 2015; Taylor et al., 2015). Further, the time and cost associated with building larger microsatellite panels and generating microsatellite data will likely be surpassed by the production of genome-wide SNPs, either by a whole-genome resequencing approach as shown in this study or by a reduced representation sequencing approaches (e.g., RAD sequencing, or genotyping-by-sequencing; Galla et al., 2016; Narum, Buerkle, Davey, Miller, & Hohenlohe, 2013). Currently, for kakī and kākāriki karaka, reduced representation sequencing is more cost-effective than whole-genome resequencing (i.e., approximately one-third of the price, depending on the genome size, as of 2019)—but we foresee more people shifting towards whole-genome resequencing in the near future, given the decreasing cost of high-throughput sequencing (Hayden, 2014) and the ability to ask more research questions using whole-genome resequencing data sets (see Future Directions below for details). This is particularly true for birds, whose genomes are small (e.g., 1.05–1.26G) relative to many vertebrates (Zhang et al., 2014).

We anticipate SNPs will be particularly applicable in circumstances when pedigrees are the least reliable. For instance, when the founders of a conservation breeding population have no ancestry data available and are likely to be related, SNP-based relatedness estimates between individuals can be used to avoid highly related matings (Hogg et al., 2019). This situation may not only coincide with the original founding event of a captive population, but iteratively when individuals are sourced from wild or translocated populations to augment the captive population, as suggested in Frankham (2008) and Hogg et al. (2019). For example, in kākāriki karaka, whole-genome resequencing has been made available for all current breeding individuals in the conservation breeding programme, including individuals who are founders themselves. Because birds of unknown ancestry are being routinely sourced from highly endangered wild populations, and will also be founders, we anticipate the need for resequencing these birds as they are incorporated into the breeding programme to assess their relatedness to other individuals. In addition to traditional captive-bred populations (i.e., ex situ management), this approach is applicable to intensively managed wild populations (i.e., in situ management), where robust pedigrees are lacking, but conservation translocations can be informed by relatedness between individuals in a managed landscape (e.g., kākāpō, *Strigops*

habroptilus, Elliott, Merton, & Jansen, 2001; scimitar-horned oryx, *Oryx dammah*, Wildt et al., 2019).

While we expect SNPs will be important for pairing recommendations moving forward, we do not expect they will eclipse well-established pedigrees, as both approaches have advantages for conservation breeding. Instead, we envision a combined approach where realized relatedness from SNPs can be used to augment data-rich pedigrees. With that said, there are relatively few studies to date that effectively combine existing pedigree data with genomic estimates of relatedness to inform pairing recommendations (but see Hogg et al., 2019; Ivy, Putnam, Navarro, Gurr, & Ryder, 2016). To date, these studies are largely limited to SNPs being used for parentage reconstruction (reviewed in Flanagan & Jones, 2019), where unknown or uncertain relationships are reconstructed using empirical data and software (e.g., Whalen, Gorjanc, & Hickey, 2019), and more complete pedigrees are used moving forward. Alternatively, there is an option to produce empirical estimates of relatedness for all founders or breeding individuals in conservation breeding programmes—as suggested in Ivy et al. (2016) and practised in Hogg et al. (2019)—and use this baseline of known relatedness moving forward using pedigrees. While the programme PMx allows for the inclusion of empirical data (Lacy et al., 2012), this approach requires caution, as the calculation of pedigree-based identity by descent for subsequent generations—including kinship and gene diversity—will be affected by the addition of empirical data (Hogg et al., 2019). We acknowledge this approach requires further investigation and validation, particularly for species that receive periodic influx of wild individuals of unknown ancestry in their conservation breeding programme.

4.4 | Future directions and concluding remarks

This study has produced pedigrees and whole-genome sequences for two critically endangered species. Beyond estimating relatedness, these tools provide an exciting opportunity to explore other questions relevant to conservation, such as characterizing the genomic basis of fitness traits, including those associated with inbreeding depression (Kardos, Taylor, Ellegren, Luikart, & Allendorf, 2016; but see also Kardos & Shafer, 2019) or adaptation to captivity (e.g., Grueber et al., 2017). We also envision using the genomic resources developed here to further investigate best practice for making pairing recommendations, for example, agent-based, multigenerational simulations can be used to evaluate whether genome-wide diversity is best maximized using pedigrees, SNPs or a combination approach.

