PHYLOGEOGRAPHY OF SOUTHERN EUROPEAN REFUGIA

Phylogeography of Southern European Refugia

Evolutionary perspectives on the origins
and conservation of European biodiversity

Edited by

Steven Weiss
Karl-Franzens Universität Graz, Austria

and

Nuno Ferrand
CIBIO, University of Porto, Portugal

Springer

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN-10 1-4020-4903-X (HB)
ISBN-13 978-1-4020-4903-3 (HB)
ISBN-10 1-4020-4904-8 (e-book)
ISBN-13 978-1-4020-4904-0 (e-book)

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

www.springer.com

Graphical assistance was provided by Gabi Kriks, and the layout was done by Gabi Kriks and Estella Weiss-Krejci. Financial support was provided by CIBIO, Portugal and Karl-Franzens University of Graz, Austria

Cover photos:
From left to right, photos 1-4 (Quercus leaf, Green frog, Wall lizard and Viper) are from Jorge Coutinho, and photo 5 (European rabbit) is from Paulo Magalhães

Printed on acid-free paper

Contents

Part I **Historical foundations and perspectives**

Part II **Broad phylogeographical studies**

Part III **Emerging phylogeographical patterns in a southern European refugium: the Iberian Peninsula**

Part IV **First contributions of islands and other peninsulas to the phylogeography of southern Europe**

Part V **Conclusions**

Contributors

Agúndez, Dolores Department of Forest Systems and Resources, CIFOR-INIA, Carretera de la Coruña km 7.5, 28040 Madrid, Spain, agundez@inia.es

Alexandrino, João Departamento de Zoologia, Instituto de Biociências, Universidade Estadual Paulista, 13506-900 Rio Claro, São Paulo, Brazil, jalex@rc.unesp.br

Alía, Ricardo Department of Forest Systems and Resources, CIFOR-INIA, Carretera de la Coruña km 7.5, 28040 Madrid, Spain, alia@inia.es

Arntzen, Jan W National Museum of Natural History, P. O. Box 9517, 2300 RA Leiden, The Netherlands, arntzen@nnm.nl

Atkinson, Rachel J National Parks and Conservation Service, Ministry of Agro-Industry and Fisheries, Réduit, Mauritius

Avise, John C Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697, USA, javise@uci.edu

Berrebi, Patrick Ecosystemes Lagunaires, UMR 5119, Equipe EFEP, Université Montpellier 2 Place E. Bataillon, 34095 Montpellier Cedex, France, berrebi@univ-montp2.fr

Branco Madalena CIBIO, Centro de Investigação e Biodiversidade e Recursos Genéticos, ICETA, Campus Agrário de Vairão, Rua do Monte, Crasto, Vairão, 4480 Vila do Conde, Portugal, msbranco@fc.up.pt

Campos, Rita CIBIO, Centro de Investigação e Biodiversidade e Recursos Genéticos, ICETA, Campus Agrário de Vairão, Rua do Monte, Crasto, Vairão, 4480 Vila do Conde, Portugal, ritacampos@fc.up.pt

Carrión, José S Department of Plant Biology, Faculty of Biology, University of Murcia, 30100 Murcia, Spain, carrion@um.es

Casula, Paolo Dipartimento di Biologia Sperimentale, Sezione di Genetica, Università di Cagliari, Viale Poetto 1, 09126 Cagliari, Italy

Crnobrnja-Isailovic, Jelka Department of Evolutionary Biology, Institute for Biological Research, Belgrade 11060, Serbia, jelka@ibiss.bg.ac.yu

Ferrand, Nuno CIBIO, Centro de Investigação e Biodiversidade e Recursos Genéticos, ICETA, Campus Agrário de Vairão, Rua do Monte, Crasto, Vairão, 4480 Vila do Conde, Portugal, nferrand@mail.icav.up.pt

Gil, Luis Unit of Anatomy, Physiology and Genetics, ETSIM, Ciudad Universitaria s/n., 28040 Madrid, Spain

Gómez, Africa Department of Biological Sciences, University of Hull, Hull HU6 7RX, UK, a.gomez@hull.ac.uk

Gómez, Aranzazu Biosafety Unit, IMIDRA, 'El Encín' A-2 km 38.200, 28800 Alcalá de Henares (Madrid), Spain

González-Martínez, Santiago C. Department of Forest Systems and Resources, CIFOR-INIA, Carretera de la Coruña km 7.5, 28040 Madrid, Spain, santiago@inia.es

Grill, Andrea Institute for Biodiversity and Ecosystem Dynamics, Faculty of Science, University of Amsterdam, P. O. Box 94766, 1090 GT Amsterdam, The Netherlands, grill@science.uva.nl

Kotlík, Petr Department of Vertebrate Evolutionary Biology and Genetics, Institute of Animal Physiology and Genetics, Academy of Sciences of the Czech Republic, 277 21 Libichov, Czech Republic, kotlik@iapg.cas.cz

Lecis, Roberta Institut de Recerca i Tecnologia Agroalimentàries (IRTA), Genètica de la Conservació, Centre de Cabrils, Carretera de Cabrils s/n, 08348 Cabrils (Barcelona), Spain, roberta.lecis@irta.es

Lunt, David H Department of Biological Sciences, University of Hull, Hull HU6 7RX, UK, d.h.lunt@hull.ac.uk

Menken, Steph Institute for Biodiversity and Ecosystem Dynamics, Faculty of Science, University of Amsterdam, P. O. Box 94766, 1090 GT Amsterdam, The Netherlands, menken@science.uva.nl

Petit, Rémy J Directeur de Recherche, INRA, UMR Biodiversité, Gènes & Ecosystèmes, Equipe de Génétique, 69 Route d'Arcachon, 33612 Cestas Cedex, France, remy@pierroton.inra.fr

Randi, Ettore Istituto Nazionale per la Fauna Selvatica (INFS), Via cá Fornacetta 9, 40064 Ozzano dell' Emilia (BO), Italy, met0217@iperbole.bo.it

Rokas, Antonis The Broad Institute of MIT & Harvard, 320 Charles Street, Cambridge, MA 02141, USA, arokas@mit.edu

Stone, Graham N Institute of Evolutionary Biology, Edinburgh University, The Kings Buildings, West Mains Road, Edinburgh EH9 3JT, UK, graham.stone@ed.ac.uk

Teixeira, José CIBIO, Centro de Investigação e Biodiversidade e Recursos Genéticos, ICETA, Campus Agrário de Vairão, Rua do Monte, Crasto, Vairão, 4480 Vila do Conde, Portugal, jteixeira@mail.icav.up.pt

Vargas, Pablo Real Jardín Botánico de Madrid (CSIC), Plaza de Murillo 2, 28014 Madrid, Spain, vargas@ma-rjb.csic.es

