Landscape approaches to historical and contemporary gene flow in plants

Victoria L. Sork, John Nason, Diane R. Campbell and Juan F. Fernandez

Growing interest in metapopulation dynamics and dispersal at a landscape level is promoting new approaches to the study of contemporary gene flow. These approaches have been fostered by the development of new genetic markers and statistical methods, as well as an awareness that contemporary gene flow cannot be reliably estimated by conventional methods based on genetic structure. Estimation of the spatial and temporal dynamics of pollen and seed movement with respect to extant landscape features can aid evolutionary and conservation biologists in predicting the demographic and genetic responses of species to naturally occurring or human-mediated population subdivision.

Victoria Sork is at the Dept of Biology, University of Missouri-St Louis, St Louis, MO 63121-4499, USA (sorkv@umsl.edu); John Nason is at the Dept of Biological Sciences, University of Iowa, Iowa City, IA 52242-1324, USA (john-nason@uiowa.edu); Diane Campbell is at the Dept of Ecology & Evolutionary Biology, University of California, Irvine, CA 92697, USA (drcampbe@uci.edu); Juan Fernandez is at the Dept of Biology, University of Missouri-St Louis, St Louis, MO 63121-4499, USA (s997022@admiral.umsl.edu).

Recently, biologists have become increasingly interested in gene flow that takes place across heterogeneous landscapes and on ecological timescales (V.L. Sork *et al.*, unpublished; http://www.nceas.ucsb.edu/papers/geneflow). The conventional approach to quantifying gene flow has been to transform measures of population structure[1] into indirect estimates of the average number of migrants exchanged per generation among a set of populations[2], most commonly using an island model (Fig. 1a). Because this model assumes all populations are equal sources of migrants and yields estimates that are historical, it does not reflect contemporary variation in gene exchange among populations or current changes to dispersal processes. Moreover, the increasing realization that many species exhibit metapopulation characteristics[3] has drawn attention to the potential genetic differences between recolonization and migration as sources of gene flow[4]. Metapopulation models incorporate information about the spatial location of populations[3] (Fig. 1b), but they do not necessarily incorporate information about the landscape context of populations. By landscape context, we mean the type of habitat in which an individual population is located and the qualities of the intervening landscape that can influence gene exchange (e.g. degree of isolation, size of fragment and successional status of surrounding landscape) (Fig. 1c). Gene-flow models that gain realism by incorporating details of the landscape where populations occur can provide insights about the ecological factors influencing immigration into individual populations and, ultimately, the dynamics of interpopulation gene exchange.

The move to incorporate an ecological timescale in studies of gene flow has been motivated, in part, by the need of conservation biologists to determine whether habitat alteration is causing population isolation. Concomitantly, evolutionary biologists, forest geneticists and others have become interested in contemporary gene flow as a way to understand the dynamics of pollen and seed movement more precisely than can be learned from indirect measures. Recent advances in genetic marker technology[5] and statistical models[6,7] mean that allozyme and microsatellite loci can provide sufficient genetic resolution to monitor current gene flow. The direct detection of gene flow provides population-specific estimates of immigration rates that can be used to analyse the influence of landscape context when several sites are studied. Moreover, recent progress in computer modeling allows us to quantify interpopulation patterns of pollen or seed movement with respect to the spatially explicit ecological context in which populations occur. This approach provides an analysis of contemporary patterns of gene flow within a landscape context.

Indirect methods for estimating historical gene movement

Historical rates of gene flow have traditionally been estimated indirectly from the fixation index of population subdivision relative to the total population (F_{ST}) according to Wright's island model of population genetic structure[1] (Box 1). This model assumes equilibrium between the homogenizing effects of gene flow and the disruptive effects of genetic drift and it also assumes that parameters such as population size and migration rate are uniform and constant over space and time. Because of the relative ease of obtaining data on gene frequencies, F_{ST} is routinely used to estimate the effective migration rate ($N_e m$) for a set of subpopulations. However, because F_{ST} is a summary statistic for a set of subpopulations, it provides little insight into the landscape-level processes that differentially affect spatial patterns of genetic structure and rates of gene flow among individual subpopulations.