Given that SNPs have been successfully used to estimate relatedness for different purposes across a wide diversity of taxonomic groups outside of this study (as reviewed in Attard et al., 2018), we anticipate a SNP-based approach for estimating relatedness and making subsequent pairing recommendations will be applicable beyond birds. In the meantime, for poorly pedigreed populations worldwide, we recommend a SNP-based approach to estimate relatedness for subsequent pairing recommendations. It should be noted that many approaches used to date have used de novo reduced representation approaches (e.g., genotyping-by-sequencing, RADseq; Narum et al., 2013) for SNP discovery, which typically have more missing data, lower depth and fewer SNPs than the reference-guided whole-genome resequencing approach used here. While these factors may contribute to bias in relatedness estimates (but see Dodds et al., 2015), research still indicates that fewer SNPs, with more missing data and lower depth, than those presented here provide more precision than microsatellites (Attard et al., 2018). We expect reduced representation approaches will persist in the short term, especially for species with large and complex genomes (e.g., some fish, amphibians and invertebrates) that otherwise cannot yet be affordably resequenced across entire conservation breeding programmes. With that said, we also expect whole-genome resequencing projects like ours will gain momentum in the years to come, as these data can be better leveraged to address multiple questions related to conservation genetic management (Harrisson, Pavlova, Telonis-Scott, & Sunnucks, 2014; see also above). In the meantime, we look forward to seeing more poorly pedigreed conservation breeding programmes for taxonomically diverse species from around the world incorporate SNPs for estimating relatedness to inform pairing decisions.

ACKNOWLEDGEMENTS

We are grateful for the support of Te Rūnanga o Ngāi Tahu, Te Ngāi Tūāhuriri Rūnanga, Te Rūnanga o Arowhenua, Te Rūnanga o Waihao and Te Rūnanga o Moeraki. We thank all members of the Kakī Recovery Programme and the Orange-fronted Parakeet Recovery Programme for their assistance with sample collection and their ongoing support. This research has been funded by the Ministry of Business, Innovation and Employment (MBIE) Endeavour Fund (UOCX1602 awarded to TES), the Brian Mason Scientific and Technical Trust (awarded to SJG and TES), and the Mohua Charitable Trust (awarded to TES). We thank the three anonymous reviewers for their thoughtful and constructive feedback to this manuscript.

CONFLICT OF INTEREST

None declared.

DATA AVAILABILITY STATEMENT

Genomic data provided in this manuscript are available through a password protected server on the Conservation, Systematics and Evolution Research Team's website (http://www.ucconsert.org/data/). Kakī and kākāriki karaka are taonga (treasured) species. For Māori (the indigenous people of Aotearoa), all genomic data obtained from taonga species have whakapapa (genealogy that includes people, plants and animals, mountains, rivers and winds) and are therefore taonga in their own right. Thus, these data are tapu (sacred) and tikanga (customary practices, protocols and ethics) determine how people interact with it. To this end, the passwords for the genomic data in this manuscript will be made

available to researchers on the recommendation of the kaitiaki (guardians) for the iwi (tribes) that affiliate with kakī and kākāriki karaka.

ORCID

Stephanie J. Galla https://orcid.org/0000-0002-4650-8067
Roger Moraga https://orcid.org/0000-0001-7806-1135
Liz Brown https://orcid.org/0000-0002-4245-3900
Marc P. Hoeppner https://orcid.org/0000-0002-0279-2270
Richard F. Maloney https://orcid.org/0000-0002-1699-763X
Lyndon Slater https://orcid.org/0000-0001-6832-2821

REFERENCES

Andrews, B. J. (2013). *Conservation genetics of the orange-fronted kākāriki (*Cyanoramphus malherbi*)*. Masters Thesis. University of Canterbury.

Andrews, B. J., Hale, M. L., & Steeves, T. E. (2013). Characterisation of microsatellite loci in the critically endangered orange-fronted kākāriki (*Cyanoramphus malherbi*) isolated using genomic next generation sequencing. *Conservation Genetics Resources*, *5*(1), 235–237. https://doi.org/10.1007/s12686-012-9777-4

Andrews, S. (2010). *FastQC: A quality control tool for high throughput sequence data* [Computer Software]. Retrieved from https://github.com/s-andrews/FastQC.

Attard, C. R. M., Beheregaray, L. B., & Möller, L. M. (2018). Genotyping-by-sequencing for estimating relatedness in nonmodel organisms: Avoiding the trap of precise bias. *Molecular Ecology Resources*, *18*(3), 381–390. https://doi.org/10.1111/1755-0998.12739

Attard, C., Möller, L. M., Sasaki, M., Hammer, M. P., Bice, C. M., Brauer, C. J., … Beheregaray, L. B. (2016). A novel holistic framework for genetic-based captive-breeding and reintroduction programs: Holistic framework for reintroductions. *Conservation Biology*, *30*(5), 1060–1069. https://doi.org/10.1111/cobi.12699

Ballou, J. D. (1983). Calculating inbreeding coefficients from pedigrees. In C. Schonewald-Cox, S. Chambers, B. MacBryde, & L. Thomas (Eds.), *Genetics and conservation* (pp. 509–516). Menlo Park, CA: Benjamin/Cummings Publishing Co.

Ballou, J. D., Earnhardt, J., & Thompson, S. (2001). *MateRx: Genetic management software*. Washington, DC: Smithsonian National Zoological Park.

Ballou, J. D., & Lacy, R. C. (1995). Identifying genetically important individuals for management of genetic variation in pedigreed populations. In J. D. Ballou, M. E. Gilpin, & T. J. Foose (Eds.), *Population management for survival and recovery* (pp. 76–111). New York, NY: Columbia University Press.

