Efficient genetic markers for population biology

Paul Sunnucks

Population genetics has come of age. Three important components have come together: efficient techniques to examine informative segments of DNA, statistics to analyse DNA data and the availability of easy-to-use computer packages. Single-locus genetic markers and those that produce gene genealogies yield information that is truly comparable among studies. These markers answer biological questions most efficiently and also contribute to much broader investigations of evolutionary, population and conservation biology. For these reasons, single-locus and genealogical markers should be the focus of the intensive genetic data collection that has begun owing to the power of genetics in population biology.

Paul Sunnucks is at the Dept of Biological Sciences, Monash University, VIC 3168, Australia (paul.sunnucks@sci.monash.edu.au).

In the past decade, several key advances in molecular genetics have greatly increased the impact of population genetics on biology. Most important have been: (1) the development of polymerase chain reaction (PCR), which amplifies specified stretches of DNA to useable concentrations; (2) the application of evolutionarily conserved sets of PCR primers (e.g. Ref. 1); (3) the advent of hypervariable microsatellite loci[2,3]; and (4) the advent of routine DNA sequencing in biology laboratories. These innovations, coupled with the recent explosion of powerful analyses and relatively user-friendly computer programs[4], have meant that much of the power inherent in molecular genetic data can be tapped, revealing otherwise unobtainable information at all levels of biotic hierarchy[3,5–7]. The profound contribution of molecular genetics is reflected in *Molecular Ecology* being among the most cited primary ecological and/or evolutionary journals (Institute for Scientific Information).

'Natural history of DNA' results in a trail of information

Individual organisms differ in the DNA sequences comprising their genomes. This genetic variation can be considered at the level of individual genes (genic) or of genotypes (genotypic). The fate of a given genetic variant in time and in space will be influenced by the biology and circumstances of the individuals through which it passes, including reproductive success, migration, population size, natural selection and historical events. Population genetic models investigate the connection between these demographic features and the distribution of molecular genetic variants[7,8]. By measuring genetic variation and by applying population genetic models, we can make inferences about the biology of organisms. Processes that affect individuals ultimately accumulate into effects on populations, which, in turn, influence speciation, and so on up the taxonomic hierarchy[8]. Thus, by examining genetic markers with appropriate rates of change, and, therefore, suitable signals, information can be obtained about almost any population and evolutionary process through the hierarchy of life.

The rates of change of the distributions of different genetic markers vary owing to the differential action of fundamental processes, including recombination, mutation and selective constraint (Box 1). Selecting appropriate genetic analysis is vital to the success of applying molecular genetics in population biology. It is a surprisingly common mistake just to use techniques that are available in a laboratory (or, indeed, just to collaborate with a convenient laboratory), rather than choosing the suite of genetic markers that would best answer the question and then obtaining those data. We might consider three levels of molecular change that would provide information at different levels of population biology (Box 1). First, the most sensitive genetic signals are genotypic arrays, most commonly encountered in the form of multiple microsatellite loci scored in samples of individuals. In sexual species, these arrays are reshuffled at each generation, and, therefore, are useful for the shortest- and finest-scale population processes, such as individual identification and tracking, parentage and relatedness of interacting individuals[9–12]. However, arrays can be recognizable for longer in organisms without frequent genetic recombination[13]. Second, microsatellites, mitochondrial DNA (mtDNA) and other single-locus markers (definitions and details in Boxes 2 and 3) can also be analysed as individual genes with frequencies and geographic distributions (genic analyses). These properties change on larger spatial and temporal scales than genotypic arrays, and are effective markers of gene flow and population history, even in species with limited genetic variation[14,15]. Third, slower again is the creation of new alleles by mutation, thus the analysis of their evolutionary relationships (allele and/or gene genealogies, or phylogenies) is informative about the longer-term processes of phylogeography, speciation and deeper taxonomic phylogenetic reconstruction[8].