Other longstanding models of gene flow include the stepping stone model[8] and the isolation by distance model[1]. In these models, the rate of gene flow is expected to decline monotonically with increasing geographic distance between spatially discrete and continuous populations, respectively (Box 1). Like the island model, these models can be employed to measure the cumulative effects of gene flow indirectly. However, unlike the island model, they incorporate spatial information and thus permit tests of hypotheses about relationships between the effective migration rate and spatial patterns of interpopulation connectivity. In particular, Slatkin's[9] use of F_{ST} to estimate $N_e m$ for pairwise combinations of subpopulations ($\hat{M}$) provides a test for isolation by distance that can be used to take features of the landscape into account. Using Slatkin's approach,

 0169-5347/99/$ – see front matter PII: S0169-5347(98)01585-7

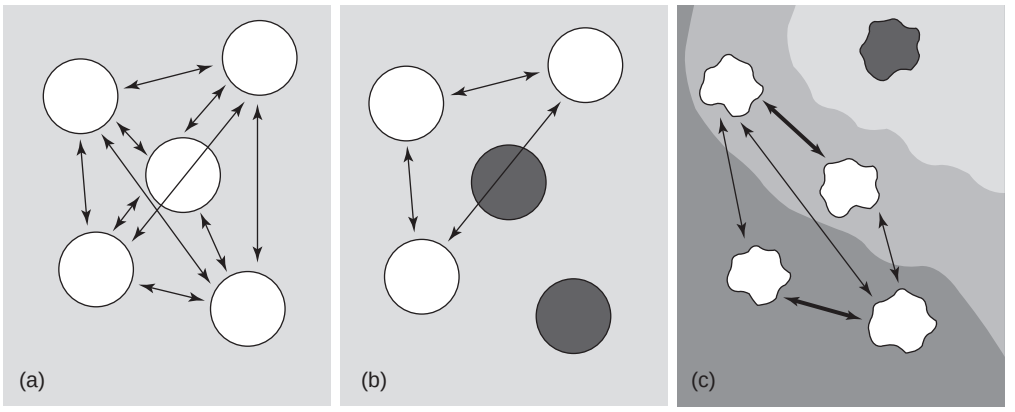


Fig. 1. Three examples of landscape-scale models that can be used to evaluate gene flow (double-headed arrows). (a) The infinite island model of Sewall Wright[1] is the conventional population genetic model of gene exchange among populations (circles). In this model, the effective number of migrants, N_em, is estimated from *F*-statistics for a set of *n* populations. The model assumes equilibrium between migration and genetic drift among all populations. There is equal gene exchange among populations. (b) Metapopulation models assume a set of populations connected by gene flow but dominated by extinction and colonization dynamics. The shaded circles represent extinct populations. (c) Landscape context models view a set of populations in a landscape mosaic (indicated by different shading). Spatially explicit information about the size and shape of populations and the quality of the landscape combined with estimates of gene flow can allow analysis of which populations are contributing and/or receiving migrants or genes. The thickness of the arrows indicates the probability of gene exchange, which is a function of the landscape context and distance. The enclosed areas represent populations (unfilled, extant; shaded, extinct). This type of model is sometimes referred to as a spatially realistic metapopulation model[3].

Kudoh and Whigham[10] examined the relationship between $\hat{M}$ values and geographic distance in *Hibiscus moscheutos* (Malvaceae), a species restricted to riparian habitats, under three alternative models of geographic interpopulation distance: (1) Euclidean distance, (2) distance along river drainages or (3) proximity to the tidal stream. They found strong statistical support for the third model, indicating that gene flow in this species has historically depended on water dispersal of seeds. Although, in some cases, F_{ST}-based data might be subject to alternative interpretations[11], in *H. moscheutos* and other species[12,13] where landscape features influence patterns of dispersal, combining pairwise estimates of gene flow with information about the landscape allows a more accurate assessment of the historical patterns of dispersal.

Another use of indirect estimates of gene flow has been to study the historical metapopulation dynamics of plant species. For example, using maternally and biparentally inherited genetic markers, McCauley *et al.*[14] examined the relative roles of colonization and migration events in determining the genetic structure of the weedy herb *Silene alba* (Caryophyllaceae). Few studies of plants have taken a metapopulation approach because extinction, colonization and migration rates are difficult to measure[12], and also because many plant populations do not possess metapopulation structure dominated by frequent extinction and recolonization events. Nonetheless, indirect methods might be useful for understanding the evolution of genetic structure in those plant species with a history of metapopulation dynamics.

In summary, for studies of evolutionary history, the *F*-statistic approach (as well as other indirect approaches, such as coalescent theory[15]) can provide valuable insights about the cumulative effects of gene flow and its effects on population genetic structure, especially when these approaches also incorporate landscape information. Indirect measures based on F_{ST} are not appropriate for studies of contemporary gene movement, however, because they measure historical gene flow and are relatively insensitive to small changes in gene frequencies[16] (V.L. Sork *et al.*, unpublished; http://www.nceas. ucsb.edu/papers/geneflow).

