

## Tropical Forestry

---

## **Tropical Forestry**

Volumes Already Published in this Series

*Tropical Forest Genetics* by Finkeldey, R., Hattemer, H. 2007

ISBN: 3-540-37396-9

*Sampling Methods, Remote Sensing and GIS Multiresource Forest Inventory* by Köhl, M. Magnussen, S., Marchetti M. 2006

ISBN: 3-540-32571-9

*Tropical Forest Ecology - The Basis for Conservation and Management* by Montagnini, F., Jordan C. 2005

ISBN: 3-540-23797-6

Reiner Finkeldey · Hans H. Hattemer

---

# Tropical Forest Genetics

With 44 Figures and 19 Tables

 Springer

Professor Dr. Reiner Finkeldey  
Professor Dr. Hans H. Hattemer  
Institute of Forest Genetics and Forest Tree Breeding  
Büsgenweg 2  
D-37077 Göttingen  
Germany

ISSN: 1614-9785  
ISBN-10 3-540-37396-9 Springer-Verlag Berlin Heidelberg New York  
ISBN-13 978-3-540-37396-4 Springer-Verlag Berlin Heidelberg New York

Library of Congress Control Number: 2006931487

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

Springer-Verlag is a part of Springer Science+Business Media

springer.com

© Springer-Verlag Berlin Heidelberg 2007

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Editor: Dr. Dieter Czeschlik, Heidelberg  
Desk Editor: Anette Lindqvist, Heidelberg  
Production: SPi  
Typesetting: SPi  
Cover Design: Design & Production, Heidelberg

Printed on acid-free paper                      3/3152-HM                      5 4 3 2 1 0

---

## Preface

This book is intended to provide information on fundamental genetic processes in tropical forests. It is based on lecture notes by the authors prepared for post-graduate students, mainly from tropical countries, of the M.Sc. course “Tropical and International Forestry” at the Faculty of Forest Sciences and Forest Ecology, Georg August University Göttingen, Germany. The intended readership is students, researchers and practitioners interested in the genetic variation of species, in particular forest trees, living in complex forest ecosystems in the tropics. Particular emphasis is placed on the human impact on forest genetic resources in the tropics.

Readers should be familiar with basics of classical and molecular genetics such as the structure and function of DNA (double-helix structure, replication), polypeptide synthesis (transcription and translation), and the transmission of genetic information during sexual reproduction (“Mendel’s rules”). This knowledge is easily available from recently published genetics textbooks; we recommend the book by Griffiths et al. (2000).

The development of biochemical and molecular marker techniques and their application to species of tropical forests have greatly improved our knowledge of genetic variation patterns of “wild” plants in the tropics during the last two decades. Many examples in this book are based on gene marker techniques. We discuss the results of selected molecular studies and their implications for the development of sustainable management strategies for forest genetic resources. However, we refrain from providing details on laboratory techniques and statistical data analyses. Readers should consult the cited references or the book by Weising et al. (2005) if they require additional information in this regard. No advanced knowledge concerning modern molecular or statistical methods for the analysis of genetic variation patterns is necessary to follow the argumentation of this book.

Comprehensive data on DNA sequences are becoming available for more and more plants, including forest trees. The first project aimed at sequencing the full genome of a forest tree (*Populus trichocarpa*) has recently been completed (Brunner et al. 2004). However, only very few genomic data are currently

available for tropical forest trees; tropical forest genetics is still in a “pre-genomic era.” Thus, the focus of this book is on population genetics processes in natural and managed tropical forests rather than on genomics and related, newly emerging fields of research in forest genetics.

---

### **Acknowledgements**

The authors are grateful to Oliver Gailing for critically reading Sect. 12.5.4 to Barbara Vornam and Martin Ziehe for assistance in literature search and in word processing, and to Wolfgang Tambour for preparing Figures 7.1, 7.2, 7.5, and 12.4.

