

## Genetic signatures of interpopulation dispersal

The Skjern River drains the largest watershed in Denmark and is the last Danish river to contain Atlantic salmon (*Salmo salar*). But the Skjern salmon population has declined drastically this century, and it has even been suggested that salmon in the Skjern are no longer Danish, but rather immigrants from salmon populations elsewhere in Scandinavia. Whether or not this is a hot topic among Danish politicians, the genetic identity of Skjern salmon is relevant to the design of supportive breeding programs.

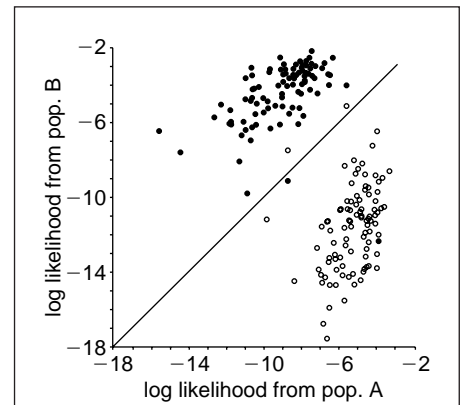
Einar Nielsen, along with Michael Hansen and Volker Loeschcke, have now used DNA microsatellites to demonstrate that Skjern salmon are indeed the descendants of old Danish stock<sup>1</sup>. Their analyses add to the growing list of microsatellite applications in population biology<sup>2-6</sup>, and hint at ways that these hypervariable genetic markers may help us estimate one of the most enigmatic parameters in population and conservation biology, interpopulation dispersal rate. Nielsen *et al.* compared the genotypes of present-day fish to those of 'historic' Danish salmon, by amplifying DNA

from scales collected by anglers during the 1930s. But they also compared present-day genotypes with those of fish from possible source populations in Sweden and Scotland. To do this they used two simple but powerful approaches that take advantage of the enormous amount of information available in individual microsatellite genotypes – the 'assignment test' and a 'shared-allele' index.

The assignment test has conceptual origins in forensic applications, for example where the object is to determine whether an animal trophy originates from a protected population rather than the legal source its owner may claim. One types the trophy at multiple loci and calculates the expected frequency of the trophy's genotype in each putative source population. The genotype is assigned to the population where its expected frequency is highest, that is, where it has the greatest probability of occurrence (Box 1). The test is readily extended to situations, like that of the Danish salmon, in which genotypes are sampled by biologists rather than by poachers. Nielsen *et al.*<sup>1</sup> showed that 28 of 36 present-day Skjern salmon had genotypes more likely to occur in the historic Danish population than either the Swedish or Scottish populations.

Assigning individuals to particular source populations is a surprisingly useful thing to do. For example, Haig *et al.*<sup>8</sup> have recently used the test to assign individual shorebirds sampled during migration to particular breeding populations. Several of the nine species they sampled breed in geographically disjunct populations whose differentiation, migratory routes and wintering grounds are unknown. While banding can be used to associate migratory or wintering birds to particular breeding populations, the assignment test could prove far more useful: very few birds have bands, but all have genotypes.

Other situations in which assignment might be useful include linking fish to the hatcheries that produced them, locating the source habitats that sustain migratory birds living in habitat sinks, and tracing parasites identified in their intermediate hosts to their source populations in definitive hosts. Reed *et al.*<sup>9</sup> have used the assignment test for an even more imaginative purpose. In a study investigating the impact of seals on commercial fisheries, they have documented seal diets from scat collected on haul-out sites. The only



**Fig. 1.** 'Misassigned' genotypes and interpopulation dispersal. This scatterplot illustrates the application of the assignment test to two simulated populations linked by an average of one disperser every other generation ( $N=100$  animals in each population; probability that any individual disperses,  $m=0.005$ ). Log likelihoods are calculated as in Box 1 based on eight loci with 12 possible alleles each (because of drift, not all alleles exist in each population). In the particular generation illustrated here, animals sampled from population A have genotypes represented by open circles; most genotypes have a higher probability of occurrence in population A, so the circles representing them fall below the 45° line. However, two 'A' genotypes in this generation are misassigned, falling above the line. Similarly, two of the 'B' genotypes (solid circles) are misassigned, one of them occurring on the right side of the cloud of 'A' genotypes. That individual is an immigrant into population B: its genotype indeed belongs to population A because it was born there. The other three misassigned genotypes are either descendants of dispersers, or are misassigned due to sampling error.

