

Methods and prospects for using molecular data in captive breeding programs: an empirical example using parma wallabies (*Macropus parma*)

Jamie A. Ivy^{1,#,*}, Adrienne Miller², Robert C. Lacy¹, and J. Andrew DeWoody³

¹Department of Conservation Science, Chicago Zoological Society, 3300 Golf Road, Brookfield, IL 60513, USA; ²Audubon Zoo, 6500 Magazine St, New Orleans, LA 70118, USA; ³Department of Forestry and Natural Resources, Purdue University, 195 Marsteller St., West Lafayette, IN 47907, USA

#Current Address: Collections Department, San Diego Zoo, PO Box 120551, San Diego, CA 92112, USA

*Corresponding Author: Jamie A. Ivy
Phone: 619-557-3905
Fax: 619-232-4117
Email: jivy@sandiegozoo.org

Running Title: Using molecular data to improve captive breeding programs

Abstract

Zoo and aquarium breeding programs rely on accurate pedigrees to manage the genetics and demographics of captive populations. Breeding recommendations are often encumbered, however, by unknown parentage. If an individual has any amount of unknown ancestry, the relationships between that individual and all other individuals in a population are ambiguous and breeding recommendations cannot be tailored to maximize genetic diversity and minimize inbreeding. In those situations, breeding program management might be improved by the incorporation of molecular data. We developed microsatellite markers for the parma wallaby (*Macropus parma*) and investigated how genetic data might be used to improve the management of the captive population. The parma wallaby is a small marsupial found in fragmented forests near the coast of New South Wales, Australia. Because the species is of conservation concern, the captive population in North America is managed by recurring breeding recommendations. The effectiveness of the population's management is hampered however, because over half of the individuals have some amount of unknown ancestry. We used microsatellite data to resolve unknown parentage, described how molecular estimates of relatedness might inform future breeding recommendations, and used computer simulations to investigate how molecular estimates of relatedness among founders might contribute to the genetic management of the population. Our results indicated that microsatellite appraisals of parentage were useful with respect to clarifying pedigrees, but that molecular assessments of founder relatedness provided very marginal benefits with regard to the preservation of genetic diversity and the avoidance of inbreeding.

Keywords:

captive breeding, founders, microsatellites, relatedness, simulation, unknown parentage

Introduction

Captive breeding programs sponsored by zoos and aquaria generally aim to maintain demographically stable, self-sustaining populations that retain genetic diversity and accumulate limited inbreeding (Ballou and Lacy 1995; Foose and Ballou 1988; Hedrick and Miller 1992; Lacy 1994). These goals are achieved through recurring breeding recommendations, which are based upon each population's unique history and future needs. Pedigree analyses, which are the foundation of most captive breeding programs, are used to produce breeding recommendations intended to meet a population's demographic and genetic goals. The best breeding strategies for captive populations to meet genetic goals are those that minimize a population's average kinship (Ballou and Lacy 1995; Fernandez and Toro 1999; Sonesson and Meuwissen 2001), with the kinship (f) between two individuals being the probability that two alleles at a given locus, one randomly drawn from each individual, are identical by descent from a common ancestor (Falconer 1981). An individual's mean kinship (mk ; Ballou and Lacy 1995) is a measure of an individual's genetic distinctiveness, and is the average of f s between that individual and all living

individuals in the population, including itself ($mk_x = \frac{\sum_{y=1}^N f_{xy}}{N}$). To minimize a population's average kinship, captive breeding programs breed genetically underrepresented individuals with low mks . Because captive breeding programs rely on pedigree analyses for developing breeding recommendations, captive population management is hampered if a population's pedigree is only partially known and neither f s nor mks can be calculated accurately. Algorithms for estimating f s and mks from only the known parts of incomplete pedigrees have been developed (Ballou and Lacy 1995), but breeding recommendations based on those estimates are considered suboptimal because the estimates are less accurate than equivalent measures of relationship calculated from complete pedigrees. Still, those algorithms currently provide the only option for conducting pedigree analyses on incomplete pedigrees, short of either assuming undocumented parents were unrelated population founders or excluding all animals with incompletely known ancestries from the analyses.

The captive parma wallaby (*Macropus parma*) population in North America is a typical example of a captive breeding program being managed by pedigree analyses. The population is currently managed through a Population Management Plan (PMP) by the Association of Zoos and Aquariums (AZA) and, like most actively managed captive breeding programs, PMPs receive regular demographic and genetic management. Parma wallabies are small marsupials native to Australia, and the species is currently found in fragmented forest areas near the coast of New South Wales. Due to continuing habitat degradation and low census size, the parma wallaby is listed under various threat categories by the government of New South Wales (Vulnerable; Threatened Species Conservation Act (NSW)), the U.S. Fish and Wildlife Service (Endangered; Endangered and Threatened Wildlife and Plants), and the IUCN (Low Risk: Near Threatened; IUCN Red List of Threatened Species). The parma wallaby was declared extinct in 1957 (Maynes 1974), but an introduced population of parma wallabies was subsequently discovered on Kawau Island, New Zealand in 1965 (Ride 1970; Wodzicki and Flux 1971) and remnant populations of the species were eventually re-discovered in Australia in 1966 (Ride 1970; Wodzicki and Flux 1971). Although the species remains of conservation concern in its native Australia, the exotic parma wallabies on Kawau Island are considered invasive pests. The captive parma wallaby population in North America was founded in the late 1960s by animals

originating from Kawau Island. Periodic imports to North America from the island have continued to occur and to date, no population founders have been imported from Australia.

The pedigree of the captive parma wallaby population in North America is incomplete and the relationships among its founders are unknown. As of January 2007, the PMP population included 157 animals and over half of those individuals (90 total) had varying amounts of unknown ancestry. Some of the missing information is deep in the pedigree and stems from a time when zoological records were inconsistently maintained. Unknown parentage also has been perpetuated throughout the recent pedigree partly because the species is occasionally housed in groups of multiple males, making paternity in those situations uncertain. Additionally, in October 2003, the PMP population received 22 new potential founders (wild-caught animals that have not yet contributed descendants to the living, captive population; Lacy 1995) from Kawau Island. The population of parma wallabies on the island is currently of moderate size, but was founded by only a few individuals. Because nothing is known about the genetics of the Kawau Island population, and there was little information regarding how the 22 new animals were captured, PMP managers were concerned that some of the new founders might be closely related (e.g., parent-offspring pairs or full-siblings). Captive breeding programs generally assume founders are unrelated, but the newly acquired parma wallaby founders may well violate this assumption.

Molecular markers can contribute to captive breeding programs in a number of ways (Table 1), but genetic data can be particularly valuable for managing populations with incomplete pedigrees. One of the simplest ways genetic data can improve the management of captive populations is to resolve unknown parentage. Contemporary pedigrees can be reconstructed from genetic data when individuals with unknown genealogy and their potential parents are available for genetic testing. For living animals with unknown ancestry deep in their pedigrees (Figure 1), molecular markers can be used to estimate the pairwise relatedness between themselves and the rest of the population. However, the accuracy of molecular relatedness estimates suffers from high sampling variances (Glaubitz et al. 2003, Lynch 1988, Lynch and Ritland 1999) and the best methods for incorporating these values into captive breeding programs remain unclear.

Most captive breeding programs assume founders are unrelated. Research suggests that this assumption generally has little effect on a breeding program's ability to retain genetic diversity over the long-term, but a marked increase in inbreeding in the early generations of captive breeding programs is observed if unidentified full-siblings are present (Rudnick and Lacy 2008). Thus, the best use of molecular estimates of founder relatedness may be to identify first-order relatives (Rudnick and Lacy 2008). This information can be used to avoid breeding recommendations that would pair close relatives, reducing the amount of inbreeding that would accumulate in a population relative to assuming founders were unrelated.

We developed microsatellite markers for the parma wallaby and investigated how genetic data could improve the overall management of the captive population in North America. Our goals were to use molecular data to 1) resolve unknown parentage among contemporary breeders; 2) describe how molecular estimates of kinship could inform future breeding recommendations; and 3) use computer simulations to investigate how marker-based kinship estimates among founders might contribute to the genetic management of the population. Our expectations were that the incorporation of molecular data into the parma wallaby breeding program would improve genetic management over time by preserving more genetic variation and more effectively limiting inbreeding than might be expected otherwise.

Methods

Sample Collection and DNA Extraction

Blood samples were opportunistically collected during routine physical examinations of 71 captive parma wallabies. Tissue samples from an additional five individuals were collected during post-mortem exams. All sampled individuals were part of the PMP population held and managed throughout North American zoos. A total of 20 individuals with some degree of unknown ancestry were sampled, as were 20 out of the 22 wild caught individuals acquired in October 2003. As of January 01, 2007, the captive parma wallaby population was known to be descended from at least 37 founders imported from Kawau Island, New Zealand (exact numbers of founders were not available due to the high degree of unknown ancestry present in the pedigree). Four additional, potential founders were still living in the population but had yet to produce any offspring. A total of 20 established founders were sampled, as were all four prospective founders (Figure 2). Because the true number of population founders was unknown, it was difficult to determine how representative our samples were of the population's total founder variation.