Ballou, J. D., Lees, C., Faust, L. J., Long, S., Lynch, C., Bingaman Lackey, L., & Foose, T. J. (2010). Demographic and genetic management of captive populations. In D. G. Kleiman, K. V. Thompson, & C. K. Baer (Eds.), *Wild Mammals in Captivity: Principles and Techniques for Zoo Management*, (pp. 219). Chicago, IL: University of Chicago Press.

Balloux, F., Amos, W., & Coulson, T. (2004). Does heterozygosity estimate inbreeding in real populations? *Molecular Ecology*, *13*(10), 3021–3031. https://doi.org/10.1111/j.1365-294X.2004.02318.x

Bérénos, C., Ellis, P. A., Pilkington, J. G., & Pemberton, J. M. (2014). Estimating quantitative genetic parameters in wild populations: A comparison of pedigree and genomic approaches. *Molecular Ecology*, *23*(14), 3434–3451. https://doi.org/10.1111/mec.12827

Bergner, L. M., Jamieson, I. G., & Robertson, B. C. (2014). Combining genetic data to identify relatedness among founders in a genetically depauperate parrot, the kakapo (*Strigops habroptilus*). *Conservation Genetics*, *15*(5), 1013–1020. https://doi.org/10.1007/s10592-014-0595-y

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

Chan, C., Ballantyne, K. N., Lambert, D. M., & Chambers, G. K. (2005). Characterization of variable microsatellite loci in Forbes' parakeet (*Cyanoramphus forbesi*) and their use in other parrots. *Conservation Genetics*, *6*(4), 651–654. https://doi.org/10.1007/s10592-005-9021-9

Chapman, J. A., Ho, I., Sunkara, S., Luo, S., Schroth, G. P., & Rokhsar, D. S. (2011). Meraculous: De novo genome assembly with short paired-end reads. *PLoS ONE*, *6*(8), e23501. https://doi.org/10.1371/journal.pone.0023501

Crow, J. F., & Kimura, M. (1970). *An introduction to population genetics theory*. New York, NY: Harper & Row, Publishers.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., & DePristo, M. A., … 1000 Genomes Project Analysis Group (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

De Fraga, R., Lima, A. P., Magnusson, W. E., Ferrão, M., & Stow, A. J. (2017). Contrasting patterns of gene flow for Amazonian snakes that actively forage and those that wait in ambush. *Journal of Heredity*, *108*(5), 524–534. https://doi.org/10.1093/jhered/esx051

de Villemereuil, P., Rutschmann, A., Lee, K. D., Ewen, J. G., Brekke, P., & Santure, A. W. (2019). Little adaptive potential in a threatened passerine bird. *Current Biology*, *29*(5), 889–894. e3. https://doi.org/10.1016/j.cub.2019.01.072

Dodds, K. G., McEwan, J. C., Brauning, R., Anderson, R. M., van Stijn, T. C., Kristjánsson, T., & Clarke, S. M. (2015). Construction of relatedness matrices using genotyping-by-sequencing data. *BMC Genomics*, *16*(1). https://doi.org/10.1186/s12864-015-2252-3

Elliott, G. P., Merton, D. V., & Jansen, P. W. (2001). Intensive management of a critically endangered species: The kakapo. *Biological Conservation*, *99*(1), 121–133. https://doi.org/10.1016/S0006-3207(00)00191-9

Escoda, L., González-Esteban, J., Gómez, A., & Castresana, J. (2017). Using relatedness networks to infer contemporary dispersal: Application to the endangered mammal *Galemys pyrenaicus*. *Molecular Ecology*, *26*(13), 3343–3357.

Faust, L. J., Bergström, Y. M., Thompson, S. D., & Bier, L. (2018). *PopLink Version 2.5* [Computer Software]. Retrieved from https://www.lpzoo.org/conservation-science/projects/poplink

Flanagan, S. P., & Jones, A. G. (2019). The future of parentage analysis: From microsatellites to SNPs and beyond. *Molecular Ecology*, *28*, 544–567.

Forstmeier, W., Schielzeth, H., Mueller, J. C., Ellegren, H., & Kempenaers, B. (2012). Heterozygosity-fitness correlations in zebra finches: Microsatellite markers can be better than their reputation: Molecular measures of inbreeding. *Molecular Ecology*, *21*(13), 3237–3249. https://doi.org/10.1111/j.1365-294X.2012.05593.x

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

Frankham, R. (2008). Genetic adaptation to captivity in species conservation programs. *Molecular Ecology*, *17*(1), 325–333. https://doi.org/10.1111/j.1365-294X.2007.03399.x

Fridolfsson, A.-K., & Ellegren, H. (1999). A simple and universal method for molecular sexing of non-ratite birds. *Journal of Avian Biology*, 116–121. https://doi.org/10.2307/3677252

