Estimated genotype error rates from bowhead whale microsatellite data

Phillip A Morin, Richard G LeDuc, Eric Archer, Karen K Martien, Barbara L Taylor
Southwest Fisheries Science Center
8604 La Jolla Shores Dr., La Jolla, CA 92037, USA

Ryan Huebinger
Texas A&M University
College Station, TX 77843, USA

John W. Bickham
Center for the Environment, Purdue University
503 Northwestern Avenue, West Lafayette, IN 47907-2966, USA

Key Words: Genetics, Biomarkers, Stock identity

ABSTRACT
We calculate genotype error rates using opportunistic replicate samples in the microsatellite data for bowhead whales. The estimated rate (1%/genotype) falls within normal ranges reviewed in this paper. The results of a jackknife analysis identified five individuals that were highly influential on estimates of Hardy-Weinberg equilibrium for four different markers. In each case, the influential individual was homozygous for a rare allele. Although these individuals have not been re-genotyped in the laboratory to determine whether the initial homozygous allele calls were correct, our result demonstrates that Hardy-Weinberg p-values are very sensitive to homozygosity in rare alleles for single individuals. This raises the possibility that even small, normal levels of laboratory errors can result in an overestimate of the degree to which markers are estimated to be out of Hardy-Weinberg equilibrium and hence overestimate the potential to infer stock structure. To avoid such bias, we recommend routine identification of influential individuals and multiple replication of those samples.

Introduction

Microsatellite genotypes are the most common type of marker used in studies of phylogeography, population structure, social structure, individual identification and paternity. Several characteristics of microsatellites make them nearly ideal markers for many of these types of studies, including their ubiquitous presence in almost all organisms, high numbers of alleles and levels of heterozygosity, and rapid mutation rates. However, some of these characteristics also represent some of the most significant limitations. Microsatellite alleles can range substantially in size, resulting in the possibility of mis-scoring due to marker characteristics and sizing technology limitations. Replication slippage across repeats can result in 'allelic stutter' patterns that can then lead to incorrect genotype scoring. Reliance on electrophoresis to determine allele sizes introduces variation that can make it difficult to match allele sizes from samples

genotyped in different laboratories, over time and technologies in the same laboratory, and even with fluctuating climatic conditions (Davison & Chiba 2003; LaHood et al. 2002).

The impact of genotyping error has been described for paternity (Gagneux et al. 1997; Hoffman & Amos 2005), individual identification (Bonin et al. 2004; Skaug & Øien 2004), and population size estimates based on genetic mark-recapture (McKelvey & Schwartz 2004). We are not aware of any publications to date that examine the potential effects of genotyping error on detection of population structure, though effects on associated parameters such as homozygosity (Gagneux et al. 1997; Taberlet et al. 1996), departure from Hardy-Weinberg equilibrium (HWE) , overestimates of inbreeding have been noted (reviewed in Bonin et al. 2004; Broquet & Petit 2004; McKelvey & Schwartz 2004).

In this paper, we review some of the types of errors that can commonly occur with microsatellites and what error rates are commonly reported in published studies. We then estimate the error rates for the current bowhead whale (*Balaena mysticetus*) microsatellite genotypes, and evaluate how they may affect population analyses. Additional effort to more accurately estimate error rates in these samples and genetic markers is ongoing, and will be presented as a working paper (Huebinger and Bickham).

Causes and types of microsatellite genotyping errors:

Genotyping errors can be broadly classified into three types:

1) Errors due to marker characteristics, such as allelic stutter, short allele dominance, and null alleles. The genotype patterns on gels or in capillary electrophoresis instruments can vary substantially among markers, and some are inherently more difficult to genotype accurately.
2) Errors due to technological limitations. Allele sizing is based on migration of amplified PCR products through a matrix, propelled by an electric current, and visualized relative to a size standard to estimate or calculate allele sizes. Electrophoretic migration can be affected by both size and nucleotide composition of the alleles, plus the addition of fluorescent molecules for visualization. Inferred allele size is, therefore, not always perfectly correlated with the actual size of the amplified alleles; allele sizes can differ by more or less than the size of the microsatellite repeat unit (e.g, a CA repeat can have alleles that differ on average by 1.8-2.2bp; Amos et al. 2007). In addition, electrophoresis is itself variable, and can cause allelic size differences of up to 7bp across time, technologies, and instruments (LaHood et al. 2002).
3) Errors introduced due to sample handling or data transcription. These are the errors that are most difficult to detect, and which can affect few or many samples, as in cases where samples are switched, or whole data sets get scrambled during manipulation in spreadsheets.