July 2006

Reiner Finkeldey  
Hans H. Hattmer

---

# Contents

<b>1</b>	<b>Introduction – Genetics of Tropical Forests</b>	<b>1</b>
<b>Part A</b>	<b>Genetic Processes in Tropical Forests</b>	<b>3</b>
<b>2</b>	<b>Population Genetics – an Overview</b>	<b>5</b>
2.1	Introduction	5
2.2	The Population	5
2.3	Variation at Gene Loci	6
2.3.1	The Molecular Basis of Genetic Variation	6
2.3.1.1	Molecular Markers	7
2.3.1.2	Biochemical Markers – Isozymes	12
2.3.2	The Gene As a Unit of Heredity	13
2.3.3	The Mode of Inheritance	14
2.3.4	Definition and Classification of Gene Markers	15
2.4	Genetic Structures Within Populations	16
2.4.1	Allelic and Genotypic Structures	16
	Example 2.1: Genetic Structures at an Isozyme Gene Locus in <i>Dalbergia sissoo</i>	17
2.4.2	Variation at Uniparentally Inherited Markers	18
	Example 2.2: Diversity of cpDNA Haplotypes in <i>D. sissoo</i>	18
2.5	Evolution and Evolutionary Factors	19
2.6	Phenotypic Variation	21
2.7	Recommended Literature	22
<b>3</b>	<b>Genetic Variation of Tropical Forest Plants</b>	<b>23</b>
3.1	Introduction	23
3.2	Genetic Inventories	23
3.3	Measurement of Genetic Variation	24
3.3.1	Genetic Variation Within Populations	24
3.3.1.1	Genetic Multiplicity	25
3.3.1.2	Genetic Diversity	25
3.3.2	Genetic Differentiation Among Populations	27
3.3.2.1	Genetic Distances	27
3.3.2.2	Genetic Differentiation	27

---

3.4	Genetic Variation in Tropical Forest Species – General Trends	28
3.5	Case Studies on Genetic Variation Patterns	31
	Example 3.1: Genetic Variation at Isozyme Gene Loci in Natural Populations of <i>Acacia auriculiformis</i>	31
	Example 3.2: Genetic Differentiation Among Populations of <i>Eugenia dysenterica</i> from the Brazilian Cerrado	33
	Example 3.3: Genetic Variation of <i>Swietenia macrophylla</i> Across the Brazilian Amazon	33
	Example 3.4: Genetic Variation of Endangered Australian Species of the Genus <i>Fontainea</i> Assessed with RAPD Markers and by DNA Sequences	34
	Example 3.5: Genetic Variation of <i>Shorea leprosula</i> and <i>Shorea parvifolia</i> in Indonesia Assessed at AFLP Loci	35
	Example 3.6: Genetic Variation of <i>Cedrela odorata</i> at cpDNA	37
3.6	Recommended Literature	39
<b>4</b>	<b>Sexual and Asexual Reproduction in Tropical Forests</b>	41
4.1	Introduction	41
4.2	Sexual Reproduction	41
4.2.1	Sexual Types and Sexual Systems	42
4.2.2	Sexual Systems	44
4.2.3	Sexual Structures and Sexual Function	47
4.3	Asexual Reproduction	49
4.3.1	Vegetative Reproduction	50
4.3.2	Apomixis	51
4.4	Recommended Literature	52
<b>5</b>	<b>Gene Flow and Migration</b>	53
5.1	Introduction	53
5.2	Gene Flow Through Pollen	54
5.2.1	Pollination and Fertilization	54
5.2.2	Pollen Vectors in Tropical Forests	54
5.2.2.1	Pollination by Wind (Anemogamy)	54
5.2.2.2	Pollination by Animals (Zoogamy)	55
5.2.3	Pollen Dispersal Assessed by Marker-Based Studies	60
5.2.3.1	Dispersal of Rare Alleles	60
	Example 5.1: Gene Dispersal in a Clonal Seed Orchard of Teak ( <i>Tectona grandis</i> )	60
5.2.3.2	Paternity Analyses	61
	Example 5.2: Gene Flow Through Pollen in <i>Neobalanocarpus</i> <i>heimii</i> and <i>Dipterocarpus tempehes</i> (Dipterocarpaceae)	62
	Example 5.3: Pollen Dispersal in Three Neotropical Tree Species with Different Spatial Distribution Patterns on Barro Colorado Island	63



