Evaluating the ability of Bayesian clustering methods to detect hybridization and introgression using an empirical red wolf data set

JUSTIN H. BOHLING,* JENNIFER R. ADAMS† and LISETTE P. WAITS†

Department of Ecosystem Science and Management, Penn State University, University Park, PA 16802, USA, †Department of Fish and Wildlife Sciences, University of Idaho, Moscow, ID 83844-1136, USA

Abstract

Bayesian clustering methods have emerged as a popular tool for assessing hybridization using genetic markers. Simulation studies have shown these methods perform well under certain conditions; however, these methods have not been evaluated using empirical data sets with individuals of known ancestry. We evaluated the performance of two clustering programs, BAPS and STRUCTURE, with genetic data from a reintroduced red wolf (*Canis rufus*) population in North Carolina, USA. Red wolves hybridize with coyotes (*C. latrans*), and a single hybridization event resulted in introgression of coyote genes into the red wolf population. A detailed pedigree has been reconstructed for the wild red wolf population that includes individuals of 50–100% red wolf ancestry, providing an ideal case study for evaluating the ability of these methods to estimate admixture. Using 17 microsatellite loci, we tested the programs using different training set compositions and varying numbers of loci. STRUCTURE was more likely than BAPS to detect an admixed genotype and correctly estimate an individual's true ancestry composition. However, STRUCTURE was more likely to misclassify a pure individual as a hybrid. Both programs were outperformed by a maximum-likelihood-based test designed specifically for this system, which never misclassified a hybrid (50–75% red wolf) as a red wolf or vice versa. Training set composition and the number of loci both had an impact on accuracy but their relative importance varied depending on the program. Our findings demonstrate the importance of evaluating methods used for detecting admixture in the context of endangered species management.

Keywords: BAPS, *Canis*, coyote, genetic introgression, pedigree, STRUCTURE

Introduction

In both evolutionary and conservation biology, there is increasing interest in interspecific hybridization and its frequency in natural systems (Rhymer & Simberloff 1996; Arnold 1997; Allendorf *et al.* 2001). However, the most basic challenge associated with studying hybridization is identifying hybrid individuals within a population. Recently, the development of clustering methods and assignment tests to analyse multilocus genotypes has revolutionized the way biologists study hybrid

Correspondence: Justin H. Bohling, Fax: 814-863-4710;
E-mail: jhb24@psu.edu

systems (Pearse & Crandall 2004; Manel *et al.* 2005). These methods provide a statistical basis to evaluate the contribution of various genetic groups to an individual's ancestry (Rannala & Mountain 1997; Cornuet *et al.* 1999; Pritchard *et al.* 2000). Often within a Bayesian framework, researchers can determine the number of genetic groups present in a system and estimate the fraction of an individual's genotype that originated from the identified groups (Pritchard *et al.* 2000; Corander *et al.* 2003). Insights provided by these methods aid in the characterization of hybrid systems and implementation of conservation programs designed to limit genetic introgression in endangered species (Manel *et al.* 2005).

However, these programs can Possess several important limitations. Their precision can depend on the level of genetic differentiation between parental populations; low levels of divergence in allele frequencies between the parental groups can limit their ability to identify the correct number of groups and assign individuals to the correct group (Latch *et al.* 2006; Vähä & Primmer 2006; Kalinowski 2010). The amount of data utilized, such as number of loci, individual genotypes, and reference populations, can impact their efficacy (Rosenberg *et al.* 2005; Latch *et al.* 2006; Vähä & Primmer 2006; Sanz *et al.* 2009). Some of these clustering programs allow the input of prior population information so that predefined individuals are used to generate allele frequencies, which can alter their performance compared with 'free clustering' methods (Piry *et al.* 2004; Hauser *et al.* 2006; Barilani *et al.* 2007; Pritchard *et al.* 2007; but see Vähä & Primmer 2006).

Interpreting the results of these programs also poses a challenge because the ancestry coefficients (*q*-values) they generate can be difficult to relate to true genetic ancestry (Vähä & Primmer 2006). To our knowledge, there is no standard, objective way to interpret these values: most studies simply set an arbitrary probability threshold in classifying an individual as 'pure' vs. 'admixed'. Several recent studies have incorporated simulations of hybrid genotypes to determine thresholds of ancestry values that are indicative of hybrid ancestry (Lancaster *et al.* 2006; Barilani *et al.* 2007). Another important consideration is that similar ancestry coefficients can be produced for different generations of hybrids, making it difficult to differentiate F1 hybrids from later generational backcrosses (Anderson & Thompson 2002; Lancaster *et al.* 2006; Barilani *et al.* 2007). Finally, many software programs have been developed that exploit the same basic Bayesian philosophy of estimating ancestry; yet, they can produce incongruent results for the same data set (Vähä & Primmer 2006; Pritchard *et al.* 2007; Burgarella *et al.* 2009; Sanz *et al.* 2009).

Past studies have compared the ability of these programs to detect hybridization using simulated data sets (Vähä & Primmer 2006; Sanz *et al.* 2009). Simulation studies allow researchers to manipulate different variables to evaluate the factors that influence the performance of these programs. However, simulation studies often cannot reflect the complexity exhibited by natural systems. Ideally, researchers should test these methods using data sets from the study system to develop specific criteria to assess hybridization. Our goal in this study was to compare the accuracy of two commonly used clustering programs, STRUCTURE (Pritchard *et al.* 2000; Falush *et al.* 2003) and BAPS (Corander *et al.* 2003, 2006; Corander & Marttinen 2006), in estimating ancestry coefficients using an empirical red wolf (*Canis rufus*) data set. Here, we utilize over 20 years of pedigree data from a reintroduced red wolf population in eastern North Carolina that has experienced hybridization with coyotes (*C. latrans*). Using this pedigree, we calculated the known contribution of red wolf ancestry to an individual's genotype and compared those to the estimates produced by STRUCTURE and BAPS. The results of both programs were also compared with the Canid Assignment Test (CAT), which is a maximum-likelihood assignment program designed specifically for this system. We also tested these programs while altering the number of loci and the composition of training sets to evaluate how those variables affect the performance of these programs. Our results provide an empirical test of the ability of these popular programs to estimate the proportion of ancestry for hybrid individuals under a range of conditions.

Study system

Historically, red wolves were distributed across eastern North America, but the species was driven to near extinction by the middle part of the 20th century due to habitat destruction, overharvesting by humans and hybridization with coyotes (McCarley 1962; Paradiso & Nowak 1972; USFWS 1990). Hybrids between red wolves and coyotes are fertile and reproduce with both parental species (Adams *et al.* 2003; Miller *et al.* 2003). In 1987, the US Fish and Wildlife Service (USFWS) reintroduced four pairs of red wolves into the Albemarle Peninsula in northeastern North Carolina, USA (Phillips & Parker 1988). Today, due to natural reproduction and subsequent releases, the population has grown to 100–120 adult individuals (USFWS 2007). In 1993, the first known hybridization event between a member of the reintroduced red wolf population and a coyote occurred (Kelly *et al.* 1999). In response, the USFWS, in collaboration with the Red Wolf Recovery Implementation Team (RWRIT, Stoskopf *et al.* 2005), developed an aggressive adaptive management plan to limit genetic introgression from coyotes into the red wolf gene pool.