We will review only error types 1 and 2, which are typical of all data sets and for which error rates have been estimated in other studies. Table 1 summarizes error rates and the publications from which they were obtained. Error rates are typically calculated as the ratio of differing replicated genotypes to the total number of genotypes in the replication study (the per genotype error rate), though some studies calculate the error

rate per allele, when multiple miss-calls at a locus can be detected. The latter is the most precise, and is how we have calculated genotype error rates. Because of this variation in the way error rates have been presented in the literature, it must be kept in mind that per genotype error rates will be roughly twice the per allele error rates in most cases. The data in Table 1 represent the range of per genotype errors observed in microsatellite data sets, and include data obtained from non-invasive samples (e.g., feces and hair), which are expected to produce higher error rates than high quality samples (e.g. preserved skin tissue) because of the quantity and quality of extracted DNA. It is important to note that almost all studies that evaluated error rates reported a non-zero error rate, regardless of sample types used.

Table 1
Published microsatellite per genotype error rates. The total error rate was calculated for those studies where error rates were reported separately for allelic dropout and false alleles. The composite genotype error is the theoretical frequency of incorrect composite genotypes assuming 10, 20 , or 30 loci with equal probability of error per locus (Bonin et al. 2004). Several of the cited papers review error rates from multiple other studies, summarized in this table.

			Composite genotype error probability (%)		
Source	Error rate (%)	Sample source	10 loci	20 loci	30 loci
Paetkau 2003	4.9	hair	63	87	95
Paetkau 2003	6.2	hair	72	92	98
Paetkau 2003	4.1	hair	57	81	92
Paetkau 2003	4.6	hair	61	85	94
Paetkau 2003	2.8	hair	43	68	82
Paetkau 2003	3.9	hair	55	80	91
Paetkau 2003	1.9	hair	32	54	68
Paetkau 2003	16.7	hair	97	100	100
Paetkau 2003	4.3	hair	58	83	93
Paetkau 2003	5.2	hair	66	88	96
Paetkau 2003	2.1	hair	35	57	72
Paetkau 2003	4.8	hair	63	86	95
Paetkau 2003	5.9	hair	70	91	97
Paetkau 2003	1.2	hair	21	38	52
Paetkau 2003	1.7	hair	29	50	64
Paetkau 2003	0.0	hair	-	-	-
Paetkau 2003	7.8	hair	80	96	99
Bonin et al. 2004	0.8	**tissue**	15	27	38
Bonin et al. 2004	1.2-2.0	feces	21-22	38-55	52-70
Broquet & Petit 2004	8.0	feces	81	96	99
Broquet & Petit 2004	16.7	feces	97	100	100
Broquet & Petit 2004	2.1	feces	35	57	72
Broquet & Petit 2004	8.0	feces	81	96	99
Broquet & Petit 2004	2.0	feces	33	55	70
Broquet & Petit 2004	1.5	feces	26	45	60
Broquet & Petit 2004	35.0	feces	100	100	100
Broquet & Petit 2004	36.9	hair	100	100	100
Broquet & Petit 2004	0.4	hair	8	15	22
Broquet & Petit 2004	7.2	feces	78	95	99
Broquet & Petit 2004	15.2	feces	96	100	100
Broquet & Petit 2004	18.6	feces	98	100	100
Broquet & Petit 2004	18.3	feces	98	100	100
Broquet & Petit 2004	6.8	feces	76	94	99
Broquet & Petit 2004	18.3	feces	98	100	100
Broquet & Petit 2004	24.0	feces	100	100	100
Broquet & Petit 2004	1.0	feces	18	33	45
Broquet & Petit 2004	0.0	hair	-	-	-
Broquet & Petit 2004	0.0	hair	-	-	-
Broquet & Petit 2004	48.0	feces	100	100	100
Broquet & Petit 2004	0.0	feces	-	-	-
(Hoffman et al. 2006)†	0.8	**tissue**	15	27	38
Hoffman & Amos 2005	0.1-0.7	**tissue**	2-13	4-24	6-34
Hoffman & Amos 2005**	0.1-12.7	**tissue**	2-93	4-100	6-100