A 1-2 cc aliquot of whole blood was collected from each individual in a standard liquid EDTA collection tube. Samples were refrigerated or kept on ice for 24-36 hours prior to DNA extraction. DNA was isolated from 300 μ L of whole blood with the Wizard® Genomic DNA Purification Kit (Promega), following the manufacturer's instructions. Tissue samples collected during post-mortem exams were placed in lysis buffer (100 mM Tris-HCl (pH 8.0), 100 mM EDTA, 10 mM NaCl, 2% SDS) and stored at room temperature. DNA was isolated from 1-2 mm tissue samples that were incubated overnight at 55°C in 700 μ L of extraction buffer (86 mM NaCl, 43 mM Tris base, 21 mM EDTA, 80 mM Tris-HCl (pH 8.0), 0.01 mg Proteinase K, 3% SDS, 7 mM dithiothreitol). Protein was precipitated by adding 233 μ L of 7.45 M ammonium acetate, extractions were centrifuged, and the supernatant was recovered. DNA was isolated from the supernatant with an isopropanol precipitation (Sambrook and Russell 2001), washed with 70% ethanol, and resuspended in 100 μ L purified water.

Microsatellite Characterization and Genotyping

Microsatellites were characterized specifically for the parma wallaby by a modified version (Williams and DeWoody 2004) of the microsatellite enrichment protocols described by Hamilton et al. (1999) and Hauswaldt and Glenn (2003). Following microsatellite library construction, primers were initially designed for 10 microsatellite loci and fluorescent chromatide rhodamine green 5'dUTP was used as a label to screen those loci for allelic variation on an ABI377 automated sequencer (Applied Biosystems). Three of the candidate loci were ultimately genotyped for the study (MP03, MP04, MP06; Table 2), because of their polymorphism and genotyping reliability.

Parma wallaby samples were genotyped at six additional microsatellite loci (Table 2): one locus derived from an allied rock-wallaby (*Petrogale assimilis*) library (Pa595; Spencer et al. 1995), one locus from a yellow-footed rock-wallaby (*P. xanthopus*) library (Y175; Zenger et al. 2002), two loci from a tammar wallaby (*M. eugenii*) library (Me14, Me17; Taylor and Cooper 1998), and two loci from an eastern grey kangaroo (*M. giganteus*) library (G19-1, G31-1; Zenger and Cooper 2001). PCRs for all loci were performed in a final volume of 10 μ L and contained 0.5 units *Taq*, 1x Thermopol Buffer, and 0.2 mM each dNTP. PCRs for MP03, MP04, and MP06 contained 0.3 μ M forward and reverse primers, while PCRs for the remaining loci

contained 0.75 μM forward and reverse primers. Annealing temperatures for all loci are listed in Table 2. Previously described thermal profiles were used for PA595, Me14, Me17, G19-1 and G31-1 (Spencer et al. 1995, Taylor and Cooper 1998, Zenger and Cooper 2002), with the exception that the six 'touchdown' cycles for G19-1 and G31-1 were excluded. The thermal profile for Y175 included an initial denaturation step of 94 °C for 2 min 30 s, followed by 30 cycles of 94 °C for 35 s, 58 °C for 60 s, and 72 °C for 30 s, with a final extension step at 72 °C for 10 min. The thermal profile for MP03, MP04, and MP06 included an initial denaturation step of 94 °C for 2 min, followed by 30 cycles of 94 °C for 30 s, 50 °C for 30 s, and 72 °C for 30 s, with a final extension step at 72 °C for 5 min. All forward primers were fluorescently labeled and PCR products were electrophoresed in 4.75% denaturing polyacrylamide gels using an ABI377 automated sequencer (Applied Biosystems). Alleles were sized with respect to an internal size standard using the GENESCAN and GENOTYPER softwares (Applied Biosystems).

Estimating Basic Genetic Parameters

The software program MICROSATELLITE ANALYZER (Version 4.05; Dieringer and Schlotterer 2003) was used to calculate the number of alleles, allele frequencies, and both the observed and expected heterozygosities for each microsatellite locus used in the study. The software program GENEPOP (Version 4.0; Rousset 2008) was used to test for linkage disequilibrium between pairs of loci and to test for deviations from Hardy-Weinberg equilibrium. Population-level allele frequencies were used to calculate multi-locus parental exclusion probabilities (Selvin 1980).

Parentage Analyses and Relatedness Estimation

Among the parma wallabies sampled for this study, there were 13 individuals for which at least one parent (as identified from studbook records) also was sampled. Both the sire and dam were sampled for five individuals, only the sire was sampled for three individuals, and only the dam was sampled for five individuals. Genotypic profiles of the offspring and parents were compared to determine if they were consistent with the presumptive parentage. Samples also were collected from a group of two male and four female parma wallabies that were housed together. Parentage records for seven offspring produced by the group were unavailable, so parentage was determined by exclusion (Selvin 1980). Finally, exclusion also was used to resolve the maternity of one additional individual that was unrelated to the previously described group.

A relatedness estimator was used to calculate relatedness coefficients between all pairs of sampled individuals. Although captive breeding programs generally use kinships (f_s) to quantify relationships, molecular data are usually used to calculate relatedness. The relatedness between two individuals is the probability that, at a given locus, an allele sampled at random from one individual is present also in the other individual, due to identity by descent from a common ancestor. Furthermore, in non-inbred, diploid populations, relatedness is equal to $2f$. We used estimates of relatedness to describe how individuals with unknown ancestry might be related to other sampled individuals, and to determine if some of the new founders were close relatives. Because the performance of different relatedness estimators varies depending on the numbers and allele frequency distributions of the microsatellite loci employed (Russello and Amato 2004; Van de Castele et al. 2001), computer simulations were used to estimate sampling variances for three estimators across a range of known relationship categories. The three relatedness

estimators tested were r_{xyQG} (Queller and Goodnight 1989), r_{xyLR} (Lynch and Ritland 1999), and r_{xyWang} (Wang 2002). A simulation written in the C programming language was used to create microsatellite genotypes for 1000 pairs of individuals from each of the following relationship categories: unrelated (UR), parent-offspring (PO), full-siblings (FS), and half-siblings (HS). Genotypes for unrelated individuals were assigned based on the observed allele frequency distributions from the captive parma wallaby population. Pairs of individuals with known relationships were created by sampling alleles from unrelated individuals in a manner consistent with the desired relationship category. The three relatedness estimators were used to calculate relatedness coefficients for all pairs of individuals within a given relationship category, then means and variances of the relatedness coefficients were calculated for each estimator. The relatedness estimator that displayed the smallest sampling variance across the greatest number of relationship categories was selected to calculate relatedness. The software program SPAGeDi (Version 1.2; Hardy and Vekemans 2002) was used to calculate relatedness coefficients for both the simulated dataset and the empirical dataset.

Investigating Founder Relationships

We used a series of computer simulations to investigate how molecular estimates of relatedness among the newly acquired parma wallaby founders might be used in the genetic management of the captive population. Breeding recommendations for PMP populations, as well as many other types of zoo breeding programs, are based on mean kinships (*mks*); animals with low *mks* are generally preferentially bred because they are genetically underrepresented. Population founders are usually assumed to be unrelated and their *fs*, both with other founders and with all other individuals in the population, are consequently zero. Molecular estimates of pairwise relatedness can provide information about the true relationships among founders. However, because empirical estimates of relatedness suffer from high sampling variances (Blouin et al. 1996), the use of the values themselves in *mk* calculations can result in inaccurate estimates of relationship. Instead, we tested two methods for using molecular relatedness coefficients to identify first-order relatives (i.e., parent-offspring pairs or full-siblings). The methods were tested through computer simulations that investigated the long-term genetic impacts of designating some pairs of founders as first-order relatives, with *fs* assigned to be 0.25, rather than assuming all founders to be unrelated.

The first method we tested for identifying first-order relatives used a graph of the pairwise relatedness coefficients estimated for the sampled founders. All relatedness coefficients were graphed in order of increasing value, and we looked for a natural break in the distribution that could indicate first-order relatives. This approach for identifying first-order relatives was more qualitative than quantitative, but we felt it was reasonable because the relatedness coefficients of first-order relatives should cluster around a higher mean than the relatedness coefficients of more distant relatives. The value around which first-order relatedness coefficients should cluster is not always obvious a priori, thus, this method allowed close relatives to be identified without prior genetic knowledge of the population or preliminary data analyses.

The second method we tested for identifying first-order relatives used the distributions of relatedness generated from the simulated pairs previously described to choose a cutoff value that would produce an acceptable compromise between the misclassification of first-order relatives as unrelated pairs and the misclassification of unrelated pairs as first-order relatives (see Blouin et al. 1996, which described a similar approach). One approach to balancing these misclassification rates would be to equalize the probability of classifying unrelated pairs as first-order relatives

and the probability of classifying first-order relatives as unrelated pairs. However, if there are many more unrelated pairs than first-order relatives present, which is likely the case within groups of captive population founders, such a cutoff will lead to many more pairs of unrelated individuals being incorrectly identified as first-order relatives than first-order relatives being incorrectly identified as unrelated. We chose a cutoff of $r = 0.45$ to identify first-order relatives because this value failed to identify only 36% of full-siblings and 40% of parent-offspring pairs as first-order relatives, while incorrectly identifying only 7% of unrelated pairs and 22% of half-siblings as first-order relatives. Pairs of founders in the parma wallaby population with relatedness coefficients higher than the cutoff were identified as first-order relatives.