Vendramin, Giovanni G Consiglio Nazionale delle Ricerche, Istituto di Genetica Vegetale, Via Madonna del Piano 10, 50019 Sesto Fiorentino, Italy, giovanni.vendramin@igv.cnr.it

Weiss, Steven Institute of Zoology, Karl-Franzens University of Graz, 8010 Graz, Austria, steven.weiss@uni-graz.at

Introduction

Nuno Ferrand* and Steven Weiss**
E-mail: *nferrand@mail.icav.up.pt; **steven.weiss@uni-graz.at

This volume contains 14 contributions, many of which have been developed from presentations made at the first international symposium devoted exclusively to phylogeography, held in Vairão, Portugal, in March 2002, at what has now become the Research Center in Biodiversity and Genetic Resources (CIBIO). Approximately 150 participants from over 20 countries attended, all sharing their enthusiasm for the growing number of phylogeographic oriented studies in Europe with a particular emphasis on the circum-Mediterranean region, nominated as one of the world's 25 biodiversity hotspots. The conservation relevance of the Mediterranean basin is exemplified by the fact that it holds approximately 20% of the world's floristic diversity (see Chapter 10).

The symposium witnessed over 100 oral and poster communications including the plenary talks of John Avise (The history and development of phylogeography); Ettore Randi (Mammalian phylogeography of South Europe); Rémy Petit (Phylogeography of temperate trees and shrubs in Europe and the importance of southern refugia); and Pierre Taberlet (Comparative phylogeography of Alpine plants). Taxonomic coverage of the presentations was broad including mammals (26), plants (20), invertebrates (17), amphibians and reptiles (17), fish (15), and birds (8).

Section one opens with a chapter by John Avise, outlining 25 evolutionary insights that have arisen from what he coins a phylogeographic revolution. Key to this perspective is the notion that the multidisciplinary field of phylogeography is serving as an epistemological bridge between the formerly distinct fields of population genetics and phylogenetics. More succinctly, phylogeography has reoriented and extended the field of population genetics incorporating the gene-tree perspective into assessment of intraspecific variation and historical demography. Many of these insights stem from his early work screening variation in the mtDNA molecule and its unique properties that have served as the working horse for the development of this nascent discipline. The next chapter, authored by another invited guest Rémy Petit

S. Weiss and N. Ferrand (eds.), Phylogeography of Southern European Refugia, 1–4.

and his colleague Giovanni Vendramin is a studious and thoroughly comprehensive review of the history and development of plant organelle genetics (both mtDNA and cpDNA) and their application in population studies. Insights are provided on the structure, levels of variation, modes of inheritance, vegetative segregation, and well documented recombination of organelle genomes in plants with particular emphasis on the unique opportunities that these characteristics provide evolutionary research.

Section two contains two chapters characterized by their broad-scale coverage. Chapter 3, also authored by an invited guest, Ettore Randi, reviews the phylogeography of mammals in southern Europe in the context of our knowledge on the paleoecological conditions that have shaped current patterns of lineage distribution. Emphasis is given on the taxon-specific patterns and growing appreciation for the complexity of glacial refugia beyond the simple three Peninsula model, while highlighting the implications for taxonomy and conservation. And finally, Chapter 4 presents a comparative study assessing species richness and genetic diversity in a coevolutionary system of oaks and their obligate parasitic gallwasps. The study is additionally unique in that it assesses and compares diversity along a longitudinal rather than latitudinal axis crossing four major glacial refugia, extending over from Iberia to Asia Minor. While requiring more organisms to test, the authors postulate that the lower species richness of the Iberian Peninsula compared to other major refugia is a result of stronger demographic fluctuations stemming from a more arid climate.

The next section (Chapters 5-9) reports on a series of review perspectives and case studies on the Iberian Peninsula, the best studied refugial region in Europe. The section begins with a comparative phylogeographic review of the Iberian Peninsula outlining not only the concordant patterns but moreover an emerging and important concept of ‘refugia-within-refugia’. The message is clear and impressive in its taxonomic coverage as multiple refugia are verified for species or species complexes in plants (8), mammals (6), reptiles and amphibians (6), fish (7), and invertebrates (5). This perspective underlines the emerging research focus on complexity and dynamics of contact zones, hybridization, introgression and population diversification of Iberian biota but also warns against making potentially misleading conclusions concerning so-called northern (postglacial) refugia, before we understand the cryptic and underappreciated lineage diversity existing in southern Europe. Chapter 6 is noted for its integration of ecological, phenotypic and phylogeographic data in characterizing the historical biogeography of an endemic Iberian salamander, *Chioglossa lusitanica*, a system that is demonstrating its amenability to the current development of landscape genetics and GIS-based approaches to evolutionary and conservation oriented research. The study provides a framework for comparative phylogeographic research that can be used to designate key areas for multi-species conservation. Chapter 7 unveils a data set on protein polymorphisms (20 loci) collected over a 10-year period that provides the basis of a long-term comprehensive research program on the European rabbit, *Oryctolagus cuniculus*. These data serve as a multi-locus baseline for the study of introgression, hybridization, and selection within the explicitly defined phylogeographic context of two divergent lineages and a temporally dynamic contact zone, existing within a Peninsula refuge. Chapter 8, reports on hemoglobin polymorphisms (HBA and HBB) across the contact zone of the European rabbit in Iberia. The starkly contrasting spatial patterns of the HBA six allele system suggest that strong selective forces are operative over a large spatial scale and further reveal the presence of a hybrid allele, two observations that underlie the value of studying contact zones in a refuge that has persisted throughout the Pleistocene. Chapter 9 presents a case study of a commercially important species, the maritime pine *Pinus pinaster*, investigating spatial structure and the effect of specific landscape features on gene flow using coalescent theory. Inferences are drawn on the role of mountain ranges within refugia in serving to both allow altitudinal migration and to isolate specific populations. The biological inferences also provide the basis for specific conservation and management recommendations.