†20% of genotypes replicated in 2 labs could not be matched.
**based on review of human medical genetics studies

Error types and rates have been studied extensively, especially for non-invasive samples (Bonin et al. 2004; Broquet & Petit 2004; Hoffman & Amos 2005; Johnson & Haydon 2007; McKelvey & Schwartz 2004; Paetkau 2003). All find that genotyping errors are to be expected, but can be minimized with proper use of controls, replication, and marker selection. In particular, higher error rates are often associated with markers that have higher heterozygosity, more alleles, more stutter bands, and larger product sizes (Hoffman & Amos 2005). Error types are also not equally common. In the studies summarized in Table 1, many estimated error rates separately for allelic dropout separately from false alleles (scoring of a non-specific band) (Broquet & Petit 2004; Paetkau 2003).

Four of the most common types of genotyping errors create a bias towards increased homozygosity. These include allelic dropout, which was found to be much more common than false alleles in studies where the two were analyzed separately; null alleles; mis-interpretation of neighboring alleles as stutter; and short allele dominance, in which the larger alleles tend to have lower signal intensity, causing them to be missed in samples where the lowest signal allele falls below the detection threshold.

BCB bowhead microsatellite per allele error rates

Genotypes were generated for 409 bowhead whale samples from Alaskan and Russian waters (Givens et al. 2007). Not all samples were genotyped with every microsatellite marker; the number of loci for individual samples ranged from 22 to 33, though there were fewer completed genotypes for some samples. To examine error rates, we looked at pairs and groups of samples that represented duplicate samples of individual whales. In all cases, these samples were not known by the laboratory personnel to be duplicates prior to genotyping. Some represented multiple biopsies from free-swimming whales, and others were cases of inadvertent multiple samplings of harvested whales.

Some of the duplicates were identified as such after the samples yielded identical genotypes. In other cases, “near-matches” (samples whose genotypes differed by only one or a few loci) were revealed upon closer examination to actually be identical but to have suffered a genotyping error. The latter cases provided an estimation of the number of errors among the duplicates, which was then calculated as a fraction of the total number of allele calls among the duplicates (including both matches and near matches). The apparent causes of errors included allelic dropout, mistakes in allele sizing and transcribing errors.

The error rate given will be a slight underestimate since we did not account for a locus having missing data from one replicate but not the other; the number of scored markers in the final dataset was used for the total number of alleles called (i.e., two identical genotypes across 35 loci indicate 140 good calls). The bias of the estimate is likely to be minimal, since duplicate samples from a given individual tend to suffer from the same genotyping failures (i.e., missing data).

The matches that were detected by identical initial genotypes (no errors evident) included 7 pairs of replicates and one cluster of 8 replicates of the same individual. In aggregate, these indicated 1392 good allele calls.

There were 15 pairs of samples that initially were near matches that ended up being actual duplicates upon closer examination. These included 32 miscalled alleles over 1780 allele calls. Incorporating the correctly called replicates, this gives 32 mistakes over

3172 allele calls – very close to 1%. The mistakes did not seem to be biased toward any particular loci. The 32 mistakes were spread across 19 loci, with only one (Bmy19) including three mistakes, three including two mistakes, and 17 with only a single mistake. The total does not add up to 32 because it is simply a tally of occurrences of miscalls, not whether there was one or two miscalled alleles in a given genotype. Although 1% is not a high error rate, even this error rate led to only about half the replicates being detected (considering the group of 8 as separate replicates).

Twelve of the 32 miscalls were from homozygote/heterozygote differences, the likely candidates for allelic dropout. These were from 11 loci. The loci involved were TV7, TV14 (twice, same allele), TV20, Bmy14, Bmy16, Bmy19, Bmy26, Bmy42, Bmy44, Bmy53, and Bmy57. There did not seem to be any obvious pattern to these discrepancies with regard to allele sizes or frequencies; both common and rare alleles were involved and sizes could be either similar or very different. For 9 of the 11 loci, the allele that varied (present in one replicate but not the other) was the less common one. As a rough measure, this would seem to indicate an allelic dropout rate of 0.4% (12/3172).