A critical component of that plan was to reconstruct the pedigree of the red wolf population to monitor the genetic ancestry of individual animals. Using microsatellite markers, Adams (2006) reconstructed the pedigree of the population. The first hybridization event in 1993 resulted in introgression of coyote genetic material into the red wolf population. Two male F1 hybrids were produced from this event and bred with female red wolves. Subsequent backcrossing with the red wolf population has created a gradient of red wolf ancestry within the population. The USFWS and RWRIT established a policy that any individual determined to have

$\geq$ 87.5% red wolf ancestry through the pedigree would be considered a red wolf (Stoskopf *et al.* 2005). For the remainder of this article, any individual that is 100% red wolf (i.e. 100% of its ancestry traced back to the founders of the captive population) will be called a 'pure' red wolf; any individual with 87.5–99% red wolf ancestry (i.e. some coyote ancestry through the 1993 hybridization event) will be called a 'backcrossed' red wolf. Additional F1 hybridization events have been detected through genetic testing and pedigree analysis, but none of these events has resulted in introgression into the red wolf gene pool (Adams 2006; USFWS 2007).

Methods

Genetic testing

Over the past 24 years, the USFWS has been capturing and monitoring canids over the 6000 km^2 peninsula. Blood samples have been collected from captive individuals released into the wild, adult individuals captured during field surveys, and newborn pups. Every spring USFWS biologists locate red wolf dens and collect blood samples from pups to locate hybrid individuals and remove them from the population before they become old enough to disperse and reproduce. For this article, we examined wild canids captured from this area from 1987 to 2008, during which a total of 824 individual canids were captured and genotyped.

Genetic analysis of blood samples followed the methods outlined in Miller *et al.* (2003) and Adams (2006). In summary, each individual was genotyped at 18 microsatellite loci using a 3130*xl* ABI sequencer. One locus (CXX30) was difficult to score, and this locus was removed from the data set. Individuals were initially assigned to species or hybrid groups using the CAT developed by Miller *et al.* (2003). This maximum-likelihood assignment test was developed specifically to identify red wolves, coyotes and their hybrids and currently is the test accepted by the USFWS for management purposes (see Stoskopf *et al.* 2005). The CAT assigns individuals to one of six groups: red wolf, coyote, F1 hybrid, F2 hybrid, F1xred wolf and F1xcoyote backcross. Allele frequency distributions of the red wolf population were generated by simulating genotypes from the founders of the captive red wolf population while modelling the effects of genetic drift (Miller *et al.* 2003). Coyotes harvested in North Carolina and Virginia were used as source populations for the coyote component of this test.

Pedigree analysis followed the methods developed by Adams (2006) and was used to determine the parentage of red wolves and hybrids. To summarize, parentage was determined using a combination of field and genetic data. USFWS biologists typically identified potential parents of a newly captured red wolf or litter based upon the proximity of the various red wolf packs and knowledge of breeding pairs. The exclusion method was first used to determine whether wolves hypothesized to be potential parents could be parents based on genotypic data. We confirmed these hypothesized relationships for 301 of 303 individuals using the program Cervus (Marshall *et al.* 1998; Kalinowski *et al.* 2007). Parents were unknown or uncertain for ~39% of captured individuals. In those situations, we used Cervus to identify the most likely parents from the potential pool of reproductive individuals in the population. We required a $\geq$ 95% confidence level and allowed a maximum of one mismatch for a potential parent pair. We also checked all parentage assignments with 1 locus mismatch to confirm that the pairing was realistic based on detailed field observations and/or telemetry of wolves during the breeding season. Twenty-three per cent of identified parent-offspring relationships had 1 genotypic mismatch; the remainder had zero mismatches. Parentage of coyotes was often not evaluated.

Of the 824 documented canids, 630 were confidently placed in the red wolf pedigree and assigned some amount fo red wolf ancestry using the methods described above (Table 1). The remaining individuals were determined by the CAT to be 132 coyotes, 60 hybrids and two red wolves, none of which could reliably be placed in the pedigree. Of the 630 animals with red wolf ancestry, 287 were 'pure' red wolves while 262 qualified as 'backcrossed' red wolves (87.5–99% red wolf), meaning that 549 individuals were considered red wolves under USFWS guidelines. Eighty-one individuals were hybrids with known red wolf parent-

Table 1 Number of animals with known red wolf ancestry based on the pedigree. Values in italics indicate the ancestry values that belong to the 'Backcross red wolf ancestry' class. Underlined values indicate ancestry values classified in the 'F1× Red wolf backcross' class. Bold values indicate ancestry values classified in the 'F1 hybrid' class

% Red wolf	Count
100	287
98.45	*9*
96.875	*62*
93.75	*110*
92	*3*
90.6	*12*
87.5	*66*
<u>75</u>	<u>30</u>
<u>68.75</u>	<u>3</u>
50	**33**
43.75	**15**

age, including 48 F1 hybrids and 33 F1 x red wolf back-crosses.

A major motivation for using the programs STRUCTURE and BAPS for this study is both allow the incorporation of more than two parental populations, which is optimal for this system because there is potential for genetic introgression from domestic dogs (*C. familiaris*) and grey wolves (*C. lupus*) (Bohling & Waits 2011; Bozarth *et al.* 2011). Other software programs designed specifically to test for hybridization, such as New-Hybrids (Anderson & Thompson 2002), do not allow the incorporation of multiple species and were not included in our study.

Training set comparison

Although other studies have noted the advantages of using training sets within the clustering framework (e.g. Hauser *et al.* 2006), none have evaluated whether training set composition impacts ancestry estimation. We developed three different training sets of red wolf genotypes and evaluated their impact on estimates of ancestry. The first training set we tested (Training set-Founders) only included genotypes from 17 red wolves captured in the 1970s that were the original founders of the captive population. These 17 individuals composed the pre-defined red wolf population in both clustering programs. Every other canid captured or released into the wild from 1987 to 2008 was classified as 'unknown' origin in these analyses.

The second training set (Training set-Updated) included the genotypes of the founders and was updated each year with 'Pure' red wolves that were either born in the wild and identified through the pedigree or born in captivity and released into the wild. This training set grew in size each year as more 'Pure' red wolves were added to the population. In each year, genotypes of newly captured or released captive individuals were included as 'unknowns' in each analysis. The exception was 1987, which only included the founders in the training set because no other red wolves existed in North Carolina prior to the initial releases. As an example of the way this training set operated, the 1988 red wolf training set included the founder red wolves and individuals that had been released in 1987. By 2008, the training set had grown to include the founders ($n = 17$) and every 'Pure' red wolf that had been released or born from 1987 to 2007 ($n = 278$). We ran a separate analysis for each year from 1987 to 2008.