Simulations were used to test the impact of using different sets of founder relationships on the long-term genetic management of the PMP population. The two f matrices generated from the methods of identifying first-order relatives were the “modified” f matrices to be tested. The f matrix of the captive parma wallaby population living on January 01, 2007, which assumed all founders to be unrelated, was considered to be the “current” f matrix. Because there was no way to ascertain if the two methods we used for identifying first-order relatives correctly identified the relationships of the parma wallaby founders, we designed simulations to quantify 1) the genetic benefits of correctly identifying first-order relatives among the founders assuming such relationships existed and 2) the genetic costs, or detriments, of incorrectly identifying pairs of founders as first-order relatives when pairs were actually unrelated. To quantify these costs and benefits, it was necessary for simulations to simultaneously track two f matrices (see Supplementary Material for a detailed description of the simulation). One matrix represented the “true” relationships in a simulated population, and this matrix was used to determine the loss of genetic diversity and accumulated inbreeding in the population. The other matrix represented the relationships that were being tested as the basis for breeding program management. Both f matrices were maintained throughout the simulation, but breeding pairs were selected based on mks calculated from the f matrix being tested. Thus, while population management was driven by the relationships being tested, the relationships assumed to represent the true relationships among individuals for a given simulation scenario were being tracked and quantified to measure genetic costs and benefits. To fully quantify the possible genetic benefits and detriments of using molecular data to identify first-order relatives, four scenarios were considered for each of the modified f matrices (Figure 3): (1) The current f matrix for the population was used for both the true relationships and the tested relationships; this generated the genetic outcomes if none of the parma wallaby founders were first-order relatives. (2) The current f matrix was used for the true relationships and a modified f matrix was used for the tested relationships; this generated the genetic outcomes if molecular data identified some pairs of founders as first-order relatives when all founders were actually unrelated. (3) A modified f matrix was used for the true relationships and the current f matrix was used for the tested relationships; this generated the genetic outcomes if molecular data failed to identify first-order relatives that were actually present among the founders. (4) A modified f matrix was used for both the true relationships and the tested relationships; this generated the genetic outcomes if molecular data correctly identified the first-order relatives that were actually present among the founders. The genetic outcomes of scenarios 1 and 2 were compared to quantify the detriment to incorrectly identifying first-order relatives among the parma wallaby founders when none were present, while the genetic outcomes of scenarios 3 and 4 were compared to quantify the benefit to correctly identifying first-order relatives when some of the parma wallaby founders were indeed related.

Because unknown ancestry was present in the parma wallaby pedigree, one additional adjustment was necessary to facilitate the comparison of variables across the four described scenarios. When unknown ancestry is present, f_s can only be calculated from the known part of the pedigree according to the equations described by Ballou and Lacy (1995). These estimates of f are inaccurate, however, and can be either larger or smaller than the true values that would be calculated if the pedigree were completely known. Because the accuracies of f_s calculated in this manner are dependent on the specific distribution of unknown ancestries in a pedigree, variables that are measured as a function of f cannot be directly compared across simulation scenarios. Thus, to ensure results were comparable across the four described scenarios, only the initial f matrices used to start simulations were calculated according to the equations described by Ballou and Lacy (1995). These initial f matrices were then assumed to represent the actual relationships in the population, all individuals were assumed to have 100% known ancestry at the beginning of the simulations, and all further f_s were calculated according to the standard methods described by Falconer (1981).

Simulations were designed to model standard captive breeding programs (Wiese and Willis 1996) and were parameterized in accordance with parma wallaby biology. Although parma wallabies are not monogamous in the wild, we modeled monogamous breeding pairs to simulate a captive breeding program's ability to regulate breeding pairs in captivity. Simulations ran on a yearly timestep, and the general yearly process of the simulations was as follows: breeding pairs were selected by a method designed to minimize mean kinship, offspring were produced by the selected pairs, all individuals were aged one year, and a proportion of individuals were removed from the population to simulate stochastic mortality. Specific simulation details and parameterizations are presented in the Supplementary Material. In addition to minimizing mean kinship, breeding pair selection also avoided close inbreeding by rejecting any breeding pair that exhibited an f greater than the current average f in the population. An offspring's inbreeding coefficient (F) is equal to the f of its parents (Falconer 1981), thus, limiting breeding pairs based on f slowed the accumulation of inbreeding in the population.

For each simulation, genetic variation and inbreeding were evaluated on a per timestep basis. Inbreeding was measured as the average inbreeding coefficient (\overline{F}) and genetic variation was measured as proportional gene diversity (GD). GD was calculated as $1 - \overline{mk}$, where \overline{mk}

was the average mean kinship in the population ($\overline{mk} = \frac{\sum_{x=1}^N mk_x}{N}$; Ballou and Lacy 1995). For a modified f matrix, the impacts of using the given founder relationships were assessed by comparing the GD and \overline{F} observed for each of the four previously described simulation scenarios. Due to the stochastic nature of the simulations, each scenario was run 1000 times and results were averaged over all iterations. For each timestep, 95% confidence intervals for average GD and \overline{F} values were calculated across all iterations.

Results

The nine microsatellites assayed for this study averaged 5.3 alleles per locus and the mean observed heterozygosity across loci was 0.60 (Table 2). All exact tests for Hardy-Weinberg equilibrium were non-significant at the 0.05 significance level, except the test for locus MP06 ($p = 0.0011$). Tests for linkage disequilibrium revealed that genotypes at loci MP04 and Pa595 were linked ($p < 0.05$), even after a sequential Bonferroni correction was applied for

multiple comparisons (Rice 1989). Loci MP04 and MP06 ultimately were removed from the study so that the analyzed suite of microsatellite loci met standard assumptions. The seven remaining microsatellites exhibited a combined parental exclusion probability of 0.97 when one parent was unknown and 0.86 when both parents were unknown.

The basic microsatellite parameters reported were calculated from the entire set of sampled parma wallabies. Parameters calculated from only wild-caught individuals were similar to those reported for the entire set of samples, and preliminary analyses indicated that the results of the study were equivalent regardless of which set of samples (wild-caught individuals vs. all individuals) the microsatellite allele frequencies used to estimate relatedness were calculated from. All microsatellite loci used in the study were in Hardy-Weinberg equilibrium, indicating that inbreeding should not have impacted relatedness estimation.

Parentage Analyses and Relatedness Estimation

Genetic parentage was investigated for 13 offspring and compared to studbook records. Of the eight sires and ten dams considered, genotypic profiles suggested that the identities of two dams were incorrectly recorded. Both offspring with suspect parentage were born within the same month to a group of wallabies that included three adult females, all of which were sampled for this study. The presumptive dam of the first offspring was excluded as a parent of that individual at three out of seven microsatellite loci. Of the two other possible dams, one was excluded as a parent at four microsatellite loci. The last remaining dam could not be excluded at any of the seven microsatellite loci used in the study, thus, that individual was considered to be the true dam of the first offspring with suspect parentage. The dam of the second offspring with suspect parentage was listed as the dam that was genetically assigned to the first offspring. Because both offspring were born within the same month, one of the other females in the group must have been the dam of the second offspring. Neither of those females could be excluded as a parent of the second offspring at any microsatellite loci, thus, the true parentage of that offspring remains unclear. Genotypic profiles also were used to conclusively determine both the sires and dams for seven offspring with no parentage records. All seven offspring were produced by a single group of parma wallabies that included two males and four females. Six of the seven offspring were sired by one male, while only one of the offspring was sired by the alternate male. Finally, the maternity of one additional individual also was investigated. Studbook records indicated that there were two possible dams for the individual, but only one of the two possibilities was sampled for this study. The female that was sampled could be excluded as the dam at three microsatellite loci; thus, the unsampled female was assumed to be the individual's dam.

Sampling variances were calculated for three relatedness estimators across four categories of relationship (unrelated, UR; parent-offspring, PO; full-siblings, FS; and half-siblings, HS). The estimator that displayed the smallest sampling variance across the greatest number of relationship categories was r_{xyWang} (Table 3), so that estimator was selected to calculate relatedness among the sampled parma wallabies. Over all relationship categories, sampling variances ranged from 0.02 to 0.08 across all estimators and the smallest sampling variances were observed for parent-offspring pairs.

Investigating Founder Relationships

We tested the impact of using molecular relatedness coefficients to identify first-order relatives among founders on the long-term genetic management of the parma wallaby PMP

population. The first method we tested for identifying first-order relatives used a graph of the 190 pairwise relatedness coefficients estimated for the 20 sampled founders (Figure 4). In a non-inbred diploid population, first-order relatives should have a relatedness coefficient of 0.5. We looked for a break in the ordered distribution of relatedness coefficients either approximately equal to or greater than this value, and ultimately selected the four highest pairwise relatedness coefficients to represent first-order relatives (Figure 4). The four selected pairs of individuals were assigned f_s of 0.25, and the genetic impacts of using the resulting modified f matrix were quantified through computer simulations. Figure 5 illustrates the rates of change observed in GD and \bar{F} across 100 years as compared to random mating.

