Ecology and Evolution

The use and abuse of genetic marker-based estimates of relatedness and inbreeding

Helen R. Taylor

Allan Wilson Centre, School of Biological Sciences, Victoria University of Wellington, Kelburn Parade, Wellington, New Zealand

Abstract

Genetic marker-based estimators remain a popular tool for measuring relatedness (r_{xy}) and inbreeding (F) coefficients at both the population and individual level. The performance of these estimators fluctuates with the number and variability of markers available, and the relatedness composition and demographic history of a population. Several methods are available to evaluate the reliability of the estimates of r_{xy} and F, some of which are implemented in the program COANCESTRY. I used the simulation module in COANCESTRY since assess the performance of marker-based estimators of r_{xy} and F in a species with very low genetic diversity, New Zealand's little spotted kiwi (*Apteryx owenii*). I also conducted a review of published papers that have used COANCESTRY as its release to assess whether and how the reliability of the estimates of r_{xy} and F produced by genetic markers are being measured and reported in published studies. My simulation results show that even when the correlation between true (simulated) and estimated r_{xy} or F is relatively high (Pearson's $r = 0.66$–0.72 and 0.81–0.85, respectively) the imprecision of the estimates renders them highly unreliable on an individual basis. The literature review demonstrates that the majority of studies do not report the reliability of marker-based estimates of r_{xy} and F. There is currently no standard practice for selecting the best estimator for a given data set or reporting an estimator's performance. This could lead to experimental results being interpreted out of context and render the robustness of conclusions based on measures of r_{xy} and F debatable.

Introduction

Quantifying the degree of relatedness between individuals within a population is a key to many genetic research topics (Ritland 1996; Lynch and Ritland 1999). Estimates of relatedness have been used widely in studies of gene flow (Morin et al. 1994; Streiff et al. 1999), kin selection and cooperative breeding (Peters et al. 1999; Hatchwell et al. 2014), trait heritability (Kruuk 2004), social behavior and structure (Ward 1983; Laidlaw and Page 1984; Queller et al. 1988), and to manage conservation breeding programs (Jones et al. 2002; Kozfkay et al. 2008; Goncalves da Silva et al. 2010; Bergner et al. 2014), while accurate estimates of individual inbreeding are pivotal to studies of inbreeding depression (e.g., Grueber et al. 2010; Nielsen et al. 2012). The coefficient of relatedness (r_{xy}) measures the expected proportion of shared alleles between pairs of individuals that are identical by descent

(IBD) (Blouin 2003), while an individual's inbreeding coefficient (F) is the probability of IBD of two alleles at a locus in an individual (i.e., the probability they were inherited from a common ancestor) (Wright 1921; Malécot 1948). These metrics can be estimated at an individual level or averaged over populations. Pedigrees are often suggested as the best method for estimating r_{xy} and F (Pemberton 2008; Santure et al. 2010), but this method is problematic for three reasons. First, pedigrees assume unrelated founders, which is rarely the case and can lead to the underestimation of r_{xy} and F (Jones et al. 2002; Russello and Amato 2004). Second, pedigree-based estimates of r_{xy} and F are unable to account for the variance in IBD that occurs by chance between dyads or individuals with the same pedigree-based r_{xy} or F, respectively (Hill and Weir 2011). Finally, the data required for pedigree construction is often lacking for wild populations, especially those that have not been monitored long term,

and inaccurate pedigrees will lead to inaccurate estimates of r_{xy} and F (Pemberton 2008; Jones and Wang 2009).

The coefficients of relatedness and inbreeding can also be estimated directly from genetic markers. This is not a new concept (Morton et al. 1971), but has become more popular with the increasing availability of relatively large panels of microsatellite and, more recently, single nucleotide polymorphism (SNP) markers. Seven widely used relatedness estimators have been developed since the late 1980s. These can be divided into two types: moment estimators (that estimate the relatedness between individuals in terms of probabilities of identity by descent) (Queller and Goodnight 1989; Li et al. 1993; Ritland 1996; Lynch and Ritland 1999; Wang 2002) and likelihood methods (that calculate the probability of individuals falling into a particular relationship given the marker data available) (Anderson and Weir 2007; Wang 2007). It is also possible to calculate F for individuals using the Ritland (1996) or Lynch and Ritland (1999) moment estimators, or the Anderson and Weir (2007) or Wang (2007) likelihood estimators (Wang 2011). The performance of marker-based estimators of r_{xy} and F has been shown to be affected by the relatedness composition of the population in question (Van de Casteele et al. 2001; Csilléry et al. 2006), the number and polymorphism of loci used (Blouin 2003) and the demographic history of a population (Robinson et al. 2013). No one estimator performs best in all scenarios and it is recommended that simulations are conducted to select the most appropriate estimator for a given scenario (Van de Casteele et al. 2001; Wang 2011).

In studies relying on marker-based estimates of r_{xy} and/or F to make inferences regarding biological systems, conducting a priori simulations is important for two reasons: (1) selecting the correct estimator(s) for use with a given marker set, and (2) assessing the likely reliability of any estimates generated. Prior knowledge of these two factors is essential for evaluating the robustness of any conclusions based on marker-based estimates of r_{xy} or F. This is particularly important for species of conservation concern, which often show low genetic diversity as a result of population bottlenecks, rendering markers less informative for estimating r_{xy} and F. Where high-density SNP panels or large numbers of microsatellite markers are not available for species with low genetic diversity, marker-based estimates of r_{xy} or F may be highly biased and/or imprecise, rendering any conclusions based on these estimates potentially unsound (see Van Horn et al. 2008 for an illustration of this). The importance of a priori simulations for studies using marker-based estimators of r_{xy} and F has already been stated several times in the scientific literature (Van de Casteele et al. 2001; Wang 2011; Pew et al. 2014). The program COANCESTRY (Wang 2011)

estimates r_{xy} and F using both moment and likelihood estimators, but also facilitates a priori simulations to assess estimator performance and aid the section of the best estimator for a given data set. The release of this program might be expected to have aided the broad adoption of a priori evaluation of estimators (particularly in studies using COANCESTRY) to ensure that conclusions based on direct marker-based estimates are always as robust as possible. This is of particular concern in cases where recommendations for species management actions are based on such conclusions. However, it is unclear whether these recommendations have been heeded by scientists implementing marker-based estimates of r_{xy} and F.