Galla, S. J., Buckley, T. R., Elshire, R., Hale, M. L., Knapp, M., McCallum, J., … Steeves, T. E. (2016). Building strong relationships between conservation genetics and primary industry leads to mutually beneficial genomic advances. *Molecular Ecology*, *25*(21), 5267–5281. https://doi.org/10.1111/mec.13837

Galla, S. J., Forsdick, N. J., Brown, L., Hoeppner, M., Knapp, M., Maloney, R. F., … Steeves, T. E. (2019). Reference genomes from distantly related species can be used for discovery of single nucleotide

polymorphisms to inform conservation management. *Genes*, *10*(1), 9. https://doi.org/10.3390/genes10010009

Giglio, R. M., Ivy, J. A., Jones, L. C., & Latch, E. K. (2016). Evaluation of alternative management strategies for maintenance of genetic variation in wildlife populations. *Animal Conservation*, *19*, 380–390.

Gooley, R., Hogg, C. J., Belov, K., & Grueber, C. E. (2017). No evidence of inbreeding depression in a Tasmanian devil insurance population despite significant variation in inbreeding. *Scientific Reports*, *7*, 1–11.

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

Grueber, C. E., Fox, S., McLennan, E. A., Gooley, R. M., Pemberton, D., Hogg, C. J., & Belov, K. (2019). Complex problems need detailed solutions: Harnessing multiple data types to inform genetic management in the wild. *Evolutionary Applications*, *12*(2), 280–291. https://doi.org/10.1111/eva.12715

Grueber, C. E., Reid-Wainscoat, E. E., Fox, S., Belov, K., Shier, D. M., Hogg, C. J., & Pemberton, D. (2017). Increasing generations in captivity is associated with increased vulnerability of Tasmanian devils to vehicle strike following release to the wild. *Scientific Reports*, *7*, 1–7.

Hagen, E. N., Hale, M. L., Maloney, R. F., & Steeves, T. E. (2011). Conservation genetic management of a critically endangered New Zealand endemic bird: Minimizing inbreeding in the Black Stilt *Himantopus novaezelandiae*. *Ibis*, *153*(3), 556–561. https://doi.org/10.1111/j.1474-919X.2011.01137.x

Hammerly, S. C., de la Cerda, D. A., Bailey, H., & Johnson, J. A. (2016). A pedigree gone bad: Increased offspring survival after using DNA-based relatedness to minimize inbreeding in a captive population. *Animal Conservation*, *19*(3), 296–303.

Hammerly, S. C., Morrow, M. E., & Johnson, J. A. (2013). A comparison of pedigree- and DNA-based measures for identifying inbreeding depression in the critically endangered Attwater's Prairie-chicken. *Molecular Ecology*, *22*(21), 5313–5328. https://doi.org/10.1111/mec.12482

Harrisson, K. A., Pavlova, A., Telonis-Scott, M., & Sunnucks, P. (2014). Using genomics to characterize evolutionary potential for conservation of wild populations. *Evolutionary Applications*, *7*(9), 1008–1025.

Hayden, E. C. (2014). The $1,000 genome. *Nature*, *507*, 295–295.

He, X., Johansson, M. L., & Heath, D. D. (2016). Role of genomics and transcriptomics in selection of reintroduction source populations. *Conservation Biology*, *30*(5), 1010–1018. https://doi.org/10.1111/cobi.12674

Hedrick, P. W., & Lacy, R. C. (2015). Measuring relatedness between inbred individuals. *Journal of Heredity*, *106*, 20–25. https://doi.org/10.1093/jhered/esu072

Hellmann, J. K., Sovic, M. G., Gibbs, H. L., Reddon, A. R., O'Connor, C. M., Ligocki, I. Y., … Hamilton, I. M. (2016). Within-group relatedness is correlated with colony-level social structure and reproductive sharing in a social fish. *Molecular Ecology*, *25*(16), 4001–4013. https://doi.org/10.1111/mec.13728

Hill, W. G., & Weir, B. S. (2011). Variation in actual relationship as a consequence of Mendelian sampling and linkage. *Genetics Research*, *93*(1), 47–64. https://doi.org/10.1017/S0016672310000480

Hill, W. G., & Weir, B. S. (2012). Variation in actual relationship among descendants of inbred individuals. *Genetics Research*, *94*(05), 267–274. https://doi.org/10.1017/S0016672312000468

Hoffman, J. I., Simpson, F., David, P., Rijks, J. M., Kuiken, T., Thorne, M. A. S., … Dasmahapatra, K. K. (2014). High-throughput sequencing reveals inbreeding depression in a natural population. *Proceedings of the National Academy of Sciences*, *111*(10), 3775–3780. https://doi.org/10.1073/pnas.1318945111

Hogg, C. J., Ivy, J. A., Srb, C., Hockley, J., Lees, C., Hibbard, C., & Jones, M. (2015). Influence of genetic provenance and birth origin on productivity of the Tasmanian devil insurance population. *Conservation Genetics*, *16*, 1465–1473.