Section four (Chapters 10-13) includes studies that survey organisms or review phylogeographic patterns in non-Iberian refugia. Chapter 10 is the first multi-taxa review of the remarkable endemic floral and faunal diversity of Sardinia. The dating of the fundamental biogeographical phenomena associated with the island's formation, compared with genetic divergences suggests that the present diversity has arisen subsequent to the marine transgressions five million years ago. Characterization of the state of the island's present system of nature preserves, emphasizes the threats to these unique biota. Among the endemics covered are 14 species of butterflies, nine plants, six cave beetles, four salamanders, two lizards, two frogs, and one mammal. Chapter 11 takes an interesting phylogenetic approach to characterizing the rich floral diversity of the five volcanic archipelagos (Azores, Madeira, Selvagens, Cape Verde and Canaries) that comprise Macaronesia. Boasting some 3100 plant species, the role of Macaronesia as a lineage refuge is assessed by applying a comparative phylogenetic analysis. The analysis distinguishes various classes of relictualism and further proposes that endozoochory has played a major role in promoting multiple colonizations of Macaronesia, aiding its status as a floral refuge. Chapter 12 employs a nested clade phylogeographic analysis (NCPA) on an mtDNA data set of a common cyprinid fish, *Barbus barbus*. A number of the historical biogeographic inferences drawn *ad hoc* from a previous study were supported by the NCPA, but several others were not, exemplifying the conservative, speculation hindering tendency of the analysis. Chapter 13 reviews the diversity, levels of endemism and available genetic data on reptiles and amphibians of the Balkan Peninsula, clearly the least studied of Europe’s three major refugial Peninsulas. Emphasis is placed on the region’s historical complexity and how it serves as a mountainous crossroad for range expansions and contractions during postglacial and glacial episodes.

The book closes with an up-to-date review on the current perspectives, caveats and prospects for phylogeography as a discipline, with particular emphasis on its importance in both understanding and conserving European biodiversity. Current controversies and criticisms concerning phylogeographic data analysis and inference are discussed with an optimistic view of new methodologies in development, recognizing that the stochastic variance of the coalescent process must be more seriously taken into consideration. As in any scientific endeavor, theory and methodology will only become broadly accepted through the rigors of repeated observation and hypothesis testing within legitimate statistical frameworks. Nonetheless, as phylogeography deals with the uncertainty of history combined with the tremendous complexity of evolutionary pathways, a plea is also made for maintaining an open-minded and pluralistic approach to study design and data analysis.

Part I

Historical foundations and perspectives

Chapter 1

Twenty-five key evolutionary insights from the phylogeographic revolution in population genetics

John C. Avise
E-mail: javise@uci.edu

Abstract

An overview is provided of 25 novel perspectives that the field of phylogeography has brought to scientific studies of population genetics and speciation. A unifying theme is that microevolution can be described as an extended genealogical process played out in space and time, and reflecting the oft-idiosyncratic biological and environmental factors that have impinged on historical population demography. Most of the empirical and conceptual methods of phylogeography depart considerably from conventional equilibrium approaches, and they are helping to reorient and extend traditional population genetics in realistic directions that emphasize historical demography and genealogy.

Keywords: phylogeography, genealogy, gene trees, demography, speciation

Introduction

Phylogeography is a relatively young discipline concerned with the principles and processes governing the geographic distributions of gene lineages, especially within and among closely related species (Figure 1). The phylogeographic revolution, inspired by mitochondrial (mt) DNA analyses that were introduced nearly three decades ago, has transformed the study of population genetics and speciation in several ways. In particular, this ongoing reformation has drawn closer empirical and conceptual connections between microevolutionary genetics and phylogenetic biology.

Here I substantiate these claims by compiling more than two dozen salient insights about microevolution that seldom (sometimes never) were an explicit part of the fabric of population genetics in the pre-phylogeography era. The entries in this list appear in a sequence generally consistent with the underlying train of logic, rather than necessarily in order of importance or date of development. Many concepts in the list are nested or partially overlapping, yet each qualifies for inclusion by virtue of having been quite unorthodox when

S. Weiss and N. Ferrand (eds.), Phylogeography of Southern European Refugia, 7–21.

introduced. More importantly, most of these insights from the field of phylogeography will (I suspect) prove to be enduring truths about microevolutionary processes in nature.

This list provides a brief overview and historical backdrop, and is intended to encapsulate phylogeography's principal contributions. Thus, for each of the 25 entries, I have cited just one key reference that either was seminal in the history of ideas or is particularly informative as a more recent review. For much fuller treatments of all topics considered, readers should consult Avise (2000) and the extensive primary literature that it summarizes.

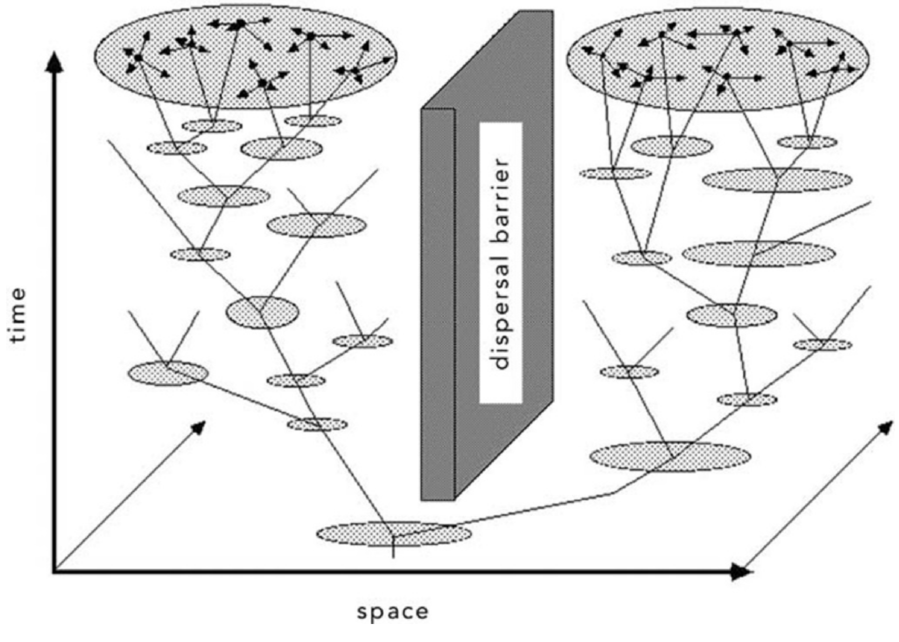

Figure 1. The axes of phylogeography are space and time, across which gene genealogies are scrutinized (modified from Avise 2000).

Twenty-five primary revelations

1. Cytoplasmic genomes add a new hierarchical level to population genetics

In diploid organisms, nuclear genomes typically exist as two copies per somatic cell and are transmitted across generations in a single-copy (haploid) molecular fashion. By contrast, mitochondrial and other cytoplasmic genomes [such as chloroplast (cp) DNA (see also Petit & Vendramin, this volume)] exist as populations of molecules within each somatic and germline cell, and usually are transmitted from parent to offspring as multiple copies. This realization led to a novel research arena directed toward the intra-individual population dynamics of cytoplasmic genomes in somatic cells and germlines – a newly recognized level in the population genetic hierarchy (Birky *et al.* 1983).