When results were graphed across all 100 years of the simulations, differences between the scenarios could not be discerned on Figure 5. Therefore, the results from only the last 10 timesteps were graphed to better display the comparisons between scenarios 1 and 2 (Figure 6) and scenarios 3 and 4 (Figure 7). Table 4 shows the final values at the end of 100 years for GD and \bar{F} from all scenarios. If the four selected founder pairs we designated as first-order relatives were actually unrelated (the contrast between scenarios 1 and 2), both the loss of GD and the increase in \bar{F} were approximately 0.0004. If the four selected founder pairs were indeed related and we designated them as such (the contrast between scenarios 3 and 4), both the increase in GD and the reduction in \bar{F} were approximately 0.0009. Thus, when the distribution of pairwise relatedness coefficients was used to identify first-order relatives, the possible benefits of using molecular data to correctly assign founder relationships was slightly higher than the possible detriments to incorrectly identifying founders as first-order relatives when they were actually unrelated.

For the second method of identifying first-order relatives, we used the distributions of relatedness values generated from simulated relationship pairs to choose a relatedness cutoff value that indicated first-order relatives. We selected relatedness coefficients of 0.45 and higher to represent first-order relatives, as simulated data suggested that this cutoff would correctly identify more than 60% of first-order relatives while misclassifying only 7% of unrelated pairs as related. Seven of the pairwise relatedness coefficients among the parma wallaby founders were higher than 0.45, and the seven selected pairs of individuals were assigned f_s of 0.25. The genetic impacts of using the resulting modified f matrix were similar to those described for the first method of identifying first-order relatives (full results not shown, but final values are given in Table 4). Using this method for modifying the f matrix to account for assumed founder relationships, if we designated the seven selected founder pairs as first-order relatives when they were actually unrelated (the contrast between scenarios 1 and 2), both the loss of GD and the increase in \bar{F} was approximately 0.0009. If we correctly designated the seven selected founder pairs as first-order relatives when they were indeed related (the contrast between scenarios 3 and 4), both the gain of GD and the decrease in \bar{F} was approximately 0.0020. In summary, like the results observed for the first method of identifying first-order relatives, the possible benefits to using molecular data to identify closely related founders were higher than the possible detriments.

Discussion

The resolution of unknown parentage is one of the primary ways molecular data can be used to improve captive population management. Unknown parentage plagues many contemporary pedigrees, especially in species housed in groups with multiple males. This

problem is exemplified in parma wallabies; 57% of individuals living in the PMP population as of January 2007 had varying amounts of unknown parentage, and nine of those individuals had either one or both parents listed as unknown. We successfully resolved parentage for seven of those nine individuals, as well as for one additional individual that died prior to 2007 (DNA samples were not available for the two remaining individuals with unknown parentage). Although a considerable amount of unknown ancestry persists deeper in the parma wallaby pedigree, the resolution of the majority of unknown parentage in the contemporary pedigree still improves the accuracy of *mk* calculations and, thus, the genetic management of the parma wallaby PMP. For species held in groups with multiple males, unknown paternity will nearly always continue to be perpetually generated. Furthermore, as demonstrated by the pedigree records investigated here, maternity is not always known with certainty. We demonstrated that molecular markers can help resolve unknown parentage in captive pedigrees, which increases the effectiveness of genetic management by improving the accuracy of *mk* calculations used for breeding recommendations.

Studbook records indicated that a total of 18 known parent-offspring pairs were included in the group of parma wallabies sampled for this study. Microsatellite data for 16 of these pairs (89%) were consistent with the recorded pedigree, indicating that parentage records for this population were generally accurate. Still, the identities of two presumptive dams (11%) appear to have been incorrectly recorded. Further analyses indicated that both individuals with misassigned dams were born within the same month to a group of parma wallabies that were being housed together, which means the inaccuracies were not independent of each other. There are several explanations for how these errors could have arisen. One possibility is that the dam identification numbers were incorrectly recorded in the parma wallaby studbook. If this was the case, the parentage records should be changed to accurately reflect the true pedigree. A second explanation however, is that the blood samples used for the parentage analyses were incorrectly labeled at the time of sample collection; both offspring with misassigned maternity and two of their three possible dams were sampled at the same location on the same date. In this case, parentage records should not be adjusted. The collection of additional blood samples is necessary to either confirm or discount the possibility that the original samples were incorrectly labeled and, at this time, the parentage of the two offspring with misassigned maternity should be considered to be provisional.

In the absence of a complete pedigree, molecular estimates of relatedness can provide information about the relationships that exist among individuals in a population. We tested two methods of using molecular relatedness coefficients to identify first-order relatives among parma wallaby founders, and quantified the possible genetic benefits and detriments to using that information for managing the parma wallaby PMP. Both methods exhibited similar trends and results. The benefit to correctly identifying first-order relatives among the founders was small; after 100 years the potential gain in *GD* and the reduction in the accumulation of \bar{F} was 0.0009 or 0.0020, depending on the method used to identify first-order relatives. The detriment to identifying some founder pairs as first-order relatives when all founders were actually unrelated also was small; after 100 years the potential gain in *GD* and the reduction in the accumulation of \bar{F} was 0.0004 or 0.0009, again depending on the method used to identify first-order relatives. As expected, the method that identified the largest number of relatives exhibited both the largest potential benefit and the largest potential detriment. For both methods, the potential genetic benefits of using molecular data to identify related founders were greater than the potential costs (in terms of *GD* and \bar{F}). However, the results suggest that incorporating molecular estimates of

founder relatedness into the genetic management of the parma wallaby PMP would be neither particularly beneficial nor detrimental over the next 100 years of management.

The two methods that we tested for identifying first-order relatives among founders identified four and seven possible first-order relatives, respectively. The identification of four or seven pairs of presumptive first-order relatives among the founders is plausible given the history of this population, and may be typical of other managed populations with some uncertainty caused by missing records or uncertain parentage due to group housing. We used the population of 157 parma wallabies living on January 01, 2007 to start our simulations. Thus, when founders were identified as first-order relatives in the initial simulation timestep, we modified f_s for only four or seven pairs of founders. Because those founders had already produced some offspring in the population, the new founder relationships also were carried through the pedigree to modify the f_s of those founder's descendants: f_s were modified for the same 30 descendants when both four and seven pairs of founders were identified, but the specific f modifications associated with each set of assumed founder relationships produced different mean kinships among those descendants (Figure 8). Collectively, these modifications were simply too minor in the context of the total population to affect significant change in genetic management. Our previous work on founder relationships, which demonstrated a measurable benefit to identifying any full-siblings present among founders, was based solely on simulated populations (Rudnick and Lacy 2008). Furthermore, although a few studies have suggested that molecular estimates of relatedness might be incorporated into captive breeding programs (e.g., Jones et al. 2002, Russello and Amato 2004), none have quantified the potential impact. Our work on the parma wallaby indicates that additional work on other captive populations may be necessary to determine if molecular data should be incorporated into their genetic management.

In addition to quantifying the impacts of using molecular estimates of founder relatedness to genetically manage the parma wallaby PMP, we also investigated how molecular relatedness coefficients might be incorporated into the parma wallaby breeding program for individuals with unknown ancestry. Microsatellite data can theoretically be used to discriminate among multiple relationship categories within a population but approximately 40 independently segregating loci are required to differentiate full-siblings from half-siblings with greater than 90% reliability (Blouin et al. 1996). Furthermore, individuals from more distant relationship categories are even harder to classify. Rather than using inaccurate estimates of relatedness for mk calculations, we suggest identifying a series of relatedness ranges that represent different degrees of relationship. Then, if a pair of individuals with ambiguous ancestry has a relatedness coefficient that falls within a given range, the pair's f could be assigned a pre-determined value intended to represent that relationship category (e.g., full-siblings, half-siblings, unrelated, etc). For a given suite of microsatellites, means and sampling variances calculated for simulated pairs of individuals that belong to specific relationship categories could be used to select both appropriate relatedness ranges and the f_s that those ranges represent. Here, we tested this approach for the parma wallaby. Because our microsatellites provided little resolution between simulated relationship categories (Table 3), we chose only a single relatedness range for evaluation. We selected relatedness coefficients equal to or greater than 0.45 to indicate "close relatives", because simulated data indicated that this value should identify a majority of parent-offspring and full-sibling pairs as first-order relatives while keeping the probability of misclassifying unrelated pairs as first-order relatives low. We then selected 0.25 to replace the f values of all pairs with relatedness coefficients equal to or greater than the cutoff.

Parma wallabies with studbook numbers 617 and 623 both have unknown ancestry that originated before the value of accurate pedigree records had been recognized. The amount of an individual's ancestry that is known is the proportion of its genome that can be traced back to known founders and, for an individual, this value can range from zero to one (Ballou and Lacy 1995). The ancestry of parma wallaby 617 is 0.0 known, because none of the individual's ancestors can be traced back to wild founders, and the ancestry of parma wallaby 623 is 0.09 known, because the individual's pedigree can be traced back to four wild founders and nine ancestors with unknown parents. Because nothing is known about the ancestry of 617 and very little is known about the ancestry of 623, a pedigree cannot be used to determine how those two individuals are related and their pairwise f is undefined. Our molecular estimate of relatedness for the pair, however, was 0.4948. Thus, in accordance with the methods we have described, the pairwise f should be changed from undefined to 0.25. This change in f would raise both mk values, making each individual less genetically valuable. These changes in mk are desirable if the individuals are truly closely related, because captive breeding programs preferentially breed genetically under-represented individuals (i.e., individuals with low mks) to equalize founders contributions and retain high levels of genetic diversity.