The third training set (Training set-12 years) was restricted to 'Pure' red wolves that had been born or released within 12 years of the year being examined. Due to the potential for genetic drift and nonrandom breeding, we hypothesized that the allele frequencies of the red wolf population changed over time. We were interested if this potential shift in allele frequencies would affect assignment accuracy, especially for current red wolves. Twelve years was selected as the time frame for each training set because it represents the potential maximum age of reproductive red wolves in the wild. Thus, the third training set was designed to only include potentially reproductive individuals when each cohort of red wolves entered the population. As before, the year the 1988 red wolf training set included the founders because no other red wolves existed in the wild prior to those releases. In 1988, however, the training set included only red wolves released in 1987 ($n = 13$). By 2008, the training set only included red wolves born from 1996 to 2007 ($n = 170$). A separate analysis was completed for each year from 1987 to 2008.

Each clustering analysis also included genotypes from coyotes, dogs and grey wolves to compose pre-defined populations. The training sets for coyotes, dogs and grey wolves remained constant and were not altered across the analyses. We assumed the allele frequencies for these three groups remained constant over time because of their larger effective population sizes. These training sets consisted of 61 coyotes from North Carolina and Virginia (Miller *et al.* 2003); 37 grey wolves from Idaho and Alaska; and 27 domestic dogs. These individuals had previously been identified as having >95% posterior probability assignment to their respective group using STRUCTURE (Bohling & Waits 2011) and did not contain any admixed individuals.

For the four populations present in the 'Founders' training set, we calculated average allelic richness, H_0, F_{IS} and number of private alleles using GenAlEx (Peakall & Smouse 2006) and pairwise F_{ST} between all four species using FSTAT (Goudet 2002). To examine the effects of genetic drift on the red wolf population, we calculated pairwise F_{ST} values between 'pure' red wolves and the other three canid species for the 'Update' and '12 years' training sets. F_{ST} was estimated between the red wolf training set for each year from 1988 to 2008, and the other species used in the training sets. We also examined the change in allelic richness, H_0, and F_{IS} for the 'pure' red wolf population by estimating those measures for each year in the '12 years' training set. The values of these variables produced by the 'Founders' were used as a starting point for these comparisons.

In the STRUCTURE analyses, each known individual for all four training sets was assigned to its putative species with the designation POPFLAG = 1 (Pritchard *et al.* 2003), and the 'Use Population Information' option was chosen as the ancestry model with the default parameter settings. Unknown individuals were

given the POPFLAG = 0 designation, meaning STRUCTURE only updated allele frequencies with the genotypes of known individuals. The number of populations was set to four with a burn-in period of 10^5 reps followed by 10^6 MCMC reps. We ran one iteration of STRUCTURE for each separate analysis: preliminary analyses found little variation in ancestry values between multiple runs at the same value of K (see Bohling & Waits 2011). For each run, STRUCTURE estimated 90% credibility intervals around ancestry coefficients using the default parameters.

For BAPS, the fixed K clustering option was enabled, and a 'Trained clustering' analysis was conducted using known individuals as the prior information and unknown individuals as the sampling units (Corander *et al.* 2006). The upper bound on the number of clusters was set to $K = 4$. These results were used to conduct an 'Admixture based on mixture clustering' analysis to estimate ancestry coefficients. The simulations consisted of 100 iterations to estimate ancestry coefficients for our genotypes, 200 simulated reference individuals from each population, and 20 iterations to estimate ancestry coefficients for these reference individuals.

Three separate methods were utilized to test the accuracy and precision of these estimates. To evaluate precision, we used one-way ANOVAs to examine the variability in the ancestry coefficients produced by each program. For these analyses, individuals were grouped into four categories: 'Pure' red wolves (100% red wolf), 'Backcrossed' red wolves (87.5–99%), 'F1x red wolf backcrosses' (68.5–75% red wolf) and 'F1 hybrids' (43.5–50% red wolf) based on their pedigree assignments (Table 1). An ANOVA was performed for each training set/program combination (six in total) and Tukey's post hoc test examined pairwise significance between the estimates for each category. Precision was also assessed by calculating the average difference between the proportions of ancestry based on the pedigree vs. those estimated by the programs.

Accuracy was tested in two ways: first, we estimated the correlation coefficient between the proportion of known red wolf ancestry and the ancestry estimate for each training set/program combination (six total) Second, we estimated misclassification rates for each program/training set combination. For calculating misclassification rates, we set 0.875 as the ancestry coefficient threshold for considering an individual a 'Pure' red wolf. If the estimated ancestry coefficient was less than this value, then an individual was classified as a hybrid. The CAT does not produce ancestry coefficients; therefore, we based misclassification rates on whether a legal red wolf (87.5–100%) was classified as a hybrid (either F1 hybrid or F1 x red wolf backcross) or a hybrid as a red wolf.

Varying the number of loci

We chose four subsets from our full suite of loci to examine the impact of loci number on ancestry analysis: 15, 12, 9 and 6 loci. Loci were selected using a measure of loci informativeness (I) implemented in GenAlEx (Peakall & Smouse 2006). This measure is similar to the Shannon index of diversity used in ecological studies, except this measure uses allelic diversity as the variable of interest. For each subset of loci, we included the loci that had the highest I-values to mimic study designs where researchers pick the most informative loci for their analyses. We ran these different subsets in both programs using only the 'Founders' training set. All individuals captured by the USFWS in North Carolina ($n = 824$) served as the unknown samples in this analysis. However, we only included those individuals that could be placed in the pedigree ($n = 630$) in our statistical analyses. To test the effect of loci number, we ran the same statistical analyses as we did for the training set comparison. Unknown individuals were grouped into the four ancestry classes described above, and ANOVAs were used to test for differences in the means. We calculated absolute average difference and correlation coefficients between real and estimated ancestry for each individual. Misclassifications rates were based on the same ancestry classes as the training set comparison, except the subsets were not employed in the CAT.

Results

Training set comparison

Pairwise F_{ST} values between the four species in the 'Founders' training set were moderate (0.10–0.17), with the red wolf founders demonstrating the highest levels of differentiation (Table S1, Supporting information). The red wolf founders had the lowest allelic richness (4.6), whereas coyotes had the highest (7.3); however, red wolves had moderate H_o (0.67) compared with the other species (Table S2, Supporting information). F_{IS} values suggest both the grey wolf and dog groups demonstrated homozygote excess; a similar but weak signal was detected in the coyotes, and the red wolf founders had a weak heterozygote excess. Coyotes had by far the most private alleles of any of the groups, although the other three species also possessed several private alleles.