2. Germline bottlenecks in mtDNA numbers attend mitochondrial inheritance

Notwithstanding concept #1 (above), the vast majority of mtDNA variation is apportioned among (not within) individual animals, even in local populations. Thus, an individual typically displays a single predominant cytoplasmic genotype that often differs clearly in DNA sequence from other conspecifics. Coupled with experimental findings from pedigree analyses (Hauswirth & Laipis 1982), this observation indicates that relatively small effective population sizes often characterize the intracellular pool of mtDNA molecules that transmits from one generation to the next through animal germlines. Thus, significant heteroplasmy (the joint appearance of two or more mtDNA genotypes within an individual) is normally a transient condition lasting only a small number (i.e. tens or perhaps hundreds) of organismal generations. This discovery carries a huge pragmatic benefit: within-individual sequence heterogeneity seldom seriously compromises mtDNA's utility for genealogical assessments at proximately higher levels in the biological hierarchy (e.g. local demes and geographic populations).

3. DNA repair mechanisms can influence molecular evolutionary rates

A traditional paradigm of molecular evolution is that genes with conserved function evolve slowly. The mitochondrial genome, with its central role in cellular energy metabolism, was thought to be a paradigm if not the epitome of functional conservatism. Thus, early reports that animal mtDNA evolves rapidly (about 5-10x faster than single-copy nuclear DNA) came as a great surprise (Brown *et al.* 1979). Subsequent studies showed that mtDNA sequence evolution is concentrated at synonymous sites and non-coding regions of the molecule, as might be expected. However, a totally unexpected factor contributing to mtDNA's rapid evolution was also intimated and later confirmed: a severe deficiency of DNA repair mechanisms within the mitochondrion.

4. Some DNA sequences in sexual species show asexual inheritance

This notion was not entirely novel because mammalian Y-chromosomes were long known to be paternally inherited, and cytoplasmic genomes such as mtDNA were assumed to be transmitted maternally. Nonetheless, explicit analyses of cytoplasmic gene sequences and their transmission genetics through organismal pedigrees soon confirmed mtDNA's maternal inheritance at the molecular level. This validation compelled the field of population genetics to reconsider the ramifications of asexual inheritance in sexually reproducing organisms. In particular, it prompted the field to more fully embrace the notion that asexual genomes exist and are transmitted across successive generations without the normal complications of intermolecular recombination otherwise attendant with sex (Hutchinson *et al.* 1974).

5. Matrilineal histories within species can be recovered

The rapid pace of mtDNA sequence evolution in animals, coupled with the molecule's maternal inheritance, meant that species display a wealth of non-recombining markers suitable for deciphering the matriarchal component of an extended organismal pedigree (Wilson *et al.* 1985). Population geneticists thereby were afforded unprecedented access to genealogical information at the intraspecific level.

6. A gene tree is a recognizable component of a population pedigree

Considerations of matrilineal ancestry pioneered the gene-tree concept. Animal mtDNA consists of about 37 functional genes, but the entire non-recombining mitochondrial genome can be considered a single locus from a genealogical perspective. Like the traditional phylogenies of higher-level systematics, an intraspecific gene tree for mtDNA is hierarchically branched and non-reticulate. The gene-tree notion can also be extended to particular sequences in the nuclear genome, at least in principle (Tajima 1983). Thus, the term gene tree gained a generalized definition: the genealogical history of any defined segment of DNA. Such gene trees are real and discrete components of population pedigrees (Figure 2).

7. Gene trees in sexual species are multitudinous and non-isomorphic

For sexually reproducing organisms, a matrilineal gene tree represents only a minuscule fraction of a species' hereditary history. More than 99% of that total history resides instead in nuclear genes whose alleles have been transmitted along multi-generation genealogical pathways involving both genders. Due to Mendelian segregation and independent assortment, the realized transmission histories of unlinked DNA sequences inevitably differ from locus to locus. Thus, gene trees for unlinked loci are highly unlikely to be strictly isomorphic (identical in branching structure). This insight led to the notion that any pictured cladogram summarizing historical relationships of populations or species is actually a much-simplified representation of an