Effects of errors on Hardy-Weinberg equilibrium

Of the 33 loci genotyped for 213 samples from Barrow, nine were found to be out of HWE, with a significant heterozygote deficiency. In order to determine if particular individuals were more influential than average on HWE, we conducted a jackknife analysis. In the jackknife, each individual was sequentially removed from the dataset and HWE was calculated again across all loci using the remaining 212 individuals.

There were 40 cases where removal of a single sample changed a locus from being out of HWE to being in HWE ($p > 0.05$, Table 2). Figure 1 shows that there were five extreme jackknife replicates (log-odds ratio between jackknife and observed p-values < 0.7). In these cases, four different loci were involved (Bmy41, TV14, TV19, and Bmy18). Each of the samples removed in these replicates was homozygous for a rare allele (frequencies < 6%) at the locus under consideration. Additionally, the degree of the difference in the HWE p-value between the original and jackknife replicates was directly related to the frequency of this rare allele. This further suggests that these individuals are largely responsible for the findings of significant heterozygote deficiency.

The other 35 replicates where there was a change in HWE (log-odds ratios between 0.8 and 1) all occurred on loci Bmy18 and Bmy42. Of these replicates, 22 of the samples were homozygotes and 13 were heterozygotes at the locus. In most of these replicates the change was from a p-value that was just below the critical value of 0.05 to a p-value just above the critical value, suggesting that the change in HWE in these cases may be due to stochasticity either in sampling or the MCMC assessment of HWE.

Table 2. Jackknife replicates where significant Hardy-Weinberg equilibrium was found to be non-significant ($p > 0.05$).

Table 2: Loci for which non-Jackknife and Jackknife replicates indicated contradictory results for Hardy-Weinberg equilibrium ($p > 0.05$). Extreme Log Odds Ratios (<0.7) are shown in bold.