For both the 'Update' and '12 years' training sets, pairwise F_{ST} values with the other three species increased over time, although they increased more rapidly and to higher levels in the '12 years' set (Fig. S1, Supporting information). Both allelic richness (Fig. S2, Supporting information) and H_0 (Fig. S3, Supporting

information) decreased over time in the '12 years' training set with the largest drop between the red wolf founders and the first individuals released into the wild. Both values showed periodic increases over the course of 21 years: these coincide with instances where captive individuals were released into the wild. F_{IS} values varied between positive and negative values with no discernible trend over time (data not shown).

Regardless of the training set or program, ancestry estimates were more variable for individuals with

(a)

(b)
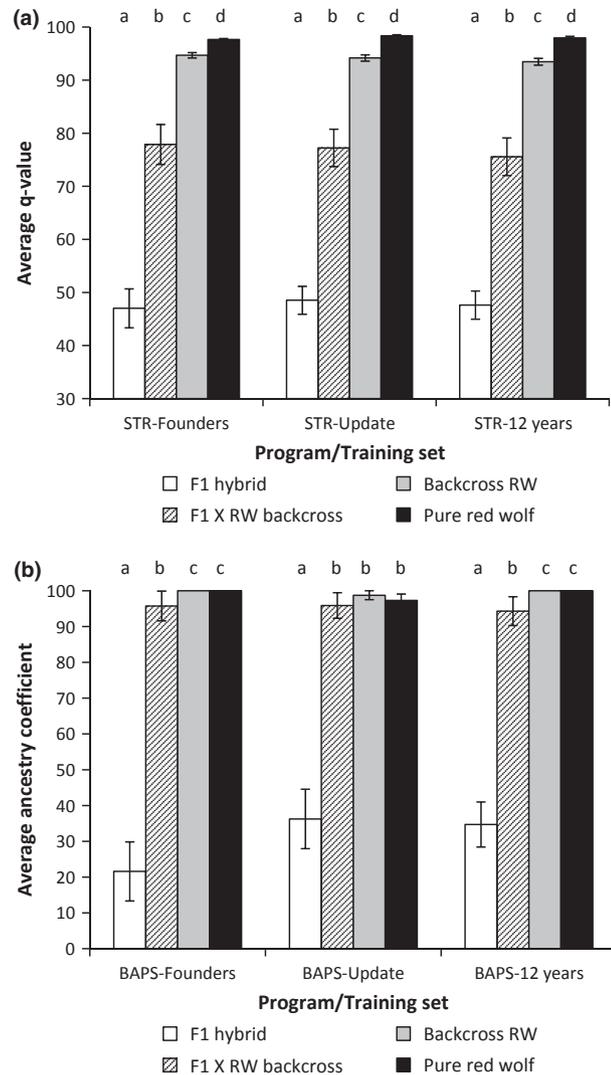
Fig. 1 Average ancestry coefficient and 95% confidence intervals estimated for each ancestry class using each program/training set combination. These were calculated by averaging the ancestry coefficients for all individuals within each ancestry class from a single run of each program. Lower case letters indicate significant deviations detected between ancestry classes using Tukey's post hoc test. The letters 'STR' refer to program STRUCTURE and 'RW' to red wolf. (a) contains the results for STRUCTURE and (b) the results from BAPS.

higher amounts of coyote ancestry (Fig. 1). A significant effect of ancestry class on ancestry coefficients was detected for every program/training set combination ($P < 0.0005$, Fig. 1). The post hoc tests showed that the mean ancestry values produced by each STRUCTURE/training set combination for all four ancestry classes were significantly different from each other. For BAPS, using the 'Founders' and '12 years' training set, all ancestry groups were significantly different ($P < 0.05$) except for the 'Pure' and 'Backcross' red wolf groups ($P > 0.9$). Using the 'Update' training set in BAPS, estimates from the 'Backcross', 'Pure' and 'F1xRW backcross' classes were not significantly different ($P > 0.6$).

The highest correlations between q-values and true ancestry values were found using STRUCTURE regardless of the training set ($r > 0.90$) (Fig. 2). None of the training sets when used in BAPS produced correlations $r > 0.90$ between ancestry coefficients and known ancestry values. The 'Founders' and '12 years' training sets in BAPS generated correlations $r > 0.8$ (0.834 and 0.853, respectively); the lowest correlation was observed for the 'Update' training set ($r = 0.665$).

There was a large range in the ancestry estimates produced by both programs for individuals with similar levels of ancestry. Across all training sets, the range in estimated red wolf ancestry produced by STRUCTURE for 'F1 hybrids' stretched from 23.3% to 72.8%. BAPS produced a range of estimated red wolf ancestry for 'F1 hybrids' from 0% to 100%. The range in values for 'F1xRW backcrosses' produced by STRUCTURE went from 58.5% to 99.4% and by BAPS from 65% to 100%. The average absolute difference between the known ancestry value and the estimated value from the clustering programs was lowest for individuals with high red wolf ancestry (Fig. S4, Supporting information). For STRUCTURE, individuals with the smallest level of difference between real and estimated ancestry values had 98% red wolf ancestry: this reflects the fact that STRUCTURE never classified an individual as having 100% ancestry for any genetic group. 'Pure' red wolves had the smallest difference between real and estimated ancestry from BAPS (Fig. S4, Supporting information).

For the credibility intervals (CIs) surrounding the q-values produced by STRUCTURE, we examined whether they overlapped the expected value of red wolf ancestry based on the pedigree. Considering the 'Founders' training set, over 96% of individuals had 90% CIs that contained the expected ancestry value. For 'Pure' red wolves, the average width of the CIs around the red wolf point ancestry value was 11.1 (SD ± 5.2). The average CI width for 'Backcrossed' red wolves was 18.4 (SD ± 8.8), for the 'F1 × hybrid backcross' individuals was 39.7 (SD ± 9.1), and 'F1 hybrid' individuals was 53.9 (SD ± 7.4).
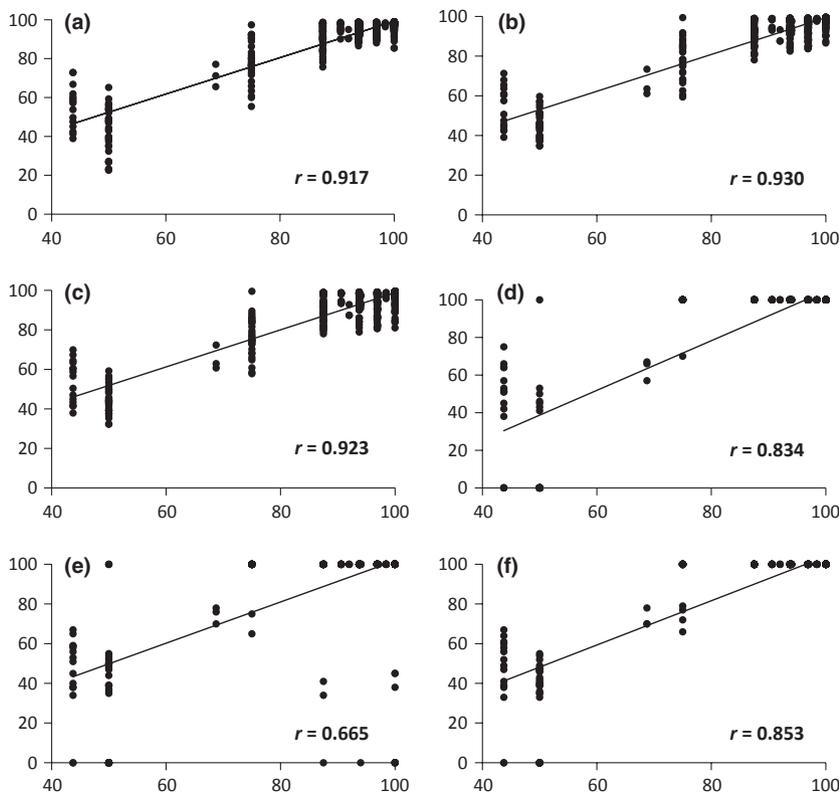
Fig. 2 Correlation between known red wolf ancestry and estimated red wolf ancestry for all six of the program/training set combinations. The x-axis indicates the known red wolf ancestry based on the pedigree (%) for each individual. The y-axis is the estimated ancestry coefficient produced by the clustering programs. Each pane is a separate program/training set combination. (a) STRUCTURE-Founders, (b) STRUCTURE- Update, (c) STRUCTURE-12 years, (d) BAPS- Founders, (e) BAPS-Update, (f) BAPS-12 years.