Direct methods for estimating contemporary gene movement

Determining how environmental heterogeneity and anthropogenic changes in a landscape influence ongoing gene flow requires methods that directly estimate gene flow. Early direct methods relied on observations of distances moved by animal pollinators between flower visits, or on tracking the movement of pollen grains or seeds. In the 1980s, however, statistical, paternity-analysis models were developed to measure realized gene movement, usually by employing multilocus genotypes to infer the fathers of seed offspring from known mothers (Box 2). Although it is sometimes possible to exclude all but one candidate if the pool of potential fathers is small enough, estimating the level of gene flow does not require unambiguous assignment of paternity to individual offspring. Rather, recent multilocus models have used likelihood methods to estimate gene flow. Multilocus likelihood models have been used to estimate distances of pollen movement within populations, as well as the rate of pollen immigration into populations. Multilocus genotypes have also been used in mating-system analysis to estimate the proportion of self-fertilized progeny produced by an individual mother or the population[17] (the most extreme case of short distance gene movement; Box 3).

Most studies of within and among population pollen movement have used allozyme genetic markers because they are often cost effective and relatively easy to obtain. With large samples of offspring or access to full-sib progeny arrays[18], allozymes provide sufficient statistical power to address many questions about interpopulation gene movement. The recent development of highly polymorphic microsatellite markers offers the promise of even greater power to detect gene flow[5]. As for allozymes, allelic variation at microsatellite loci is codominant, which means that these markers can be analysed using the available statistical models. Because of their exceptional allelic diversity, microsatellites should minimize the estimation variance in measuring dispersal. However, although the situation is improving, microsatellites have generally proven more difficult and expensive to develop and assay than allozymes, as evidenced by the greatly reduced number of progeny assayed in most microsatellite compared with allozyme studies of plants. Also, because the variance of parameter estimates is a function of both the level of assayable genetic variation and the number of progeny scored, it should not be assumed that microsatellite studies always provide more precise dispersal estimates than those obtained in allozyme studies.

The emphasis of many studies to achieve completely unambiguous paternity assignment where each progeny can be assigned to a single father has led to a bias towards the study of small, isolated populations. From a landscape perspective, however, the objective of any study should

not be to establish 100% unambiguous paternity assignment. Instead, the objective should be to take advantage of available statistical models to extend the use of the genetic data for estimating dispersal processes over as large a population or set of population subdivisions as possible. Large numbers of potential fathers introduce the problem of cryptic gene flow[7]. Cryptic gene flow is defined as that fraction of total gene flow that cannot be identified using the available genetic markers because of overlap in the sets of multilocus pollen gametes produced by potential fathers (occurring inside and outside the study population). If unaccounted for, cryptic pollen flow will result in the underestimation of total gene flow. To minimize this problem, investigators can use highly variable genetic markers (where possible) and statistical models that adjust for cryptic gene flow[7] (Box 2).

Total gene movement can be separated into components occurring within and between populations. At the within-population level, paternity analysis is used to measure patterns of pollen-mediated gene movement based on likelihood probabilities of male fertility on each of several mothers[19,20] (Box 2). Likelihood estimates of male fertility can be used to fit a pollen dispersal curve relating mating probabilities to the distance between mates[21]. This relationship can then be used to estimate the genetic neighborhood area featured in Wright's isolation by distance models[22]. Although paternity analysis of pollen-mediated gene movement is more common, maternity analysis models of seed-mediated gene dispersal have also been developed[23], and will become an especially useful approach as maternally inherited mitochondrial DNA (mtDNA) and chloroplast DNA (cpDNA) markers become more available.

At the interpopulation level, paternity analysis has been used to estimate rates of pollen immigration into spatially continuous and isolated populations (Box 2, methods 2–3). By comparing immigration into several populations, it is possible to explore the impact of landscape features on contemporary gene flow. Goodell *et al.*[24] used experimental populations of the herb *Raphanus* (Brassicaceae) to show that population size and isolation distance interact in their influence on rates of total pollen immigration. In the insect-pollinated Neotropical canopy tree *Spondias mombin* (Anacardiaceae), Nason and Hamrick[25] found rates of total pollen immigration into small, insular forest fragments to be 100%, despite a range of patch population sizes (one to four trees) and physical isolation distances (80–1000 m). In contrast, pollen flow in continuous forest was approximately 45% over distances of 100–200 m. These results suggest that forest fragments, while physically isolated, might not be genetically isolated. Although few studies do so, assessing the impact of that gene flow also requires information on fitness of migrants. In *S. mombin*, both fruit set and germination rates were significantly lower in populations in small fragments than in continuous forest populations. These two cases demonstrate that direct estimates of contemporary gene flow are sufficiently sensitive to detect differences in gene flow owing to landscape factors.