Basic properties of genetic markers in population biology

There is an apparently bewildering array of genetic techniques available for population genetic analysis[5,7,16]. However, ignoring technical details and focusing on important properties helps to make sense of the methods. Genetic markers are simply heritable characters with multiple states at each character. Typically, in a diploid organism, each individual can have one or two different states (alleles) per character (locus). All genetic markers reflect differences in DNA sequences, usually with a trade-off between precision and convenience. Separate loci can provide independent tests of hypotheses, thus using many together can yield extreme sensitivity. Genetic variation is usually organized hierarchically (for example, two alleles within an individual, *N* individuals within a subpopulation and *Y* subpopulations within populations). Thus, data

 0169-5347/00/$ – see front matter PII: S0169-5347(00)01825-5

Box 1. Rates of change of the major categories of genetic variants in population biology

Genotypic

Genotypic arrays (composite genotypes of multiple loci): individual genotypes are labile – a single round of sexual recombination usually destroys a genotype. They are quantified most rigorously using multiple single-locus nuclear markers (typically microsatellites) (see Boxes 2–4 and Table 1 for definitions concerning, and attributes of, genetic markers).

- Application: the individual level, including movements of individuals, noninvasive sampling, parentage and relatedness of interacting individuals[7,9,11,12,17,28]. Attributes of individual genotypes can address questions such as the relationship between inbreeding and/or outbreeding and fitness[4].

Interlocus allelic correlations (linkage disequilibrium, LD): associations of alleles at pairs of nuclear loci will persist according to their recombination rate, the effective population size (N_e) and natural selection. Textbook population genetics holds that newly generated correlations of unlinked loci in ideal populations decay to zero in approximately ten generations, but associations among linked loci might persist for hundreds or thousands of generations.

- Application: population level including N_e estimation, understanding metapopulation dynamics, and recognizing recent colonizations and introductions[7,29].

Genic

Allele and/or haplotype frequencies are population statistics that can be changed by genetic drift, founder effect, gene flow and selection. They are estimated most accurately from multiple, separate, nuclear loci and mitochondrial DNA (mtDNA).

- Application: includes estimating gene flow and population subdivision[8].

Gene genealogies

Sequence of mtDNA or nuclear genetic regions [microsatellites and single copy nuclear (scn) markers] will evolve as determined by the mutation rate, selection and population parameters, such as N_e and changes in N_e.

- Application: among populations or species, including intraspecific phylogeography and population history, systematics and biodiversity assessment[1,5,8,19,30].

Box 2. Favourable attributes of genetic markers in population biology

Assayable by polymerase chain reaction (PCR): PCR can use low quantities of even degraded DNA and can target specific DNA regions.

Comparability ('connectibility'): PCR primers that amplify homologous regions over a wide taxonomic range generate directly comparable data, thus facilitating profound insights by meta-analysis (deriving conclusions from the results of many studies)[8,22,23,25].

DNA rather than protein: unlike allozymes (Box 3), DNA can be extracted from old material, is more convenient for collection and storage of samples, is more variable, is PCR-assayable and can yield gene genealogies.

Gene genealogies and frequency data: molecular genealogies uniquely can untangle current and historical processes, and can yield information on demographic trends[19]. Markers giving both allele and/or haplotype frequency and sequence data are informative over a range of scales[5,16,19].

Many separate loci available: the use of multiple markers has two main advantages: (1) overcoming stochastic biological sampling that can cause different patterns in loci with similar histories; and (2) detection of nonconcordant characters that are likely to be biologically interesting (e.g. affected by natural selection).

Rapid development and screening: ideal markers would be useable in new taxa with little further development and amenable to rapid screening.

Single-locus as opposed to multilocus markers: multilocus approaches (i.e. they can visualize many anonymous genes simultaneously, for example, RAPD and AFLP) are technically convenient but imprecise, and have many major technical and/or analytical drawbacks, such as dominance (only one allele identified)[31–33]. Data are of limited comparability among studies. By contrast, single-locus, codominant (both alleles identified) or haploid organellar markers supply robust data for input into precise analyses. Data are comparable among studies, thus contribute globally.

are amenable to analysis by approaches closely related to mainstream statistics familiar to ecologists, including analyses of variance, spatial autocorrelation, Mantel testing and diverse multivariate analyses.