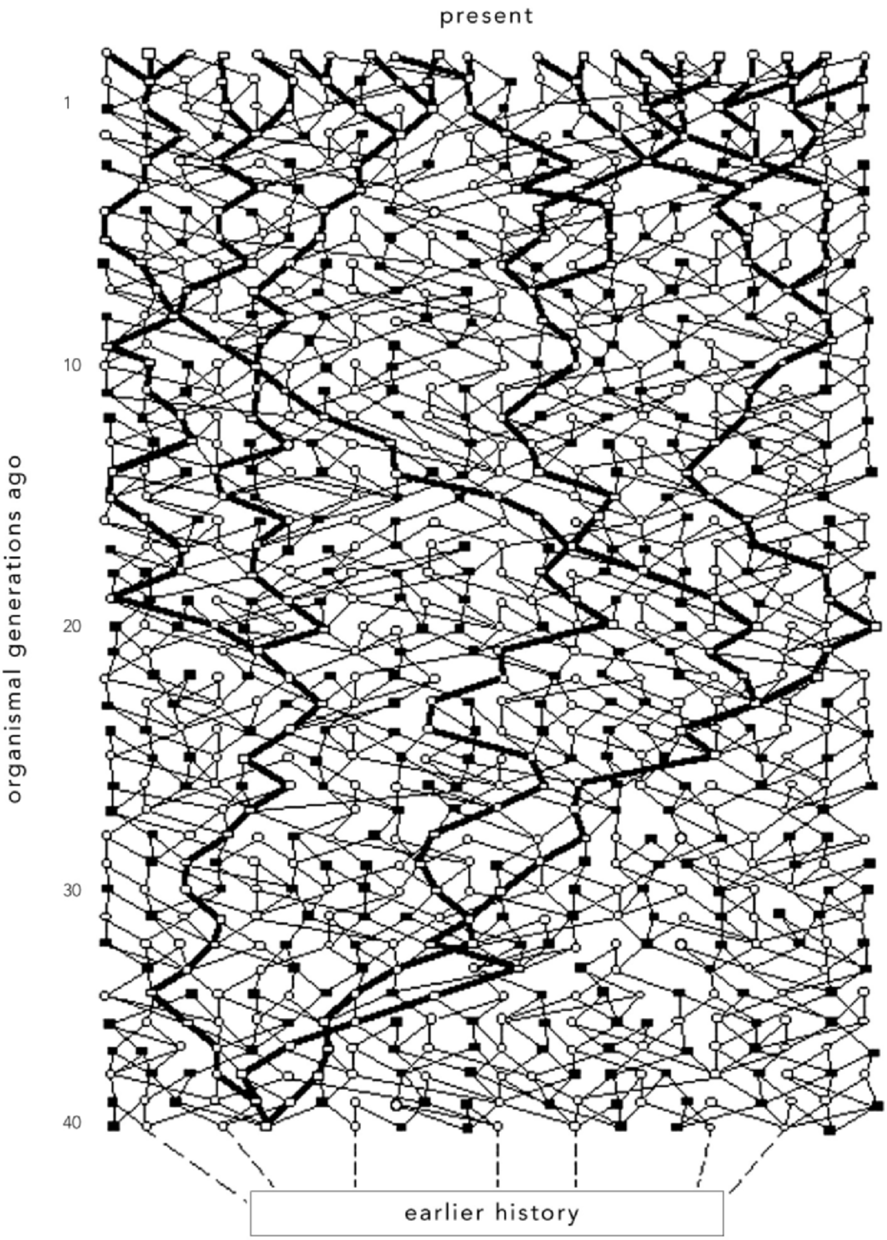

***Figure 2**. Highlighted in heavy lines is one (among multitudinous) gene trees within an organismal pedigree across 40 generations. Note how the lineages coalesce to shared ancestors (modified from Avise 2000).*

underlying statistical ‘cloudogram’ of quasi-independent gene trees with a variance (Maddison 1997).

8. Phylogenetic reasoning is quite relevant at the intraspecific level

Traditional wisdom in systematics had been that phylogenetic principles apply only above the taxonomic rank of species, i.e. that they have no meaning in the supposedly reticulate realm of intraspecific evolution. The gene-tree concept, prompted by studies of mtDNA, challenged this dogma by clarifying the principle that particular DNA sequences do in fact have genealogical (phylogenetic) histories of transmission within a species pedigree also (Avise 1989a). Furthermore, gene trees (or at the least, unrooted genealogical networks) can often be empirically recovered when the complications of sex-mediated recombination are absent or minimal. So, historical gene genealogies can be estimated from suitable molecular data using phylogenetic algorithms, and such historical representations aptly lend themselves to description by traditional macro-phylogenetic concepts such as clades, outgroups, and synapomorphic (shared-derived) characters (provided that these terms are now interpreted to apply explicitly to features of the gene tree *per se*).

9. Individuals can be treated as ‘operational taxonomic units’ (OTUs)

The basic data of traditional population genetic analysis consist of allelic or genotypic frequencies in population samples, with the populations themselves often prespecified by criteria such as geography or suspected reproductive relationships. Although such collective empirical data can always be used to estimate genetic relationships among suites or assemblages of individuals, an undesirable element of circular reasoning underlies the exercise, and much useful information on particular specimens is lost. Both the circularity and the information loss are removed entirely when *individuals* are treated as the basic units of analysis in genealogical reconstructions (Avise *et al.* 1979). Since the advent of phylogeography, this ‘individual as OTU’ approach is now included routinely in phylogenetic appraisals of mtDNA (and some nuclear loci).

10. Intraspecific genealogy and historical demography are intertwined

At the intraspecific level, concepts of genealogy and historical population demography are inextricably associated. Precise mating relationships of individuals, coupled with generation-by-generation means and variances in individual reproductive success, describe the extended pedigree of a population, thereby defining the genealogical pathways that were available for allelic transmission. Any gene tree is one realized subset, or historical sample, from this constellation of pathways (Figure 2). As such, each gene tree is an ineluctable reflection of historical population demography. This realization gave rise to 'coalescent theory' (Hudson 1990), a burgeoning discipline in mathematical population genetics that seeks to uncover and formalize the relationships between historical population demography and the structure of intraspecific gene genealogies.

11. Evolutionary effective population sizes of most animal species are relatively small

In most surveyed animal species (especially those that are relatively abundant today, and whose populations are characterized by high historical levels of gene flow), estimated evolutionary effective population sizes (N_e values over the long term) have proved to be orders-of-magnitude smaller than contemporary census numbers (N). This conclusion stems from coalescent theory as applied to the surprisingly shallow intraspecific gene trees for such species, as evidenced in empirical mtDNA data sets (Avise *et al.* 1988). Two explanations are likely, the first probably being of greater importance: a) population-demographic histories *per se*, such as occasional bottlenecks in population size, or large variances among females in reproductive success; and b) rare 'selective sweeps' that purge existing variation as selectively advantageous mutations course through a species to fixation. Either way, gene lineages that survived for current observation have been historically squeezed through many fewer ancestors than otherwise might have been supposed, thereby constraining what would otherwise be greater temporal depths in intraspecific mtDNA gene trees.

12. Cytonuclear associations matter

The joint availability of molecular data from nuclear and cytoplasmic (e.g. mitochondrial or chloroplast) loci prompted an important new research area dealing with 'cytonuclear' patterns. Of special interest is how natural selection and other biological factors interact to produce the non-random associations (cytonuclear disequilibria) often observed between uni-parentally and bi-parentally inherited alleles in particular populations or species (Asmussen *et al.* 1987).

13. Key behavioral and demographic parameters can differ between the genders

Several parameters relevant to intraspecific gene genealogies often show fundamental asymmetries between the genders. For example, males are the primary dispersers in many animal species, females so in others. In many plant species, seed propagules to which cytoplasmic genomes are confined may be far less dispersive than pollen granules that typically carry nuclear genes only. In both animal and plant taxa, variances in individual reproductive success often differ between the sexes. In general, matrilineal genetic markers in conjunction with those from nuclear loci have opened many novel opportunities to empirically assess the population genetic consequences of such gender-associated biological asymmetries (Melnick & Hoelzer 1992).

14. Conspecific populations are genealogically allied yet often highly distinctive from one another

Molecular data from mtDNA are especially useful in revealing phylogeographic structure of populations within a species. An important realization is that these historical population structures can range along a continuum from evolutionarily (temporally) shallow to deep (the latter being especially true for species that have had severe restrictions on historical gene flow). Within a species, the most distinctive deeper units (the major matrilineal branches) sometimes are referred to as intraspecific ‘phylogroups’ (Avise & Walker 1998). Such phylogroups often, but not invariably, are also apparent in appropriate assays of nuclear genes, in which case they may warrant potential recognition as evolutionarily significant units for purposes of taxonomy or conservation efforts (see concept # 20 below).

15. Principles of genealogical concordance assess the depth and strength of phylogeographic structure

Not all phylogeographic population structures are equal in magnitude. To distinguish the historically deep from the shallow population separations, four distinct aspects of phylogeographic concordance are employed (Avise & Ball 1990). Each examines the level of agreement or consensus among multiple classes of information: across multiple sequence characters within a single gene tree (aspect 1 of genealogical concordance); across multiple gene trees within a species (aspect 2); across multiple species within a regional biota (aspect 3); and across multiple categories of data, such as molecular genetics and historical geography (aspect 4). By hard criteria, only when concordance has been demonstrated in at least some (preferably several) of these various aspects is it proper to conclude that the available data register salient evolutionary separations among the conspecific populations examined.

