POPULATION GENETICS OF FOREST TREES

FORESTRY SCIENCES

Volume 42

Population Genetics of Forest Trees

Proceedings of the International Symposium
on Population Genetics of Forest Trees
Corvallis, Oregon, U.S.A., July 31–August 2, 1990

Edited by

W. T. ADAMS
Oregon State University

STEVEN H. STRAUSS
Oregon State University

DONALD L. COPES
USDA Forest Service, Pacific Northwest Research Station, Oregon

and

A. R. GRIFFIN
Shell International Petroleum Co., London

Printed on acid-free paper

Contents

Introduction

This book contains 19 invited papers from the International Symposium on Population Genetics of Forest Trees, held in Corvallis, Oregon, USA, July 31 – August 2, 1990. The theme of this joint meeting of the International Union of Forestry Research Organizations (IUFRO) Working Parties on Biochemical Genetics (S2.04.05) and Population and Ecological Genetics (S2.04.01) was "the contribution of biochemical markers to the understanding of population genetics of forest trees". Over 140 individuals from 15 countries attended. In addition to IUFRO, sponsors included the Forest Science Department of Oregon State University, and the USDA Forest Service, Pacific Northwest Research Station.

The increased availability of biochemical markers has resulted in extensive research directed to the population genetics of forest trees during the past 15 years. These studies have added greatly to our understanding of the genetics of tree species, especially in the areas of population genetic structure, evolutionary biology, and mating systems. The goals of this symposium were to summarize and review the results of these efforts and to identify fruitful areas for future research.

The papers in this volume address all of the major applications of biochemical markers to forest genetics research. These include description of the patterns of genetic diversity within and between species, and assessment of their underlying causes (8 papers); investigation of mating systems, gene dispersal, and genetic structure within populations (4 papers); and, application of biochemical markers to problems in forest management (4 papers).

While the emphasis in the above papers is on studies employing isozymes or terpenes, two additional papers specifically address the potential of DNA as a biochemical marker for forest genetics research. The final paper was written by invitation as a commentary on the symposium. This paper does not attempt, however, to summarize the major conclusions of the individual papers, but rather offers personal critiques and insights on current approaches to population genetics research in forest trees.

We thank the symposium sponsors for helping to make this symposium a success. We are also grateful to Dr. Mary Duryea, Editor-in-Chief of *New Forests*, and to Mr. Ad C. Plaizier of Kluwer Academic Publishers, for their encouragement in this project.

Special Editors

W. T. Adams
S. H. Strauss
D. L. Copes
A. R. Griffin

Dedication

We dedicate these proceedings to our colleague and friend, Dr. James Hanover, who passed away on February 21, 1992. In a distinguished career of more than 35 years as teacher and scholar, Jim contributed greatly to the fields of forest genetics and tree physiology.
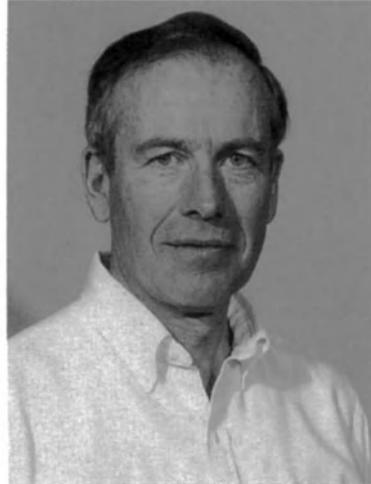
Review paper

Genetic diversity—seeing the forest through the trees

M. THOMPSON CONKLE

Institute of Forest Genetics, Pacific Southwest Forest and Range Experiment Station, USDA Forest Service, Box 245, Berkeley, California, 94701, USA

Application. Processing biochemical markers from individual trees provides a means to resolve and measure genetic diversity. Identifying that variation and interpreting that rich source of genotypic information leads to effective means for managing populations to maintain forest health. Looking ahead, research in molecular forestry will extend toward discovering how the biology of forest species is effected by variation in individual genes.