Recent major advances in analysis in molecular population biology

After some considerable lags behind theory (mostly because suitable data were not readily available before the advent of microsatellites), population genetic analysis is making great use of maximum likelihood, Bayesian statistics and Markov chain Monte Carlo simulation techniques to extract more information from genetic data[4,17,18]. A fundamental advance is the development of analyses of gene genealogies focusing on evolutionary relationships among individual alleles sampled from populations within species[3,19] (Box 4). These techniques (notably 'coalescence approaches'[3,4,8,19] and 'nested clade analysis'[19]) allow clear testing of historical and spatial hypotheses, such as distinguishing current restricted gene flow from past gene flow, and investigating the direction and relative timing of events, such as range expansions. Finally, models of molecular evolutionary change are becoming sophisticated and can greatly improve genealogical inference[16]. These developments have resulted in fundamental advances in what can be learnt about populations. Luikart and England[4] cite 26 different analyses in four categories of population biology [relatedness and parentage, dispersal and migration, inbreeding, and effective population size (N_e)] that can provide previously unattainable information, and this counts only those primarily involving microsatellites. The new approaches are making great strides in analysing nonequilibrium situations that predominate in conservation biology and in studies of invading organisms[15,19,20].

Choosing genetic markers for a given question

A population genetic survey must start with a decision regarding appropriate genetic markers. The main issues are outlined below, and attributes of specific markers are summarized in Boxes 2 and 3, and Table 1.

Sensitivity

A marker must have the correct sensitivity for the question (Box 1 and Table 1). It is possible to have too much information (if entities are too different there is nothing to link them) or too little information (no signal). Accumulated data have given us a good idea about what sorts of markers are probably informative at a given level in the biotic hierarchy, but pilot studies are usually a good idea. Among markers with suitable resolution, choices can be made on more pragmatic bases (Box 2).

Multilocus or single-locus?

Usually, there is a trade-off between practicality and accuracy of genetic markers. One manifestation of this is the dichotomy between multilocus DNA techniques [usually RAPDs (randomly amplified polymorphic DNA) and AFLP (amplified fragment length polymorphic DNA)] and single-locus techniques [commonly microsatellites and single copy nuclear (scn)DNA regions] (Boxes 2 and 3). Multilocus approaches are technically convenient, but have some marked weaknesses and limitations, including that a substantial proportion of the variation they detect can be nonheritable or not even derived from the target organism. These drawbacks have been shown in organisms including plants, nematodes, flies, birds and mammals (Box 2). A fundamental limitation is dominant inheritance – DNA fragments can be scored only as present or absent, in contrast to codominant inheritance where each of the two alleles at a locus in an individual can be identified and thus analysed more precisely. As a consequence of simultaneous visualization of many dominant markers, multilocus data typically

are analysed as pairwise comparisons of complex patterns that only have meaning relative to others in the same study. Thus, multilocus data can be compared only superficially among studies. Allele frequencies are rarely available, making many of the powerful analyses discussed here impossible. By contrast, single-locus markers are far more flexible, informative and connectible, because they can be analysed as genotypic arrays, as alleles with frequencies and as gene genealogies. It is sometimes asserted that multilocus techniques are more economical, but this is doubtful, especially per unit information[21]. In that study, single-locus microsatellites detected much that RAPD markers had overlooked: a highly divergent cryptic species, interspecific gene flow and high levels of genetic recombination, as well as some more complex aspects of biology. Single-locus markers are even more economical when the value of comparing data sets is considered[22,23].

It is no coincidence that most of the important recent advances in population genetic analysis cannot use multilocus data (Box 4); they depend on comparisons of attributes of alleles within and among loci, which are not provided by multilocus techniques. Notwithstanding, multilocus techniques can generate many variable bands, and, consequently, can be powerful in applications such as gene mapping and analysis of quantitative traits. In one elegant example, a variable RAPD band was cloned and converted to a single-locus codominant marker of asexuality in an aphid[24].

Gene genealogies and frequencies

Markers capable of yielding gene genealogies present some enormous benefits over those that do not (Box 4). Analysis of the relationships between demographics and genealogies are a major growth area leading to some previously unimaginable advances in what can be deduced about population processes and history[3,4,7,8,15,19]. For example, molecular phylogenies can help untangle current structure from the effects of historical events[19] and might be the only way of obtaining long-term demographic information for conservation planning[6,8]. Nested clade analysis[19] is formulated in a particularly clear statistical and hypothesis-testing framework, and can be implemented by the computer program GEODIS (D. Posada, http://bioag.byu.edu/zoology/crandall_lab/geodis.htm). Worked examples, including inference of spatial and temporal details of postglacial range expansion in gophers (*Geomys bursarius*), can be found in Ref. 19. As well as contributing to such developments, another global advantage of the acquisition of genealogies is to facilitate the increasingly profound inferences becoming possible using data sets from comparable ('connectible' *sensu*[5]) genetic markers. These include bringing together data from diverse organisms to uncover the major evolutionary impacts of phenomena such as glaciation, and use of standard measures of genetic divergence to integrate biological and phylogenetic species concepts[22,23,25].