This study has two aims: (1) to illustrate the necessity of using a priori simulations to thoroughly evaluate performance of marker-based estimators of r_{xy} and F at a population and individual level, and (2) to quantify how often a priori simulations are used to assess estimator performance and select the best estimator of r_{xy} and F in scientific studies. As such, this study is divided into two sections. First, I conduct simulations using empirical allele frequencies to select the best estimator and evaluate the performance of marker-based estimators of r_{xy} and F in a species with very low genetic diversity, New Zealand's little spotted kiwi (*Apteryx owenii*) (LSK). Second, I review the scientific literature for studies that have used COANCESTRY to estimate r_{xy} and/or F and ask whether and how researchers select specific estimators and evaluate the power of the genetic marker sets available to them. Specifically I ask: (1) Can accurate estimates of r_{xy} and/or F be generated via marker-based estimators for species with very low genetic diversity, such as LSK? (2) How are marker-based estimators of r_{xy} and F selected by the studies that use them? (3) Is the likely accuracy of marker-based estimates of r_{xy} and/or F assessed by the majority of studies employing them?

Methods

Marker-based relatedness and inbreeding simulations

Study species

Little spotted kiwi are a flightless, nocturnal ratite endemic to New Zealand. Although once widespread throughout New Zealand, all mainland LSK populations had been extirpated by introduced predators by the late 1980s. The species survived solely due to a successful population on Kapiti Island that was founded by, at most, five birds in 1912 (Ramstad et al. 2013). Between 1982 and 2010, individuals from Kapiti Island were translocated to found new LSK populations on six other islands and in one

mainland island sanctuary, leading to a current population of ~1700 birds (H. Robertson, unpubl. data.). The new populations were founded with between two and 40 individuals and these secondary bottlenecks, combined with the original bottleneck of ≤5, have left LSK with very low genetic diversity and an elevated risk of inbreeding (Ramstad et al. 2013). Thus, measuring r_{xy} and F within LSK populations is of interest to the future protection and management of this species, but is rendered challenging due to a lack of pedigree information for any LSK population. As there are currently 21 polymorphic microsatellite markers characterized for LSK and two of the extant populations (those on Long Island and in Zealandia ecosanctuary) have been extensively sampled for DNA, marker-based estimates of r_{xy} and F represent a potentially useful tool for the management of this species.

Simulations

Simulations in the program COANCESTRY were used to select the best estimator from the seven implemented in the program and to assess the performance of this best estimator using empirical allele frequencies across 21 microsatellite markers from the Long Island and Zealandia populations of LSK. Simulated populations for relatedness testing consisted of 600 dyads spread equally across six categories of relatedness: parent–offspring ($r_{xy} = 0.5$), full siblings ($r_{xy} = 0.5$), half siblings/avuncular/grandparent–grandchild ($r_{xy} = 0.25$), first cousins ($r_{xy} = 0.125$), second cousins ($r_{xy} = 0.03125$), and unrelated ($r_{xy} = 0$). For inbreeding coefficients, the simulated data set consisted of 2100 individuals with inbreeding coefficients that varied from 0 to 1 at intervals of 0.05, and 100 individuals in each inbreeding category. Both approaches were modeled after those taken by Brekke et al. (2010). In both relatedness and inbreeding simulations, the allele frequencies, missing data and error rates for simulated microsatellite loci genotypes were based on those found in two different LSK populations: Long Island and Zealandia (Table S1). These populations were selected because they were founded with the lowest (two) and highest (40) numbers of individuals and thus represent the minimum and maximum amount of genetic diversity for any recently translocated LSK populations. Both populations have also been subject to extensive genetic sampling, with genotypes available for 43 Long Island and 113 Zealandia birds (86 and 94% of the current estimated population sizes, respectively).

Estimates produced by the triadic likelihood (TrioML) method were the most closely correlated with the simulated true relatedness and inbreeding coefficients for both the Long Island and Zealandia marker sets. Thus, estimates from the other six estimators were excluded from further analysis. Wilcoxon sign rank tests were used to test for significant differences between simulated and TrioML-estimated relatedness values, and the coefficient of variation (Abdi 2010) was calculated for estimates of each category of relatedness. Linear regression of simulated inbreeding coefficients against TrioML estimates of inbreeding coefficients was conducted to assess the bias and precision of estimates of F. All statistical analyses were conducted in R (R Development Core Team 2013).

Literature review

A Web of Science search was conducted in September 2014 for all papers that had cited COANCESTRY since its publication in 2011. Marker-based estimators have been in use for several decades prior to the release of COANCESTRY, but COANCESTRY facilitates simulation-based evaluation of the discriminatory power of the markers being used. Thus, it was of interest to quantify which studies using COANCESTRY to estimate relatedness or inbreeding had also used it to select the best estimator and determine the likely performance of that estimator for their data set. This search resulted in a total of 82 peer-reviewed publications, seven of which were species specific or wider topic reviews, or theoretical papers and were thus removed from further analysis. The remaining 75 papers were analysed to determine the purpose of each study and the metrics being estimated, the type and number of molecular markers used to estimate the relevant metrics, and the use of simulations to select appropriate estimators and assess their reliability. Papers are not always listed on Web of Science immediately on publication. Thus, it is acknowledged that some recent studies using COANCESTRY may have been omitted from this review (e.g., Bergner et al.'s study on relatedness in kākāpō (*Strigops habroptilus*) (2014)). A full list of all the studies reviewed can be found in the Supplementary Information for this paper (Table S2).

Results

Marker-based relatedness and inbreeding simulations

In general, marker-based estimators of both coefficients performed worse when using the empirical allele frequencies available for Long Island than those for Zealandia (Table 1, Figs. 1 and 2). However, even in the more variable Zealandia population, estimates of r_{xy} and F ranged widely, particularly for distantly related individuals and those that were in the middle range of inbreeding coefficients simulated (Figs. 1 and 2).