Abstract. Forest trees, populations, races, species, and taxonomic groups above the species level display rich variation in biochemical markers. The variation stems from inherited modifications that trace back in time, through converging ancestries, towards common progenitors. Past movements of continents, mountain building events, and climate changes isolated forest populations one from another and provided critical challenges to the lineages that survived to the present day. A wealth of molecular variants in forest trees characterize these widely-distributed, large, long-lived, outbreeding, organisms. Forest trees have an abundance of rare variants and over one third of all the alleles (different forms of one gene) occur only rarely (< 2% frequency) in a few trees of a species sample. Those rare alleles may either represent new variation or persistent forms of genes that have low adaptive value under present conditions. From another perspective, however, the largest share of genetic variation in forest trees is due to the presence of multiple alleles found at intermediate frequencies for only a small percentage of all the genes, and those alleles are commonly widespread throughout species areas. These common alleles may mark genes that track historical events in lineages or mark genes with adaptive significance in present populations. Evidence from enzyme studies supports the conclusion that highly comparable functional genes are common to different forest taxa. Future research will be toward understanding the phenotypic expression of particular genes and revealing the relative importance of genetic variants to adaptation and growth.

Introduction

"You can't see the forest for the trees!" That familiar phrase is often intended as light hearted criticism of work on narrow problems. It carries a double meaning for forest geneticists who gathered at a symposium in

Corvallis, Oregon, to reflect, report, and speculate on a broad spectrum of information about forest diversity. This opening contribution to the symposium addressed a mix of three subjects ranging from the origins of forest taxa, through examination of diversity in a sample of species, to concluding remarks about the construct of forest species genomes. It covered general subject matter and intentionally avoided the specific themes in molecular forestry that were addressed by the other participants.

The timing of the symposium was appropriate because a sizable volume of findings had become available since July, 1979, when a similar group convened for discussions on isozymes of North American forest trees and insects (Conkle 1981). The presentations in 1979 demonstrated the high value of biochemical markers (primarily mobility variants for numerous, specific enzymes coded by nuclear genes) for evaluating variation in forest tree populations. Since then, the necessity to conserve and manage genetic variation has become a worldwide priority. And with few efficient means available to evaluate diversity, the rapid acquisition of precise genetic information using biochemical markers plays a central and critical roll in assessing populations and monitoring species in native forest ecosystems.

A complete consideration of forest diversity would include mention of variation in natural morphology, progeny growth in plantation trials, and survival responses in stress tests. Those measurements, trials, and tests are essential for evaluating whole plant responses but the differences observed in those trials are controlled by variation in unidentified genes. Here, consideration focuses on knowledge about variation in alleles of identified genes. Currently, these alleles and variation in these genes appear to have only minor predictive value of specific growth responses in whole plants.

Clues about diversity from the past

Skilled analysis of morphological characteristics has provided reasonably dependable and time proven methods for classifying organisms. Phylogenetic analyses of extant and fossil taxa when combined with events in paleohistory gives a basis for appreciating the ancestral relationships among modern species and speculating about the origins of their various adaptations. It is not trivial to observe that the relationships among forest species are best pictured as a tree with branches and twigs (clades) that trace back to a single root, a progenitor common to all plants.

Forest seed-plant origins are in the Coniferopsida (a subdivision of the gymnosperms which includes orders of ginkgos, conifers, and yews) and date from the Carboniferous, roughly 300 million years ago (Ma). Con-

7

stance Millar (Millar 1991; Millar and Kinloch 1991) evaluated informa-
tion about key events in the history of pines, a genus with about 100
species, wide distribution in the northern hemisphere, and an informative
example of forest tree species evolution. She used evidence from the
oldest pine fossil (dated about 130 Ma) and from the wide abundance of
other fossil pines shortly thereafter to estimate the derivation of the genus
Pinus at about 200 Ma. The first pines evolved in the middle latitudes of
the northern hemisphere at a time when continents formed a single great
land mass (Pangaea) and ice blanketed extensive areas (Fig. 1a) of the
south polar masses (Harrington 1987; Parrish 1987). Continental colli-
sions during formation of that land mass raised mountains near points of
impact (Appalachian and Atlas Mountains, and elevated topography in
Greenland and Norway) that gave variety to habitats for the evolving
lineages.