16. The number of phylogroups per species usually is small

In most vertebrate species and many invertebrate and plant species surveyed to date, the number of highly distinctive intraspecific phylogroups is small or modest – typically only about 1-5 per taxonomic species (Avise & Walker 1999). This observation, coupled with the finding that cytoplasmic gene sequences even in closely related biological species usually tend to be readily distinguishable, suggests that historically distinctive units identified in molecular-genetic analyses often conform quite well both in composition and number (at least within an order-of-magnitude) to the arrays of taxonomic species recognized in more traditional biological classifications. Thus, when judged from the newer vantages and criteria of molecular phylogeography, traditional non-molecular systematists generally seem to have done an excellent job in identifying and classifying salient historical discontinuities in the biological world.

17. Intraspecific phylogroups are nearly always allopatric

Because individuals can be considered OTUs in gene-genealogical analyses (concept #9), there is no logic demanding that major branches in gene trees must be allopatric. Empirically, however, most such intraspecific phylogroups have proved to be non-overlapping or nearly so in geographic distribution (and when this is not the case, secondary overlap often seems to be the most plausible explanation). Furthermore, these phylogroups are often spatially arrayed in coherent regional patterns such that they can be thought of as corresponding roughly to what was implied under the traditional concepts of subspecies, incipient species, or (in more recent literature) 'evolutionarily significant units' (Moritz 1994).

18. The geographic distributions of intraspecific phylogroups usually make biogeographic sense

In specific instances, the spatial arrangements of major branches in intraspecific gene trees usually orient well with known or suspected biogeographic agents, such as obvious environmental barriers to historical gene flow, or the locations of Pleistocene refugia. Indeed, the primary aim of most phylogeographic studies has been to employ gene-tree data to help recover and interpret the genealogical history of conspecific populations and closely related species in the context of historical geography and other relevant factors. In recent years, this approach often has been extended to multiple codistributed species, thereby revealing the composite histories of regional biotas. This type of endeavor has blossomed into a new subdiscipline in its own right that can be termed comparative phylogeography (Avise 1992).

19. Species' natural histories also impact phylogeographic patterns

In addition to vicariant historical factors associated with changes in the physical environment, endogenous biological factors – species' ecologies, behaviors, and natural histories – play key roles in shaping phylogeographic patterns (Avise 2000). To mention just one example, highly dispersive marine fishes, as a rule, have proved to show far less phylogeographic population structure than most of their freshwater counterparts when sampled across ranges of comparable size.

20. Phylogeographic units have primary importance for taxonomy and conservation

Biodiversity (which in the final analysis is genetic diversity) is what taxonomy seeks to name and conservation biology seeks to preserve. By describing the spatial distributions of genealogical variety within and among related species, the data of phylogeography can help tremendously in recognizing historical biotic partitions that should be of central relevance both to microevolutionary systematics and to biodiversity preservation (Avise 1989b).

21. Gene trees can differ in topology from population trees and species trees

This statement applies to sexually reproducing species, but not to strictly asexual taxa (where, in principle, one-and-the-same historical transmission pathway characterizes all loci). The fundamental distinction between a gene tree and a species tree in sexual species was unappreciated until fairly recently (Hey 1994). For example, an earlier paradigm in systematics stated that even one synapomorph (shared-derived character) is enough to define a clade. This is patently false (unless ‘clade’ refers solely to a branch in the particular gene tree in question). It is for such reasons that principles of genealogical concordance (concept #15) are important in deciding whether or not deep historical partitions in particular gene trees accurately register genome-wide partitions that should distinguish long-separated populations, intraspecific phylogroups, or species.

22. The phylogenetic status of sister populations or species can itself be evolutionarily dynamic

With respect to gene genealogies, it is no longer adequate to consider recently separated populations or species as having a fixed phylogenetic relationship to one another (Neigel & Avise 1986). Due to lineage sorting across the generations mediated by demographic turnover (organismal reproduction and death), extant populations at any point in time carry only a subset of the lineage diversity of their ancestors, plus newly arisen lineage diversity postdating the vicariant separations. Thus, a common phylogenetic progression for gene trees in recently separated sister taxa is initial polyphyly or paraphyly, only eventually followed by reciprocal monophyly (Figure 3). The rate at which this genealogical transition proceeds (under neutrality) is a function of population demographic events immediately preceding, during, and following the vicariant split. The transition also takes longer in principle, all else being equal, for autosomal than for cytoplasmic genes, due to a four-fold larger effective population size for genes that are diploid (as opposed to haploid) and inherited bi-parentally (as opposed to uni-parentally).

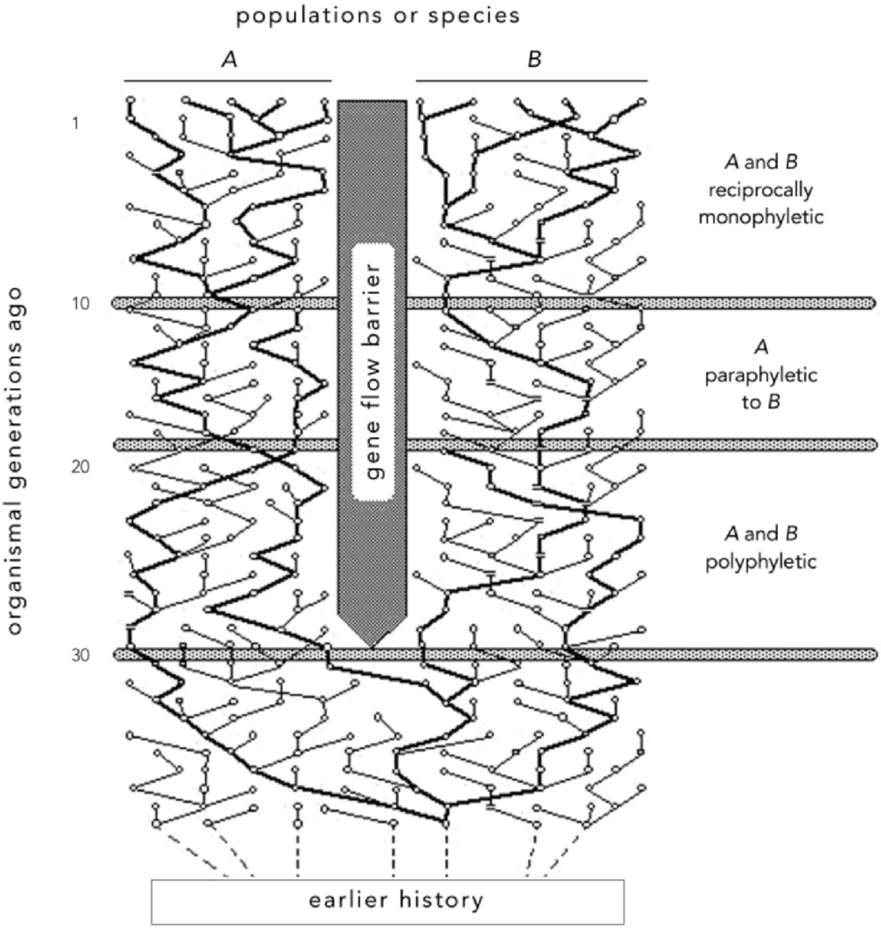

Figure 3. *Illustration of the shifting phylogenetic status of a gene tree (heavy lines) through time in two recently separated populations or species.*