Table 1. Differences between relatedness coefficients estimated using TrioML in COANCESTRY and true r_{xy} for simulated dyads in six relationship categories. Simulated genotypes of dyads were based on the Long Island and Zealandia microsatellite marker sets.

Allele frequencies used	True relationship	Actual r_{xy}	TrioML mean estimated r_{xy} (±95% CIs)	Wilcoxon V	P	Coefficient of variation
Long Island	Parent–offspring	0.5	0.48 (±0.02)	1059	NS	20%
	Full siblings	0.5	0.48 (±0.03)	2116	NS	31%
	Half siblings	0.25	0.25 (±0.04)	2412	NS	70%
	First cousins/avuncular	0.125	0.21 (±0.04)	1603	***	89%
	Second cousins	0.01325	0.14 (±0.03)	1028	***	117%
	Unrelated	0	0.15 (±0.04)	1891	***	124%
Zealandia	Parent-offspring	0.5	0.47 (±0.02)	914	**	23%
	Full siblings	0.5	0.46 (±0.03)	2371	*	37%
	Half siblings	0.25	0.25 (±0.04)	2490	NS	72%
	First cousins/avuncular	0.125	0.16 (±0.03)	2145	NS	98%
	Second cousins	0.01325	0.13 (±0.03)	1031	***	115%
	Unrelated	0	0.08 (±0.02)	2145	***	144%

* = $P < 0.05$. ** = $P < 0.01$. *** = $P < 0.001$. NS = Not significant.
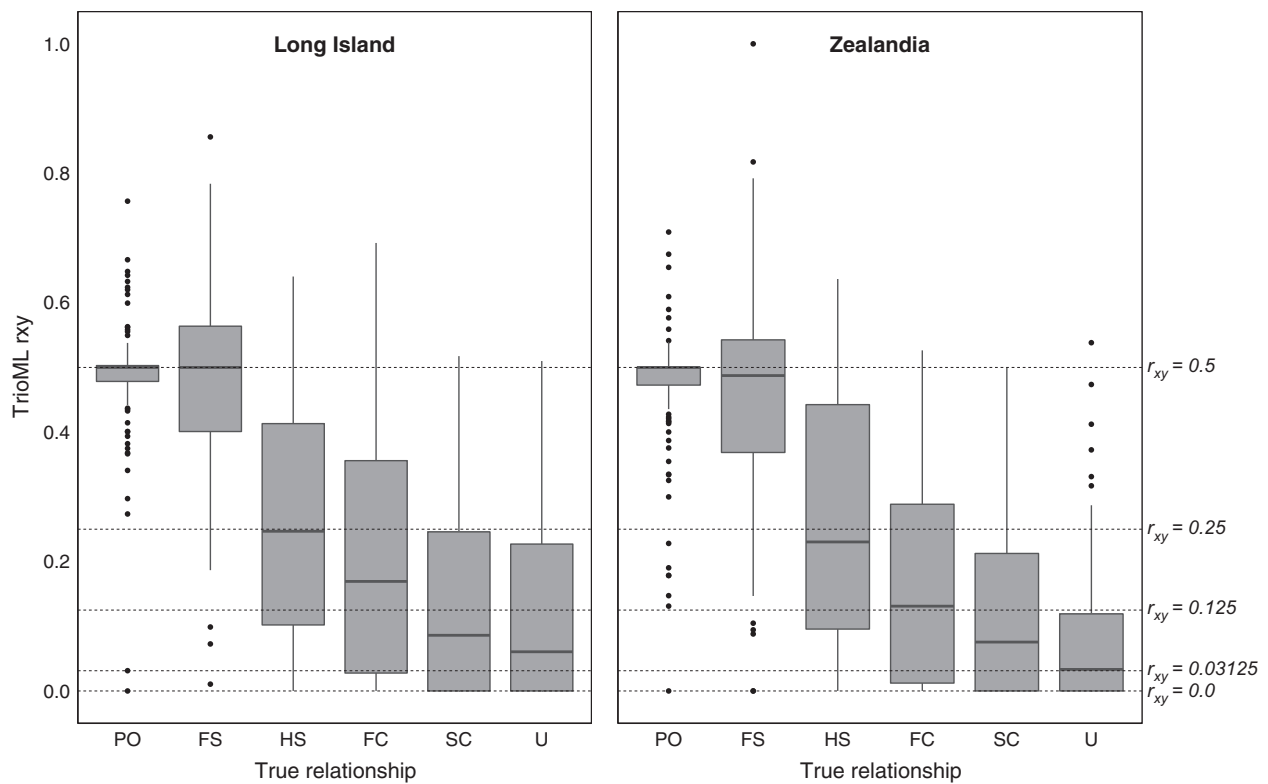
Figure 1. Spread of relatedness coefficients estimated by TrioML in COANCESTRY for simulated dyads in six true relationship categories using simulated genotypes based on the Long Island and Zealandia microsatellite marker sets. Boxes represent the upper and lower quartiles, divided by the median. The 10 and 90 percent quartiles are depicted by lines and dots represent the outliers. Dashed horizontal lines mark true r_{xy} coefficients of 0.5 (parents-offspring (PO) and full siblings (FS)), 0.25 (half siblings/avuncular/grandparent-grandchild (HS)), 0.125 (first cousin (FC)), 0.01325 (second cousin (SC)) and 0 (unrelated (U)).

COANCESTRY reported Pearson's correlation coefficients of 0.66 and 0.72 between TrioML-estimated and simulated values of r_{xy} for Long Island and Zealandia, respectively. In spite of these relatively strong correlations, estimates of r_{xy} ranged widely in all six categories of kinship tested (Fig. 1 and Table 1). For the Long Island simula-

tions, mean estimates of r_{xy} were not significantly different from simulated values for parent–offspring, full-sibling and half-sibling dyads, but were significantly different for first cousin, second cousin, and unrelated dyads (Table 1). When the Zealandia allele frequencies were used in simulations, mean r_{xy} estimates for half sibling dyads and first cousin dyads were not significantly different from the simulated values, but those for all four other categories were (Table 1). For both populations, the coefficient of variation (CV) of estimates increased with decreasing simulated r_{xy} (Table 1). CV ranged from 20% (Long Island) and 23% (Zealandia) for parent–offspring dyads to 124% (Long Island) and 144% (Zealandia) for unrelated dyads.