The 90% CIs produced for the 'Update' training set contained the expected ancestry value for 92.5% of individuals. The average CI width for 'Pure' red wolves was 7.5 (SD ± 4.6), for 'Backcrossed' red wolves was 16.1 (SD ± 8.4), for the 'F1 × hybrid backcross' individuals was 34.2 (SD ± 7.7), and 'F1 hybrid' individuals was 45.8 (SD ± 5.8). For the '12 years' training set, the 90% CIs contained the expected ancestry value for 92.9% of individuals. The average CI width for 'Pure' red wolves was 8.6 (SD ± 6.2), for 'Backcrossed' red wolves was 17.3 (SD ± 8.6), for the 'F1 × hybrid backcross' individuals was 34.6 (SD ± 7.3), and 'F1 hybrid' individuals was 44.9 (SD ± 5.3).

When using the USFWS 87.5% ancestry threshold as a benchmark for a legal red wolf, both programs misclassified individuals, although it tended to be in opposite directions (Table 2). STRUCTURE was more likely to classify a legal red wolf (≥87.5% red wolf) as a hybrid, whereas BAPS was more likely to classify hybrids (≤ 75% red wolf) as red wolves. Training set composition impacted misclassification rates produced by both programs, but it had less impact on BAPS. The 'Founders' training set in STRUCTURE produced the fewest misclassifications of true red wolves compared with other training sets; however, it also misclassified more 'F1 x red wolf backcross' individuals as red wolves than the others. BAPS never classified a true red wolf as a hybrid

except when the 'Update' training set was used. Both programs classified several 'F1 x red wolf backcrosses' as red wolves, although BAPS misclassified 3–9 times as many as STRUCTURE, depending on the training set. No 'F1 hybrid' was misclassified as a red wolf by STRUCTURE, but several were classified as red wolves by BAPS. Interestingly, BAPS using the 'Update' Training Set classified a large number of known hybrids ($n = 15$) and several red wolves ($n = 8$; both 'Pure' and Backcross') as pure coyotes with no hybrid ancestry.

The CAT never misclassified a hybrid as a red wolf, but did classify 117 legal red wolves (87.5–98% RW ancestry) as hybrids: all of these individuals were 'Backcross' red wolves. However, it should be noted that the CAT was not designed to provide greater resolution than classification as 100% or 75% RW ancestry.

Varying the number of loci

Reducing the number of loci had little impact on the correlation between real and estimated values of red wolf ancestry (Table S3, Supporting information). Using BAPS, all correlations were >0.8 regardless of loci set; in STRUCTURE, the correlations were >0.9 except for the nine and six loci sets, which produced correlation coefficients of 0.887 and 0.872, respectively. Lower numbers of loci had no impact on the STRUCTURE results: all four

ancestry categories were found to be significantly different regardless of the loci set ($P < 0.001$). For BAPS, the overall ANOVA models were statistically significant ($P < 0.001$), but the post hoc tests revealed that the 'Pure' and 'Backcross' ancestry groups could not be differentiated regardless of the number of loci, similar to the findings for the full suite of 17 loci. The one exception was the 12 loci set: in BAPS the 'Pure', 'Backcross' and 'F1 × red wolf backcross' groups were not significantly different. The average absolute difference between real and estimated ancestry values did not increase much for true red wolves (87.5–100%) with decreasing number of loci in BAPS, although for hybrids (50–75%), the difference did increase slightly (Fig. S4, Supporting information). These differences did increase in STRUCTURE, especially for hybrids, although it was not a strong trend.

Misclassification rates were almost identical for each loci subset compared with the full suite of loci in BAPS (Fig. S5, Supporting information). The patterns of misclassification rates in STRUCTURE displayed several trends (Fig. S6, Supporting information). Initially, at 17 loci, misclassification rates were highest for 'F1 × red wolf backcrosses', followed by 'Backcrossed' red wolves. However, as loci number was reduced, misclassification rates for both 'Pure' and 'Backcross' red wolves increased, while misclassification rates for 'F1 × red wolf backcrosses' declined. At 9 loci, misclassification rates were highest for 'Backcross' red wolves and nearly the same for 'F1 × red wolf backcrosses' and 'Pure' red wolves. The trend reversed itself at 6 loci. No 'F1 hybrids' were misclassified at any loci set indicating that these methods have the power to distinguish F1 hybrids with only 6 loci.

Discussion

Tests of clustering programs using simulated raw data (Anderson & Thompson 2002; Vähä & Primmer 2006) and simulations based on data from natural populations (Lancaster *et al.* 2006; Barilani *et al.* 2007; Sanz *et al.* 2009) have contributed to our understanding of how these programs perform in various circumstances. Still, simulated data cannot replicate the complexity of natural systems and require researchers to make assumptions about biological processes that may not be representative of reality (Huelsenbeck 1995). Using this empirical data set provides additional insights into how these programs respond to conditions that are not ideal but common in real populations.

The risk of performing tests with a real data set is it may violate the assumptions inherent to the methods or lack the conditions needed for the programs to optimally operate. Low levels of divergence and a limited

number of loci limit the efficacy of these programs (Pritchard *et al.* 2000). Vähä & Primmer (2006) tested STRUCTURE with simulated data that had a similar number of loci and F_{ST} values as those in our data set. Interestingly, the misassignment rates produced by their simulated data in STRUCTURE were comparable to ours, although they did not use prior population information. Other studies have used these programs on populations with fewer loci and lower levels of divergence than this red wolf system (e.g. Beaumont *et al.* 2001; Pierpaoli *et al.* 2003). Thus, we feel confident that the data from this system are appropriate for use with these programs and provide a rigorous test to examine the performance of these programs in a natural system.