### Box 1. Using the assignment test

The assignment test assumes that the population is in Hardy–Weinberg equilibrium and that there is no linkage disequilibrium. It has been used with microsatellite, RAPD and allozyme loci, and with combinations of these. Assigning a genotype to a population proceeds in the following steps: (1) Remove the test individual's genotype from the population it was sampled in and estimate allele frequencies at each locus ( $p_i, p_j, \dots$  for alleles  $i, j, \dots$  at locus  $l$ ); (2) determine the genotype's expected frequency in that population at each locus ( $p_i^2$  for homozygotes,  $2p_i p_j$  for heterozygotes); (3) multiply across loci and log-transform the product; (4) perform the same calculations to estimate the genotype's frequency in other putative source populations; and (5) assign the genotype to the population in which it has the highest log likelihood of occurrence. In calculating allele frequencies, a problem arises when the test genotype contains an apparently unique allele. Because this situation might arise from sampling error, the unique allele is added to all test populations at a small frequency, e.g.  $1/2N$  copies. This convention differs slightly from that originally proposed<sup>7</sup> and assigns individuals with less bias when populations differ in size. In this form, the assignment test is essentially the 'leave one out' technique of crossvalidation in discriminant analysis<sup>15</sup>. An assignment test calculator is on the web at <http://www.biology.ualberta.ca/jbrzusto/Doh.html>.

problem with this approach comes in identifying which seal species contributed the scat! Reed *et al.* amplified microsatellite loci common to gray seals *Halichoerus grypus* and harbour seals *Phoca vitulina* from their scat samples, and used the assignment test to attribute individual scats to species.

While the inherent variability of microsatellite loci can make the assignment test highly reliable as a means of identifying an individual's population of origin, in other cases it is the phenomenon of 'misassignment' that makes the test valuable. Paetkau *et al.*<sup>7</sup> developed the test to estimate the degree of differentiation between polar bear (*Ursus maritimus*) populations: pairs of populations among which animals were frequently misassigned were those separated by short genetic distances. In their cross-species comparisons, Haig *et al.*<sup>8</sup> found that shorebird species with fewer birds misassigned between populations were characterized by higher interpopulation  $F_{ST}$  scores and larger genetic distances. Similarly, Kjaer *et al.*<sup>10</sup> have shown that misassignments of teak (*Tectona grandis*) genotypes are

frequent among populations that are little-differentiated by traditional population genetic measures (low  $F_{ST}$ ). Andersen *et al.*<sup>11</sup> applied the assignment test to harbor porpoises (*Phocoena phocoena*), finding that North Sea porpoise genotypes were frequently misassigned to the Danish population, and vice versa, but that it was much less common to misassign either North Sea or Danish porpoises to the Greenland population. They argued that this result reflected a closer relationship between the Danish and North Sea populations.

Work by Favre *et al.*<sup>12</sup> on European shrews (*Crocidura russula*) suggests what may eventually become an even more important application of the assignment test. These authors showed that the log likelihoods of individuals' belonging (genetically) to the population they were trapped in were on average lower, more variable, and more left-skewed for female than male shrews. Shrews have female-biased dispersal, and individuals with a low probability of being assigned to the population they were trapped in were mostly immigrants. Favre *et al.* therefore suggested that differences in the male and female log likelihood distributions can be used to detect sex-biased dispersal.

It is a small step from thinking of the assignment test as a descriptor of population differentiation to thinking of it as a means of estimating interpopulation dispersal rates. In their polar bear study, Paetkau *et al.*<sup>7</sup> found that animals were most often misassigned between populations that are adjacent to each other and suspected of interchange based on radiotracking individual bears. A genotype is 'misassigned' if it has a lower likelihood in the population it was sampled in than in some other sampled population. Paetkau *et al.* therefore suggested that 'misassigned' genotypes might not be misassigned at all. Instead, they might belong to immigrants that originated in their assigned population (Fig. 1).

The assignment test is not the only simple but powerful approach that Nielsen *et al.*<sup>1</sup> have applied to Danish salmon. They also use a second microsatellite-based statistic, the proportion of shared alleles, to assess the probability that Skjern salmon are Danish. At any locus, two individuals' genotypes share either 0, 1 or 2 alleles; across large numbers of loci, the proportion of shared alleles becomes almost continuously distributed and is an index of genetic similarity. The shared-allele indices of pairs of fish can be used to illustrate the pattern of genetic similarity by using a neighbor-joining dendrogram. This approach makes no assumptions about the number or identity of separate populations the individuals are drawn from, but sure enough nearly all present-day Skjern sal-

mon cluster with each other and with historic Danish fish. Only about 10% of the fish cluster with the 'wrong' populations in the dendrogram.