Organelle and nuclear DNA

Cells from most eukaryotes contain biparentally inherited nuclear DNA, as well as DNA in organelles (mitochondria; and chloroplasts in plants) that is usually inherited uniparentally. This difference in transmission, and some major differences in patterns of evolution, causes organellar DNA and nuclear DNA gene genealogies to reflect different aspects of population biology and history. Mitochondrial DNA has a lower N_e than nuclear markers, and, consequently (under most demographic scenarios), mtDNA variants become diagnostic of taxa more rapidly. Comparison of nuclear and mitochondrial genotypes can help recognize hybrid individuals, asymmetrical mating preferences and stochastic effects on variants for which ancestral taxa were polymorphic. These phenomena can cause phylogenetic trees of some genes to not match those of the taxa that carry them[5]. Accordingly, it is generally preferable to use a suite of markers capable of detecting such phenomena.

Rapid development and screening

Major gains in efficiency in new research can arise if genetic markers have already been developed or can be transferred from earlier work, and the possibility of rapid screening, such as single-stranded conformation polymorphism (SSCP)

Box 3. Attributes of specific genetic markers most useful in population biology

Anonymous single copy nuclear (scn)DNA: several methods have been used to develop PCR-assayable scnDNA regions[34]. These techniques can be technically convenient and can yield sensitive data, but the actual markers are transferable only among similar taxa. The technique is too recent for its full potential to be known.

Conserved primers for amplifying specific variable nuclear regions (e.g. introns): primers in evolutionarily conserved DNA regions that flank more variable regions can be technically convenient, and can yield connectible frequency and genealogical data[15,35]. Loci tend to be too invariant for the shortest-term population processes. The technique is too recent for its full potential to be known.

Protein electrophoresis (allozymes): modest numbers of variable, codominant, nuclear allozymes are available with minimal development. The technique is inexpensive, but requires high-quality samples, often reveals little variation and gives limited genealogical information. It is used extensively in studies of gene flow.

Sequences from known mtDNA and nuclear regions amplified by conserved or specific primers: is technically convenient for most organisms, and is the mainstay of intraspecific phylogeography and of systematics.

Single-locus microsatellites: numerous hypervariable, codominant nuclear markers are available, assayed via PCR to reveal length variation among alleles. These provide sensitive, connectible data from individual identification through to shallow phylogeny. Loci have a wide range of evolutionary rates, thus examine different timescales. Development of loci is laborious, but once developed they can have moderate taxonomic breadth of application, thus suitable markers might be available in the literature[36] (unpublished information can be obtained via e-mail: micro-sat@sfu.ca). Microsatellites are the mainstay of modern population genetics other than systematics.

Box 4. Important recent developments in genetic population biology analysis

Using low probabilities of individual genotypes (multiple single loci)

Parentage and/or relatedness analysis: these find probable parents from pools of candidates, test individual parentage and/or relatedness hypotheses and estimate kinship, with adjustable assumptions and giving reliability estimates[4,10,17].

Migration and nonequilibium populations – assignment tests and mixed stocks analysis: assignment tests estimate the population of origin of individuals, thus identifying individual migrants (and even their descendants)[11,20] and uncovering attribute-biased dispersal. Related tests (e.g. mixed stocks analysis) examine population admixture[4,15,28]. These analyses estimate the same phenomena as trapping and/or marking, but have the enormous advantage that 'recapture' is not required: '…very few birds have bands, but all have genotypes'[28].

Using gene genealogies and distributions of other allelic properties

Microsatellite allele length data (in the light of molecular genetic models) and DNA sequences yield gene genealogies[3,4,7,8]. Genealogies illuminate population processes, phylogeographic events and speciation, because they add the dimension of evolutionary (thus temporal) relationships among alleles, and these can be related to spatial organization[3,4,8,19]. Other allele properties (lengths, frequencies and distribution shapes) yield sensitive information[4]. These procedures are becoming well validated using large data sets, simulation and experiments[3,4,7,8]. Gene genealogies and marker attributes can be used to estimate current, past, changes in, timing of changes in, and even rate of changes in N_e, thus uncovering population history and overcoming previously intractable problems[3,4,15,28]. For example, BOTTLENECK (Ref. 4) infers recent bottlenecks without prebottleneck data, by examining observed relationships between allelic diversity and heterozygosity compared with molecular genetic models.