Only a few studies have used microsatellite loci to measure interpopulation gene movement[26–28], and most of these examined just a single, isolated site. In general, the findings suggest that pollen-mediated immigration can be substantial. To date, only one study has used microsatellites to measure the impact of landscape-level change on both pollen and seed-mediated gene movement. In a Neotropical canopy tree *Symphonia globulifera* (Clusiaceae) pollinated by hummingbirds, Aldrich *et al.*[29] found significantly higher rates of selfing (26%) for trees isolated in pastures than for trees occurring in fragmented (10%) or continuous forest (9%). As a result of the foraging behavior of seed-dispersing bats, most seedlings found in forest remnants were produced

Box 1. *F*-statistics and indirect estimates of gene flow

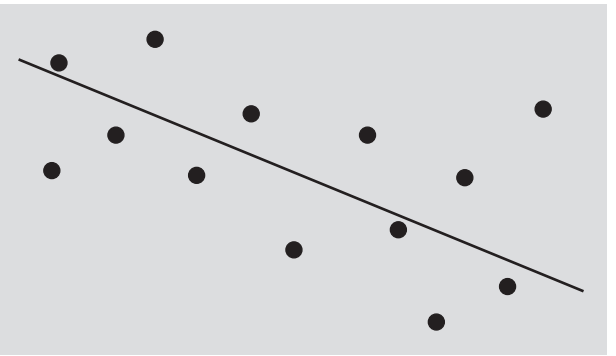
F_{ST} is the correlation between random gametes within subpopulations relative to gametes of a total population. Under Wright's island model of population genetic structure[1], with subpopulations at migration–drift equilibrium, F_{ST} is a function of the average effective number of migrants per generation, $N_e m$, where N_e is the effective subpopulation size and m is the proportion of migrants per generation drawn at random from the whole population. At equilibrium F_{ST} can be estimated for a set of populations as:

$$F_{ST} \approx \frac{1}{4N_e m + 1} \quad (1)$$

Thus, the effective number of migrants can be estimated from F_{ST} as:

$$N_e m \approx \frac{1}{4}\left(\frac{1}{F_{ST}} - 1\right) \quad (2)$$

To test hypotheses concerning spatial patterns of gene migration at various spatial scales, we can use Eqn 2 to estimate the migration between pairs of populations ($\hat{M}$)[9]. If gene exchange among populations conforms to a process of isolation by distance then the effective number of migrants exchanged between populations should be a monotonically decreasing function of interpopulation geographic distances (Figure). (Note: Because these observations lack independence, one should use a Mantel test based on randomization methods to test whether the observed correlation is significantly different from random[33].)



(Online: Fig. I)

The concept of isolation by distance was originally coined by Sewall Wright to describe the decreasing probability of mating as distance between parents increases in a continuous population[1]. In this scenario, migration is estimated as the variance of the parent–offspring distance, which in turn sets the genetic neighborhood area. Application of the term has expanded, however, to include the negative correlation between gene exchange and geographic distance among subpopulations regardless of population structure. The relationship between estimates of migration and distance has been used to examine migration pathways and landscape context by constraining measurements of geographic distance to follow specific pathways in response to features of the landscape (e.g. Ref. 10).

The availability of molecular markers from biparentally (nuclear) and maternally (organellar) inherited genomes in angiosperms allows us to distinguish historical contributions of seed and pollen to estimated gene flow[32] by estimating the ratio of pollen flow to seed flow as:

$$\frac{\text{Pollen flow}}{\text{Seed flow}} \approx \frac{\left(\frac{1}{F_{ST_b}} - 1\right) - 2\left(\frac{1}{F_{ST_m}} - 1\right)}{\left(\frac{1}{F_{ST_m}} - 1\right)} \quad (3)$$

This equation uses separate fixation indices for biparentally and maternally inherited genomes as indicated by sub-subscripts, b and m, respectively. In general, for studies in which nuclear and organellar genetic structure has been compared, rates of gene flow via pollen have greatly exceeded that via seed[34,35].

Box 2. Three methods of paternity analysis for obtaining multilocus estimates of ongoing pollen gene flow

(1) Within population pollen movement: several likelihood-based models have been developed for estimating patterns of individual male reproductive success (RS) and pollen movement onto individual maternal plants[19,36]. The state-of-the-art ones, by Roeder *et al.*[20] and Smouse and Meagher[37], provide unbiased, maximum-likelihood estimates of male RS using iterative fractional paternity-analysis techniques. Potential constraints on the application of this method are: (1) that it requires sampling more progeny per female than there are potential fathers in the study population; (2) that potential fathers have distinguishable multilocus genotypes; and (3) that there is no cryptic pollen immigration (V.L. Sork *et al.*, unpublished, discuss biases in this model resulting from cryptic gene flow; http://www.nceas.ucsb.edu/papers/geneflow). These constraints have generally limited the application of these models to relatively small, spatially isolated populations that might not reflect the landscape-level dispersal dynamics of natural populations.