Although the method we have described for incorporating molecular estimates of relatedness into captive breeding programs is quite plausible, additional research on both simulated and real populations is needed to fully quantify the impacts this, or a similar method, would have on a population's long-term genetics. Willis (2001) demonstrated that overestimating kinship among individuals with unknown ancestry was generally more detrimental to genetic diversity than underestimating kinship to an equivalent degree. Molecular estimates of relatedness can either overrepresent or underrepresent kinship, and the direction of error for a specific estimate is unknowable. Thus, any change in the f matrix used to manage a captive breeding program should be made with caution and be consistent with any ancillary information that may be available. For example, Gautschi et al. (2003) combined molecular relatedness data, mitochondrial DNA sequence data, and studbook information to investigate founder relationships for a captive bearded vulture (*Gypaetus barbatus*) population. Likelihood tests using the molecular relatedness data indicated 40 pairs of founders were first-order relatives. However, when all available information was considered together, it was determined that 33 of those pairs (82.5 %) could not be closely related. For the parma wallabies, the limited amount of pedigree information that is available for 617 and 623 does not preclude a close relationship; both the sire of 617 and the sire of 623 were born to unknown parents around the same time, at the same location. Thus, if their sires were closely related, 617 and 623 would be closely related as well. In a case such as this, where both molecular and pedigree data are compatible with the supposition that two individuals are closely related, the f for the individuals could be modified to reflect this relationship.

Data on parma wallabies 617 and 788 demonstrate the risks of using modest molecular data sets for identifying relatives. The pair's relatedness coefficient was estimated as 0.4746, which was greater than the 0.45 cutoff we selected for identifying "close relatives". Parma wallaby 788 was one of the new founders acquired from the wild in 2003. Although the grandparents of 617 are unknown, studbook records indicate that both the sire and dam of 617 were born to parents within the captive population around 1989. Thus, it seems highly unlikely that 617 and 788 could be closely related. These results exemplify the need for large suites of molecular markers to accurately estimate genetic parameters such as heterozygosity or relatedness (Blouin et al. 1996; DeWoody and DeWoody 2005; Glaubitz et al. 2003).

Conclusions

We used the captive parma wallaby population in North America to investigate the methods and prospects of incorporating molecular data into zoo and aquarium breeding programs. We evaluated the efficacy of molecular markers in resolving captive pedigrees, and quantified the impact of the improved pedigrees on the standard genetic goals of captive breeding programs. Specifically, we described a suite of microsatellites sufficient for resolving contemporary unknown parentage in the captive parma wallaby population, and used those microsatellites to 1) successfully resolve parentage for eight individuals and 2) identify two dams that may have been incorrectly recorded in the parma wallaby studbook. Although previous research suggested molecular estimates of relatedness may be useful for identifying close relatives among population founders, results from our computer simulations suggested the incorporation of this information into the genetic management of the parma wallaby PMP provides little benefit. This result is likely due to both the small number of related founders we identified and the relatively small impact of those initial relationships on kinships in the descendant population. Thus, the benefit of using molecular data to ascertain founder relationships would be greater for populations that have a higher proportion of related founders or more descendants from those founders that were inter-related. Extensive analysis of such pedigrees would be necessary to determine how rapidly the benefits of incorporating molecular data on founder relationships accrued as the amount of kinship among founders increased. Finally, we used the parma wallaby to demonstrate a plausible method for incorporating molecular relatedness coefficients into captive breeding programs that have unknown ancestry deep in their pedigrees. This method can be applied to a broad range of captive breeding programs, but additional research on both simulated and real populations is needed to fully quantify long-term genetic impacts.

Funding

Molecular data collection was funded by the Marsupial and Monotreme Taxon Advisory Group for the Association of Zoos and Aquariums and donations from holding institutions to the Parma Wallaby Population Management Plan. Funding for JI during the completion of this project was provided by a Conservation Project grant from the Institute of Museum and Library Services, which was received by the Chicago Zoological Society.

Acknowledgements

We thank the following institutions for kindly providing parma wallaby samples: Bermuda Zoo, Brevard Zoo, Cape May Zoo, Cleveland Metroparks Zoo, Disney's Animal Kingdom, Fort Worth Zoo, Happy Hollow Zoo, Henry Doorly Zoo, Miami Metrozoo, North Carolina Zoo, Oklahoma City Zoo, Riverbanks Zoo, Roger Williams Park Zoo, San Antonio Zoo, San Diego Zoo, and Sunset Zoo. We also thank Guha Dharmarajan for his help in constructing the parma wallaby microsatellite library and Joseph Busch for his help in the final stages of microsatellite data collection. The DeWoody lab group provided many insightful comments on a previous draft of this paper.

References

- Ballou JD, Lacy RC, 1995. Identifying genetically important individuals for management of genetic diversity in pedigreed populations. In: Ballou JD, Gilpin M, Foose TJ, editors. Population management for survival and recovery. New York: Columbia Press; p. 76-111.
- Blouin MS, Parsons M, Lacille V, Lotz S, 1996. Use of microsatellite loci to classify individuals by relatedness. *Mol Ecol.* 5:393-401.
- Bowling AT, Zimmermann W, Ryder O, Penado C, Peto S, Chemnick L, Yasinetskaya N, Zharkikh T, 2003. Genetic variation in Przewalski's horses, with special focus on the last wild caught mare, 231 Orlitza III. *Cytogenet. Genome Res.* 101:226-234.
- DeWoody YD, DeWoody JA, 2005. On the estimation of genome-wide heterozygosity using molecular markers. *J Hered.* 96:85-88.
- Dieringer D, Schlotterer C, 2003. MICROSATELLITE ANALYSER (MSA): a platform independent analysis tool for large microsatellite data sets. *Mol Ecol Notes.* 3:167-169.
- Ely JJ, Dye B, Frels WI, Fritz J, Gagneux P, Khun HH, Switzer WM, Lee DR, 2005. Subspecies composition and founder contribution of the captive U.S. chimpanzee (*Pan troglodytes*) population. *Am J Primatol.* 67:223-241.
- Endangered and Threatened Wildlife and Plants, 2005. 50 CFR 17.11 and 17.12. U.S. Fish and Wildlife Service.
- Falconer DS, 1981. Introduction to quantitative genetics, 2nd edn. New York: Longman Inc.
- Fernandez J, Toro MA, 1999. The use of mathematical programming to control inbreeding in selection schemes. *J Anim Breed Genet.* 116:447-466.
- Foose TJ, Ballou JD, 1988. Population management: theory and practice. *Int Zoo Yb.* 27:26-41.
- Gautschi B, Jacob G, Negro JJ, Godoy JA, Muller JP, Schmid B, 2003. Analysis of relatedness and determination of the source of founders in the captive bearded vulture, *Gypaetus barbatus*, population. *Conserv Genet.* 4:479-490.
- Geyer CJ, Ryder OA, Chemnick LG, Thompson EA, 1993. Analysis of relatedness in the California condors, from DNA fingerprinting. *Mol Biol Evol.* 10:571-589.
- Glaubitz JC, Rhodes EO Jr, DeWoody JA, 2003. Prospects for inferring pairwise relationships with single nucleotide polymorphisms. *Mol Ecol.* 12:1039-1047.
- Haig SM, Ballou JD, Casna NJ, 1994. Identification of kin structure among Guam rail founders: a comparison of pedigrees and DNA profiles. *Mol Ecol.* 3:109-119.

Haig SM, Ballou JD, Casna NJ, 1995. Genetic identification of kin in Micronesian kingfishers. *J. Hered.* 86:423-431.

Hamilton MB, Pincus EL, Di Fiore A, Fleischer RC, 1999. Universal linker and ligation procedures for construction of genomic DNA libraries enriched for microsatellites. *Biotechniques.* 27:500-507.

Hardy OJ, Vekemans X, 2002. SPAGeDi: a versatile computer program to analyze spatial genetic structure at the individual or population levels. *Mol Ecol Notes.* 2:618-620.

Hauswaldt SJ, Glenn TC, 2003. Microsatellite DNA loci from the diamondback terrapin (*Malaclemys terrapin*). *Mol Ecol Notes.* 3:174-176.

Hedrick PW, Miller P, 1992. Conservation genetics: theory and management of captive populations. In: Sandlund OT, Hindar K, Brown AHD, editors. Conservation of biodiversity for sustainable development. Oslo: Scandinavian University Press; p. 70-87.

IUCN 2007. 2007 IUCN Red List of Threatened Species. <<http://www.iucnredlist.org>>. Downloaded on 15 February 2008.

Jones KL, Glenn TC, Lacy RC, Pierce JR, Unruh N, Mirande CM, Chavez-Ramirez F, 2002. Refining the Whooping Crane studbook by incorporating microsatellite DNA and leg-banding analyses. *Conserv Biol.* 16:789-799.

Lacy RC, 1994. Managing genetic diversity in captive populations of animals. In: Bowles ML, Whelan CJ, editors. Restoration and recovery of endangered plants and animals. Cambridge : Cambridge University Press; p. 63-89.

Lacy RC, 1995. Clarification of genetic terms and their use in the management of captive populations. *Zoo Biology* 14:565-577.

Lynch M, 1988. Estimation of relatedness by DNA fingerprinting. *Mol Bio Evol.* 5:584-599.

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

Maynes GM, 1974. Occurrence and field recognition of *Macropus parma*. *Aust Zool.* 18:72-87.