Table 1. Attributes of markers commonly used in molecular population biology[a]

	PCR assay	Single locus	Codominant	Allele genealogy feasible	No. loci readily available	Connectibility of data among studies	Rapid transfer to new taxa	Overall variability[g]
Mitochondrial (and chloroplast)								
Sequence	Yes	Yes	Yes[d]	Yes	Single	Direct	Yes	Low–high
RFLP	No, large	Yes	Yes[d]	Yes	Single	Direct	Yes	Low–moderate
Multilocus nuclear								
Mini- and/or micro-satellite 'fingerprints'	No, large	No	No	No	Many	Limited	Yes	High
RAPD[b]	Yes	No	No	No	Many	Limited	Yes	High
AFLP[b]	Yes	No	No	No	Many	Limited	Yes	High
rDNA[c]	Yes	No	No	No	Few	Limited	Yes	Moderate–high
Single-locus nuclear (single copy nuclear, scn)								
Allozymes	No, protein	Yes	Yes	Rarely	Moderate	Direct	Yes	Low–moderate
Minisatellites	Few	Yes	Yes	Rarely	Moderate	Indirect[e]	Few	High
Microsatellites	Yes	Yes	Yes	Yes	Many	Indirect[e]	Some	High
Anonymous scn	Yes	Yes	Yes	Yes	Many	Indirect[e]	No?[f]	Moderate?[f]
Specific scn	Yes	Yes	Yes	Yes	Moderate	Direct	Yes?[f]	Moderate?[f]
rDNA[c]	Yes	In effect	Yes	Yes	Few	Direct	Yes	Low–moderate

[a]More details in Boxes 2 and 3.
[b]Some RAPD (randomly amplified polymorphic DNA) and AFLP (amplified fragment length polymorphic DNA) bands can be converted to single-locus markers, in which case they behave like 'anonymous scn'or 'specific scn' categories.
[c]rDNA consists of tandem arrays of a few regions. In some taxa the arrays are effectively identical and regions act as single loci, but in some taxa there can be many different sequences within individuals, in which case rDNA acts more like a multilocus system.
[d]mtDNA and chloroplast DNA are haploid and show one of a range of alternative positive states, in contrast to dominant markers that are either present or absent.
[e]Data from these markers are indirectly, but meaningfully, connectible given adequate models of molecular evolution.
[f]Insufficient research effort has been put into these markers.
[g]Variability depends on variation per marker and number of markers obtained readily. The assessment here approximates the outcome of a typical marker system.

(Refs 7,8,13,16), can yield important savings in resources. These issues of technical convenience should be included in assessments of relative suitability of candidate marker systems for each project.

Prospects

The advances outlined here depend on the application of sophisticated statistics to the sensitive data that can be obtained from microsatellites and from other single-locus markers, and the development of methods for extracting demographic information from gene genealogies. Many computer programs and refinements or new applications of statistics can be found on the Internet (Ref. 4). Although many sorts of patterning in genetic data (e.g. allele frequency distributions and evolutionary divergences among alleles) are coming into common use, others are underutilized, a good example being allelic correlations among loci (linkage disequilibium). Although genotypic approaches have made in-roads into modelling nonequilibrium demographic scenarios, much progress remains to be made[15]. For example, assignment tests assume linkage and Hardy–Weinberg equilibria, but these are unlikely to hold during colonizations or in recently perturbed populations, such as those of conservation concern. Violations of these assumptions might have important effects on the accuracy of current tests (B. Rannala, pers. commun.). On a technical note, approaches that detect high levels of sequence variation in nuclear genome regions, such as introns and variable regions of ribosomal RNA (rRNA) genes, have not yet supplied sufficient markers for many taxa. Impressive efficiencies in marker development and screening have been made in biomedicine and agriculture, particularly codominant AFLPs and microchip-based screening of single nucleotide polymorphisms (SNPs). However, these require substantial input of resources and it remains to be seen what impact they will have in studies of species that are not of economic interest.