First angiosperm fossils appeared at about 120 Ma and differentiation

200 Ma
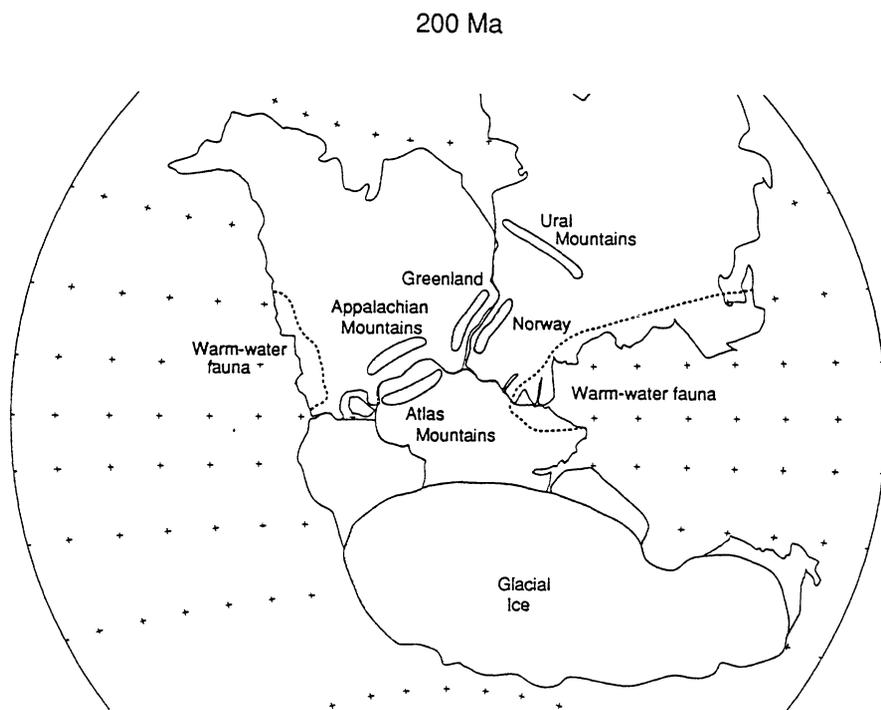
Fig. 1a

Fig. 1. Diagrammatic representation of land masses at three periods important to the
history of temperate forest species, *a*) 200 million years ago (Ma), *b*) 54 Ma, and *c*) 18
thousand years ago (ka)(Newell 1972).

8

into dicotyledon and monocotyledon divisions virtually coincided with their first appearance (Friis et al. 1987). An explosive radiation of angiosperms from low latitudes established them world wide by 100 Ma as the overwhelming dominant land plant group (Crane 1987). Conifers, by that time, had spread east and west throughout North America and Eurasia (Laurasian land mass) and they equaled all the other land plants combined in absolute species diversity (Nichols et al. 1985).

At about 100 Ma a warm sea separated the northern from the southern land masses. The North American and Eurasian masses were drifting apart then, but they were still joined in their far northern areas. A group of continents (South America, Africa, India, and Australia joined to Antarctica) were drifting apart from the southern conglomerate (Gondwana). North—south interior seaways bisected North America (from the Gulf of Mexico to the Arctic) and Eurasia (from the eastern Mediterranean to the Arctic). Waters of those interior seaways isolated substantial regions of both of the northern land masses.