5.2.3.3	Model-Based Estimates of Pollen Dispersal	64
	Example 5.4: Pollen Dispersal in <i>Dinizia excelsa</i> (Fabaceae)	64
5.2.4	Efficiency of Pollen Vectors for Gene Flow	65
5.3	Migration of Seeds	65
5.3.1	Seed Vectors	65
5.3.1.1	Abiotic Seed Dispersal	65
5.3.1.2	Biotic Seed Dispersal	66
5.3.2	Efficiency of Seed Dispersal	66
5.4	Long-Distance Gene Flow and Migration in Tropical Forest Species	67
5.5	Recommended Literature	68
<b>6</b>	<b>Mating Systems</b>	<b>69</b>
6.1	Introduction	69
6.2	Random Mating and Panmixis	70
6.2.1	Heterogeneity of Pollen Allele Frequencies	70
6.2.2	Panmixis and Hardy–Weinberg Structures	72
6.3	Selfing and Outcrossing Rates	73
6.3.1	Estimates of Selfing Rates Based on Rare Alleles	74
	Example 6.1: Estimates of Selfing Rates in Teak ( <i>Tectona grandis</i> )	
	Populations	74
6.3.2	Estimates of Selfing Rates Based on a Mixed Mating Model	76
6.3.3	Estimates of Selfing Rates Based on Nonmaternal Alleles	76
6.4	Inbreeding and Inbreeding Depression	78
6.4.1	Genetic Consequences of Inbreeding	78
6.4.2	Inbreeding Depression	80
	Example 6.2: Inbreeding Depression in Eucalypts	80
6.5	Incompatibility and Self-Sterility	82
6.5.1	Incompatibility	82
6.5.2	Self-Sterility	84
6.6	Environmental Effects on Mating	84
6.7	Recommended Literature	85
<b>7</b>	<b>Adaptation and Coevolution</b>	<b>87</b>
7.1	Introduction	87
7.2	Physiological and Evolutionary Adaptation	87
	Example 7.1: Viability Selection During Early Life Stages	
	of <i>Platypodium elegans</i>	89
7.3	Species Interactions and Coevolution	90
	Example 7.2: Figs and Their Pollinators	91
7.4	Recommended Literature	97
<b>8</b>	<b>Phylogenies and Evolution Above the Species Level</b>	<b>99</b>
8.1	Introduction	99
8.2	The Evolution of Species Diversity in the Tropics	99

8.2.1	Species Diversity in Tropical Forests	99
8.2.2	Evolution Above the Species Level	101
8.3	Molecular Phylogenies	102
	Example 8.1: A Molecular Phylogeny of Indonesian Dipterocarpoideae	103
	Example 8.2: Evolution of the Genus <i>Inga</i>	104
8.4	Recommended Literature	106
<b>Part B Applications of Genetics to Tropical Forestry</b>		109
<b>9</b>	<b>Fragmentation of Forests</b>	111
9.1	Introduction	111
9.2	The Genetic Status of Fragmented Tree Populations	113
	Example 9.1: Genetic Impact of Fragmentation on a Rain Forest Canopy Tree	116
	Example 9.2: Connectivity of Population Fragments of a Medium-Sized Dry-Forest Tree	119
9.3	Genetic Preconditions for Restoration and Persistence	123
9.4	Recommended Literature	127
<b>10</b>	<b>Genetic Aspects of the Management of Natural Tropical Forests</b>	129
10.1	Introduction	129
10.2	Selective Logging	130
10.2.1	Effective Population Density After Logging	131
10.2.2	The Question of Dysgenic Selection	135
10.3	Natural Regeneration	136
10.4	Genetic Aspects of the Manipulation of Dynamic Processes in Natural Forests	137
10.5	Genetic Aspects of Sustainability in Natural Tropical Forests	140
10.6	Recommended Literature	146
<b>11</b>	<b>Provenance Research</b>	147
11.1	Introduction	147
11.2	Definitions	148
11.3	Historical Development	150
11.4	Provenance Tests	151
11.4.1	Types of Field Experiments in Provenance Research	153
11.4.2	Traits Studied in Field Trials	155
11.4.3	Design, Conduct, and Analysis of Provenance Trials	156
11.4.3.1	Collection of Material for Provenance Trials	156
11.4.3.2	Production of Planting Stock	159
11.4.3.3	Experimental Design	159