Much of the current boom in molecular ecology has resulted from user-friendly computer packages such as GENEPOP (Ref. 26) (version 3.1d can be obtained from http://www.cefe.cnrs-mop.fr/). However, although the principle of interconnection among packages embodied in GENEPOP greatly facilitates development of the field, it still has a long way to go. Hours can be spent converting data from one format to another, and trying to find elusive typographical errors that cause programs to crash. Finally, greater understanding of molecular evolution is central to increased precision in inferences about population processes. Two important examples are refining the assumptions and approaches to absolute dating of population events using molecular clocks, and a clearer understanding of the occurrence and impact of convergent evolution of microsatellite allele lengths (length homoplasy)[3,7,8,23,24].

By providing solutions to difficult problems, molecular ecology is an invaluable part of modern population biology; it is most powerful when integrated with whole animal biology[7,27]. Technical and analytical developments have facilitated each other, and we are rapidly achieving great resolution for interesting biological questions. Primarily this will be via the intelligent application of multiple, single-locus, codominant, genealogy-yielding genetic markers.

Acknowledgements

I apologize to the many excellent researchers whose work I have not been able to cite directly in this review. Tristan Marshall, Chris Simon, Gordon Luikart, Andrea Taylor, Alex Wilson, Luciano Beheregaray, Steve Jordan, Lou Rodgerson and an anonymous reviewer made valuable comments on this article. Dave Briscoe coined 'the natural history of DNA'.

Thanks to Bruce Rannala for useful discussions. I thank Mike Bruford, Bob Wayne and Andrea Taylor for their lasting effects on my way of thinking about molecular population biology. This is publication 308 of the Key Centre for Biodiversity and Bioresources at Macquarie University, Sydney.

References

1 Simon, C. *et al.* (1994) Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Ann. Entomol. Soc. Am.* 87, 651–701

2 Jarne, P. and Lagoda, J.L. (1996) Microsatellites, from molecules to populations and back. *Trends Ecol. Evol.* 11, 424–429

3 Goldstein, D.B. and Schlötterer, C., eds (1999) *Microsatellites: Evolution and Applications*, Oxford University Press

4 Luikart, G. and England, P.E. (1999) Statistical analysis of microsatellite data. *Trends Ecol. Evol.* 14, 253–256

5 Avise, J.C. (1994) *Molecular Markers, Natural History and Evolution*, Chapman & Hall

6 Schierwater, B. *et al.*, eds (1994) *Molecular Ecology and Evolution: Approaches and Applications*, Birkhäuser

7 Smith, T.B. and Wayne, R.K., eds (1996) *Molecular Genetic Approaches in Conservation*, Oxford University Press

8 Burke, T. *et al.*, eds (1998) Special Issue – Phylogeography. *Mol. Ecol.* 7, 367–545

9 Taylor, A.C. *et al.* (1997) Unusual relatedness structure detected by microsatellite analysis, and parentage analysis in an endangered marsupial, the Northern Hairy-nosed wombat, *Lasiorhinus krefftii. Mol. Ecol.* 6, 9–20

10 Marshall, T.C. *et al.* (1998) Statistical confidence for likelihood-based paternity in natural populations. *Mol. Ecol.* 7, 639–656

11 Paetkau, D. *et al.* (1998) Gene flow between insular, coastal and interior populations of brown bears in Alaska. *Mol. Ecol.* 7, 1283–1292

12 Taberlet, P. *et al.* (1999) Non-invasive genetic sampling: look before you leap. *Trends Ecol. Evol.* 14, 323–327

13 Wilson, A.C.C. *et al.* (1999) Microevolution, low clonal diversity and genetic affinities of parthenogenetic *Sitobion* aphids in New Zealand. *Mol. Ecol.* 8, 1655–1666

14 Taylor, A.C. *et al.* (1994) Genetic variation of microsatellite loci in a bottlenecked species: the northern hairy-nosed wombat *Lasiorhinus krefftii. Mol. Ecol.* 3, 277–290

15 Davies, N. *et al.* (1999) Determining the source of individuals: multilocus genotyping in non-equilibrium population genetics. *Trends Ecol. Evol.* 14, 17–21

16 Hillis, D.M. *et al.*, eds (1996) *Molecular Systematics*, Sinauer

17 Goodnight, K.F. and Queller, D.C. (1999) Computer software for performing likelihood tests of pedigree relationship using genetic markers. *Mol. Ecol.* 8, 1231–1234