Events beginning about 54 Ma (Fig. 1b) profoundly influenced the

54 Ma
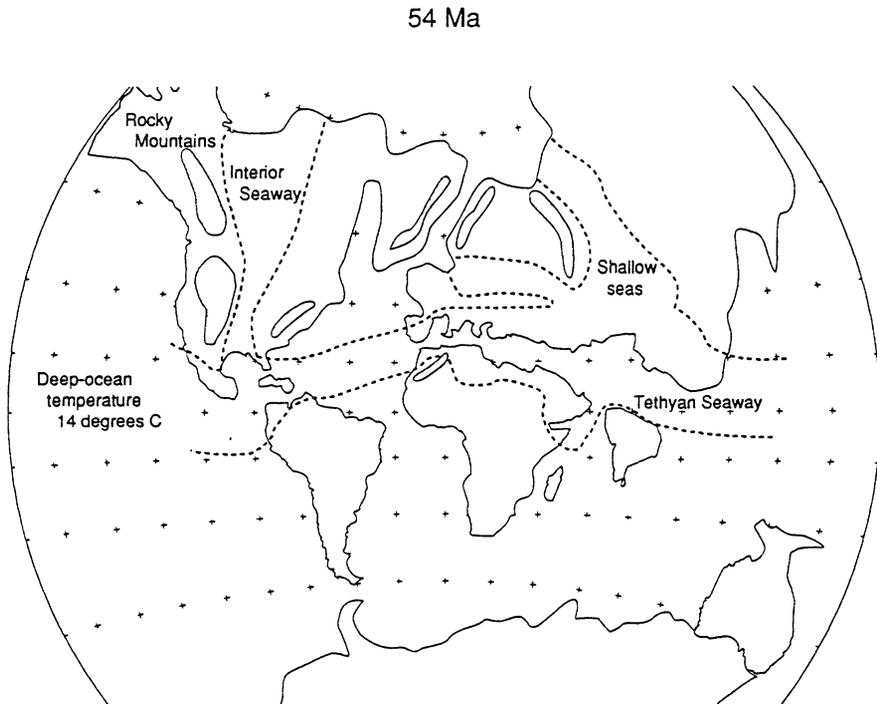
Fig. 1b

evolution of conifers (evidence from pines, Millar 1991) and deciduous forest genera (Critchfield 1980). During a period of warm humid climates, sub-tropical floristic associations became established in relatively high latitudes (centered 75 °N), fracturing and displacing the temperate species into either more northern, more southern, or more upland areas. Pine species adaptations and geographic patterns associated with modern sub-sections of *Pinus* largely reflect evolution in refugial regions for the period from about 54 Ma to 37 Ma (Millar 1991). Appalachian forests displayed a mix of moist tropical and deciduous hardwood vegetation suggesting that some seasonally dry and warm-temperate climates prevailed there during the same period (Critchfield 1980).

The continued westward drift of North and South Americas further isolated them from the eastern continents and their lineages from their eastern vegetation counterparts. The genetic differentiation of hard pines continued on separate continents and now genetic barriers to crossing precludes hybridizations between species native to the Eastern and Western hemispheres (Critchfield 1986).

Dramatic topographic and climatic events took place between 54 and 37 Ma. Rocky Mountain topography (with a 100 million year period of
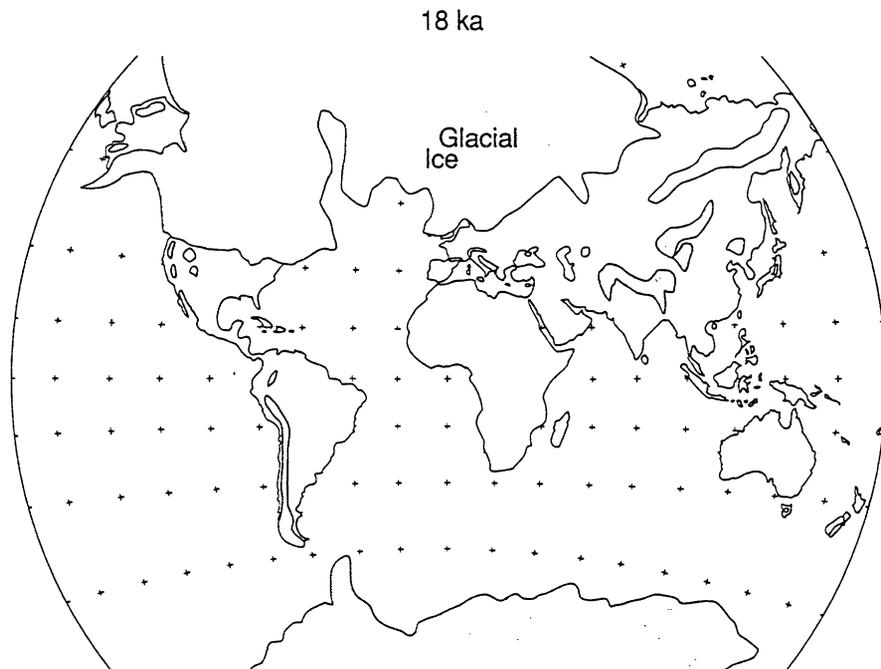
18 ka

Fig. 1c.

uplift) emerged in full development. India collided with Asia and raised the Himalayas. Australia separated from Antarctica. Ocean currents, then shunted through cold polar seas in a new seaway south of Australia, brought about significant ocean cooling (Kennett 1977). A corresponding severe drop in temperatures on land (along with increased seasonality) brought the warm periods to a close. Tropical and subtropical plants were lost from mid and upper latitudes, and the previously displaced taxa reclaimed the temperate mid latitudes. The present differentiation among subsections of *Pinus* points to their evolution in the isolated refugial regions they occupied during the warm periods of the Eocene (Millar 1991). Further differentiation of lineages radiating from those refugia may have established many modern day species.