(2) Pollen immigration: Devlin and Ellstrand[7] developed a model for estimating total pollen gene flow into a population that takes into account the contribution of cryptic gene flow that would go undetected in a conventional paternity analysis. The primary assumption of this model is that gene-flow gametes can be treated as being drawn at random from a large source population of known allele frequency. Although this model has been applied mostly to relatively small, spatially isolated populations, experimental designs have been devised to apply this model effectively to large continuously distributed populations[25].

(3) Simultaneous estimators of intra- and interpopulation pollen flow: the neighborhood model of Adams and Birkes[6] groups fathers by distance and fits a within-population pollen-dispersal function to the paternity data instead of estimating individual male RS. This model provides estimates of selfing and the probability of within-population dispersal as a function of intermate distance, while simultaneously estimating the rate of total pollen gene flow into a study population. Constraints on the application of this model are that it works best for species with even spatial distributions and, at the present time, is restricted to species with assessable megagametophytes (e.g. conifers but not angiosperms).

Recently, based on Roeder *et al.*[20], Nason has developed a general paternity analysis model of pollen gene movement that simultaneously estimates selfing, movement within populations, and immigration. The method provides unbiased estimates of relative male reproductive success that are adjusted for cryptic pollen gene flow. Like the neighborhood model[6], this approach is useful for generating a pollen dispersion curve as well as for estimating gene movement from outside a circumscribed area. In addition, like the model of Kaufman *et al.*[38], it can estimate rates of pollen immigration originating from more than one source population.

Current versions of the models of Nason and Kaufman *et al.*[38] assume that all gene flow source populations can be identified and sampled for allele frequencies and multilocus genotypes, respectively.

(The models of Adams and Birkes, and Nason are available at http://www.nceas.ucsb.edu/papers/geneflow/software)

by a relatively few number of adult trees located in surrounding pasture. Despite seed immigration into forest fragments, higher selfing rates and fertilities skewed towards pasture trees suggest that *Symphonia* populations in fragmented habitat will be subject to genetic bottlenecks in the future.

In summary, paternity- (and maternity-) analysis methods are effective approaches for estimating contemporary gene flow. Studies conducted so far indicate that ongoing rates of pollen-mediated gene movement into fragmented populations can be extensive in spite of the expectation that physically isolated populations should experience reduced gene flow. Nevertheless, more studies are needed that compare populations under different landscape conditions (e.g. fragmented versus continuous, and isolated versus nonisolated, populations), and that do so for a wider variety of pollen and seed vectors, before we can understand the role of landscape change on the genetic structure of future populations. Such studies will require the comparison of multiple populations (and often, to this end, the sampling of progeny from multiple source-plants per population).

Box 3. Outcrossing rate: a special case of contemporary gene movement

The outcrossing rate, defined as the proportion of outcrossed progeny produced by an individual mother or population[17], is the most localized form of within-population pollen gene movement. This aspect of the mating system has been evaluated for hundreds of plant populations, but few studies have used multilocus outcrossing-rate estimates to address how species respond to landscape factors. Because estimating outcrossing rate requires fewer loci and fewer progeny per mother than paternity analysis, it is feasible to sample a number of sites, as required by a landscape approach. For many species, an increase in selfing will be accompanied by inbreeding depression and, possibly, by deleterious demographic consequences for the population. Therefore, the outcrossing rate is a critical indicator of the effects of habitat change on population fitness, particularly for historically outcrossing species. Empirical studies have shown that the landscape context, such as habitat disturbance, of a population can influence its outcrossing rate. In a study of a tropical Dipterocarp tree (*Shorea megistophylla*) in Sri Lanka, Murawski *et al.*[39] observed lower outcrossing rates in disturbed than in undisturbed populations. For a Costa Rican tree species (*Pithellobium elegans*, Fabaceae), Hall *et al.*[40] found that fragmentation reduces the outcrossing rate. Given the feasibility of measuring outcrossing rates for several populations, and the potential consequences of changes in levels of inbreeding to population fitness, we anticipate that future studies will pay more attention to landscape influences on outcrossing rate.

The relatively large progeny sample sizes needed for accurately estimating gene movement within and among individual populations, however, might limit the number of populations that can be analysed in a single study. Given this constraint, careful consideration should be made to choose a paternity- (or maternity-) analysis method, experimental design and genetic marker system that strikes a balance between population coverage and individual population accuracy for a particular study system.