Pearson's correlation coefficients for TrioML-estimated and simulated values of F were 0.81 for Long Island and 0.85 for Zealandia. Again, in each case, TrioML estimates of inbreeding coefficients varied widely for a given true value of F (Fig. 2). This variation was greatest between simulated F values of ~0.4–0.65, regardless of which population's allele frequencies were used, with less variation in estimates of F for individuals with very low (0–0.1) or very high (0.9–1.0) F. Linear regression analyses indicated a slight bias in estimates of F for both populations (Long Island $\beta = 0.85$; Zealandia $\beta = 0.88$) with marker-based measures tending to overestimate F when simulated F was low and underestimate F when it was high (Fig. 2). Zero bias only occurred at simulated F of ~0.4–0.5 (Fig. 2).

Linear regression also illustrated a lack of precision of marker generated estimates of F for both populations. Although precision was higher for the Zealandia population ($r^2 = 0.67$) than the Long Island population ($r^2 = 0.61$), neither set of estimates were especially precise (Fig. 2).

Literature review

A total of 75 papers citing COANCESTRY were reviewed for this study (Table S2). These papers spanned 33 different peer-reviewed journals and included studies on species of mammals (41%), birds (19%), fish (9%), insects (9%), reptiles (8%), plants (8%), gastropods (1%), amphibians (1%), maxillopods (1%), and malacostracans (1%). The majority (77%) of papers that have cited COANCESTRY to date have been solely concerned with estimating r_{xy}, with 12% of studies focussed on estimating just F, and 9% on a combination of the two metrics. Two of the papers used COANCESTRY to check for identical individuals in a sample or to simulate a set of individual genotypes rather than estimating r_{xy} or F and were discarded from further analysis, leaving 73 papers. A total of 23% of the papers reviewed estimated r_{xy} of dyads and/or F of individuals, 53% estimated population means, and 23% estimated both.

The purpose of estimating r_{xy} and/or F varied widely from study to study. Only three papers (4%) used mar-
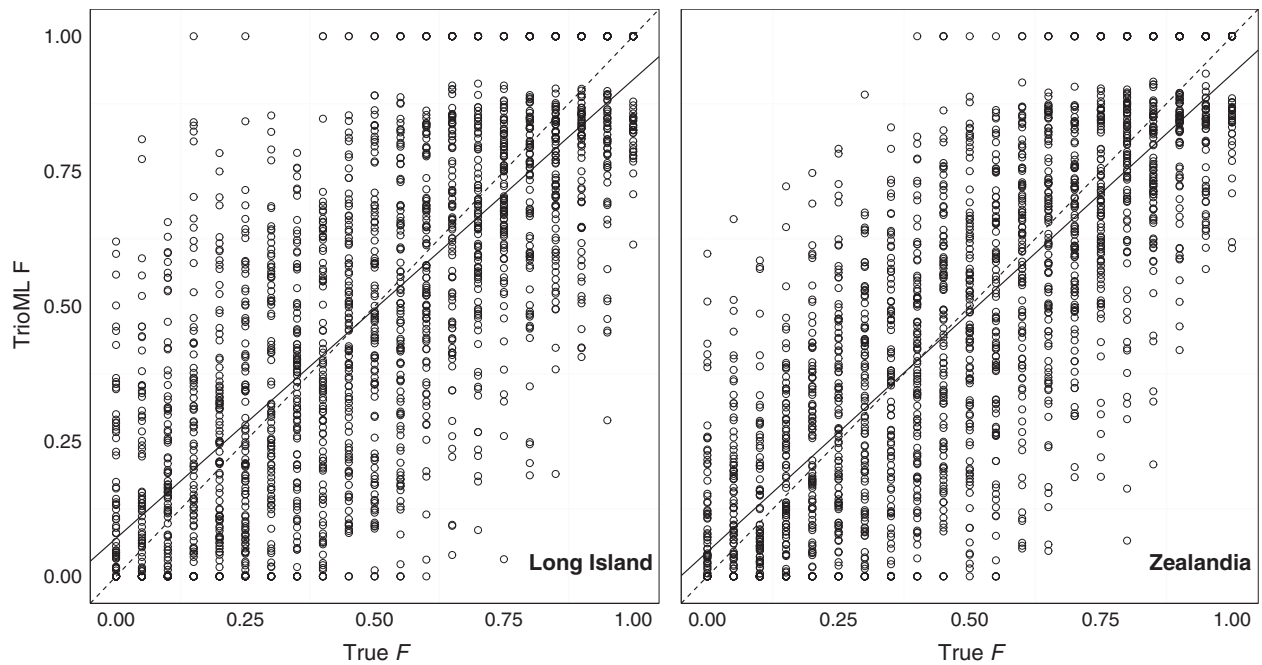
Figure 2. Regression line (solid) versus 1:1 line (dashed) for regressions of inbreeding coefficients estimated using TrioML in COANCESTRY versus true inbreeding coefficients for simulated individuals with genotypes based on the Long Island and Zealandia microsatellite marker sets. Long Island $\beta = 0.85$, $r^2 = 0.61$, $F = 3260$, $P < 0.001$. Zealandia $\beta = 0.88$, $r^2 = 0.67$, $F = 4257$, $P < 0.001$.

ker-based estimates to detect inbreeding depression. Other uses included assessing social organization, excluding related individuals from downstream analyses, investigating co-operative breeding, detecting sibling-related cannibalism, and facilitating the development of SNP panels. Microsatellite markers were the most commonly used tool for estimating r_{xy} and/or F (96% of studies), with anywhere from 4 to 33 markers implemented.

The vast majority (95%) of studies reported which of the seven available estimators was used. However, only 28% of the studies that reported the estimator justified their choice in any respect and even fewer (14%) used a simulation-based approach to select the appropriate estimator for the markers used in the study. Of those that did conduct simulations to select an estimator, 70% reported the performance of their chosen estimator in some form. This means that of 73 studies estimating relatedness and/or inbreeding directly using genetic markers, 9% (seven papers) tested and reported the performance of the implemented estimator. These seven studies variously used Pearson's r (four studies), the raw variance of estimates (one study), the r^2 of estimates (one study), and statistical power (PW_R) calculated in the program KinInfor (Wang 2006) (one study). Of the 22 studies that estimated individual r_{xy} and/or F (alongside or instead of population means), 91% specified the estimator used. Of these, 32% justified their choice of estimator, 19% used a simulation-based approach to select the best estimator for the marker set available, and 13% (four papers) reported the performance of their chosen estimator.