Hogg, C. J., Wright, B., Morris, K. M., Lee, A. V., Ivy, J. A., Grueber, C. E., & Belov, K. (2019). Founder relationships and conservation management: empirical kinships reveal the effect on breeding programmes when founders are assumed to be unrelated. *Animal Conservation*, *22*(4), 348–361.

Ivy, J. A., & Lacy, R. C. (2012). A comparison of strategies for selecting breeding pairs to maximize genetic diversity retention in managed populations. *Journal of Heredity*, *103*(2), 186–196. https://doi.org/10.1093/jhered/esr129

Ivy, J. A., Miller, A., Lacy, R. C., & DeWoody, J. A. (2009). Methods and prospects for using molecular data in captive breeding programs: An empirical example using parma wallabies (*Macropus parma*). *Journal of Heredity*, *100*(4), 441–454. https://doi.org/10.1093/jhered/esp019

Ivy, J. A., Putnam, A. S., Navarro, A. Y., Gurr, J., & Ryder, O. A. (2016). Applying SNP-derived molecular coancestry estimates to captive breeding programs. *Journal of Heredity*, *107*(5), 403–412. https://doi.org/10.1093/jhered/esw029

Jiménez-Mena, B., Schad, K., Hanna, N., & Lacy, R. C. (2016). Pedigree analysis for the genetic management of group-living species. *Ecology and Evolution*, *6*(10), 3067–3078. https://doi.org/10.1002/ece3.1831

Jamieson, I. G. (2015). Significance of population genetics for managing small natural and reintroduced populations in New Zealand. *New Zealand Journal of Ecology*, *39*(1), 1–18.

Jones, A. G., & Ardren, W. R. (2003). Methods of parentage analysis in natural populations. *Molecular Ecology*, *12*(10), 2511–2523. https://doi.org/10.1046/j.1365-294X.2003.01928.x

Kardos, M., Luikart, G., & Allendorf, F. W. (2015). Measuring individual inbreeding in the age of genomics: Marker-based measures are better than pedigrees. *Heredity*, *115*(1), 63–72. https://doi.org/10.1038/hdy.2015.17

Kardos, M., & Shafer, A. B. A. (2019). The peril of gene-targeted conservation. *Trends in Ecology and Evolution*, *33*(11), 827–839.

Kardos, M., Taylor, H. R., Ellegren, H., Luikart, G., & Allendorf, F. W. (2016). Genomics advances the study of inbreeding depression in the wild. *Evolutionary Applications*, *9*, 1205–1218. https://doi.org/10.1111/eva.12414

Kearvell, J. C., & Legault, A. J. (2017). Orange-fronted parakeet. In Miskelly, C.M. (ed.) New Zealand Birds. Retrieved from www.nzbirdsonline.org.nz

Keller, M. C., Visscher, P. M., & Goddard, M. E. (2011). Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data. *Genetics*, *189*(1), 237–249. https://doi.org/10.1534/genetics.111.130922

Kielbasa, S. M., Wan, R., Sato, K., Horton, P., & Frith, M. C. (2011). Adaptive seeds tame genomic sequence comparison. *Genome Research*, *21*(3), 487–493. https://doi.org/10.1101/gr.113985.110

Knief, U., Hemmrich-Stanisak, G., Wittig, M., Franke, A., Griffith, S. C., Kempenaers, B., & Forstmeier, W. (2015). Quantifying realized inbreeding in wild and captive animal populations. *Heredity*, *114*(4), 397–403. https://doi.org/10.1038/hdy.2014.116

Knief, U., Kempenaers, B., & Forstmeier, W. (2017). Meiotic recombination shapes precision of pedigree- and marker-based estimates of inbreeding. *Heredity*, *118*, 239–248. https://doi.org/10.1038/hdy.2016.95

Krueger, F. (2019). *Trim Galore* [Computer Software]. Retrieved from https://github.com/FelixKrueger/TrimGalore.

Lacy, R. C. (1995). Clarification of genetic terms and their use in the management of captive populations. *Zoo Biology*, *14*(6), 565–577. https://doi.org/10.1002/zoo.1430140609

Lacy, R. C. (2009). Stopping evolution: Genetic management of captive populations. In G. Amato, R. DeSalle, O. A. Ryder, & H. C. Rosenbaum (Eds.), *Conservation genetics in the age of genomics* (pp. 58–81). New York: Columbia University Press.