Future directions

Several research areas are likely to take advantage of landscape approaches to study ongoing gene movement. First, landscape studies that couple studies of pollinator behavior with ongoing gene movement are needed. Ecologists have been aware for some time that landscape change can affect movement of pollen via pollinators. Reductions of pollinator species are postulated to accompany habitat fragmentation, and these reductions might negatively affect not only seed production, but also outcrossing rates[30]. Changes in species composition of pollinators might also accompany habitat fragmentation, as suggested by the decline in relative pollination success of species pollinated by native insects in tropical forest fragments studied by Aizen and Feinsinger[31]. Even if habitat fragmentation does not change the abundance of pollinator species, it could alter pollinator foraging behavior and rates of pollen movement between populations. Foraging behavior of many pollinators is known to respond to plant density and other features of the landscape. However, despite increased attentiveness to the consequences of landscape disturbance in general, few studies have documented the extent to which changes in pollinator diversity, abundance or movement patterns impact outcrossing rates, much less gene flow. A similar argument could be made for studying the effects of landscape change on inter-relationships between frugivorous animals and seed-mediated gene flow.

A second area where a landscape approach to the study of ongoing gene flow will be useful is metapopulation biology. Making a link between metapopulation ecology and genetics is important because both disciplines are concerned with rates of gene exchange and colonization. Although many metapopulation genetic studies have

focused on indirect measures of gene flow, direct measures of ongoing gene flow will be more appropriate for systems where metapopulation dynamics have been recently induced by habitat alteration. Gene flow information from paternally and maternally inherited genetic markers can measure separate contributions of pollen and seed movement to genetic structure[32]. Such information can provide a basis for distinguishing between the genetic consequences of colonization by seed versus pollen-mediated gene exchange among established populations.

In conclusion, recent attention to landscape context has provided valuable insights to the study of historical gene flow. For those interested in contemporary gene flow, the incorporation of landscape information will probably be essential to understand whether populations are sources or sinks for pollen and seed movement. Regardless of whether landscape heterogeneity is natural or created by recent anthropogenic disturbance, a critical question that researchers can now address is the extent to which the landscape context of populations influences gene movement.

Acknowledgements

This work resulted from a workshop on gene flow conducted in January 1998 at the National Center for Ecological Analysis and Synthesis at the University of California-Santa Barbara, a Center funded by NSF grant DEB-92-21535, the University of California at Santa Barbara, and the State of California (http://www.nceas.ucsb.edu/papers/geneflow). VLS was supported by a sabbatical fellowship at NCEAS and funding from the Missouri Dept of Conservation. We are grateful for the input of the other workshop participants: W.T. Adams, F. Davis, R. Dyer, M. Gilpin, J. Hamrick, J. Neigel, R. Petit, O. Savolainen, P. Smouse and E. Steinberg. We thank the following individuals for comments: N. Ellstrand, J. Thompson (WSU), V. Apsit, K. Bailey, R. Dyer, W. Gram and J. Wall.

References

1 Wright, S. (1969) *Evolution and the Genetics of Populations, Vol. 2. The Theory of Gene Frequencies,* University of Chicago Press

2 Neigel, J.E. (1997) **A comparison of alternative strategies for estimating gene flow from genetic markers,** *Annu. Rev. Ecol. Syst.* 28, 105–128

3 Hanski, I. and Simberloff, D. (1997) **The metapopulation approach, its history, conceptual domain, and application to conservation,** in *Metapopulation Biology: Ecology, Genetics, and Evolution* (Hanski, I.A. and Gilpin, M.E., eds), pp. 5–26, Academic Press

4 McCauley, D.E. (1995) **Effects of population dynamics on genetics in mosaic landscapes,** in *Mosaic Landscapes and Ecological Processes* (Hansson, L., Fahrig, L. and Merriam, G., eds), pp. 179–198, Chapman & Hall

5 Chase, M., Kesseli, R. and Bawa, K. (1996) **Microsatellite markers for population and conservation genetics of tropical trees,** *Am. J. Bot.* 83, 51–57

6 Adams, W.T. and Birkes, D.S. (1991) **Estimating mating patterns in forest tree populations,** in *Biochemical Markers in the Population Genetics of Forest Trees* (Fineschi, S., Malvoti, E.M., Cannata, F. and Hattemer, H.H., eds), pp. 152–172, SPB Academic Publishing

7 Devlin, B. and Ellstrand, N.C. (1990) **The development and application of a refined method for estimating gene flow from angiosperm paternity analysis,** *Evolution* 44, 248–259

8 Kimura, M. and Weiss, G.H. (1964) **The stepping stone model of population structure and the decrease of genetic correlation with distance,** *Genetics* 49, 561–576

9 Slatkin, M. (1993) **Isolation by distance in equilibrium and nonequilibrium populations,** *Evolution* 47, 264–279

10 Kudoh, H. and Whigham, D.F. (1997) **Microgeographic genetic structure and gene flow in *Hibiscus moscheutos* (Malvaceae) populations,** *Am. J. Bot.* 84, 1285–1293

11 Bossart, J.L. and Prowell, D.P. (1998) **Genetic estimates of population structure and gene flow: limitations, lessons, and new directions,** *Trends Ecol. Evol.* 13, 202–206