Locus	GeneticID	Observed p-value	Jackknife p-value	Observed Odds	Jackknife Odds	Odds Ratio	Odds Difference	Genotype	Allele Frequencies
Bmy41_1	02B16	0.006	0.579	-5.115	0.318	**-0.062**	5.433	233 / 233	0.005
TV14_1	02B6	0.025	0.228	-3.661	-1.221	**0.334**	2.440	93 / 93	0.017
TV19_1	05B7	0.004	0.114	-5.467	-2.055	**0.376**	3.412	184 / 184	0.013
Bmy18_1	99B3	0.049	0.130	-2.959	-1.897	**0.641**	1.062	137 / 137	0.024
Bmy18_1	83B1	0.049	0.112	-2.959	-2.067	**0.699**	0.892	127 / 127	0.031
Bmy42_1	96B11	0.038	0.077	-3.226	-2.482	0.769	0.744	182 / 182	0.056
Bmy18_1	96B8	0.049	0.082	-2.959	-2.418	0.817	0.541	117 / 117	0.052
Bmy18_1	05BpB12	0.049	0.077	-2.959	-2.480	0.838	0.479	125 / 125	0.054
Bmy18_1	92B3	0.049	0.076	-2.959	-2.493	0.842	0.466	117 / 117	0.052
Bmy42_1	02B6	0.038	0.056	-3.226	-2.820	0.874	0.406	162 / 162	0.08
Bmy18_1	05B5	0.049	0.068	-2.959	-2.611	0.882	0.348	131 / 131	0.052
Bmy42_1	03B13	0.038	0.055	-3.226	-2.851	0.884	0.374	162 / 162	0.08
Bmy42_1	90B8	0.038	0.051	-3.226	-2.915	0.904	0.311	180 / 180	0.103
Bmy42_1	04B15	0.038	0.051	-3.226	-2.916	0.904	0.310	176 / 176	0.131
Bmy42_1	97B18	0.038	0.051	-3.226	-2.918	0.905	0.308	180 / 180	0.103
Bmy18_1	02B7	0.049	0.062	-2.959	-2.719	0.919	0.240	133 / 133	0.066
Bmy18_1	03B12	0.049	0.061	-2.959	-2.732	0.923	0.227	109 / 109	0.13
Bmy18_1	03B4	0.049	0.061	-2.959	-2.739	0.926	0.220	119 / 119	0.144
Bmy18_1	92B5	0.049	0.061	-2.959	-2.741	0.926	0.218	119 / 119	0.144
Bmy18_1	05B23	0.049	0.060	-2.959	-2.748	0.929	0.211	109 / 109	0.13
Bmy18_1	92B2	0.049	0.059	-2.959	-2.771	0.936	0.188	119 / 119	0.144
Bmy18_1	96B11	0.049	0.057	-2.959	-2.797	0.945	0.162	109 / 109	0.13
Bmy18_1	96B7	0.049	0.057	-2.959	-2.802	0.947	0.157	119 / 119	0.144
Bmy18_1	99B1	0.049	0.056	-2.959	-2.829	0.956	0.130	119 / 125	0.144 / 0.054
Bmy18_1	05B27	0.049	0.053	-2.959	-2.881	0.974	0.078	115 / 117	0.2 / 0.052
Bmy18_1	04B2	0.049	0.053	-2.959	-2.882	0.974	0.077	115 / 115	0.2
Bmy18_1	97B17	0.049	0.053	-2.959	-2.893	0.978	0.066	115 / 129	0.2 / 0.075
Bmy18_1	96B20	0.049	0.052	-2.959	-2.894	0.978	0.065	115 / 115	0.2
Bmy18_1	04B5	0.049	0.052	-2.959	-2.895	0.979	0.063	115 / 125	0.2 / 0.054
Bmy18_1	97B29	0.049	0.052	-2.959	-2.896	0.979	0.063	113 / 115	0.047 / 0.2
Bmy18_1	84B4	0.049	0.052	-2.959	-2.904	0.981	0.055	115 / 115	0.2
Bmy18_1	97B18	0.049	0.052	-2.959	-2.910	0.983	0.049	119 / 135	0.144 / 0.047
Bmy18_1	97B20	0.049	0.051	-2.959	-2.917	0.986	0.042	121 / 131	0.026 / 0.052
Bmy18_1	05B20	0.049	0.051	-2.959	-2.919	0.987	0.040	119 / 131	0.144 / 0.052
Bmy18_1	88B9	0.049	0.051	-2.959	-2.922	0.987	0.037	109 / 135	0.13 / 0.047
Bmy18_1	05B26	0.049	0.051	-2.959	-2.924	0.988	0.035	121 / 135	0.026 / 0.047
Bmy18_1	97B11	0.049	0.051	-2.959	-2.924	0.988	0.035	115 / 119	0.2 / 0.144
Bmy18_1	96B5	0.049	0.051	-2.959	-2.928	0.989	0.031	131 / 135	0.052 / 0.047
Bmy18_1	97B19	0.049	0.051	-2.959	-2.934	0.991	0.025	109 / 133	0.13 / 0.066
Bmy18_1	03B6	0.049	0.050	-2.959	-2.944	0.995	0.015	119 / 119	0.144