Morin PA, Ryder OA, 1991. Founder contribution and pedigree inference in a captive breeding colony of lion-tailed macaques, using mitochondrial DNA and DNA fingerprint analyses. *Zoo Biol.* 10:341-352.

Norton JE, Ashley MV, 2004. Genetic variability and population differentiation in captive Baird's tapirs (*Tapirus bairdii*). *Zoo Biol.* 23:521-531.

- O'Brien SJ, Joslin P, Smith III, GL, Wolfe R, Schaffer N, Heath E, Ott-Joslin J, Rawal PP, Bhattacharjee KK, Martenson JS, 1987. Evidence for African origins of founders of the Asiatic lion Species Survival Plan. *Zoo Biol.* 6:99-116.
- Queller DC, Goodnight KF, 1989. Estimating relatedness using genetic markers. *Evolution.* 43:258-275.
- Ramirez O, Altet L, Ensenat C, Vila C, Sanchez A, Ruiz A, 2006. Genetic assessment of the Iberian wolf *Canis lupus signatus* captive breeding program. *Conserv Genet.* 7:861-878.
- Rice WR, 1989. Analyzing tables of statistical tests. *Evolution.* 43:223-225.
- Ride WDL, 1970. A guide to the native mammals of Australia. New York: Oxford University Press.
- Rousset F, 2008. GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Mol Ecol Notes.* 8:103-106.
- Rudnick JA, Lacy RC, 2008. The impact of assumptions about founder relationships on the effectiveness of captive breeding strategies. *Conserv Genet.* 9:1439-1450.
- Ruokonen M, Andersson A, Tegelstrom H, 2007. Using historical captive stocks in conservation. The case of the lesser white-fronted goose. *Conserv Genet.* 8:197-207.
- Russello A, Amato G, 2004. *Ex situ* population management in the absence of pedigree information. *Mol Ecol.* 13:2829-2840.
- Russello MA, Hyseni C, Gibbs JP, Cruz S, Marquez C, Tapia W, Velensky P, Powell JR, Caccone A, 2007. Lineage identification of Galapagos tortoises in captivity worldwide. *Anim Conserv.* 10:304-311.
- Sambrook J, Russell DW, 2001. Molecular cloning: a laboratory manual, 3rd edn. New York: Cold Spring Harbor Laboratory Press.
- Selvin S, 1980. Probability of nonpaternity determined by multiple allele codominant systems. *Am J Hum Genet.* 32:276-278.
- Sonesson AK, Meuwissen THE, 2001. Minimization of rate of inbreeding for small populations with overlapping generations. *Genet Res Camb.* 77:285-292.
- Spencer PBS, Odorico DM, Jones SJ, March HD, Miller DJ, 1995. Highly variable microsatellites in isolated colonies of the rock-wallaby (*Petrogale assimilis*). *Mol Ecol.* 4:523-525.
- Taylor AC, Cooper, DW, 1998. A set of tammar wallaby (*Macropus eugenii*) microsatellites tested for genetic linkage. *Mol Ecol.* 7: 925-931.

Threatened Species Conservation Act, 1995 (NSW). 1995. Parliament of New South Wales, Australia

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

Wang JL, 2002. An estimator for pairwise relatedness using molecular markers. *Genetics.* 160:1203-1215.

Williams RN, DeWoody JA, 2004. Fluorescent dUTP helps characterize 10 novel tetranucleotide microsatellites from an enriched salamander (*Ambystoma texanum*) genomic library. *Mol Ecol Notes.* 4:17-19.

Willis K, 2001. Unpedigreed populations and worst-case scenarios. *Zoo Biol.* 20:305-314.

Wodzicki K, Flux JEC, 1971. The parma wallaby and its future. *Oryx.* 11:40-47.

Zenger KR, Cooper DW, 2001. A set of highly polymorphic microsatellite markers developed for the eastern grey kangaroo (*Macropus giganteus*). *Mol Ecol Notes.* 1:98-100.

Zenger KR, McKenzie LM, Cooper DW, 2002. The first comprehensive genetic linkage map of a marsupial: the tammar wallaby (*Macropus eugenii*). *Genetics.* 162:321-330.

Figure 1. An example of unknown parentage in the parma wallaby pedigree. Individual 719 was living in the population as of January 01, 2007. Squares in the pedigree represent sires, circles represent dams, individuals with “wild” parents were population founders, and individuals with ‘?’ were born in captivity to undocumented parents. Historic gaps deep in a pedigree are difficult to resolve because genetic samples usually are not available for all individuals.

Figure 2. Known and potential founders of the captive parma wallaby population, graphed in order of their representation in the living population on January 01, 2007. Black bars represent founders for which biological samples were collected. Individuals 38-41 were potential founders that had yet to contribute offspring to the population (all four of these individuals also were sampled).

Figure 3. The four simulation scenarios used to test the impact of using different sets of founder relationships on the long-term genetic management of the parma wallaby PMP population. Each scenario used two f matrices, one represented the true relationships in the population and the other represented the relationships being tested. The “current” f matrix was the parma wallaby population living on January 01, 2007. The “modified” f matrix was generated through one of two methods used to identify first-order relatives.

Figure 4. Pairwise relatedness coefficients estimated for the 20 sampled founders, graphed in order of increasing value. A dashed horizontal line is drawn at 0.5, which is the relatedness coefficient of first-order relatives in a non-inbred diploid population. Values above the dashed 0.5 relatedness line (four values total) were considered to be closely related.

Figure 5. Trends in a) GD and b) \overline{F} for simulation scenarios that used a distribution of pairwise relatedness coefficients to identify first-order relatives among parma wallaby founders. The differences between the four tested scenarios were indistinguishable at this scale, thus, data is presented for only a single scenario: the current f matrix for the parma wallaby PMP was used for both the true founder relationships and the tested founder relationships (solid line). For comparison, the outcome under random mating also is included (dotted line). Figures 5 and 6 illustrate the final ten timesteps of these graphs, where differences between the four tested scenarios are visible.

Figure 6. Graphs of a) GD and b) \overline{F} that demonstrate the impact of designating selected founder pairs as first-order relatives when all founders were actually unrelated. Data is presented for the final ten timesteps of the following simulation scenarios: the current f matrix for the parma wallaby PMP was used for both the true founder relationships and the tested founder relationships (circle); the current f matrix for the parma wallaby PMP was used for the true founder relationships and a modified f matrix was used for the tested founder relationships (diamond). The distribution of pairwise relatedness coefficients for the newly acquired parma wallaby founders was used to identify first-order relatives for the modified f matrix.

Figure 7. Graphs of a) GD and b) \overline{F} that demonstrate the impact of correctly designating selected founder pairs as first-order relatives rather than assuming all founders to be unrelated. Data is presented for the final ten timesteps of the following simulation scenarios: a modified f matrix was used for both the true founder relationships and the tested founder relationships

(triangle); a modified f matrix was used for the true founder relationships and the current f matrix for the parma wallaby PMP was used for the tested founder relationships (square). The distribution of pairwise relatedness coefficients for the newly acquired parma wallaby founders was used to identify first-order relatives for the modified f matrix.

Figure 8. Changes in individual mean kinships due to identifying first-order relatives among the captive parma wallaby founders. Mean kinships for a total of 38 individuals are graphed for three different sets of founder relationships: founders were unrelated (■), four pairs of first-order relatives were present among the founders (▣), and seven pairs of first-order relatives were present among the founders (▤). The 38 individuals for which data are presented are the founders ($n = 8$) and their descendants ($n = 30$) living on January 01, 2007 that experienced a change in mean kinship when first-order relatives were present among the founders.

Table 1. Types of contributions that molecular markers can make to captive breeding programs.

Contribution to Breeding Program	Species	Citation
assessment of founder relationships	California condor (<i>Gymnogyps californianus</i>)	Geyer et al. 1993
	Guam rail (<i>Rallus owstoni</i>)	Haig et al. 1994
	Micronesian kingfishers (<i>Halcyon cinnamomina</i>)	Haig et al. 1995
	bearded vulture (<i>Gypaetus barbatus</i>)	Gautschi et al. 2003
pedigree reconstruction	lion-tailed macaques (<i>Macaca silenus</i>)	Morin and Ryder 1991
	Przewalski's horse (<i>Equus ferus przewalskii</i>)	Bowling et al. 2003
subspecies identification	lion (<i>Panthera leo</i>)	O'Brien et al. 1987
	chimpanzee (<i>Pan troglodytes</i>)	Ely et al. 2005
identification of geographic origin	Galapagos tortoise (<i>Geochelone nigra</i>)	Russello et al. 2007
quantification of wild genetic diversity captured	Baird's tapir (<i>Tapirus bairdii</i>)	Norton and Ashley 2004
	Iberian wolf (<i>Canis lupus signatus</i>)	Ramirez et al. 2006
identification of genetically valuable individuals	whooping crane (<i>Grus americana</i>)	Jones et al. 2002
	St Vincent parrot (<i>Amazona guildingii</i>)	Russello and Amato 2004
assessment of hybridization	lesser white-fronted goose (<i>Anser erythropus</i>)	Ruokonen et al. 2007

Table 2. Characterization of 9 microsatellite loci in the parma wallaby. The DNA sequences from which the MP primers were designed have been deposited in GenBank under the accession numbers EU851732 (MP03), EU851733 (MP04), and EU851734 (MP06).