23. Discordant gene trees and species trees can also characterize ancient taxa

Suppose that two or more successive cladogenetic events occurred long ago but close together in evolutionary time. If evolutionary effective population sizes of the species traversing the relevant nodes in the phylogenetic tree were larger than the internodal times as measured in organismal generations, then lineage sorting may not have proceeded to reciprocal monophyly in all gene trees of the immediate descendents (Tateno *et al.* 1982). Then, a topological discordance between a gene tree and a species tree will become evolutionarily 'locked in' as subsequent lineage sorting results in the eventual fixations of ancestral lineages in derivative taxa. Such idiosyncratic lineage sorting is one of several ways (others include secondary hybridization and various means of horizontal gene transfer) by which gene trees can come to differ topologically from one another and also from the composite species tree.

24. Allopatric speciation is a temporally extended process

A habit in traditional systematics is to view speciation as a point event in time (i.e. as a discrete node in a phylogeny). Although this may generally be acceptable for ancient cladogenetic events (where any temporal durations for speciation are small in comparison to the total time elapsed since), it can be grossly inadequate for recent speciations. By comparing the branching structures of mitochondrial gene trees within and among extant pairs of sister species, and by applying molecular clocks, recent phylogeographic appraisals suggest that the temporal duration of allopatric speciation in many vertebrate taxa averages (albeit with a large variance) about two million years (Avise *et al.* 1998). Such lengthy timeframes cannot be neglected when appraising, for example, the impacts of Pleistocene or more recent events on patterns of biological diversification.

25. Microevolution, like macroevolution, is historical

This catch-all truism sums up many of the phylogeographic insights described above. It is a basic realization that too often was overlooked in conventional population genetics, probably due in large part to that discipline's underlying formal theoretical framework. For reasons of mathematical tractability, many derivations and formulations in traditional population genetics dealt with equilibrium expectations (e.g. between mutation and selection, or genetic drift and migration) in unrealistically simplified contemporary settings (e.g. an 'island model' in which equal-sized populations are all assumed to exchange genes at equal rates). These were always presented as simplifying assumptions, but the net result nonetheless was a discipline too seldom focused on historical idiosyncrasies and non-equilibrium outcomes that are a *sine qua non* of real-life intraspecific evolution. Phylogeographic perspectives have enriched population genetics by adding an explicit focus on historical genealogy, and thereby drawing the field much closer to allied disciplines such as population demography, biogeography, and phylogenetic biology (Avise *et al.* 1987).

Synopsis

Historical reasoning and phylogenetic analysis have long been central themes of macroevolutionary biology and higher-level systematics, but until recently they had not permeated studies of intraspecific evolution to nearly the same extent. Thus, throughout most of the 20th century, there was a major gulf between the fields of phylogenetic biology and population genetics, to the detriment of both. Phylogeography is helping to bridge this gulf.

Microevolution too is a historical-genealogical process. Indeed, all limbs, branches, and twigs in any phylogenetic tree summarizing species' relationships ultimately consist of generation-to-generation organismal pedigrees through which genes were transmitted. The tools of molecular biology can now provide explicit historical information about genealogical tracings through such extended pedigrees, within as well as among living species. The net result has been the birth and growth of phylogeographic perspectives that promise to forge a useful new synthesis of micro- and macroevolutionary thought.

Acknowledgements

The author's recent work has been supported by a Fellowship from the Pew Foundation.

References

Asmussen MA, Arnold J, Avise JC (1987) Definition and properties of disequilibrium statistics for associations between nuclear and cytoplasmic genotypes. ***Genetics*** **115**, 755-768.

Avise JC (1989a) Gene trees and organismal histories: a phylogenetic approach to population biology. ***Evolution*** **43**, 1192-1208.

Avise JC (1989b) A role for molecular genetics in the recognition and conservation of endangered species. ***Trends in Ecology and Evolution*** **4**, 279-281.

Avise JC (1992) Molecular population structure and the biogeographic history of a regional fauna: a case history with lessons for conservation biology. ***Oikos* 63**, 62-76.

Avise JC (2000) ***Phylogeography: The History and Formation of Species.*** Harvard University Press, Cambridge.

Avise JC, Ball RM Jr (1990) Principles of genealogical concordance in species concepts and biological taxonomy. ***Oxford Surveys in Evolutionary Biology* 7**, 45-67.

Avise JC, Walker D (1998) Pleistocene phylogeographic effects on avian populations and the speciation process. ***Proceedings of the Royal Society of London, B* 265**, 457-463.

Avise JC, Walker D (1999) Species realities and numbers in sexual vertebrates: perspectives from an asexually transmitted genome. ***Proceedings of the National Academy of Sciences, USA* 96**, 992-995.

Avise JC, Giblin-Davidson C, Laerm J, Patton JC, Lansman RA (1979) Mitochondrial DNA clones and matriarchal phylogeny within and among geographic populations of the pocket gopher, *Geomys pinetis.* ***Proceedings of the National Academy of Sciences, USA* 76**, 6694-6698.

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. ***Annual Review of Ecology and Systematics* 18**, 489-522.

Avise JC, Ball RM Jr, Arnold J (1988) Current versus historical population sizes in vertebrate species with high gene flow: a comparison based on mitochondrial DNA lineages and inbreeding theory for neutral mutations. ***Molecular Biology and Evolution* 5**, 331-344.

Avise JC, Walker D, Johns GC (1998) Speciation durations and Pleistocene effects on vertebrate phylogeography. ***Proceedings of the Royal Society of London, B* 265**, 1707-1712.

Birky CW Jr, Maruyama T, Fuerst PA (1983) An approach to population and evolutionary genetic theory for genes in mitochondria and chloroplasts and some results. ***Genetics* 103**, 513-527.