During a period between 37 and 2 Ma., the collision of Africa and Europe resulted in complete desiccation of the Mediterranean Sea (Hsu 1972) and development of the Pyrenees, Alps, and Carpathian Mountains. Continued uplift accompanied by volcanic activity along western margins of North, Central, and South America raised the Cascade-Sierra Nevada, elevated the Sierra Madre Occidental and the Transverse volcanic belt in Mexico, and formed the Andes in South America. Deep ocean cooling continued unabated. Deep ocean temperatures dropped from 14 °C, about 75 Ma, to 3 °C at the onset of the ice ages (Newell 1972).

Repeated cycles of the advance and retreat of glacial ice, beginning about 2.5 Ma, forced mid and upper latitude forest species into large scale migrations. The current ice ages began with about 40 full climatic cycles in 41 thousand-year rhythms. They then shifted to about 8 full cycles with long periods of glaciation, of 80 to 100 thousand years each, followed by short interglacial intervals, of 10 to 15 thousand years each (Ruddiman and Wright 1987). During each cycle, glacial ice advanced over polar and upper latitude areas and in high mountain terrain. Masses of ice at the culmination of the recent cold cycle, 18 thousand years ago, were the most extensive of all masses during the ice ages (Fig. 1c). Those great ice volumes lowered global seas about 130 m below the present levels (Cronin et al. 1981).

Forest species in mid and high latitudes shifted north and south, and higher and lower in elevation, in response to shifting habitats during the cold and warm periods of glacial cycles (Barnosky et al. 1987; Jacobson et al. 1987). The magnitude of these migrations can be appreciated by noting that northern boreal species of eastern North America were centered in southeastern United States about 18 thousand years ago (Webb et al. 1987). Vegetation in more stable environments (locations moderated by oceanic influences, arid rain-shadow locations, and equatorial regions) was effected less drastically. The migration of forest species revealed the limits of their respective adaptations. The near universal survival of forest

species plus their general lack of morphological change during extreme stress imposed by the ice age climates are evidence of their persistence and the well buffered genetic stability of their traits.

Critchfield (1984) used data from genetic markers to summarize several consequences of recent glacial events for North American conifers. He noted populations in previously-glaciated northern areas to be less variable than their southern counterparts; the migrating founders of some species carried forward only a limited sample of progenitor genomes. Critchfield identified extant species that lack or have distinctly low levels of variation, resulting presumably from being driven to near extinction. He identified geographic races that formed as populations radiated from different refugia. Some races had transient existence owing to introgression or fusion upon encountering compatible taxa. Other races, evolved during periods of isolation to become varieties, subspecies, or species. Critchfield also provided evidence that some species increased variation by gene exchange during periods of contact with compatible species.

Given the past histories of boreal and temperate lineages, taxa differ widely in their adaptations to different climatic conditions. An expected response of some successful species to repeated glacial cycles might be the accumulation of wide climatic tolerances. Since temperate forest species have long been mobile geographic units responding to climate change, they might be expected to have accumulated adaptations for colonizing new habitats along their leading fronts and for resisting elimination along their trailing boarders. Each climatic cycle would result in shifts, from advance to retreat and from leading to trailing margins. Viewed today, some noteworthy species have extremely wide adaptations, they span significantly large ranges of elevations and latitudinal distances. Other species, presumably those for whom migration and survival were severe challenges, have narrow distributions. Some narrowly distributed taxa possess moderate variation, but others are known to be genetically depauperate and have little or no capacity remaining for significant evolutionary change. Biochemical markers, whether adaptive or neutral in present day environments, reveal amounts of allelic variation, identify genetic relationships, and track patterns of differentiation within and among populations.

Summary items for past evolution

· Forest species evolved from common ancestors;
· Conifers have about a 200 million year history that began when all continents were one land mass;
· Forest angiosperms evolved about 100 million years ago;

- Major differentiation of boreal and temperate forest groups occurred about 54 to 37 million years ago when warm climates favored development of subtropical vegetation in mid and upper latitudes, forcing temperate vegetation into cooler or drier refugia;
- Climatic cooling beginning about 37 Ma favored speciation of boreal and temperate forest taxa during periods of migration into new areas;
- Ice ages of recent history caused mass migration of species and contributed to further differentiation of lineages as populations repeatedly advanced and retreated, some into and out of restricted refugia, and others across latitudes and elevations that may have resulted in the development of wide environmental tolerances;
- Amounts and distribution patterns of variation in modern taxa reflect the molecular history of lineages as well as the disequilibrium inherent in mobile groups.