11.4.3.4	Choice of Experimental Sites	161
11.4.3.5	Number and Distribution of Locations of an Experimental Series	162
11.4.3.6	Observation Period of Field Experiments	163
11.5	Provenance Differentiation and Geographic Variation Pattern	163
11.6	Choice of Provenances in Tropical Forestry	165
	Example 11.1: Provenances of <i>Eucalyptus camaldulensis</i> and Their Growth Performance	169
	Example 11.2: International Provenance Trials in Teak ( <i>Tectona grandis</i> )	170
11.7	Recommended Literature	172
<b>12</b>	<b>Domestication and Breeding of Tropical Forest Trees</b>	173
12.1	Introduction	173
12.2	Domestication	174
12.3	Genetic Controlledness of Phenotypic Traits	176
12.3.1	Individuals Related by Descent	176
12.3.2	Pair Comparisons in the Field	177
12.3.3	Response to Natural Viability Selection	179
12.4	Linear Model of Genetic Effects on a Phenotypic Trait	179
12.4.1	Phenotypic Trait Expressions	179
12.4.2	Genetic Variance Components and Heritability	182
12.5	Estimation of Breeding Parameters; Progeny Testing	183
12.5.1	Progenies of Open-Pollinated Trees	183
12.5.2	Progeny Tests After Controlled Pollination	184
12.6	Methods of Selection	186
12.6.1	Selection of Plus Trees	187
12.6.2	Selection in Progeny Tests	188
12.6.3	Selection of Clones	189
12.6.4	Marker-Assisted Selection	189
12.7	Heterosis Breeding in Forest Trees	192
	Example 12.1: Breeding Eucalypts in Aracruz, Brazil	194
12.8	Propagation of Breeding Products	196
12.8.1	Clonal Seed Orchards	196
	Example 12.2: Clonal Seed Orchards of Teak ( <i>Tectona grandis</i> )	198
12.8.2	Seedling Seed Orchards	201
12.8.3	Mass Multiplication of Clones	204
12.9	Multiple Population Breeding	205
12.10	Breeding Strategy	207
12.11	Genetic Consequences of Domestication and Breeding	209
12.11.1	General Considerations	209
12.11.2	Monitoring of Breeding Projects by Use of Genetic Markers	210
12.12	Recommended Literature	211

---

<b>13</b>	<b>Genetic Aspects of Plantation Forestry in the Tropics</b>	213
13.1	Introduction	213
13.2	Plantations of Exotic Tree Species	215
	Example 13.1: The Origin of Early Plantations of <i>Acacia mangium</i> in Sabah	217
13.3	Plantations of Indigenous Species	218
13.4	Basic and Reproductive Material	220
13.5	Production and Collection of Seed	222
13.5.1	Seed Production Areas	222
13.5.2	Provenance Resource Stands	223
13.5.3	Seed Orchards	224
13.6	Collection and Storage of Seed	226
13.7	Planting Stock Production	227
13.7.1	Seedlings	227
13.7.2	Clonal Multiplication	228
13.8	Establishment and Development of Plantations	229
13.9	Natural Regeneration of Plantations	230
13.10	Use of Breeding Products	231
13.11	Recommended Literature	235
<b>14</b>	<b>Conservation of Genetic Resources in Tropical Forest Trees</b>	237
14.1	Introduction	237
14.2	Development of Tree-Conservation Genetics	239
14.3	Defining Priorities	240
14.4	Conservation Objectives	241
14.4.1	Objective 1: Preservation of the Potential for Particular Trait Expressions	241
14.4.2	Objective 2: Preservation of Maximum Variation	242
14.4.3	Objective 3: Preservation of Adaptability	243
14.5	Selection of Genetic Resources	243
14.5.1	Inventory of Genetic Marker Loci	245
14.5.2	Inventory of Adaptive Trait Expressions and Adaptive Markers	247
14.6	Conservation Methods	250
14.6.1	Dynamic Conservation in Situ	251
14.6.2	Dynamic Conservation ex Situ in Man-Made Forests	253
14.6.3	Conservation of Seeds in Gene Banks	255
14.6.4	Vegetative Propagation for the Conservation of Forest Genetic Resources	256
14.6.5	Biotechnological Methods of Preservation	256
14.7	Regeneration of Forest Genetic Resources	257
	Example 14.1: Conservation of the Genetic Resources of <i>Pinus merkusii</i> in Thailand	259
14.8	Programs for the Protection of Forest Genetic Resources	264
14.9	Recommended Literature	265
	<b>References</b>	267
	<b>Index</b>	307

. . . the significance of forest genetics lies not only in squeezing extra gain out of marginal investments, but more importantly in understanding the genetics of forests, in appreciating what forest ecosystems are and how they operate, and in better stewards of forest resources (Namkoong 1989).