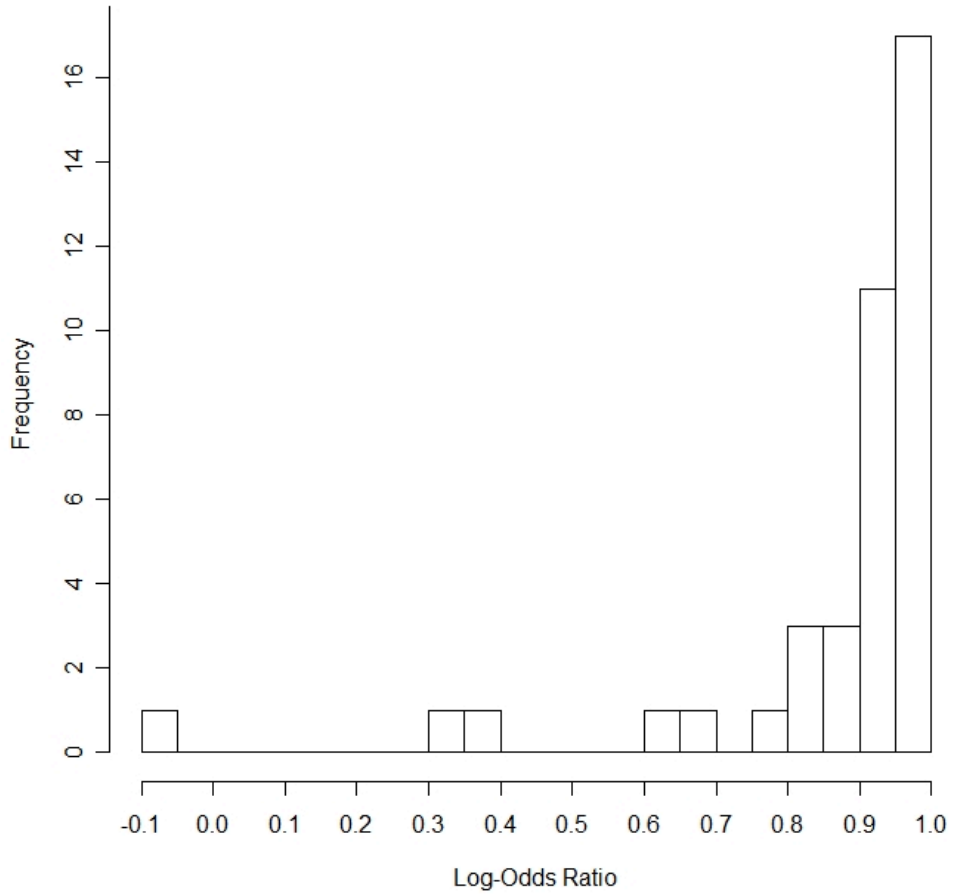

Figure 1. Frequency of the log-odds ratio of the Hardy-Weinberg *p*-values from the 40 jackknife replicates given in Table 2.

Discussion

There is no doubt that errors exist in all genetic data sets. A review of the literature and analysis of unintentionally replicated samples indicates that the observed 1% per allele error rate for the BCB bowhead microsatellite data is low and similar to the published observed error rates (e.g., ≤0.8% per genotype from several studies based on DNA from tissue). A more accurate estimate of genotyping error rates in this data set will be possible when targeted replication is complete.

The effects of these various types of genotyping errors on analyses of population structure depend heavily on the type of error and the assumptions of the particular analysis being attempted. Analyses, such as F_{st} (Weir & Cockerham 1984), that are based on assessing differences among population allelic frequency distributions would tend to be affected by allele-specific errors. This is because the primary assumption underlying these analyses is that the allelic frequency distributions in the data are a random sample and therefore accurate reflection of the frequency distributions of the population. Therefore, any error that significantly alters the frequency distribution will produce results that do not correctly reflect the true degree of population differentiation.

It is not possible to determine the magnitude or direction of any potential bias of this nature a-priori as this is the product of the specific type of genotyping error and the allelic distributions within the strata under consideration. As an example, a laboratory condition that leads to allelic dropout of larger products could cause strata to be more similar if a large allele exists at high frequency in one stratum, but not in the other. Conversely, if the same allele is common in both populations, dropout would decrease its overall frequency and could lead to an assessment of significant differentiation due to secondary differences in the relative frequencies of other rarer alleles. It has recently been shown that null alleles, which increase observed homozygosity over most alleles in populations, can lead to significantly increased estimates of population differentiation (Chapuis & Estoup 2007).

The results of our jackknife analysis have demonstrated how analyses that rely on estimates of homo- or heterozygote frequencies, such as the detection of Wahlund effect via Hardy-Weinberg disequilibrium, can be adversely affected by errors that modify these distributions. Other analyses that rely on the detection and assessment of HWE, such as the Bayesian clustering program STRUCTURE (Falush et al. 2003; Pritchard et al. 2000) may be similarly sensitive to genotyping errors. It is clear that the HWE test will be sensitive to individuals that are homozygous for rare alleles, which are, by definition, unlikely to exist in a sample from a single, randomly-mating population. The presence of these individuals could lead to an assessment of significant population subdivision which is in reality based entirely on a handful of erroneous allele calls.