F_{IS} values for our four species suggest a possible violation of Hardy–Weinberg and linkage equilibrium that may influence our results. Although this is true, we did not use these programs to identify clusters since our populations were pre-defined, which should lessen the impact of violating this assumption. Also, previous studies have found that even when assumptions of equilibrium are ignored, STRUCTURE (but importantly not BAPS) still performed well compared to distance-based methods (Hauser *et al.* 2006; Rodríguez-Ramilo *et al.* 2009). The greatest deviations in F_{IS} belong to the grey wolf and dog clusters, which were relatively unimportant in this system.

Several general insights can be drawn from the overall performance of these programs with this data set. The first is that the basic statistical framework implemented by each program dramatically influences ancestry estimates. The simultaneous calculation of genetic clusters and proportion of ancestry assigned to each cluster causes STRUCTURE to produce ancestry estimates that tend to be suggestive of admixture. STRUCTURE estimates a fraction of ancestry for each cluster, which typically results in ancestry estimates <1 even for pure individuals. This complicates interpretation of q-values and identification of pure individuals (Vähä & Primmer 2006). BAPS, by first assigning an individual to a genetic cluster and then using simulations to detect significant levels of admixture, was less likely to detect low levels of admixture. With our data set, STRUCTURE was better at identifying hybrid genotypes and predicting an individual's expected ancestry. The one drawback was STRUCTURE's tendency to classify pure individuals as admixed, which was similarly found by Vähä & Primmer (2006). The lack of a standard, quantitative method for setting a q-value threshold for hybrid ancestry contributes to this problem. Simulations of hybrid genotypes from known individuals to determine thresholds for significant ancestry coefficients may provide the best method to avoid this problem (Lancaster *et al.* 2006; Barilani *et al.* 2007; Sacks *et al.* 2011). Using a pedigree that

Table 2 Percentage of individuals misclassified by the program/training set combinations developed in this study. The initials 'STR' stand for STRUCTURE; all of the other labels refer to program/training set combinations that were tested. The letters 'RW' stand for red wolf. The number inside the () indicates the number of individuals that were misclassified

Program- training set	'Pure' RW classified as hybrid (N = 287)	'Backcrossed' RW classified as hybrid (N = 262)	'F1 x RW backcross' classified as RW (N = 33)	'F1 hybrid' classified as RW (N = 48)
STR- Founders	0.7% (2)	6.1% (16)	24.2% (8)	0% (0)
STR- Update	0.3% (1)	11.5% (30)	15.2% (5)	0% (0)
STR- 12 years	1.7% (5)	16.4% (43)	9.1% (3)	0% (0)
BAPS- Founders	0% (0)	0% (0)	87.9% (29)	2.1% (1)
BAPS- Update	3.0% (9)	1.5% (4)	84.8% (28)	6.3% (3)
BAPS- 12 years	0% (0)	0% (0)	78.8% (26)	0% (0)
CAT	0% (0)	44.7% (117)	0% (0)	0% (0)

contains hybrid individuals such as ours can also achieve this goal, allowing researchers to quantitatively assess the likely range of q-values to be produced for hybrid individuals.

Utilizing the credibility intervals produced by STRUCTURE provides a more quantitative means to assess an individual's ancestry than relying on point estimates. For most of our individuals STRUCTURE produced CIs that contained the ancestry value predicted by the pedigree with slight differences between training sets. However, the widths of these CIs, especially for admixed individuals, were large and frequently overlapped multiple ancestry classes. In addition, for several individuals with known hybrid ancestry, the CIs for the coyote cluster overlapped zero (data not shown). For some hybrids, the CIs for the dog and grey wolf clusters had upper bounds in the range of 20–30%, with one over 50%. The large widths of these CIs would make them impractical for management purposes.

Perhaps the most important observation from our study is the utility a detailed pedigree can provide in describing hybrid systems. Although creation of such a pedigree requires identification and sampling of the parental source populations, it provides a direct means to track genetic ancestry oppose to indirect estimation that relies on allele frequencies. Given the challenges associated by clustering methods, multidimensional scaling, and assignment tests, researchers should consider developing genetic pedigrees to monitor hybridization in small populations (see Pemberton 2008).

The use of prior information

The use of training sets can circumvent problems associated with these clustering programs identifying spurious clusters that do not reflect biological system. Several studies have noted that these programs can produce spurious clusters that do not reflect population genetic structure under different scenarios, such as isolation by distance (Frantz *et al.* 2009; Schwartz &

McKelvey 2009) and the presence of family groups (Anderson & Dunham 2008). Using our data set, the 'free clustering' method implemented in STRUCTURE divides the red wolves into three groups and combined coyotes, grey wolves, and dogs in a single cluster at $K = 4$ (data not shown). Conversely, BAPS at $K = 4$ divided the coyotes into two groups, combined grey wolves and dogs into a single cluster and identified the red wolves as a unique cluster (data not shown). The problem of different clustering programs identifying different patterns of genetic structure has been noted by other researchers (Jensen *et al.* 2005; Fjellheim *et al.* 2009; Kalinowski 2010).

Recently, hybridization studies have begun utilizing training sets in clustering tests to identify hybrid individuals (e.g. Pritchard *et al.* 2007; Oliveira *et al.* 2008; Trigo *et al.* 2008; Bohling & Waits 2011). However, within this framework, the clustering methods function like assignment tests in that populations are predefined. Several studies have compared the performance of clustering methods vs. traditional assignments tests in hybrid identification and have found mixed results (Hauser *et al.* 2006; Pritchard *et al.* 2007). When prior population knowledge is known, assignment tests may be used in conjunction with clustering methods to provide multiple tests of admixture. The maximum-likelihood assignment test implemented in the CAT outperformed both clustering methods in identifying admixed individuals: the CAT never classified a 'Pure' red wolf as a hybrid or a hybrid (50–75% red wolf) as a red wolf. The only individuals misclassified were 'Backcross' red wolves, which was expected as the program was originally not designed to accommodate small levels of coyote introgression. Also, the CAT includes additional information regarding the change in allele frequencies in the released red wolf population and private alleles present only in coyotes. When the CAT was designed, the goal was to never misidentify a red wolf (Miller *et al.* 2003). From the perspective of USFWS management, the greatest concern is distinguishing

pure red wolves from hybrids rather than determining the exact ancestry of admixed individuals. The CAT's reliability in identifying pure red wolves while never misidentifying hybrids demonstrates its value to the red wolf recovery program.