Profiles of species diversity

Diversity data from a number of species in California are available (Fig. 2). Species that are included all had genetic analyses performed using similar analytical and interpretative techniques and many grow together in natural mixtures (sympatry strengthens the comparisons among them). Wide spread species that span a notable range of elevations are: ponderosa (*Pinus ponderosa* Dougl. ex Laws.), Jeffrey (*P. jeffreyi* Grev. & Balf.) , and sugar pines (*P. lambertiana* Dougl.), Douglas-fir (*Pseudotsuga menziesii* [Mirb.] Franco*)*, incense cedar (*Calocedris decurrens* Torr.) and California white fir (*Abies concolor* var. *lowiana* [Grod.] Lemm.). Others — Bishop (*Pinus muricata* D. Don), knobcone (*P. attenuata* Lemm.), Monterey (*P. radiata* D. Don), Coulter (*P. coulteri* D. Don), and Digger (*P. sabiniana* Dougl.) pines — are restricted to relatively narrow elevational bands and occupy well defined habitats. A third group consists of Monterey cypress (*Cupressus macrocarpa* Hartw.), Torrey (*Pinus torreyana* Parry ex Carr.) and Washoe (*P. washoensis* Mason & Stockwell) pines that seem near extinction because they grow in only a few diminutive native populations.

Three pines in the California sample — ponderosa, Washoe, and Jeffrey — are members of the *Ponderosae* subsection and stem from a common lineage. All members of their subsection inhabit western North America, and differentiation of the lineage traces to the period between 54 and 37 million years ago. These three species have northern distributions and northern ecological affinities but they are close relatives with a large group of western yellow pines endemic to Mexico and Central America (Conkle and Critchfield 1988). Jeffrey pine crosses to a moderate extent with
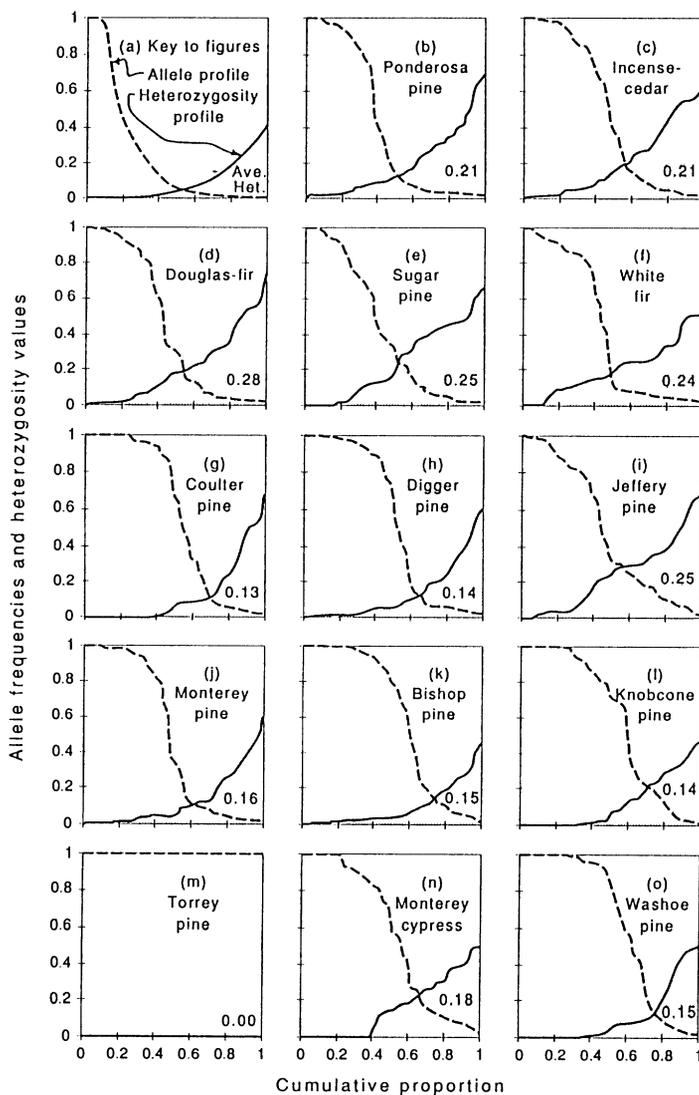
Fig. 2. Diversity profiles for 14 conifer species displaying cumulative distributions for: 1) alleles of different frequencies, irrespective of locus of origin, plotted in descending order, 2) expected heterozygosities by locus given in ascending order, and 3) mean expected heterozygosities (shaded areas under curves) whose species values are given: *a*) key to figures, *b*) ponderosa pine (*Pinus ponderosa*), *c*) incense-cedar (*Calocedris decurrens,* Harry 1984), *d*) Douglas-fir (*Pseudotsuga menziesii), e*) sugar pines (*Pinus lambertiana*), *f*) white fir (*Abies concolor*), *g*) Coulter pine (*Pinus coulteri*), *h*) Digger pine (*Pinus sabiniana*), *i*) Jeffrey (*Pinus jeffreyi,* Furnier 1984), *j*) Monterey pine (*Pinus radiata,* Millar et al. 1988), *k*) bishop pine (*Pinus muricata,* Millar et al. 1988), *l*) knobcone pine (*Pinus attenuata,* Millar et al. 1988), *m*) Torrey pine (*Pinus torreyana,* Ledig and Conkle 1983), *n*) Monterey cypress (*Cupressus macrocarpa* Conkle 1987), and *o*) Washoe Pine (*Pinus washoensis,* Niebling and Conkle 1990).