The importance of forest ecosystems for environmental protection, global biodiversity, and human welfare can hardly be overestimated. Genetic resources are widely recognized as the basis of biodiversity and the main components of ecosystems. Conservation and utilization of genetic resources are crucial aspects of sustainable forest management practices. However, our current understanding of forest genetic resources is poor at best. Genetic variation patterns have been studied for a small fraction of forest trees only. Evolutionary processes resulting in variation within species are difficult to study for long-living organisms such as forest trees, and human impact on forest genetic resources is often subtle.

Thus, it is hardly surprising that misconceptions about forest genetic resources and implications of human impact on genetic resources are commonplace. This holds specifically true for tropical forests, which deserve particular attention for the following reasons. First, tropical forests are centers of biological diversity at least at the level of species diversity. Second, human impact on forest ecosystems is particularly severe in the tropics. Alterations of land use, notably deforestation, endanger forest genetic resources in most tropical countries. Common silvicultural practices such as selective cutting systems based on target diameters threaten genetic resources in managed forests. On the other hand, many tropical countries have launched ambitious reforestation and afforestation programs. The choice of reproductive material used for plantation establishment is a matter of vital importance for the success of reforestation programs. Utilization of genetic resources (tree improvement) plays an important role in this context.

This book is aimed at familiarizing readers with the fundamental principles related to genetic resources in tropical forests. We will concentrate our discussion on forest trees since trees are by definition the main structural and functional

components of forest ecosystems. However, other organisms will also be considered. This holds in particular for animals interacting with trees by distributing their genetic information through the dispersal of pollen or seeds. Priority is given to species of considerable economic importance, mainly trees grown in plantations for wood production (e.g., eucalypts) and keystone species serving vital functions in particular ecosystems (e.g., figs in many Asian forests).

The analysis of patterns of genetic variation is in the center of our discussion. The dynamics of genetic structures of selected forest taxa without human interference is discussed in Part A of this book. The role of the evolutionary factors for shaping genetic variation patterns within and among populations is described with emphasis on the reproduction of forest trees (sexual systems, gene flow, and mating systems).

Part B is devoted to the human impact on forest genetic resources in the tropics. Both deliberate changes of genetic structures owing to tree improvement and genetic implications of other human activities, including forest destruction, fragmentation, and silvicultural practices, are considered.

We describe mutual dependencies between environmental conditions and evolutionary change in primary and secondary forests as well as in plantations. Genetic processes are not seen in isolation but are discussed in an ecological context following the fundamental work of Klaus Stern, the founder of the authors' institute (Stern and Roche 1974). Genetic variation is regarded as a key aspect of biodiversity and as the basis for evolutionary adaptability. The development of management and conservation strategies for natural and man-made forests in the tropics requires the consideration of genetic aspects.

The main focus of the book is on mutual dependencies between environmental conditions in the tropics and genetic variation patterns of selected species, in particular forest trees. The species richness of many tropical forests has numerous implications in this regard.

The enormous species richness of tropical forests is mirrored by multiple strategies to ensure the survival of species in tropical forests. Survival implies adaptedness to prevailing environmental conditions, but also, at least in the long run, adaptability to environmental change. Genetic variation is the basis for evolutionary adaptive change. In this part our current understanding of the dynamics of genetic variation in time and space is reviewed for species living in tropical forests.

Basics of population genetics and the concept of gene marker loci are briefly introduced to readers not familiar with the analysis of genetic variation (Chap. 2). Surprisingly high levels of genetic variation within populations have been reported for the majority of, but not all, tropical forest tree species (Chap. 3). The maintenance of genetic variation during reproduction (Chap. 4) is due to particular features of gene flow and migration mechanisms (Chap. 5), and the mating system (Chap. 6) of tropical forest trees. Changes of genetic structures due to selection are difficult to monitor for long-living organisms growing in complex forest ecosystems, but are crucial for patterns of adaptive variation (Chap. 7). Finally, evolution and genetic differentiation above the species level, i.e., within genera, families, or even higher taxonomic units, are discussed (Chap. 8).