Lacy, R. C. (2012). Extending pedigree analysis for uncertain parentage and diverse breeding systems. *Journal of Heredity*, *103*(2), 197–205. https://doi.org/10.1093/jhered/esr135

Lacy, R. C., Ballou, J. D., & Pollak, J. P. (2012). PMx: Software package for demographic and genetic analysis and management of pedigreed populations. *Methods in Ecology and Evolution*, *3*, 433–437. https://doi.org/10.1111/j.2041-210X.2011.00148.x

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

Lemopoulos, A., Prokkola, J. M., Uusi-Heikkilä, S., Vasemägi, A., Huusko, A., Hyvärinen, P., … Vainikka, A. (2019). Comparing RADseq and microsatellites for estimating genetic diversity and relatedness—Implications for brown trout conservation. *Ecology and Evolution*, *9*(4), 2106–2120.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., & Homer, N. … 1000 Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li, M. H., Strandén, I., Tiirikka, T., Sevón-Aimonen, M.-L., & Kantanen, J. (2011). A Comparison of approaches to estimate the inbreeding coefficient and pairwise relatedness using genomic and pedigree data in a sheep population. *PLoS ONE*, *6*(11), e26256. https://doi.org/10.1371/journal.pone.0026256

Li, N. (2010). *identity: Jacquard condensed coefficients of identity* [R Programme]. Retrieved from https://cran.r-project.org/web/packages/identity/index.html.

Long, S., Dorsey, C., & Boyle, P. (2011). Status of association of zoos and aquariums cooperatively managed populations. *WAZA Magazine*, *12*, 15–18.

Marçais, G., & Kingsford, C. (2011). A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics*, *27*(6), 764–770. https://doi.org/10.1093/bioinformatics/btr011

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, *17*(1), 10–12. https://doi.org/10.14806/ej.17.1.200

McLennan, E. A., Wright, B. R., Belov, K., Hogg, C. J., & Grueber, C. E. (2019). Too much of a good thing? Finding the most informative genetic data set to answer conservation questions. *Molecular Ecology Resources*, *19*(3), 659–671. https://doi.org/10.1111/1755-0998.12997

Mitchell, A. A., Lau, J., Chemnick, L. G., Thompson, E. A., Alberts, A. C., Ryder, O. A., & Gerber, G. P. (2011). Using microsatellite diversity in wild Anegada iguanas (*Cyclura pinguis*) to establish relatedness in a captive breeding group of this critically endangered species. *Conservation Genetics*, *12*(3), 771–781. https://doi.org/10.1007/s10592-011-0184-2

Moore, C. T., Converse, S. J., Folk, M. J., Runge, M. C., & Nesbitt, S. A. (2012). Evaluating release alternatives for a long-lived bird species under uncertainty about long-term demographic rates. *Journal of Ornithology*, *152*(2), 339–353. https://doi.org/10.1007/s10336-010-0592-y

Moraga, R. (2018a). *concatenate_genome.pl* [Custom Perl Script]. Retrieved from https://github.com/sgalla32/concatenate_genome.

Moraga, R. (2018b). *split_bamfile_tasks.pl* [Custom Perl Script]. Retrieved from https://github.com/Lanilen/SubSampler_SNPcaller.

Narum, S. R., Buerkle, C. A., Davey, J. W., Miller, M. R., & Hohenlohe, P. A. (2013). Genotyping-by-sequencing in ecological and conservation genomics. *Molecular Ecology*, *22*(11), 2841–2847. https://doi.org/10.1111/mec.12350

Nietlisbach, P., Keller, L. F., Camenisch, G., Guillaume, F., Arcese, P., Reid, J. M., & Postma, E. (2017). Pedigree-based inbreeding coefficient explains more variation in fitness than heterozygosity at 160 microsatellites in a wild bird population. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1850), 20162763–https://doi.org/10.1098/rspb.2016.2763

Overbeek, A. L., Hauber, M. E., Brown, E., Cleland, S., Maloney, R. F., & Steeves, T. E. (2017). Evidence for brood parasitism in a critically endangered Charadriiform with implications for conservation. *Journal of Ornithology*, *158*(1), 333–337. https://doi.org/10.1007/s10336-016-1375-x

O'Grady, J. J., Brook, B. W., Reed, D. H., Ballou, J. D., Tonkyn, D. W., & Frankham, R. (2006). Realistic levels of inbreeding depression strongly affect extinction risk in wild populations. *Biological Conservation*, *133*, 42–51.

Paradis, E., & Schliep, K. (2018). ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, *35*, 526–528. https://doi.org/10.1093/bioinformatics/bty633

Peakall, R. O. D., & Smouse, P. E. (2006). GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, *6*(1), 288–295.

Pemberton, J. (2004). Measuring inbreeding depression in the wild: The old ways are the best. *Trends in Ecology & Evolution*, *19*(12), 613–615. https://doi.org/10.1016/j.tree.2004.09.010

Pemberton, J. M. (2008). Wild pedigrees: The way forward. *Proceedings of the Royal Society B: Biological Sciences*, *275*(1635), 613–621. https://doi.org/10.1098/rspb.2007.1531

Pierce, R. J. (2013). Black stilt. In C. M. Miskelly (ed.) *New Zealand Birds Online*. Retrieved from www.nzbirdsonline.org.nz.

Premachandra, H. K. A., Nguyen, N. H., & Knibb, W. (2019). Effectiveness of SNPs for parentage and sibship assessment in polygamous yellowtail kingfish *Seriola lalandi*. *Aquaculture*, *499*, 24–31. https://doi.org/10.1016/j.aquaculture.2018.09.022

Putnam, A. S., & Ivy, J. A. (2014). Kinship-based management strategies for captive breeding programs when pedigrees are unknown or uncertain. *Journal of Heredity*, *105*(3), 303–311.