Locus	Source or Primer Sequences (5'-3')	T _a (°C)	Size (bp)	No. of Alleles	H _O	H _E
Pa595	Spencer et al. (1995)	58	316-346	7	0.63	0.72
Me14	Taylor and Cooper (1998)	65	156-176	2	0.52	0.49
Me17	Taylor and Cooper (1998)	63	133-143	4	0.55	0.56
Y175	Zenger et al. (2002)	58	275-291	4	0.52	0.55
G19-1	Zenger and Cooper (2002)	65	169-201	6	0.77	0.77
G31-1	Zenger and Cooper (2002)	62	118-122	3	0.42	0.52
MP03	F: CCATGATTTGACTAGCCTGGA R: TCTACTTCTTTTCCTCAACTTGAAC	50	196-240	7	0.80	0.79
MP04	F: TCTCTACAAAATAGAGATGTCCGTGT R: AGAACTTCTTTGCATATTTGACTTT	50	143-231	7	0.42	0.50
MP06	F: TGATAGATCGATTGACCGATTG R: GAAGCCAGGTATGCCGATTA	50	235-271	8	0.76	0.75

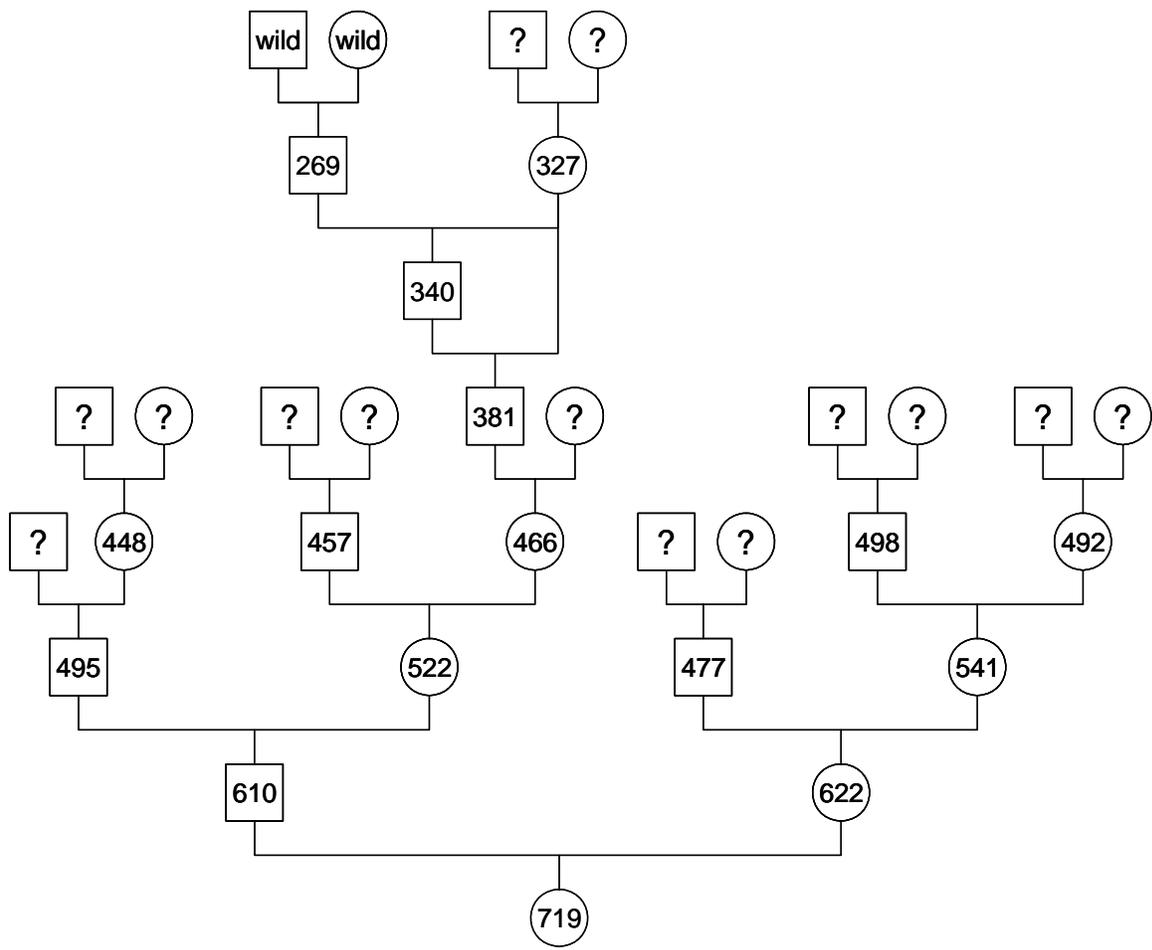
T_a, annealing temperature; H_O, observed heterozygosity; H_E, expected heterozygosity

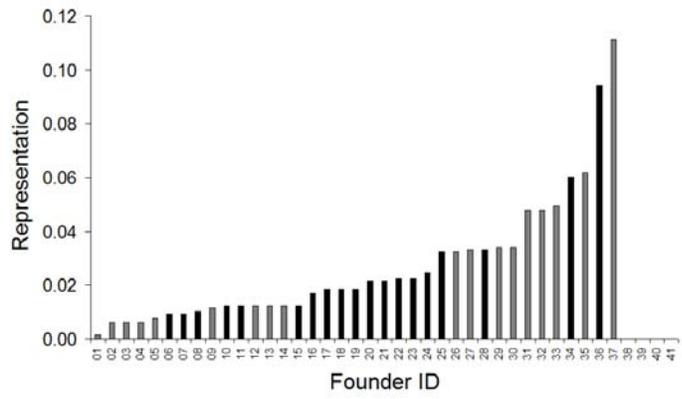
Table 3. Mean relatedness coefficients and variances for three relatedness estimators, based on allele frequencies from the captive North American parma wallaby population. Values were based on 1000 simulated pairs from each of the following relationship categories: unrelated (UR), half-siblings (HS), full-siblings (FS), and parent-offspring (PO). The smallest sampling variances per relationship category are marked with an asterisk.

	Relationship			
	UR	HS	FS	PO
r_{xyQG}	0.0167 (0.0704)	0.2510 (0.0668)	0.5065 (0.0573)	0.5030 (0.0304)
r_{xyLR}	0.0159 (0.0480)*	0.2509 (0.0674)	0.5052 (0.0643)	0.4952 (0.0437)
r_{xyWang}	0.0199 (0.0784)	0.2425 (0.0654)*	0.5063 (0.0527)*	0.4998 (0.0198)*

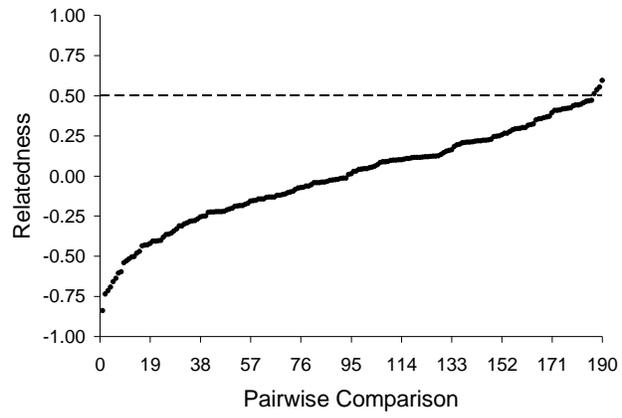
Table 4. Final gene diversity (GD) and mean inbreeding (F) values after 100 simulation timesteps for the four scenarios that were designed to quantify the impacts of using different sets of founder relationships on the long-term genetic management of a captive parma wallaby population. First order relatives were identified among the founders by either an observed break in the distribution of all empirically generated pairwise kinships, or by a specified kinship cutoff value. In all cases, the 95% confidence intervals around mean GD and F at the end of the 100 timesteps were approximately 0.0001.

Scenario	Distribution Break		Specified Cutoff	
	GD	F	GD	F
1: no first-order relatives	0.9334	0.0529	0.9334	0.0529
2: incorrect identification of first-order relatives	0.9330	0.0533	0.9325	0.0538
3: failure to detect first-order relatives	0.9310	0.0533	0.9291	0.0572
4. correct identification of first-order relatives	0.9319	0.0544	0.9311	0.0552