Population geneticists spend most of their time doing one of two things: describing the genetic structure of populations or theorizing on the evolutionary forces acting on populations. On a good day, these two activities mesh and true insights emerge (Gillespie 1998).

---

### 2.1 Introduction

Population genetics provides the framework for an understanding of the dynamics of genetic structures or evolution. The following considerations are neither specific for nor confined to forest species. Most aspects mentioned in this chapter are as relevant for tropical forest species as for any other higher-plant or animal species. However, the population genetics approach taken throughout this book to discuss genetic processes in tropical forests requires a basic understanding of the most important fundamentals concerning genetic structures and their dynamics in time and space, and introduces the concept of genetic markers as well as the most important molecular and biochemical gene markers.

---

### 2.2 The Population

The genetic information or genotype of an organism does not change throughout its lifetime. It is necessary to observe and to compare genetic information of many plants or animals in order to describe genetic variation and its dynamics in time. The fundamental units for studies of genetic variation are collectives of individuals exchanging their genetic information among each other for the sexual production of the next generation. Such units of (sexual) reproduction are defined as populations.

A species rarely consists of one single population. It is usually composed of several more or less isolated populations. Isolation means the absence of mating



events between plants (or animals) from different populations in this context. However, an assessment of reproductive isolation usually requires a fairly detailed understanding of the reproduction system of a species. Furthermore, the definition of a population is operational only if an occasional (but rare) exchange of genetic information among populations is taken into consideration. Thus, interfertile, spatially clustered plants are often regarded as belonging to the same population, while spatially separated plants are regarded as parts of different populations.

The delineation of populations is particularly difficult for immobile organisms with poorly described spatial distribution patterns and largely unknown means of gene dispersal through seed and pollen. In most cases, it is far from obvious which trees in a species-rich tropical forest belong to the same population and how big tree populations are. This has far-reaching consequences for levels of genetic variation of tropical forest trees, as will be discussed in later chapters.

Analyses of spatial patterns of genotypes without a priori information on the delineation of populations or partially isolated subpopulations have recently been proposed as an alternative to the “traditional” approach of dividing the overall variation into components within and among demes or populations (Diniz-Filho and Telles 2002; Manel et al. 2003). Spatially explicit analyses of genetic variation patterns might be particularly informative for plant species occurring in low density in tropical forests; however, currently available information is mainly confined to comparatively small areas (Ng et al. 2004).

---

## 2.3 Variation at Gene Loci

### 2.3.1 The Molecular Basis of Genetic Variation

The genetic information of all organisms is stored as a sequence of the four bases or nucleotides adenine, thymine, cytosine, and guanine in their DNA. It is multiplied by the process of replication and translated to a sequence of amino acids by the processes of transcription (from DNA to messenger RNA, mRNA) and translation (from mRNA to polypeptides) (Griffiths et al. 2000). A cistron is a part of the DNA coding for a particular polypeptide, for example, an enzyme. From this point of view, the variation at a particular gene or gene locus can be assessed by observing differences in the DNA sequence within a cistron. The total number of “genes” is unknown for any tropical forest plant; however, it may be assumed to be at least on the order of more than 25,000 estimated for the first fully sequenced flowering plant, the herbaceous

*Arabidopsis thaliana* (The Arabidopsis Genome Initiative 2000). The number of genes of the first sequenced tree species (*Populus trichocarpa*) is estimated to be even higher (<http://genome.jgi-psf.org/Poptr1/Poptr1.home.html>).

The observation of variation at the most basic level of DNA sequences is routine for tropical trees just as for other organisms, but it is time-consuming and costly if sequences of hundreds of organisms need to be compared. To date, only a few population studies have been conducted based on the observation of variation of DNA sequences within a single or a few closely related tropical forest species (but see, e.g., Ishiyama et al. 2003). Sequencing is more frequently applied to clarify taxonomic subdivisions and phylogenies (Chap. 8).

Numerous alternatives to sequencing are available. The choice of the most appropriate method depends on several factors, including available resources and the purpose of a study. A comprehensive overview of currently available molecular tools to assess patterns of genetic variation in plants has been given by Weising et al. (2005).