12 Husband, B.C. and Barrett, S.C.H. (1996) **A metapopulation perspective in plant population biology,** *J. Ecol.* 84, 461–469

13 Ruckelshaus, M.H. (1998) **Spatial scale of genetic structure and an indirect estimate of gene flow in eelgrass, *Zostera marina*,** *Evolution* 52, 330–343

14 McCauley, D.E. (1997) **The relative contributions of seed and pollen movement to the local genetic structure of Silene alba,** *J. Hered.* 88, 257–263

15 Slatkin, M. (1991) **Inbreeding coefficients and coalescence times,** *Genet. Res.* 58, 167–175

16 Steinberg, E.K. and Jordan, C.E. (1997) **Using molecular genetics to learn about the ecology of threatened species: the allure and the illusion of measuring genetic structure in natural populations,** in *Conversation Biology* (Fiedler, P.L. and Kareiva, P.M., eds), pp. 440–460, Chapman & Hall

17 Ritland, K. and Jain, S. (1981) **A model for the estimation of outcrossing rate and gene frequencies using n independent loci,** *Heredity* 47, 35–52

18 Nason, J.D., Herre, E.A. and Hamrick, J.L. (1998) **The breeding structure of a tropical keystone plant resource,** *Nature* 391, 685–687

19 Meagher, T.R. (1986) **Analysis of paternity within a natural population of *Chamaelirium luteum*. 1. Identification of most-likely male parents,** *Am. Nat.* 128, 199–215

20 Roeder, K., Devlin, B. and Lindsay, B.G. (1989) **Application of maximum likelihood methods to population genetic data for the estimation of individual fertilities,** *Biometrics* 45, 363–379

21 Devlin, B. and Ellstrand, N.C. (1990) **Male and female fertility variation in wild radish, a hermaphrodite,** *Am. Nat.* 136, 87–107

22 Crawford, T.J. (1984) **The estimation of neighborhood parameters for plant populations,** *Heredity* 52, 273–283

23 Schnabel, A., Nason, J.D. and Hamrick, J.L. (1998) **Understanding the population genetic structure of *Gleditsia triacanthos* L.: seed dispersal and variation in female reproductive success,** *Mol. Ecol.* 7, 819–832

24 Goodell, K. *et al.* (1997) **Gene flow among small populations of a self-incompatible plant: an interaction between demography and genetics,** *Am. J. Bot.* 84, 1362–1371

25 Nason, J.D. and Hamrick, J.L. (1997) **Reproductive and genetic consequences of forest fragmentation: two case studies of Neotropical canopy trees,** *J. Hered.* 88, 264–276

26 Dow, B.D. and Ashley, M.V. (1996) **Microsatellite analysis of seed dispersal and parentage of saplings in bur oak, *Quercus macrocarpa*,** *Mol. Ecol.* 5, 615–627

27 Streiff, R. *et al.* (1998) **Within population genetic structure in *Quercus robur* L. and *Quercus petraea* (Matt.) Liebl. assessed with isozymes and microsatellites,** *Mol. Ecol.* 7, 317–328

28 Chase, M.R. *et al.* (1996) **Distant gene flow in tropical trees,** *Nature* 383, 398–399

29 Aldrich, P.R. *et al.* (1998) **Microsatellite analysis of demographic genetic structure in fragmented populations of the tropical tree *Symphonia globulifera*,** *Mol. Ecol.* 7, 933–944

30 Kearns, C.A., Inouye, D.W. and Waser, N.M. (1998) **Endangered mutualism: the conservation of plant–pollinator interactions,** *Annu. Rev. Ecol. Syst.* 29, 83–112

31 Aizen, M.A. and Feinsinger, P. (1994) **Forest fragmentation, pollination, and plant reproduction in a Chaco dry forest, Argentina,** *Ecology* 75, 330–351

32 Ennos, R.A. (1994) **Estimating the relative rates of pollen and seed migration among plant populations,** *Heredity* 72, 250–259

33 Smouse, P.E., Long, J.C. and Sokal, R.R. (1986) **Multiple regression and correlation extensions of the Mantel test of matrix correspondence,** *Syst. Zool.* 35, 627–632

34 El Mousodik, A. and Petit, R.J. (1996) **Chloroplast DNA phylogeography of the argan tree of Morocco,** *Mol. Ecol.* 5, 547–555

35 Latta, R.G. *et al.* (1998) **Direct and indirect estimates of seed versus pollen movement within a population of ponderosa pine,** *Evolution* 52, 61–67