Coulter pine, and Jeffrey pine appears to be a species linking the *Pondero-sa* and *Sabinianae* subsections (Conkle and Critchfield 1988). Coulter, Digger, and Torrey pines comprise the *Sabinianae* subsection. Their differentiation from yellow pines in Mexico and Central America traces to about 65 million years ago (Millar 1991). California closed cone pines — Monterey, Bishop, and knobcone — have historical differentiation times and geographic origins similar to the *Sabinianae*. They are the northern most members of the *Oocarpae* and are related to a group of extant species in Mexico and Central America. All pine species, just mentioned, are hard pines (subgenus *Pinus*).

Sugar pine is the only white pine (subgenus *Strobus*) in this sample. Its origins are obscure, but other members of subsection *Strobi* are distrib-uted throughout North America and Eurasia. Douglas-fir, white fir, and Monterey cypress likewise are related to species distributed throughout North America and Eurasia. Northern hemisphere distributions for *Pinus*, *Pseudotsuga*, *Abies*, and *Cupressus* support the conclusion that these conifers evolved on northern land masses after separation from the southern land masses. Of the species sampled, only incense-cedar with its range centered in Oregon and California appears to have differentiated to a high degree prior to the break up of the universal land mass (Pangaea). Other Northern hemisphere incense-cedars (*Calocedrus*) are in China and Taiwan. Southern hemisphere incense-cedars (*Libocedrus*) are native to Chile in South America (also included under *Austrocedrus*) and to New Zealand and New Caledonia in the South Pacific.

Protein samples from forest trees, whether from seeds or vegetative tissues, when separated during electrophoresis and stained, reveal the presence of bands for specific enzymes. The majority of mobility variants found in enzyme studies of genetic diversity segregate in equal proportions with 50% of germ cells having one form and 50% having the alternative. Equal segregation of different alleles for genes tracing to egg or pollen cells provides strong evidence that they are alleles coded by nuclear genes (allozymes). Data from laboratory studies of allelic mobility variants for particular enzymes in species are often summarized in the form of allele frequencies and expected heterozygosities for numerous genes.

Species diversity characteristics can be generated from allozyme data by ignoring gene identities and by considering enzyme genes to be random samples of a large class of functional genes. Cumulative distribution frequencies characterize allele frequencies (frequency of occurrence) and provide a visual description of species heterozygosity (average proportion of loci with different alleles per gene in an individual, and also the average proportion of individuals in populations with different alleles per gene). Figure 2 pictures three comparisons for each species:

1. allele profiles based on a descending sorted sequence of allele frequencies for all genes;
2. heterozygosity profiles based on ascending sorted sequences of expected heterozygosities for all genes; and
3. average heterozygosity values (Ave. Het.).