Reed, C. E. M. (1998). *Management plan for captive black stilts*. Wellington, New Zealand: Biodiversity Recovery Unit, Department of Conservation. https://www.doc.govt.nz/globalassets/documents/science-andtechnical/tsop11.pdf

Robertson, B. C., & Gemmell, N. J. (2006). PCR-based sexing in conservation biology: Wrong answers from an accurate methodology? *Conservation Genetics*, *7*(2), 267–271. https://doi.org/10.1007/s10592-005-9105-6

Robinson, S. P., Simmons, L. W., & Kennington, W. J. (2013). Estimating relatedness and inbreeding using molecular markers and pedigrees: The effect of demographic history. *Molecular Ecology*, *22*(23), 5779–5792. https://doi.org/10.1111/mec.12529

Rudnick, J. A., & Lacy, R. C. (2008). The impact of assumptions about founder relationships on the effectiveness of captive breeding strategies. *Conservation Genetics*, *9*(6), 1439–1450. https://doi.org/10.1007/s10592-007-9472-2

Sanders, M. D., & Maloney, R. F. (2002). Causes of mortality at nests of ground-nesting birds in the Upper Waitaki Basin, South Island, New Zealand: A 5-year video study. *Biological Conservation*, *106*(2), 225–236. https://doi.org/10.1016/S0006-3207(01)00248-8

Santure, A. W., Stapley, J., Ball, A. D., Birkhead, T. R., Burke, T., & Slate, J. (2010). On the use of large marker panels to estimate inbreeding and relatedness: Empirical and simulation studies of a pedigreed zebra finch population typed at 771 SNPs. *Molecular Ecology*, *19*(7), 1439–1451. https://doi.org/10.1111/j.1365-294X.2010.04554.x

Slate, J., David, P., Dodds, K. G., Veenvliet, B. A., Glass, B. C., Broad, T. E., & McEwan, J. C. (2004). Understanding the relationship between the inbreeding coefficient and multilocus heterozygosity: Theoretical expectations and empirical data. *Heredity*, *93*(3), 255–265. https://doi.org/10.1038/sj.hdy.6800485

Slotta-Bachmayr, L., Boegel, R., Kaczensky, P., Stauffer, C., & Walzer, C. (2004). Use of population viability analysis to identify management priorities and success in reintroducing Przewalski's horses to southwestern Mongolia. *The Journal of Wildlife Management*, *68*(4), 790–798. https://doi.org/10.2193/0022-541X(2004)068[0790:UOPVAT]2.0.CO;2

Smilkov, D., Thorat, N., Nicholson, C., Reif, E., Viégas, F. B., & Wattenberg, M. (2016). Embedding projector: Interactive visualization and interpretation of embeddings. *ArXiv*, *abs/1611.05469*.

Smith, D. H. V., Converse, S. J., Gibson, K. W., Moehrenschlager, A., Link, W. A., Olsen, G. H., & Maguire, K. (2011). Decision analysis for conservation breeding: Maximizing production for reintroduction of whooping cranes. *The Journal of Wildlife Management*, *75*(3), 501–508. https://doi.org/10.1002/jwmg.88

Smouse, R. P. P., & Peakall, R. (2012). GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics*, *28*(19), 2537–2539.

Soulé, M. E. (1985). What is conservation biology? *BioScience*, *35*(11), 727–734.

Speed, D., & Balding, D. J. (2015). Relatedness in the post-genomic era: Is it still useful? *Nature Reviews Genetics*, *16*(1), 33–44. https://doi.org/10.1038/nrg3821

Spielman, D., Brook, B. W., & Frankham, R. (2004). Most species are not driven to extinction before genetic factors impact them. *PNAS*, *101*(42), 15261–15264.

Steeves, T. E., Hale, M. L., & Gemmell, N. J. (2008). Development of polymorphic microsatellite markers for the New Zealand black stilt (*Himantopus novaezelandiae*) and cross-amplification in the pied stilt (*Himantopus himantopus leucocephalus*). *Molecular Ecology Resources*, *8*(5), 1105–1107. https://doi.org/10.1111/j.1755-0998.2008.02170.x

Steeves, T. E., Maloney, R. F., Hale, M. L., Tylianakis, J. M., & Gemmell, N. J. (2010). Genetic analyses reveal hybridization but no hybrid swarm in one of the world's rarest birds. *Molecular Ecology*, *19*(23), 5090–5100. https://doi.org/10.1111/j.1365-294X.2010.04895.x

Sutton, J. T., Robertson, B. C., & Jamieson, I. G. (2011). Dye shift: A neglected source of genotyping error in molecular ecology. *Molecular Ecology Resources*, *11*(3), 514–520. https://doi.org/10.1111/j.1755-0998.2011.02981.x