36 Devlin, B., Roeder, K. and Ellstrand, N.C. (1988) **Fractional paternity assignment: theoretical development and comparison to other methods,** *Theor. Appl. Genet.* 76, 369–380
37 Smouse, P.E. and Meagher, T.R. (1994) **Genetic analysis of male reproductive contributions in *Chamaelirium luteum* (L.) gray (Liliaceae),** *Genetics* 136, 313–322
38 Kaufman, S.R., Smouse, P.E. and Alvarez-Buylla, E.R. (1998) **Pollen-mediated gene flow and differential male reproductive success in a tropical pioneer tree, *Cecropia obtusifolia* Bertol. (Moraceae),** *Heredity* 81, 164–173
39 Murawski, D.A., Nimal-Gunatilleke, I.A.U. and Bawa, K.S. (1994) **The effects of selective logging on inbreeding in *Shorea megistophylla* (Dipterocarpaceae) from Sri Lanka,** *Conserv. Biol.* 8, 997–1002
40 Hall, P., Walker, S. and Bawa, K. (1996) **Effect of forest fragmentation on genetic diversity and mating system in a tropical tree, *Pithellobium elegans*,** *Conserv. Biol.* 10, 757–768

Nematode community structure as a bioindicator in environmental monitoring

Tom Bongers and Howard Ferris

Four of every five multicellular animals on the planet are nematodes. They occupy any niche that provides an available source of organic carbon in marine, freshwater and terrestrial environments. Nematodes vary in sensitivity to pollutants and environmental disturbance. Recent development of indices that integrate the responses of different taxa and trophic groups to perturbation provides a powerful basis for analysis of faunal assemblages in soil as *in situ* environmental assessment systems.

Tom Bongers is at the Laboratory of Nematology, Wageningen University, PB 8123 NL-6700 ES, Wageningen, The Netherlands (tom.bongers@medew.nema.wau.nl); Howard Ferris is at the Dept of Nematology, University of California, Davis, CA 95616, USA (hferris@ucdavis.edu).

Particles vary in size, nature and aggregation throughout the horizontal and vertical dimensions of the soil. Variability in the solid state affects volume, exchange, flow and diurnal and seasonal fluctuations of the liquid and gas phases of soil. The spatial and temporal heterogeneity provides myriad habitats for a vast diversity of organisms. Larger organisms live in natural channels in the soil or create tunnels and chambers. Smaller organisms, including nematodes, are primarily aquatic and live in the water films between soil particles. Their size allows movement through pore necks between particles or between aggregates of particles without tunneling activity. The ecological functions of soil organisms include organic matter decomposition, mineralization of nutrients, degradation of toxicants and population regulation of plant disease agents.

Soil organisms depend on each other for carbon and energy. The structure and function of below-ground food webs are disrupted by hydrocarbon and heavy-metal contaminants, mineral fertilizers and pesticides, and by physical disturbance. However, the results of such disruptions are unpredictable because they are influenced by the heterogeneity of the soil, fluctuations in abiotic conditions, chemical and physical buffering capacity, and by other biotic and abiotic interactions. Because nematodes occupy key positions as primary and intermediate consumers in soil food webs, evaluation and interpretation of the abundance and function of their faunal assemblages or community structure offers an *in situ* assessment of disruptive factors. An individual assessment at a site provides a snapshot of current conditions, whereas sequential assessments allow analysis of environmental degradation or remediation.

Nematodes were used in water quality assessment in the 1970s (Refs 1,2) during which the informational value of terrestrial nematode community structure became apparent[3]. During the 1980s, as concerns increased about the functioning and vulnerability of the soil ecosystem, researchers turned to assessment of *in situ* nematode faunae in environmental studies[4,5]. The characteristics of nematodes (Box 1) allow analysis of presence and abundance of individual taxa. Their faunal composition has emerged as a useful monitor of environmental conditions and ecosystem function in the soil[6–8].

Recent research indicates that simple analyses of *in situ* nematode faunae at family level provide a wealth of information on the nature of decomposition pathways and soil nutrient status[7,9] (Box 2). The analyses also indicate effects of agricultural practices and contaminants on the functioning of the soil food web[10–15]. They provide a basis for environmental management, remediation and conservation decisions.

Classification of nematodes by feeding behavior and life-history strategy

Nematodes are an evolutionarily successful group of organisms. They are ubiquitous in all habitats that provide available organic carbon sources; they are the planet's most abundant metazoa. In soil, some nematodes feed on higher plants, some on fungi or bacteria; others are carnivores or omnivores. They range in reproductive potential from explosive opportunists to conservative survivalists. They vary in sensitivity to pollutants and environmental disturbance. Recognition of the approximately 20 000 described species is based primarily on morphological and anatomical features supplemented by ecological function[16], although molecular markers are increasingly important for identification and classification[17–19]. Nematode form and function, especially concerning feeding behavior, are closely linked and consequently the feeding behavior of unknown species can be inferred from their morphology[20,21].

 0169-5347/99/$ – see front matter PII: S0169-5347(98)01583-3