Discussion

The simulation results presented here underline the importance of conducting a priori tests of estimator accuracy for a given marker set before estimating r_{xy} and F. They also illustrate that, even when relatively strong correlations between true and estimated values of r_{xy} and F are predicted, individual estimates can still vary widely for a given value, causing precision to be low. In light of this, it is troubling that so many papers reviewed here failed to investigate and report the performance of their selected estimator of r_{xy} or F. Unless standard practice for selecting and assessing the performance of marker-based estimators of r_{xy} and F is implemented, the conclusions based on such estimates remain open to debate and should be treated with caution.

Using marker-based estimates of r_{xy} and F in little spotted kiwi

In the relatedness simulation tests using empirical allele frequencies and the existing LSK microsatellite marker set, the precision of estimates decreased in tandem with relatedness. This is likely due to the increase in variance of IBD between dyads with the same r_{xy} as relatedness decreases. It is thought that, on average, more than twice as many loci are required to precisely discriminate second degree relatives from unrelated dyads than first degree (Blouin 2003). Variance in IBD is likely also the cause of the higher precision seen here for parent–offspring dyads versus full-sibling dyads. Although average r_{xy} for both kinds of dyads is 0.5, the pattern of allele sharing is different. Offspring will almost always inherit 50% of their alleles from each parent whereas full siblings will share 50% of their alleles on average, but with more variance from dyad to dyad (Weir et al. 2006).

All marker-based estimators measure r_{xy} and/or F relative to a reference population, which is assumed to contain noninbred and unrelated individuals. In the majority of studies (as here), the current population also serves as the reference, resulting in relatedness being estimated relative to all individuals in the sample rather than to an separate unrelated sample (Wang 2014). When marker diversity is low, there will be little difference in the genetic similarity of unrelated and highly related individuals due to increased identity by state (IBS). Thus, for both populations, estimates of r_{xy} were (on average) underestimates of closely related dyads, overestimates of loosely or unrelated dyads and there was a roughly equal amount of under and overestimation for dyads with a true r_{xy} of 0.25, which is halfway between the minimum and maximum level of relatedness in the simulated population. A similar pattern was seen in the bias of individual inbreeding estimates.

The microsatellite markers currently available for LSK have low power to directly estimate pairwise relatedness or individual inbreeding, even in Zealandia, which has some of the highest allelic diversity of any LSK population (Taylor 2014). The relatively high allelic diversity in the Zealandia population resulted in a closer overall correlation between estimated and simulated values of pairwise relatedness and individual inbreeding than that seen for Long Island. The overall bias of relatedness and inbreeding estimates was not severe for either marker set, but the variability in estimates of either metric would render estimates highly unreliable at an individual level, especially with the relatively small sample sizes available for this and many other threatened species. Marker-based estimates of relatedness and inbreeding are expected to be more robust when averaged over large numbers of individuals (Rollins et al. 2012); thus, these estimators could potentially be used to generate mean values of r_{xy} and F for LSK populations, but not for dyads or individuals. Conclusions regarding inbreeding depression using individual estimates generated in this fashion would be highly

questionable. It would also be unadvisable to use such estimates to select individuals for use in translocation-based management. The poor performance of the estimators in LSK at an individual level is not immediately apparent when viewing COANCESTRY generated Pearson's r statistic in isolation. This highlights the need for a more comprehensive assessment of estimator performance, especially when the intention is to estimate r_{xy} for individual dyads or F for individuals.

The low reliability of marker-based estimates of r_{xy} and F at the individual level in LSK is unfortunate as there is currently no pedigree information for any population of this species. As the extant population of LSK is descended from, at most, five birds (Ramstad et al. 2013) and all eight subpopulations within this species were founded with between two and 40 individuals, inbreeding depression and the selection of individuals for future translocations are topics of interest for ongoing management of LSK. However, the lack of pedigree data and extremely low genetic variation seen in LSK will render attempts to accurately quantify r_{xy} or F in this species challenging. The scenario exhibited by LSK is not uncommon, with many wild populations lacking pedigree data (Pemberton 2008) and many threatened species exhibiting low genetic variation (e.g., Haig and Avise 1996; Leonard et al. 2005; Schultz et al. 2009; Miller et al. 2011; Chen et al. 2012). Clearly, new tools are required to tackle the issues of estimating r_{xy} and F in such species; high-density SNP panels (e.g., Santure et al. 2010; Saura et al. 2013) and runs of homozygosity (e.g., McQuillan et al. 2008; Purfield et al. 2012; Prado-Martinez et al. 2013) are currently the most promising new techniques, and these will become more feasible for nonmodel threatened species as the cost of next-generation sequencing continues to fall. However, as the review conducted here illustrates, microsatellites are still the prevalent molecular tool in use for estimating r_{xy} and F, and assessing and reporting the power of these marker sets for estimating r_{xy} and F remains an important issue.

Current uses and reliability assessments for marker-based estimators of r_{xy} and F

Marker-based estimates of r_{xy} and F are in use across a variety of study areas in a diverse array of taxa. The results from the COANCESTRY-based literature review show that, currently, genetic markers are more often used to estimate r_{xy} than F. This is possibly due to the fact that COANCESTERY is promoted as a relatedness estimation tool, with inbreeding estimates as an added bonus, and the fact that pedigree analysis is still widely encouraged as the best way of estimating F. As a result, marker-based estimates of F from COANCESTRY are not currently used

extensively to detect inbreeding depression, but COANCESTRY is used heavily in behavioral studies to estimate r_{xy}.