---

### 2.3.1.1

#### Molecular Markers

The most important methods to assess patterns of genetic variation within and among species rely on the amplification of short DNA fragments by means of polymerase chain reaction (PCR). PCR allows short fragments of the DNA of a target organism to be selectively amplified (multiplied) for further analyses. The sequence of two short oligonucleotides (primers), each usually between ten and 25 base pairs in length, is decisive for the amplified DNA region. The following marker types, all based on the PCR method, have gained particular importance in the study of genetic variation patterns of tropical forest trees.

#### (Partial) Gene Sequences and Single Nucleotide Polymorphisms

Selective primers can be designed to amplify the DNA coding for a particular gene (cistron). DNA regions translated into a polypeptide (exons) and non-translated regions (introns) of structural genes as well as regulatory genes can be analyzed. It is possible to sequence the amplified PCR products, often after a cloning step. For example, Kamiya et al. (2005) investigated the phylogeny of the species-rich genus *Shorea* (Dipterocarpaceae) on the basis of a (partial) sequence of the *PgiC* gene (Example 8.1).

An investigation of genetic variation is particularly rewarding at gene loci with known or putative function and a directly observable effect on phenotypic traits responsible for the adaptation to particular environmental conditions (candidate genes). Populations can be screened for single nucleotide polymorphisms (SNPs) without the need for repeated sequencing (Morin et al. 2004).

The characterization of expressed sequence tags, studies of molecular variation at the level of expressed genes, and analyses of variation at SNPs are expected to greatly improve our understanding of adaptive genetic variation and its dynamics in space and time. However, these recently developed marker types are not discussed in detail owing to the lack of published reports for tropical forest species (but see Sect. 7.2).

### Microsatellites (Simple Sequence Repeats)

Microsatellites or simple sequence repeats (SSRs) are short DNA fragments of usually only two or three base pairs in length which are repeated several times in a particular location of the DNA (Fig. 2.1). SSRs can be studied by the development of primers in conserved DNA regions flanking the microsatellite. The development of primers is rather costly and time-consuming (Squirrell et al. 2003). Nuclear microsatellites developed for a particular species can only be transferred to closely related species, usually within the same genus.

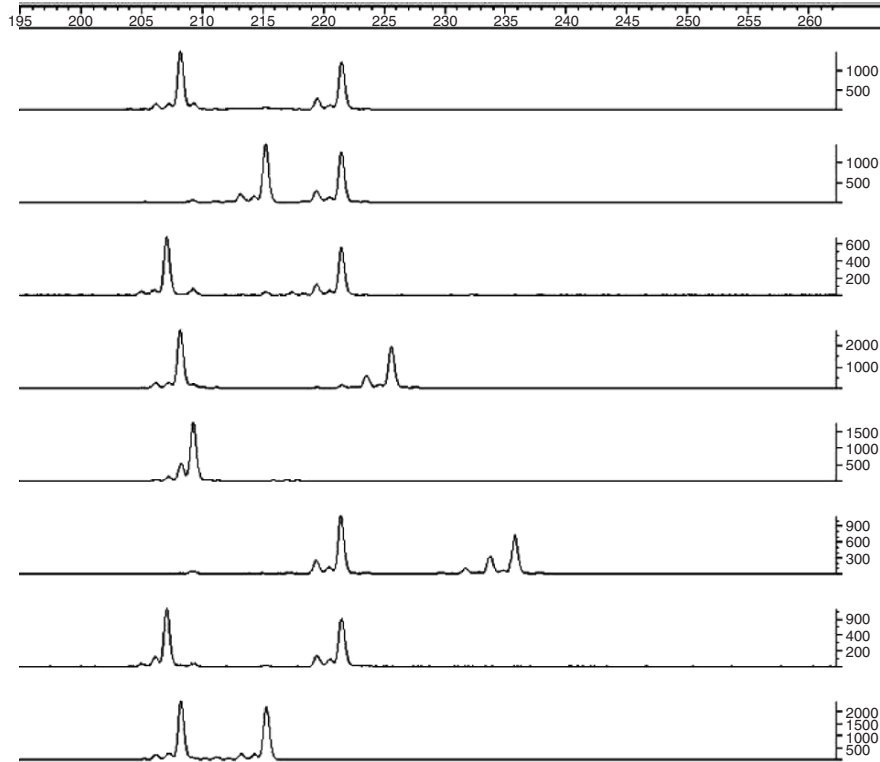
Nuclear microsatellites are important genetic markers owing to their usually high variability and codominant mode of inheritance (Sect. 2.3.2, Chap. 3).