The shared-allele index uses less of the information potentially available in microsatellite genotypes than the assignment test does, but it may be useful when populations are not well defined *a priori*. For example, Bowcock *et al.*<sup>13</sup> have used it with human microsatellite loci to examine an analog of the 'mitochondrial Eve' hypothesis based on nuclear DNA. And Blouin *et al.*<sup>14</sup> have used computer simulation to show that, at least for 'virtual' offspring of known-genotype wild house mice (*Mus musculus*), the tree constructed from the shared-allele index perfectly mirrors known parent-offspring relationships.

Many of the advantages of both the assignment test and the shared-allele index arise because, unlike traditional measures of interpopulation differentiation (for example, those based on Wright's  $F_{ST}$ ) they are based on *individual* genotypes. It is the existence of individual, rather than population-wide descriptors that makes estimating dispersal rates between particular populations, and even identifying individual immigrants, imaginable. The power and the conceptual simplicity of individual-based genetic descriptors become possible because of the huge amount of information present in large numbers of highly variable loci.

The possibility that the assignment test or related statistics might be used to identify individual dispersers, estimate rates of dispersal between populations, or reconstruct genetic lineages in the wild seems a long way from the beginnings that Nielsen *et al.*'s paper exemplify. But their salmon were typed at only four microsatellite loci, while many studies already use eight or more<sup>7,9,12-14</sup>. Because the power of these approaches will increase with the number of available loci, such applications may be closer to realization than we think.

#### Peter M. Waser

Dept of Biological Sciences, Purdue University, West Lafayette, IN 47905, USA

#### Curtis Strobeck

Dept of Biological Sciences, University of Alberta, Edmonton, AL, Canada T6G 2E9

#### References

- 1 Nielsen, E.E., Hansen, M.M. and Loeschcke, V. (1997) **Analysis of microsatellite DNA from old scale samples of Atlantic salmon *Salmo salar*: a comparison of genetic composition over 60 years**, *Mol. Ecol.* 6, 487-492

- 2 Ashley, M.V. and Dow, B.D. (1994) **The use of microsatellite analysis in population biology: background, methods, and potential applications**, in *Molecular Ecology and Evolution: Approaches and Applications* (Schierwater, B. *et al.*, eds), pp. 185-201, Birkhäuser-Verlag
- 3 Dhondt, A.A. (1996) **Molecular techniques in conservation and evolutionary biology: a quantum leap?** *Trends Ecol. Evol.* 11, 147-148
- 4 Jarne, P. and Lagoda, P.J.L. (1996) **Microsatellites, from molecules to populations and back**, *Trends Ecol. Evol.* 11, 424-428
- 5 Kohn, M.H. and Wayne, R.K. (1997) **Facts from feces revised**, *Trends Ecol. Evol.* 12, 223-227
- 6 Queller, D.C., Strassmann, J.E. and Hughes, C.R. (1993) **Microsatellites and kinship**, *Trends Ecol. Evol.* 8, 286-288
- 7 Paetkau, D. *et al.* (1995) **Microsatellite analysis of population structure in Canadian polar bears**, *Mol. Ecol.* 4, 347-354
- 8 Haig, S.M. *et al.* (1997) **Population identification of western hemisphere shorebirds throughout the annual cycle**, *Mol. Ecol.* 6, 412-427
- 9 Reed, J.Z. *et al.* (1997) **Molecular scatology: the use of molecular genetic analysis to assign species, sex and individual identity to seal faeces**, *Mol. Ecol.* 6, 225-234
- 10 Kjaer, E.D., Siegismund, H.R. and Suangtho, V. (1996) **A multivariate study on genetic variation in teak (*Tectona grandis* L.)**, *Silvae Genet.* 45, 361-368
- 11 Andersen, L.W. *et al.* (1997) **A combined DNA-microsatellite and isozyme analysis of the population structure of the harbour porpoise in Danish waters and West Greenland**, *Heredity* 78, 270-276
- 12 Favre, L. *et al.* (1997) **Female-biased dispersal in the monogamous mammal *Crocidura russula*: evidence from field data and microsatellite patterns**, *Proc. R. Soc. London Ser. B* 264, 127-132
- 13 Bowcock, A.M. *et al.* (1994) **High resolution of human evolutionary trees with polymorphic microsatellites**, *Nature* 368, 455-457
- 14 Blouin, M.S. *et al.* (1996) **Use of microsatellite loci to classify individuals by relatedness**, *Mol. Ecol.* 5, 393-401
- 15 McLachlin, J.J. (1992) *Discriminant Analysis and Statistical Pattern Recognition*, John Wiley & Sons

## TREE is no. 1!

The latest Impact Factor tables published by the Science Citation Index place *TREE* in front of all other ecological and evolutionary journals in terms of frequency of citation, with the highest impact factor (6.252) recorded in recent years in ecology or evolutionary biology.

Don't forget to renew your subscription for 1998, to continue receiving your regular supply of high-quality reading in ecology and evolution.