Training set composition

By altering the composition of our red wolf training set, we demonstrated the potential impact this component of the study design can have on ancestry estimation. The training set in STRUCTURE that produced the lowest misclassification rates for legal red wolves ('Founders') also produced the highest misclassification rate for hybrids. Conversely, the training set that misclassified the fewest hybrids ('12 years') misclassified the most pure red wolves as hybrids. This introduces an important paradox: the founders of the population, as would be expected, are most representative of the genetic diversity present in the total population, especially since individuals from the captive population are released into the wild. However, genetic drift has caused the current red wolf population to become more differentiated from the other hybridizing canid species (Fig. S1), which would improve hybrid identification (Vähä & Primmer 2006). Thus, the use of recent vs. historical samples as reference samples can impact conclusions made about hybrid systems when genetic drift has impacted allele frequencies.

Combining samples from multiple time periods may be the least accurate method, as suggested by our results using BAPS. Researchers utilizing training sets should be cautious about combining samples from multiple generations due to the potential impact of genetic drift on allele frequencies. However, it is important to note that this training set performed well in STRUCTURE, suggesting the relative performance of different training sets depends on the method.

Impact of number of loci

Reducing the number of loci had a variable impact on the performance of these programs. Although STRUCTURE continued to outperform BAPS when the number of loci was reduced, it appeared to suffer more than BAPS as loci were removed. Similar effects of loci number on the performance of STRUCTURE have been found by other studies (Evanno et al. 2005; Rosenberg et al. 2005; Vähä & Primmer 2006); however, they did not test BAPS under similar conditions.

We were surprised to find that overall performance of the programs did not decline much with reduced numbers of loci. Misclassification rates did increase with STRUCTURE, but otherwise the loci subsets produced results similar to the full suite of loci. In BAPS, altering the training set had a greater impact on these correlations than lowering the number of loci, whereas the opposite was found in STRUCTURE. Using simulated data sets, Vähä & Primmer (2006) found reducing the number of loci had a greater impact on admixture estimation than our study. One potential explanation is the genetic differentiation between our populations (F_{ST} >0.1) was much higher than those simulated by other studies that have tested clustering methods (e.g. Latch et al. 2006; Vähä & Primmer 2006). Also, we purposely selected for loci with the highest informativeness from an original set of markers that were previously selected due to their highly polymorphic nature, meaning these loci had increased power for efficient ancestry assignment. Regardless, a valuable insight from our study is that training set composition can be as important to assessing ancestry as the number of loci in the data set, highlighting the importance of training set composition.

Conclusion

Despite their utility, the accuracy of clustering methods can vary depending on the data set and the statistical properties of specific software programs. STRUCTURE outperformed BAPS in terms of identifying hybrid individuals and detecting lower levels of introgression. Training set composition has not received much attention as an influential factor in the performance of these methods; however, our results suggest the different training sets can produce markedly different conclusions about the same data set. In fact, in our study, training set composition had a larger impact on program accuracy than loci number. Even with the training sets, the performance of these programs may not be adequate for managers attempting to identify hybrids in an endangered population. In that case, system-specific analysis methods and/or maximum-likelihood-based assignment tests, such as the CAT, may be more appropriate. For identification of hybrids in the context of red wolf recovery, we recommend the continued use of the CAT as the main method for evaluating ancestry of unknown individuals. However, the inaccuracies associated which each program stresses the value of reconstructing pedigrees from wild populations.

Acknowledgements

members of the Waits laboratory provided valuable advice and comments that helped direct this study. We appreciate the valuable comments provided by the handling editor and several anonymous reviewers.

References

Adams JR (2006) A multi-faceted molecular approach to red wolf (*Canis rufus*) conservation and management. PhD Dissertation, University of Idaho, Idaho.

Adams JR, Kelly BT, Waits LP (2003) Using faecal DNA sampling and GIS to monitor hybridization between red wolves (*Canis rufus*) and (*Canis latrans*). *Molecular Ecology*, **12**, 2175–2186.

Allendorf FW, Leary RF, Spruell P, Wenburg JK (2001) The problems with hybrids: setting conservation guidelines. *Trends in Ecology and Evolution*, **16**, 613–622.

Anderson EC, Dunham KK (2008) The influence of family groups on inferences made with the program Structure. *Molecular Ecology Resources*, **8**, 1219–1229.

Anderson EC, Thompson EA (2002) A model-based method for identifying species hybrids using multilocus genetic data. *Genetics*, **160**, 1217–1229.

Arnold ML (1997) Natural hybridization and evolution. Oxford University Press, New York.

Barilani M, Sfougaris A, Giannakopoulos A, Mucci N, Tabarroni C, Randi E (2007) Detecting introgressivehybridisation in rock partridge population (*Alectorisgraeca*) in Greece through Bayesian admixture analyses of multilocus genotypes. *Conservation Genetics*, **8**, 343–354.

Beaumont M, Barratt EM, Gottelli D *et al.* (2001) Genetic diversity and introgression in the Scottish wildcat. *Molecular Ecology*, **10**, 319–336.

Bohling JH, Waits LP (2011) Assessing the prevalence of hybridization between sympatric *Canis* species surrounding the red wolf (*Canis rufus*) recovery area in North Carolina. *Molecular Ecology*, **20**, 2142–2156.

Bozarth CA, Hailer F, Rockwood LL, Edwards CW, Maldonado JE (2011) Coyote colonization of northern Virginia and admixture with Great Lakes wolves. *Journal of Mammalogy*, **92**, 1070–1080.

Burgarella C, Lorenzo Z, Jabbour-Zahab R *et al.* (2009) Detection of hybrids in nature: application to oaks (*Quercus suber* and *Q. ilex*). *Heredity*, **102**, 442–452.

Corander J, Marttinen P (2006) Bayesian identification of admixture events using multilocus molecular markers. *Molecular Ecology*, **15**, 2833–2843.

Corander J, Waldmann P, Sillanpaa MJ (2003) Bayesian analysis of genetic differentiation between populations. *Genetics*, **163**, 367–374.

Corander J, Marttinen P, Mantyniemi S (2006) A Bayesian method for identification of stock mixtures from molecular marker data. *Fisheries Bulletin*, **104**, 550–558.

Cornuet J-M, Siry S, Luikart G, Estoup A, Solignac M (1999) New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics*, **153**, 1989–2000.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–1587.

Fjellheim S, Jorgensen MH, Kjos M, Borgen L (2009) A molecular study of hybridization and homoploid hybrid speciation in *Argyranthemum* (Asteraceae) on Tenerife, the Canary Islands. *Botanical Journal of the Linnean Society*, **159**, 19–31.

Frantz AC, Cellina S, Krier A, Schley L, Burke T (2009) Using Bayesian methods to determine the genetic structure of a continuously distributed population: clusters or isolation by distance? *Journal of Applied Ecology*, **46**, 493–505.

Goudet J (2002) FSTAT: computer package for PCs which estimates and tests gene diversities and differentiation statistics from codominant genetic markers. Version 2.9.3.2.Available fromwww2.unil.ch/popgen/softwares/fstat.htm.