Once individuals such as these have been identified, it is imperative that their genotypes be replicated in the laboratory to ensure that they are valid prior to continuing with other analyses. In the example we have presented with the 213 samples from Barrow, this would mean verifying at most 27 genotypes for 25 samples. If the original genotype is determined to be correct, then it is left to the researcher to decide whether the sample should remain in the dataset based on the circumstances of its collection and what is known about genetic diversity overall. An alternative would be to eliminate that particular locus from analyses of population subdivision.

Acknowledgments

We are grateful to Nick Kellar and Tomo Eguchi for helpful suggestions on the manuscript.

References

Amos, W., Hoffman, J. I., Frodsham, A., Zhang, L., Best, S., and Hill, V. S. 2007. Automated binning of microsatellite alleles: problems and solutions. Molecular Ecology Notes **7**:10-14.

Bonin, A., Bellemain, E., Bronken Eidesen, P., Pompanon, F., Brochmann, C., and Taberlet, P. 2004. How to track and assess genotyping errors in population genetics studies. Molecular Ecology **13**:3261-3273.

Broquet, T., and Petit, E. 2004. Quantifying genotyping errors in noninvasive population genetics. Molecular Ecology **13**:3601-3608.

Chapuis, M. P., and Estoup, A. 2007. Microsatellite null alleles and estimation of population differentiation. Molecular Biology and Evolution **24**:621-631.

Davison, A., and Chiba, S. 2003. Laboratory temperature variation is a previously unrecognized source of genotyping error during capillary electrophoresis. Molecular Ecology Notes **3**:321-323.

Falush, D., Stephens, M., and Pritchard, J. K. 2003. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. Genetics **164**:1567-1587.

Gagneux, P., Boesch, C., and Woodruff, D. S. 1997. Microsatellite scoring errors associated with noninvasive genotyping based on nuclear DNA amplified from shed hair. Molecular Ecology **6**:861-868.

Givens, G. H., Huebinger, R. M., Bickham, J. W., George, J. C., and Suydam, R. 2007. Patterns of genetic differentiation in bowhead whales (*Balaena mysticetus*) from the western arctic. International Whaling Commission report No. SC/M07/AWMP1.

Hoffman, J. I., and Amos, W. 2005. Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. Molecular Ecology **14**:599-612.

Hoffman, J. I., Matson, C. W., Amos, W., Loughlin, T. R., and Bickham, J. W. 2006. Deep genetic subdivision within a continuously distributed and highly vagile marine mammal, the Steller's sea lion (*Eumetopias jubatus*). Molecular Ecology **15**:2821-2832.

Johnson, P. C. D., and Haydon, D. T. 2007. Maximum-Likelihood estimation of allelic dropout and false allele error rates from microsatellite genotypes in the absence of reference data. Genetics **175**:827-842.

LaHood, E. S., Moran, P., Olsen, J., Grant, W. S., and Park, L. K. 2002. Microsatellite allele ladders in two species of Pacific salmon: preparation and field-test results. Molecular Ecology Notes **2**:187-190.

McKelvey, K. S., and Schwartz, M. K. 2004. Genetic errors associated with population estimation using non-invasive molecular tagging: problems and new solutions. Journal of Wildlife Management **68**:439-448.

Paetkau, D. 2003. An empirical exploration of data quality in DNA-based population inventories. Molecular Ecology **12**:1375-1387.

Pritchard, J. K., Stephens, M., and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. Genetics **155**:945-959.

Skaug, H. J., and Øien, N. 2004. Genetic tagging of males in north Atlantic minke whales through comparison of mother and fetus DNA profiles. International Whaling Commission report No. SC/56/SD3.

Taberlet, P., Griffin, S., Goossens, B., Questiau, S., Manceau, V., Escaravage, N., Waits, L. P., and Bouvet, J. 1996. Reliable genotyping of samples with very low DNA quantities using PCR. Nucleic Acids Research **24**:3189-3194.

Weir, B. S., and Cockerham, C. C. 1984. Estimating F-statistics for the analysis of population structure. Evolution **38**:1358-1370.