The results from the LSK simulation experiment show the importance of proper estimator selection and a priori assessment of estimator performance in more depth than that currently facilitated by COANCESTRY. However, the literature review data illustrate that these procedures are not being followed or their results not being reported in the majority of studies using molecular markers to estimate r_{xy} and F. In some cases, the performance of the estimator was reported, but found to be low, and this was not discussed in terms of the validity of the conclusions formed (e.g., Hammerly et al. 2013). The fact that studies involving marker-based estimates of r_{xy} and F are sometimes published without stating the estimator used is surprising as it reduces the repeatability of the study. When estimators are reported, the methods of justifying their selection are wide ranging. Only 10 of the 73 studies reviewed stated that the estimator used was selected due to it outperforming the other available estimators in simulation tests. Other studies went as far as to research the best estimator based on previous studies or made generic statements regarding the performance of the estimator in different circumstances, but did not assess that estimator based on their own markers – a potentially critical error.

When estimator performance was reported, the methods of doing so also varied, with five measures employed across seven studies. Rollins et al. (2012) used a combination of Pearson's r correlation between COANCESTRY and pedigree estimates to select the best estimator and PW_R calculated in the program KinInfor (Wang 2006) to quantify the power of their marker set. This was one of the most comprehensive evaluation procedures undertaken in any of the studies reviewed and could potentially form the basis for a standardized protocol for estimator evaluation and selection. Vangestel et al. (2011) used simulations and Pearson's r to select the best estimator for r_{xy} and feature a figure similar to Figure 1 illustrating the variation in their simulation estimates. Bonin et al. (2012) used the variance of different estimators in simulations to select the appropriate method. There is currently no apparent agreed upon best practice for reporting estimator performance and this reduces the comparability of studies. A standard measure of marker power/reliability of estimates should be adopted – even if it is the potentially misleading Pearson's r correlation, but a thorough approach such as that of Rollins et al. (2012) would be preferable. As illustrated by the LSK example presented here, metrics that encompass variation in estimates at an individual level are particularly important for studies attempting to estimate r_{xy} and F for dyads and individu-

als, respectively. The program KinInfor offers several metrics for assessing the informativeness of a given marker set and these, in tandem with a simulation approach and examination of correlations between true and estimated values plus the variance of estimates could provide a more reliable and repeatable approach. More recently, an R implementation of COANCESTRY called *related* has been developed (Pew et al. 2014). This package not only retains the original simulation functions of CONACETRY, but also outputs boxplots comparing the performance of four commonly used r_{xy} estimators (Queller and Goodnight 1989; Li et al. 1993; Lynch and Ritland 1999; Wang 2002) across relatedness values. This is designed to make it even easier for researchers to reliably assess likely estimator performance and select the optimum estimator for their chosen data set.

Even with large numbers of markers, marker-based estimators are more suitable for calculating population-wide mean estimates of r_{xy} than for individual dyads (Santure et al. 2010). Indeed, pairwise relatedness estimators were never intended to classify pairs of individuals into discrete categories (Csilléry et al. 2006). In light of this, it is reassuring that the majority of studies reviewed here only use COANCESTRY to estimate mean values of r_{xy} and F across groups of individuals. This certainly does not negate the need to assess marker power and reliability of estimates, but it means that, in general, the mean estimates presented in these studies should be more reliable than those for dyads or individuals. However, within the 34 papers that did estimate r_{xy} and F on an individual basis, only four reported the power of the marker set or likely reliability of the estimates used, placing the conclusions of the remaining papers in doubt, especially given that some of these papers used as few as 7–9 microsatellites to generate their estimates.

A final issue addressed in only one of the reviewed papers (Domingos et al. 2014) is that of the reference population used when estimating relatedness using genetic markers (Wang 2014). As the estimator can only calculate relatedness based on the individuals available, the resulting estimates are relative to the reference population. Thus, if an unrelated reference population is used in r_{xy} estimation (as in Domingos et al.), then the results will be far closer to reality than if, as is often the case (as here in the LSK example), the current sample also acts as the reference (Wang 2014). This is analogous to the issue of assuming pedigree founders are unrelated when calculating pedigree inbreeding coefficients – it is usually untrue and will likely lead to underestimation of inbreeding (Jones et al. 2002; Russello and Amato 2004). If true relatedness in the reference population is high, then marker-based estimates of r_{xy} will be downwardly biased; marker-based estimates are relative rather than absolute measures and this should be acknowledged in the studies that use them.

Conclusion

Little spotted kiwi are an excellent example for demonstrating that microsatellite markers are not always sufficient to reliably estimate r_{xy} and F at an individual level in threatened species with low genetic diversity. High-density SNP panels and whole genome sequencing will go some way to addressing the issues of reliability in marker-based estimators, but even large panels of SNPs cannot always reliably estimate r_{xy} or F on an individual basis (Santure et al. 2010). The review conducted here illustrates that, currently, microsatellite markers remain the dominant tool for estimating relatedness and inbreeding. This is particularly true in conservation studies where funds are often scant and behavioral studies where genetics is not usually the main focus of the study. In both these cases, it is still more prudent financially to use an existing microsatellite marker set rather than invest in a SNP discovery process.

The use of simulation-based approaches to select the correct estimator for r_{xy} and F and to assess the likely reliability of that estimator given the available marker set has been recommended repeatedly (Van de Casteele et al. 2001; Wang 2011). Several papers have demonstrated useful simulation approaches for assessing the power of a marker set to estimate r_{xy} and F (Rollins et al. 2012; Robinson et al. 2013) and software now exists that facilitates such exploration for naïve users. In spite of this, results from marker-based r_{xy} and F estimators are regularly being used to form conclusions and recommendations without the reliability of these estimates being assessed. When they are assessed, it is often via a simple correlation metric which, as the LSK example presented here shows, can be misleading as to the reliability of estimates on an individual basis. With the absence of pedigrees for many wild populations, marker-based estimates are likely to remain popular and have the potential to be useful, especially as more markers become available. However, in order for the studies employing marker-based estimators to draw robust conclusions that withstand close scrutiny, a standard practice for evaluating these techniques must be implemented and adhered to.

Acknowledgments

manuscript. This study was funded by the Allan Wilson Centre, Victoria University of Wellington, the Centre for Biodiversity and Restoration Ecology, the New Zealand Ministry for Business, Innovation and Employment, and Kaitiaki o Kapiti Trust.