Hauser L, Seamons TR, Dauer M, Naish KA, Quinn TP (2006) An empirical verification of population assignment methods by marking and parentage data: hatchery and wild steelhead (*Oncorhynchusmykiss*) in Forks Creek, Washington, USA. *Molecular Ecology*, **15**, 3157–3173.

Huelsenbeck JP (1995) Performance of phylogenetic methods in simulation. *Systematic Biology*, **44**, 17–48.

Jensen AB, Palmer KA, Boomsma JJ, Pedersen BV (2005) Varying degrees of *Apismelliferaligustica* introgression in protected populations of the black honeybee, *Apismellifera*, in northwest Europe. *Molecular Ecology*, **14**, 93–106.

Kalinowski ST (2010) The computer program STRUCTURE does not reliably identify the main genetic clusters within species: simulations and implications for human population structure. *Heredity*, **106**, 625–632.

Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, **16**, 1099–1106.

Kelly BT, Miller PS, Seal US (1999) Population and habitat viability assessment workshop for the red wolf (*Canis rufus*). IUCN/SSC Conservation Breeding Specialist Group: Apple Valley, Minnesota.

Lancaster ML, Gemmell NJ, Negro S, Goldsworthy S, Sunnucks P (2006) Ménage a trois on Macquarie Island: hybridization among three species of fur seal (*Arctocephalus* spp.) following historical population extinction. *Molecular Ecology*, **15**, 3681–3692.

Latch EK, Dharmarajan G, Glaubitz JC, Rhodes Jr OE (2006) Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. *Conservation Genetics*, **7**, 295–302.

Manel S, Gaggiotti OE, Waples RS (2005) Assignment methods: matching biological questions with appropriate. *Trends in Ecology and Evolution*, **20**, 136–142.

Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology*, **7**, 639–655.

McCarley H (1962) The taxonomic status of wild *Canis* (Canidae) in the south central United States. *The Southwestern Naturalist*, **7**, 227–235.

Miller CR, Adams JR, Waits LP (2003) Pedigree based assignment tests for reversing coyote (*Canis latrans*) introgression into the wild red wolf (*Canis rufus*) population. *Molecular Ecology*, **12**, 3287–3301.

Oliveira R, Godinho R, Randi E, Ferrand N, Alves PC (2008) Molecular analysis of hybridization between wild and

domestic cats (*Felis silvestris*) in Portugal: implications for conservation. *Conservation Genetics*, **9**, 1–11.

Paradiso JL, Nowak RM (1972) Canis rufus. *Mammalian Species*, **22**, 1–4.

Peakall R, Smouse PE (2006) GENALEX 6: genetics analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288–295.

Pearse DE, Crandall KA (2004) Beyond F_{st}: analysis of population genetic data for conservation. *Conservation Genetics*, **5**, 585–602.

Pemberton JM (2008) Wild pedigree: the way forward. *Proceedings of the Royal Society B*, **275**, 613–621.

Phillips MK, Parker WT (1988) Red wolf recovery: a progress report. *Conservation Biology*, **2**, 139–141.

Pierpaoli M, Biró S, Herrmann M *et al.* (2003) Genetic distinction of the wildcat (*Felis silvestris*) populations in Europe, and hybridization with domestic cats in Hungary. *Molecular Ecology*, **12**, 2585–2598.

Piry S, Alapetite A, Cornuet J-M, Paetkau D, Baudouin L, Estoup A (2004) GENECLASS2: a software for genetic assignment and first-generation migrant detection. *Journal of Heredity*, **95**, 536–539.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Pritchard JK, Wen X, Falush D (2003) Documentation for structure software: version 2.2.

Pritchard VL, Jones K, Cowley DE (2007) Estimation of introgression in cutthroat trout populations using microsatellites. *Conservation Genetics*, **8**, 1311–1329.

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences*, **94**, 9197–9201.

Rhymer JM, Simberloff D (1996) Extinction by hybridization and introgression. *Annual Review of Ecology and Systematics*, **27**, 83–109.

Rodríguez-Ramilo ST, Toro MA, Fernández J (2009) Assessing population genetic structure via the maximisation of genetic distance. *Genetics Section Evolution*, **41**, 49.

Sacks BN, Moore M, Statham MJ, Wittmers HU (2011) A restricted hybrid zone between native and introduced red fox (*Vulpes vulpes*) populations suggests reproductive barriers and competitive exclusion. *Molecular Ecology*, **20**, 326–341.

Sanz N, Araguas RM, Fernandez R, Vera M, Garcia-Marin J-L (2009) Efficiency of markers and methods for detecting hybrids and introgression in stocked populations. *Conservation Genetics*, **10**, 225–236.

Schwartz ML, McKelvey KS (2009) Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. *Conservation Genetics*, **10**, 441–452.

Stoskopf MK, Beck K, Fazio BB *et al.* (2005) From the field: implementing recovery of the red wolf- integrating research scientists and managers. *Wildlife Society Bulletin*, **33**, 1145–1152.

Trigo TC, Freitas TRO, Kunzler G *et al.* (2008) Inter-species hybridization among Neotropical cats of the genus *Leopardus*, and evidence for an introgressive hybrid zone between *L. geoffroyi* and *L. tigrinus* in southern Brazil. *Molecular Ecology*, **17**, 4317–4333.

USFWS (1990) Red Wolf Recovery/Species Survival Plan.US Fish and Wildlife Service, Atlanta, Georgia.

USFWS (2007) Red wolf (Canis rufus) 5-year review: summary and evaluation. US Fish and Wildlife Service, Manteo, North Carolina.

Vähä J-P, Primmer CR (2006) Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Molecular Ecology*, **15**, 63–72.

Data accessibility

Microsatellite genotypes, STRUCTURE input and parameter files and BAPS input files: DRYAD entry doi:10.5061/dryad.8j486

Supporting information

Additional supporting information may be found in the online version of this article.

Fig. S1 Changes in pairwise F_{ST} over time between red wolves and the other canid species using the (A) 'Update' training set and (B) '12 years' training set.

Fig. S2 Change in allelic richness for the 'pure' (i.e. 100%) red wolf population over time.

Fig. S3 Change in observed heterozygosity (H_0) for the 'pure' red wolf population over time.

Fig. S4 Average absolute difference between real and estimated ancestry values using (A) STRUCTURE and (B) BAPS for each loci subset.

Fig. S5 Misclassification rate of individuals by BAPS using different sets of loci. Individuals are grouped according to the four red wolf ancestry classes and misclassification rate refers to the percentage of individuals in that ancestry class that were misclassified.

Fig. S6 Misclassification rate of individuals by STRUCTURE using different sets of loci.

Table S1 Pairwise F_{ST} values between the four species present in the 'Founders' training set.

Table S2 Measures of average allelic richness, H$_0$, *F*IS, and private alleles across all loci for the four species incorporated in the 'Founders' training set.

Table S3 Correlation values (*r*) between known red wolf ancestry and estimated red wolf ancestry using the different loci subsets.

Table S4 Number of individuals assigned to each category of red wolf ancestry from 1987–2008.