Brown WM, George, M Jr, Wilson AC (1979) Rapid evolution of animal mitochondrial DNA. ***Proceedings of the National Academy of Sciences, USA* 76**, 1967-1971.

Hauswirth WW, Laipis PJ (1982) Mitochondrial DNA polymorphism in a maternal lineage of Holstein cows. ***Proceedings of the National Academy of Sciences, USA* 79**, 4686-4690.

Hey J (1994) Bridging phylogenetics and population genetics with gene tree models. In: ***Molecular Ecology and Evolution: Approaches and Applications*** (eds. Schierwater B, Streit B, Wagner GP, DeSalle R), pp. 435-449. Birkhäuser Verlag, Basel.

Hudson RR (1990) Gene genealogies and the coalescent process. ***Oxford Surveys in Evolutionary Biology* 7**, 1-44.

Hutchinson CA III, Newbold JE, Potter SS, Edgell MH (1974) Maternal inheritance of mammalian mitochondrial DNA. ***Nature* 251**, 536-538.

Maddison WP (1997) Gene trees in species trees. ***Systematic Biology* 46**, 523-536.

Melnick DJ, Hoelzer GA (1992) Differences in male and female macaque dispersal lead to contrasting distributions of nuclear and mitochondrial DNA variation. ***International Journal of Primatology* 13**, 379-393.

Moritz CC (1994) Defining ‘evolutionarily significant units’ for conservation. ***Trends in Ecology and Evolution* 9**, 373-375.

Neigel JE, Avise JC (1986) Phylogenetic relationships of mitochondrial DNA under various demographic models of speciation. In: ***Evolutionary Processes and Theory*** (eds. Nevo, Karlin, S), pp. 515-534. Academic Press, New York,

Tajima F (1983) Evolutionary relationships of DNA sequences in finite populations. ***Genetical Research*** **123**, 437-460.

Tateno Y, Nei M, Tajima F (1982) Accuracy of estimated phylogenetic trees from molecular data. I. Distantly related species. ***Journal of Molecular Evolution*** **18**, 387-404.

Wilson AC, Cann RL, Carr SM, George M, Gyllenstein UB, Helm-Bychovski KM, Higuchi RG, Palumbi SR, Prager EM, Sage RS, Stoneking, M (1985) Mitochondrial DNA and two perspectives on evolutionary genetics. ***Biological Journal of the Linnean Society*** **26**, 375-400.

Chapter 2

Plant phylogeography based on organelle genes: an introduction

Rémy J. Petit* and Giovanni G. Vendramin
*E-mail: petit@pierroton.inra.fr

Abstract

Plants have unique biological attributes of great interest to researchers investigating population dynamics. Yet, until recently, organelle DNA had been seldom utilized for phylogeographic studies in plants. While cpDNA variation has been used extensively to construct interspecific phylogenies, many researchers have considered that the relatively low levels of intraspecific variation revealed by early studies of plant organelle DNA render these genomes of little use for intraspecific studies. In this review we introduce the field of plant phylogeography based on organelle polymorphisms by providing a detailed discussion of the processes underlying this variation. Progress in molecular organelle genetics has provided insights into the structure, variation, inheritance, vegetative segregation and recombination of organelle genomes. While some of these features (e.g. low substitution rates) may complicate phylogeographic studies, others (e.g. presence of two genomes and frequency of atypical modes of transmission) offer unique opportunities, many of which are virtually unexplored.

Keywords: cpDNA, mtDNA, population, history, structure, variation, inheritance, recombination, gene flow, dispersal, selection

Introduction

Plants have many unique features with great appeal for researchers trying to unravel spatio-temporal dynamics of populations and their consequences for evolution, the objects of phylogeography (see Avise, this volume). In particular, they stand still during most of their life. This explains the central importance of space in their study (Silvertown & Charlesworth 2001) and considerably facilitates their sampling. Despite this immobility, seed plants can move their genes in two specialized vehicles during short but critical phases of their life cycle: before fertilization, in the male gametophyte (pollen), and later in the young sporophyte (the seed). Hence, in stark contrast to many animals, the new diploid embryo is normally mobile while juveniles and adults are sedentary. Even the closest relatives of seed plants, the

S. Weiss and N. Ferrand (eds.), Phylogeography of Southern European Refugia, 23–97.

ferns, have a distinct life cycle and live in a different world with respect to migration processes (Sauer 1988). In angiosperms and gymnosperms, which will be the target of this review, pollen plays a major role in connecting extant populations with gene flow, but seeds (or other plant parts) are necessary to establish new populations of plants. Consequently, maternally inherited genes (which are not transmitted to the next generation by pollen) should be of special value for clarifying the spatio-temporal dynamics of plant populations.

Phototrophic plants are central components of ecosystems, responsible for the primary production of biomass, providing food and shelter to most animals. At the same time, plants often depend on animals for reproduction and/or dispersal. Zoologists are starting to realize that the timing and pattern of colonization by plants are of particular importance in understanding animal phylogeography, whereas the phylogeographic patterns studied by botanists have actually been shaped by the movements and behavior of animals as well as by the distribution of pathogens or symbionts.

Given their central place in the ecology of life, the vital importance of their genetic resources for sustainable agriculture and forestry, and their unique biological attributes, one would expect plants to be at the forefront of phylogeographic research. Surprisingly, however, there were until recently few explicit phylogeographic studies of plants (Schaal *et al.* 1998), in contrast with the situation for animals (Avise 2000). Fortunately, this situation is changing rapidly, and plant phylogeography is attracting increasing interest.

In this review, we provide an introduction into the field of population genetics and phylogeography using genetic variation in plant organelles. Typically, one of two approaches is used by geneticists: either they focus on the frequencies of variants (haplotypes) within and among populations, but do not attempt to take haplotype similarities into account, or they focus instead on intraspecific phylogenies but do not consider within population variation. Ideally, population sampling should be combined with intraspecific phylogenies of the variants, bridging the gap between these two approaches. Besides reviewing information useful to interpret population or phylogeographic surveys (or a combination of both), we identify some of the opportunities – many of them unexplored – that this field might provide in the future.

Part 1 summarizes some key steps in the history of plant organelle genetics, and the changing perception of the usefulness of organelle DNA as a source of markers for population and phylogeographic surveys. Part 2 examines in some detail the characteristics of plant organelle genomes that are relevant for population studies. Plants are unique among eukaryotes in possessing two DNA-containing organelles, the ubiquitous mitochondrion and the distinctive plastid. We consider in turn, genome structure and variation organelle inheritance and vegetative segregation, as well as recombination and the association between chloroplast (cp) DNA and mitochondrial (mt) DNA. Part 3 considers important parameters that can influence the geographic structure of organelle genes: intraspecific gene flow, as seen from a genetic and from