Conflict of Interest

The author declares no conflict of interest.

References

Abdi, H. 2010. Coefficient of Variation. Pp. 170–172 *in* N. Salkind, ed. Encyclopedia of research design. Sage, Thousand Oaks, CA.

Anderson, A. D., and B. S. Weir. 2007. A maximum-likelihood method for the estimation of pairwise relatedness in structured populations. Genetics 176:421–440.

Bergner, L. M., I. G. Jamieson, and B. C. Robertson. 2014. Combining genetic data to identify relatedness among founders in a genetically depauperate parrot, the Kakapo (*Strigops habroptilus*). Conserv. Genet. 15:1013–1020.

Blouin, M. S. 2003. DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. Trends Ecol. Evol. 18:503–511.

Bonin, C. A., M. E. Goebel, G. M. O'Corry-Crowe, and R. S. Burton. 2012. Twins or not? Genetic analysis of putative twins in Antarctic fur seals, *Arctocephalus gazella*, on the South Shetland Islands. J. Exp. Mar. Biol. Ecol. 412:13–19.

Brekke, P., P. M. Bennett, J. Wang, N. Pettorelli, and J. G. Ewen. 2010. Sensitive males: inbreeding depression in an endangered bird. Proc. R. Soc. B 277:3677–3684.

Chen, S.-Y., Y.-J. Zhang, X.-L. Wang, J.-Y. Sun, Y. Xue, P. Zhang, et al. 2012. Extremely low genetic diversity indicating the endangered status of *Ranodon sibiricus* (Amphibia: Caudata) and implications for phylogeography. PLoS ONE 7:e33378.

Csilléry, K., T. Johnson, D. Beraldi, T. Clutton-Brock, D. Coltman, B. Hansson, et al. 2006. Performance of marker-based relatedness estimators in natural populations of outbred vertebrates. Genetics 173:2091–2101.

Development Core Team, R. 2013. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Domingos, J. A., C. Smith-Keune, and D. R. Jerry. 2014. Fate of genetic diversity within and between generations and implications for DNA parentage analysis in selective breeding of mass spawners: a case study of commercially farmed barramundi, *Lates calcarifer*. Aquaculture 424–425:174–182.

Goncalves da Silva, A., D. R. Lalonde, V. Quse, A. Shoemaker, and M. A. Russello. 2010. Genetic approaches refine ex situ lowland tapir (*Tapirus terrestris*) conservation. J. Hered. 101:581–590.

Grueber, C. E., R. J. Laws, S. Nakagawa, and I. G. Jamieson. 2010. Inbreeding depression accumulation across life-history stages of the endangered takahe. Conserv. Biol. 24:1617–1625.

Haig, S. M., and J. C. Avise. 1996. Avian conservation genetics. Pp. 160–189 *in* J. C. Avise and J. L. Hamrick, eds. Conservation genetics. Chapman & Hall, New York.

Hammerly, S. C., M. E. Morrow, and J. A. Johnson. 2013. A comparison of pedigree- and DNA-based measures for identifying inbreeding depression in the critically endangered Attwater's Prairie-chicken. Mol. Ecol. 22:5313–5328.

Hatchwell, B. J., P. R. Gullett, and M. J. Adams. 2014. Helping in cooperatively breeding long-tailed tits: a test of Hamilton's rule. Philos. Trans. R. Soc. B-Biol. Sci. 369:20130565.

Hill, W. G., and B. S. Weir. 2011. Variation in actual relationship as a consequence of Mendelian sampling and linkage. Genet. Res. 93:47–64.

Jones, O. R., and J. Wang. 2009. Molecular marker-based pedigrees for animal conservation biologists. Anim. Conserv. 13:26–34.

Jones, K. L., T. C. Glenn, R. C. Lacy, J. R., Pierce, N. Unruh, C. M. Mirande, and F. Chavez-Ramirez. 2002. Refining the Whooping Crane studbook by incorporating microsatellite DNA and leg-banding analyses. Conserv. Biol. 16:789–799.

Kozfkay, C. C., M. R. Campbell, J. A. Heindel, D. J. Baker, P. Kline, M. S. Powell, and T. Flagg. 2008. A genetic evaluation of relatedness for broodstock management of captive, endangered Snake River sockeye salmon, *Oncorhynchus nerka*. Conserv. Genet. 9:1421–1430.

Kruuk, L. E. B. 2004. Estimating genetic parameters in natural populations using the 'animal model'. Philos. Trans. R. Soc. Lond. B-Biol. Sci. 359:873–890.

Laidlaw, H. H., and R. E. Page. 1984. Polyandry in honey bees (*Apis mellifera* L): sperm utilization and intracolony genetic relationships. Genetics 108:985–997.

Leonard, J. A., C. VilÀ, and R. K. Wayne. 2005. Legacy lost: genetic variability and population size of extirpated US grey wolves (*Canis lupus*). Mol. Ecol. 14:9–17.

Li, C. C., D. E. Weeks, and A. Chakravarti. 1993. Similarity of DNA fingerprints due to chance and relatedness. Hum. Hered. 43:45–52.

Lynch, M., and K. Ritland. 1999. Estimation of pairwise relatedness with molecular markers. Genetics 152:1753–1766.

Malécot, G. 1948. Les Mathématiques de l'hérédité. Masson et Cie, Paris.

McQuillan, R., A. L. Leutenegger, R. Abdel-Rahman, C. S. Franklin, M. Pericic, L. Bara-Lauc, et al. 2008. Runs of homozygosity in European populations. Am. J. Hum. Genet. 83:359–372.

Miller, W., V. M. Hayes, A. Ratan, D. C. Petersen, N. E. Wittekindt, J. Miller, et al. 2011. Genetic diversity and population structure of the endangered marsupial *Sarcophilus harrisii* (Tasmanian devil). Proc. Natl Acad. Sci. 108:12348–12353.

Morin, P. A., J. J. Moore, R. Chakraborty, L. Jin, J. Goodal, and D. S. Woodruff. 1994. Kin selection, social-structure, gene flow, and the evolution of chimpanzees. Science 265:1193–1201.