18 Shoemaker, J.S. *et al.* (1999) Bayesian statistics in genetics – a guide for the uninitiated. *Trends Genet.* 15, 354–358

19 Templeton, A.R. (1998) Nested clade analysis of phylogenetic data: testing hypotheses about gene flow and population history. *Mol. Ecol.* 7, 381–397

20 Rannala, B. and Mountain, J.L. (1997) Detecting immigration by using multilocus genotypes. *Proc. Natl. Acad. Sci. U. S. A.* 94, 9197–9201

21 Sunnucks, P. *et al.* (1997) Genetic structure of an aphid studied using microsatellites: cyclic parthenogenesis, differentiated lineages, and host specialization. *Mol. Ecol.* 6, 1059–1073

22 Johns, G.C. and Avise, J.C. (1998) A comparative summary of genetic distances in the vertebrates from the mitochondrial cytochrome *b* gene. *Mol. Biol. Evol.* 15, 1481–1490

23 Avise, J.C. *et al.* (1998) Speciation durations and pleistocene effects on vertebrate phylogeography. *Proc. R. Soc. London Ser. B* 265, 1707–1712

24 Simon, J-C. *et al.* (1999) Molecular markers linked to breeding system differences in segregating and natural populations of the cereal aphid *Rhopalosiphum padi* L. *Mol. Ecol.* 8, 965–973

25 Avise, J.C. and Johns, G.C. (1999) Proposal for a standardized temporal scheme of biological classification for extant species. *Proc. Natl. Acad. Sci. U. S. A.* 96, 7358–7363

26 Raymond, M. and Rousset, F. (1995) GENEPOP (version 1.2): a population genetics software for exact tests and ecumenicism. *J. Hered.* 86, 248–249

27 Moritz, C. and Lavery, S. (1996) Molecular ecology: contributions from molecular genetics to population ecology. In *Frontiers of Population Ecology* (Floyd, R.B. *et al.*, eds), pp. 433–450, CSIRO Publishing

28 Waser, P.M. and Strobeck, C. (1998) Genetic signatures of interpopulation dispersal. *Trends Ecol. Evol.* 13, 43–44

29 Tishkoff, S.A. *et al.* (1996) Global patterns of linkage disequilibrium at the CD4 locus and modern human origins. *Science* 271, 1380–1387

30 Crozier, R.C. *et al.* (1999) Towards complete biodiversity assessment: an evaluation of the subterranean bacterial communities in the Oklo region of the sole surviving natural nuclear reactor. *FEMS Microbiol. Ecol.* 28, 325–334

31 Hurme, P. and Savolainen, O. (1999) Comparison of homology and linkage of random amplified polymorphic DNA (RAPD) markers between individual trees of Scots pine (*Pinus sylvestris* L.). *Mol. Ecol.* 8, 15–22

32 Pérez, T. *et al.* (1998) An evaluation of RAPD reproducibility and nature. *Mol. Ecol.* 7, 1347–1358

33 Rabouam, C. *et al.* (1999) Features of DNA fragments obtained by random amplified polymorphic DNA (RAPD) assays. *Mol. Ecol.* 8, 493–504

34 Burt, A. *et al.* (1996) Molecular markers reveal cryptic sex in the human pathogen *Coccidioides immitis*. *Proc. Natl. Acad. Sci. U. S. A.* 93, 770–773

35 Congdon, B.C. *et al.* Rapid population expansion and peripheral isolation in marbled murrelets: historical vs contemporary evolutionary processes. *Evolution* (in press)

36 Slate, J. *et al.* (1998) Bovine microsatellite loci are highly conserved in red deer (*Cervus elaphus*), sika deer (*Cervus nippon*) and Soay sheep (*Ovis aries*). *Anim. Genet.* 29, 307–315

Letters to *TREE*

Correspondence in *TREE* may address topics raised in very recent issues of *TREE*, or (occasionally) other matters of general current interest to ecologists and evolutionary biologists. Letters should be **no more than 500 words long with a maximum of 12 references and one small figure**; original results, new data or new models are not allowed. Letters should be sent by e-mail to TREE@current-trends.com. The decision to publish rests with the Editor, and the author(s) of any *TREE* article criticized in a Letter will normally be invited to reply. Full-length manuscripts in response to previous *TREE* articles will not be considered.