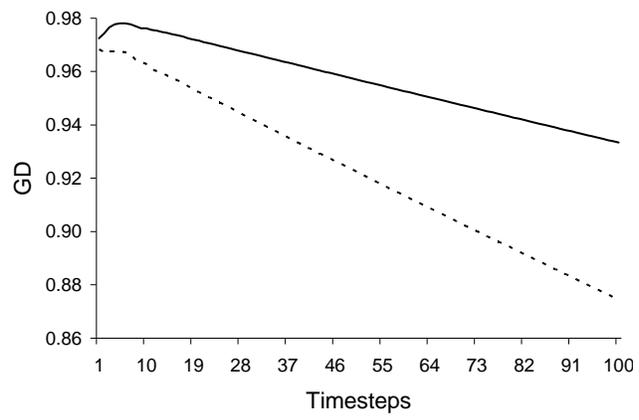




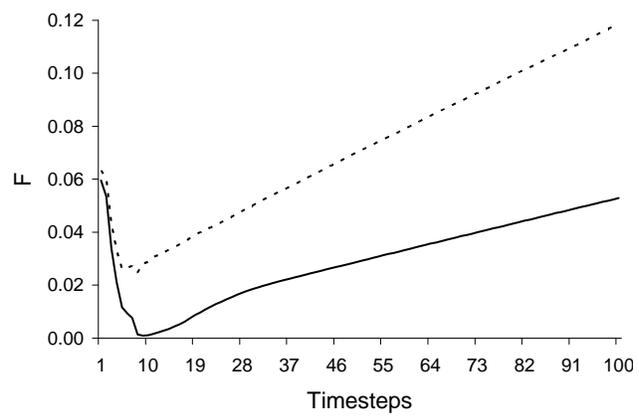
Scenario 1	<p>true relationships → current <i>f</i> matrix</p> <p>tested relationships → current <i>f</i> matrix</p>	no first-order relatives were present among the founders	<p>compared to quantify the detriments to incorrectly identifying founders as first-order relatives when they were actually unrelated</p> <p>compared to quantify the benefits to correctly identifying founders as first-order relatives</p>
Scenario 2	<p>true relationships → current <i>f</i> matrix</p> <p>tested relationships → modified <i>f</i> matrix</p>	molecular data incorrectly identified first-order relatives among the founders when no such relationships existed	
Scenario 3	<p>true relationships → modified <i>f</i> matrix</p> <p>tested relationships → current <i>f</i> matrix</p>	molecular data failed to identify first-order relatives that were actually present among the founders	
Scenario 4	<p>true relationships → modified <i>f</i> matrix</p> <p>tested relationships → modified <i>f</i> matrix</p>	molecular data correctly identified the first-order relatives that were actually present among the founders	



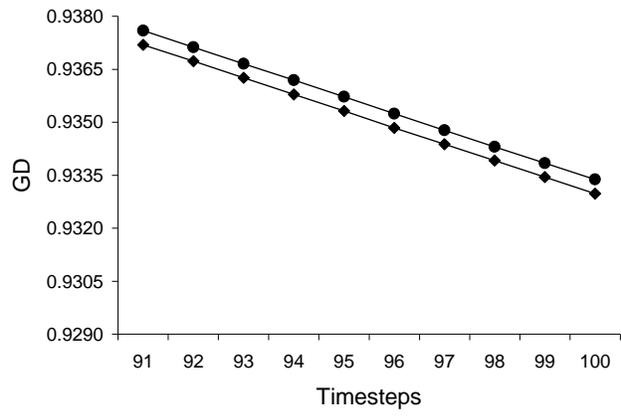
a)



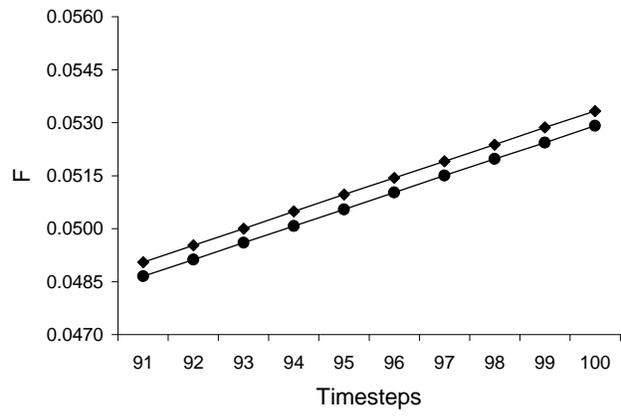
b)



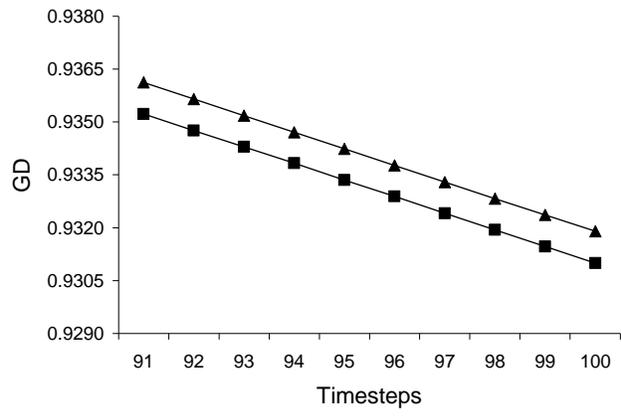
a)



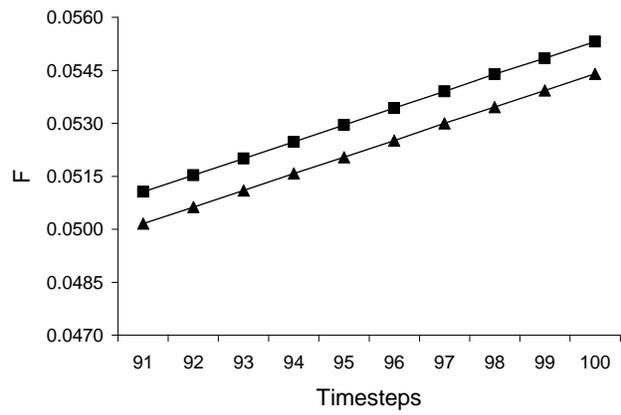
b)

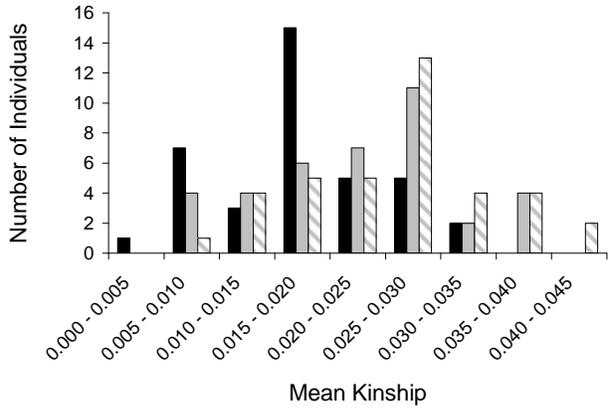


a)



b)





Supplementary Material

Specific details and parameterization of the general computer simulation used to investigate the relationships among parma wallaby founders.

1. The population of 157 parma wallabies living on January 01, 2007 was imported into the simulation with two accompanying f matrices. One f matrix represented the true relationships in the population, while the other represented the relationships being tested.
2. Breeding pairs were chosen from the available individuals in such a way as to select a group of individuals that exhibited the lowest overall mean kinship. The number of pairs selected was designed to maintain a population size of 160 and was equal to $(n + m)/p$, where p was the probability that a breeding pair produced offspring, n was the difference between 160 and the current population size, and m was the total number of individuals that were expected to be lost from the population during a given timestep due to mortality. If the specified number of breeding pairs could not be made due to the current population size, the maximum number of pairs possible was made instead. Selection of breeding pairs was as follows:
 - a. The mk of each individual was initially calculated.
 - b. The male and female with the lowest mks were selected to breed, the pair produced a single offspring that was temporarily added to the population, and all mks were recalculated. Once created, temporary offspring remained in the population throughout the pairing process and were included in all subsequent mk calculations, but they were never selected to breed.
 - c. Breeding pair selection continued until the desired number of breeding pairs was created. To avoid close inbreeding, any breeding pair that exhibited an f greater than the average f in the population was rejected. An offspring's inbreeding coefficient (F) is equal to the f of its parents (Falconer 1981), thus, limiting breeding pairs based on f slowed the accumulation of inbreeding in the population. If a pair was rejected, all females remaining to be paired were evaluated in order of increasing mk to determine if one was a suitable match for the male of the rejected pair. A pair was made if a suitable female was found. If a suitable female was not found, that male was removed from the pool of potential breeders, mks were recalculated, and pairing continued.
 - d. At the end of pair selection, all hypothetical offspring that had been created for the dynamic mk calculations were removed from the population. The breeding pairs selected throughout the pairing process became the list of breeding pairs used to continue the simulation.
3. After all breeding pairs were selected, each pair had a 0.8 probability of producing a single offspring (p). Each offspring was assigned one sex or the other with a 0.5 probability.
4. After all offspring were produced, the relationships between all individuals currently in the population were quantified and recorded. For relationships to be tracked through time, a matrix of all possible pairwise fs (including an individual's f with itself) and each individual's inbreeding coefficient (F , equal to the kinship between the individual's sire and dam; Falconer 1981) were calculated each timestep. Pairwise fs were calculated as $f_{xy} = 0.5(f_{xs} + f_{xd})$, where the subscripts s and d refer to the sire and dam of individual y (Falconer 1981). Two f matrices were calculated

each timestep; one matrix was derived from the true relationships in the population and the other was derived from the relationships being tested.

5. Individuals were aged one timestep, and animals older than 10 years of age were removed from the population. To simulate stochastic mortality, individuals 10 years of age or less were removed from the population with a 0.1 probability.

Steps 2-5 were repeated for 100 timesteps. Genetic variation and inbreeding were evaluated on a per timestep basis, immediately following step 5. Inbreeding was measured as the average inbreeding coefficient (\bar{F}) and genetic variation was measured as proportional gene diversity (GD). GD was calculated as $1 - \bar{mk}$, where \bar{mk} was the average mean kinship in the

population ($\bar{mk} = \frac{\sum_{x=1}^N mk_x}{N}$; Ballou and Lacy 1995). GD and \bar{F} were calculated from the f

matrix derived from the population's true relationships, and both values could range from zero to one.