Taylor, H. R. (2015). The use and abuse of genetic marker-based estimates of relatedness and inbreeding. *Ecology and Evolution*, *5*(15), 3140–3150. https://doi.org/10.1002/ece3.1541

Taylor, H. R., Kardos, M. D., Ramstad, K. M., & Allendorf, F. W. (2015). Valid estimates of individual inbreeding coefficients from marker-based pedigrees are not feasible in wild populations with low allelic diversity. *Conservation Genetics*, *16*(4), 901–913. https://doi.org/10.1007/s10592-015-0709-1

Tenhumberg, B., Tyre, A. J., Shea, K., & Possingham, H. P. (2004). Linking wild and captive populations to maximize species persistence: Optimal translocation strategies. *Conservation Biology*, *18*(5), 1304–1314. https://doi.org/10.1111/j.1523-1739.2004.00246.x

Thrasher, D. J., Butcher, B. G., Campagna, L., Webster, M. S., & Lovette, I. J. (2018). Double-digest RAD sequencing outperforms microsatellite loci at assigning paternity and estimating relatedness: A proof of concept in a highly promiscuous bird. *Molecular Ecology Resources*, *18*, 953–965.

Tzika, A. C., Remy, C., Gibson, R., & Milinkovitch, M. C. (2009). Molecular genetic analysis of a captive-breeding program: The vulnerable endemic Jamaican yellow boa. *Conservation Genetics*, *10*(1), 69–77. https://doi.org/10.1007/s10592-008-9519-z

Visscher, P. M., Medland, S. E., Ferreira, M. A. R., Morley, K. I., Zhu, G. U., Cornes, B. K., … Martin, N. G. (2006). Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings. *PLoS Genetics*, *2*(3), e41. https://doi.org/10.1371/journal.pgen.0020041

Walters, J. R., Derrickson, S. R., Michael Fry, D., Haig, S. M., Marzluff, J. M., & Wunderle, J. M. (2010). Status of the California Condor (*Gymnogyps californianus*) and efforts to achieve its recovery. *The Auk*, *127*(4), 969–1001.

Wang, J. (2007). Triadic IBD coefficients and applications to estimating pairwise relatedness. *Genetics Research*, *89*(3), 135–153. https://doi.org/10.1017/S0016672307008798

Wang, J. (2011). coancestry: A program for simulating, estimating and analysing relatedness and inbreeding coefficients. *Molecular Ecology Resources*, *11*(1), 141–145. https://doi.org/10.1111/j.1755-0998.2010.02885.x

Wang, J. (2016). Pedigrees or markers: Which are better in estimating relatedness and inbreeding coefficient? *Theoretical Population Biology*, *107*, 4–13. https://doi.org/10.1016/j.tpb.2015.08.006

Wang, J. (2017). Estimating pairwise relatedness in a small sample of individuals. *Heredity*, *119*(5), 302. https://doi.org/10.1038/hdy.2017.52

Waples, R. K., Albrechtsen, A., & Moltke, I. (2019). Allele-frequency-free inference of close familial relationships from genotypes or low-depth sequencing data. *Molecular Ecology*, *28*, 35–48. https://doi.org/10.1111/mec.14954

Whalen, A., Gorjanc, G., & Hickey, J. M. (2019). Parentage assignment with genotyping-by-sequencing data. *Journal of Animal Breeding and Genetics*, *136*(2), 102–112.

Wildt, D., Miller, P., Koepfli, K.-P., Pukazhenthi, B., Palfrey, K., Livingston, G., … Snodgrass, K. (2019). Breeding centers, private ranches, and genomics for creating sustainable wildlife populations. *BioScience*, *69*(11), 928–943. https://doi.org/10.1093/biosci/biz091

Willoughby, J. R., Fernandez, N. B., Lamb, M. C., Ivy, J. A., Lacy, R. C., & DeWoody, J. A. (2015). The impacts of inbreeding, drift and selection on genetic diversity in captive breeding populations. *Molecular Ecology*, *24*(1), 98–110. https://doi.org/10.1111/mec.13020

Wright, S. (1922). Coefficients of inbreeding and relationship. *The American Naturalist*, *56*(645), 330–338.

Zhang, G., Li, B., Li, C., Gilbert, M. T. P., Jarvis, E. D., & Wang, J. (2014). Comparative genomic data of the Avian Phylogenomics Project. *GigaScience*, *3*(1), 26. https://doi.org/10.1186/2047-217X-3-26

Zimin, A. V., Marçais, G., Puiu, D., Roberts, M., Salzberg, S. L., & Yorke, J. A. (2013). The MaSuRCA genome assembler. *Bioinformatics*, *29*(21), 2669–2677. https://doi.org/10.1093/bioinformatics/btt476

SUPPORTING INFORMATION

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

How to cite this article: Galla SJ, Moraga R, Brown L, et al. A comparison of pedigree, genetic and genomic estimates of relatedness for informing pairing decisions in two critically endangered birds: Implications for conservation breeding programmes worldwide. *Evol Appl*. 2020;00:1–18. https://doi.org/10.1111/eva.12916