

Supplement C: Sampling Variances of Two Approaches for Estimating the Ancestry of Offspring

Our goal was to derive the sampling variance of two methods of estimating the genetic ancestry of juvenile Westslope Cutthroat Trout. Method 1 (which we used in this study) was to genotype a set of loci from each parent; the alternative, method 2, was to genotype the same number of loci from the offspring.

Assumptions

We made three assumptions, namely, that diagnostic loci are available; the ancestry of the offspring is equal to the average ancestry of the parents; and genetic sampling is binomial. These are reasonable assumptions for populations that have experienced multiple generations of interbreeding. The formulae derived below will not apply to F₁s or other recent backcrosses.

Notation

P_1	Frequency of RBT genes in parent 1
P_2	Frequency of RBT genes in parent 2
P_{off}	Frequency of RBT genes of offspring of parents 1 and 2
\hat{P}_1	Estimate of frequency of RBT genes in parent 1
\hat{P}_2	Estimate of frequency of RBT genes in parent 2
\hat{P}_{off}	Estimate of frequency of RBT genes in offspring
$Var(\hat{P}_1)$	Variance of estimates of P_1
n	number of alleles genotypes per individual

Variance of \hat{P}_{off} Using Method 1

Using our method,

$$\hat{P}_{Off} = \frac{\hat{P}_1 + \hat{P}_2}{2}.$$

The variance of \hat{P}_{Off} is equal to

$$Var(\hat{P}_{Off}) = Var\left(\frac{\hat{P}_1 + \hat{P}_2}{2}\right).$$

This is easily rearranged to

$$Var(\hat{P}_{Off}) = \frac{1}{4} Var(\hat{P}_1) + \frac{1}{4} Var(\hat{P}_2).$$

The variance of \hat{P}_1 is well known. It is equal to

$$Var(\hat{P}_1) = \frac{P_1(1-P_1)}{n}.$$

Inserting the term above in the preceding equation gives us a formula for the sampling variance of juvenile fish using our approach:

$$(1) \quad Var(\hat{P}_{Off}) = \frac{1}{4} Var\left(\frac{P_1(1-P_1)}{n}\right) + \frac{1}{4} Var\left(\frac{P_2(1-P_2)}{n}\right)$$

Notice that this variance is less than the variance for either parent.

Variance of \hat{P}_{Off} Using Method 2

We assume that the genetic ancestry of the parents is equal to the average of the parents

$$P_{Off} = \frac{P_1 + P_2}{2}.$$

The variance of estimates of P_{Off} is the variance of an estimate of a binomial proportion (as above), namely,

$$(2) \quad Var(\hat{P}_{Off}) = \frac{P_{Off}(1-P_{Off})}{n}.$$

Comparing Variances

Comparing equation (1) with equation (2) shows that the variance for the sampling method that we used is smaller than that for the alternative of genotyping offspring. The

difference depends on the values of P_1 and P_2 . If $P_1 = P_2$, our method will have half the sampling variance of the alternative; if $P_1 \neq P_2$, our method will have a variance less than half that of the alternative.

We conclude that genotyping parents produces better estimates of the ancestry of juveniles than genotyping offspring for the same number of loci.