Morton, N. E., S. Yee, D. E. Harris, and R. Lew. 1971. Bioassay of kinship. Theor. Popul. Biol. 2:507–524.

Nielsen, J. F., S. English, W. P. Goodall-Copestake, J. L. Wang, C. A. Walling, A. W. Bateman, et al. 2012. Inbreeding and inbreeding depression of early life traits in a cooperative mammal. Mol. Ecol. 21:2788–2804.

Pemberton, J. M. 2008. Wild pedigrees: the way forward. Proc. R. Soc. B 275:613–621.

Peters, J. M., D. C. Queller, V. L. Imperatriz-Fonseca, D. W. Roubik, and J. E. Strassmann. 1999. Mate number, kin selection and social conflicts in stingless bees and honeybees. Proc. R. Soc. B Biol. Sci. 266:379–384.

Pew, J., P. H. Muir, J. Wang, and T. R. Frasier. 2014. Related: an R package for analysing pairwise relatedness from codominant molecular markers. Mol. Ecol. Resour. 15:557–561.

Prado-Martinez, J., I. Hernando-Herraez, B. Lorente-Galdos, M. Dabad, O. Ramirez, C. Baeza-Delgado, et al. 2013. The genome sequencing of an albino Western lowland gorilla reveals inbreeding in the wild. BMC Genom. 14: 363.

Purfield, D. C., D. P. Berry, S. McParland, and D. G. Bradley. 2012. Runs of homozygosity and population history in cattle. BMC Genet. 13:70.

Queller, D. C., and K. F. Goodnight. 1989. Estimating relatedness using genetic markers. Evolution 43:258–275.

Queller, D. C., J. E. Strassmann, and C. R. Hughes. 1988. Genetic relatedness in colonies of tropical wasps with multiple queens. Science 242:1155–1157.

Ramstad, K. M., R. M. Colbourne, H. A. Robertson, F. W. Allendorf, and C. H. Daugherty. 2013. Genetic consequences of a century of protection: serial founder events and survival of the little spotted kiwi (*Apteryx owenii*). Proc. R. Soc. B 280:20130576.

Ritland, K. 1996. Marker-based method for inferences about quantitative inheritance in natural populations. Evolution 50:1062–1073.

Robinson, S. P., L. W. Simmons, and W. J. Kennington. 2013. Estimating relatedness and inbreeding using molecular markers and pedigrees: the effect of demographic history. Mol. Ecol. 22:5779–5792.

Rollins, L. A., L. E. Browning, C. E. Holleley, J. L. Savage, A. F. Russell, and S. C. Griffith. 2012. Building genetic networks using relatedness information: a novel approach for the estimation of dispersal and characterization of group structure in social animals. Mol. Ecol. 21:1727–1740.

Russello, M. A., and G. Amato. 2004. Ex situ population management in the absence of pedigree information. Mol. Ecol. 13:2829–2840.

Santure, A. W., J. Stapley, A. D. Ball, T. R. Birkhead, T. Burke, and J. Slate. 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. Mol. Ecol. 19:1439–1451.

Saura, M., A. Fernandez, M. C. Rodriguez, M. A. Toro, C. Barragan, A. I. Fernandez, and B. Villarivera. 2013. Genome-wide estimates of coancestry and inbreeding in a closed herd of ancient iberian pigs. PLoS ONE 8: e78314.

Schultz, J. K., J. D. Baker, R. J. Toonen, and B. W. Bowen. 2009. Extremely low genetic diversity in the endangered Hawaiian monk seal (*Monachus schauinslandi*). J. Hered. 100:25–33.

Streiff, R., A. Ducousso, C. Lexer, et al. 1999. Pollen dispersal inferred from paternity analysis in a mixed oak stand of *Quercus robur* L-and *Q-petraea* (Matt.) Liebl. Mol. Ecol. 8:831–841.

Taylor, H. R. 2014. Detecting inbreeding depression in a severely bottlenecked, recovering species: the little spotted kiwi (*Apteryx owenii*): a thesis submitted to the Victoria University of Wellington in fulfilment of the requirements for the degree of Doctor of Philosophy in Ecology and Biodiversity. Victoria University of Welligton, New Zealand.

Van de Casteele, T., P. Galbusera, and E. Matthysen. 2001. A comparison of microsatellite-based pairwise relatedness estimators. Mol. Ecol. 10:1539–1549.

Van Horn, R. C., J. Altmann, and S. C. Alberts. 2008. Can't get there from here: inferring kinship from pairwise genetic relatedness. Anim. Behav. 75:1173–1180.

Vangestel, C., J. Mergeay, D. A. Dawson, V. Vandomme, and L. Lens. 2011. Spatial heterogeneity in genetic relatedness among house sparrows along an urban-rural gradient as revealed by individual-based analysis. Mol. Ecol. 20:4643–4653.

Wang, J. L. 2002. An estimator for pairwise relatedness using molecular markers. Genetics 160:1203–1215.

Wang, J. L. 2006. Informativeness of genetic markers for pairwise relationship and relatedness inference. Theor. Popul. Biol. 70:300–321.

Wang, J. 2007. Triadic IBD coefficients and applications to estimating pairwise relatedness. Genet. Res. 89:135–153.

Wang, J. L. 2011. COANCESTRY: a program for simulating, estimating and analysing relatedness and inbreeding coefficients. Mol. Ecol. Resour. 11:141–145.

Wang, J. 2014. Marker-based estimates of relatedness and inbreeding coefficients: an assessment of current methods. J. Evol. Biol. 27:518–530.

Ward, P. S. 1983. Genetic relatedness and colony organization in a species complex of ponerine ants, 1: Phenotypic and

genotypic composition of colonies. Behav. Ecol. Sociobiol. 12:285–299.

Weir, B. S., A. D. Anderson, and A. B. Hepler. 2006. Genetic relatedness analysis: modern data and new challenges. Nat. Rev. Genet. 7:771–780.

Wright, S. 1921. Systems of mating. Genetics 6:111–178.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. Locus information entered into COANCESTRY for the relatedness and inbreeding coefficient simulations.

Table S2. Summary of literature review of papers citing COANCESTRY for use in estimating relatedness or inbreeding coefficients.