The horizontal axis, marked in unit scale, is the same size as the vertical axis (allele frequencies and heterozygosity values) to make the figure a unit of area (the shaded area below the heterozygosity profile is equal to the heterozygosity for each species). Since alleles with frequencies less than 0.02 only have minor effects on heterozygosities, they were excluded from the allele frequency distributions (Fig. 2). For ponderosa pine, 20% of alleles have frequencies greater that 0.90, 50% of loci have expected values less than 0.10, and the average expected heterozygosity is 0.21.

Tree species commonly have minor proportions of fixed alleles (frequencies equal 1.00), and heterozygosities, when sorted, depict a linear to exponential increase in diversity from fixation (one allele, heterozygosity equals 0.00) to the most variable loci (heterozygosities > 0.6). The exception is Torrey pine populations (Fig. 2 [m]) which have all genes with fixed alleles (average heterozygosity equals 0.00).

The wide spread species — ponderosa pine, Fig. 2 (b), incense-cedar (c), Douglas-fir (d), sugar pine (e), white fir (f), and Jeffrey pine (i) — have a range of average heterozygosity values from 0.21 to 0.28 (mean 0.24). Species with somewhat limited distributions — Coulter pine (g), Digger pine (h), Monterey pine (j), Bishop pine (k), and knobcone pine (l) — on examination have comparatively more fixed alleles in their allele profiles and somewhat more concave than linear heterozygosity profiles. The latter species range from 0.13 to 0.16 (mean 0.14) for average heterozygosities.

The rare endemic species — Torrey pine Fig. 2 (m), Monterey cypress (n), and Washoe pine (o) — contrast with the previous species. Torrey pine has no measurable diversity, while Monterey cypress and Washoe pine have profiles and heterozygosities that are nearly comparable with species having large distributions. A rare event in Torrey pine's history, perhaps the present species establishment by one or only a few founders, or a near extinction drop in population size, presumably caused fixation in its genes (Ledig and Conkle 1983).

Genes in the sampled species differ in the numbers of alleles per locus (from 1 to about 10). Since rare alleles occur almost exclusively in the heterozygous condition and since alleles with intermediate frequencies are major contributors to population heterozygosities, it is informative to examine the percent of alleles at different frequencies for genes with differing number of alleles (Table 1). Table 1 was developed from allele

Table 1. Proportions of alleles at various frequencies grouped by the number of alleles that are detected for genes in a sample of 13 conifer species (see Fig. 2, and excluding Torrey pine)

Allele frequency	Number of alleles per gene					Total(%)
	1	2	3	4	> 4	
	Number of alleles (% of total)					
	65 (6)	163 (15)	283 (26)	250 (23)	326 (30)	1,087 (100)
	(%) in each class					
> 0.95	100	31	14	6	3	17
> 0.90	0	38	19	12	5	21
> 0.80	0	44	26	17	11	27
> 0.50	0	51	34	24	16	33
> 0.20	0	57	42	32	21	39
> 0.10	0	63	50	39	26	45
> 0.05	0	69	57	49	36	53
> 0.02	0	75	69	61	51	64
0.02 and lower	0	25	31	39	49	36

frequency data (total of 1,087 alleles) for all the sampled species except Torrey pine.

Overall, 6% of the genes had only one allele, 15% had 2 alleles, and close to 80% had 3 or more alleles (Table 1). About one third of all alleles had frequencies greater than 0.50. Thirty one % of genes with 2 alleles had frequencies in excess of 0.95 (> 0.95), but genes with more the 4 alleles had only 3% of alleles with frequencies in excess of 0.95. Comparisons across other classes with different numbers of alleles per gene also show declining percentages; only about 20% of the genes with greater than 4 alleles have frequencies exceeding 0.20. Rare alleles, here noted as ones in frequencies equal to 0.02 and lower, by tally are 25% of the 2 allele per gene cases, but increase to near 50% for genes with more than 4 alleles. Rare alleles are abundant when expressed as a percent of all alleles, their percentages exceed one third of all alleles.

Heterozygosity values were also tabulated by the number of alleles per gene for all genes in the sample species (Table 2). Of a total of 360 loci, 82% were polymorphic. Genes with 2 and 3 alleles made up 52% of the total. Note in the right hand total column that 50% of the genes had heterozygosity values greater than 0.10. There is a positive relationship between the number of alleles per gene and the percent values in classes