The variation of microsatellite fragment sizes is due to a different number of repeated motifs resulting in slightly different sizes of amplified fragments (Fig. 2.2). The availability of highly variable markers is particularly useful for analyses of the mating system and gene flow (Chaps. 5, 6). Microsatellites have been developed for a number of tropical forest tree species, including *Pithecellobium elegans* (Chase et al. 1996a), *Swietenia humilis* (White and Powell 1997), *Gliricidia sepium* (Dawson et al. 1997), *Shorea curtisii* (Ujino et al. 1998), *Caryocar brasiliense* (Collevatti et al. 1999), *Neobalanocarpus heimii* (Iwata et al. 2000), and *Prosopis* spp. (Mottura et al. 2005).

Microsatellite motifs are not only found in DNA of the nucleus, but also in DNA of chloroplasts (cpDNA; Sect. 2.3.2). Single nucleotide repeats are particularly common in cpDNA. Universal primers are available to investigate variation at cpSSRs in many different species of angiosperms (Weising and Gardner 1999) and gymnosperms (Vendramin et al. 1996).

```
AAGGATAAGTTAAA ACACACACACACACAC GTTGCTCCATTT
TATGATGTATGAAT ACACACACACACACAC CCCACCTGGTTTT
TTGACGTCACACAG ACACACACACACACAC TCTTCATCCACA
```

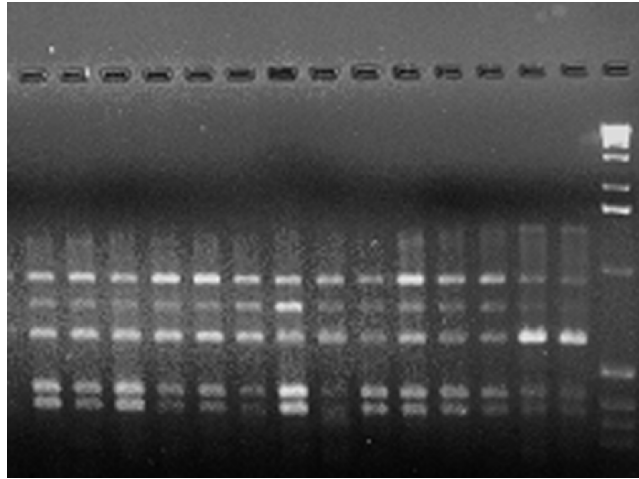
**Fig. 2.1.** Three different DNA sequences of *Prosopis chilensis* containing the microsatellite repeat motif AC in nine (*lower* and *middle*) or ten (*upper*) copies (*underlined*). A adenine, C cytosine, G guanine, T thymine. (From Mottura, unpublished)



**Fig. 2.2.** Length variation for eight *Prosopis* spp. trees at a microsatellite gene locus (Mottura et al. 2005) separated in an automated sequencer (ABI 3100). Each line refers to one tree. Different sizes of amplified fragments are visualized as peaks at different positions in the chromatogram. Trees with two amplified fragments are heterozygous; tree 5 is homozygous with only one amplified fragment. (From Mottura, unpublished)

### Random Amplified Polymorphic DNA and Amplified Fragment Length Polymorphism

Random amplified polymorphic DNA (RAPD; Newbury and Ford-Lloyd 1993; Fig. 2.3) and amplified fragment length polymorphism (AFLP; Vos et al. 1995; Fig. 2.4) are genetic fingerprinting techniques resulting in more or less complex DNA patterns. Unlike most other techniques including microsatellites no previous sequence information is necessary for RAPD and AFLP studies. The presence (+ or 1) or absence (– or 0) of a DNA fragment of a particular size is scored for each sample plant investigated. This information is the basis for a matrix of size  $n \times m$  ( $n$  is the number of plants investigated;  $m$  is the number of DNA fragments observed) with elements 1 (fragment present) or 0 (fragment absent). RAPDs and AFLPs are usually interpreted as dominant markers (see later). No



**Fig. 2.3.** Low variation at a random amplified polymorphic DNA marker in teak (*Tectona grandis*). (From Finkeldey, unpublished)

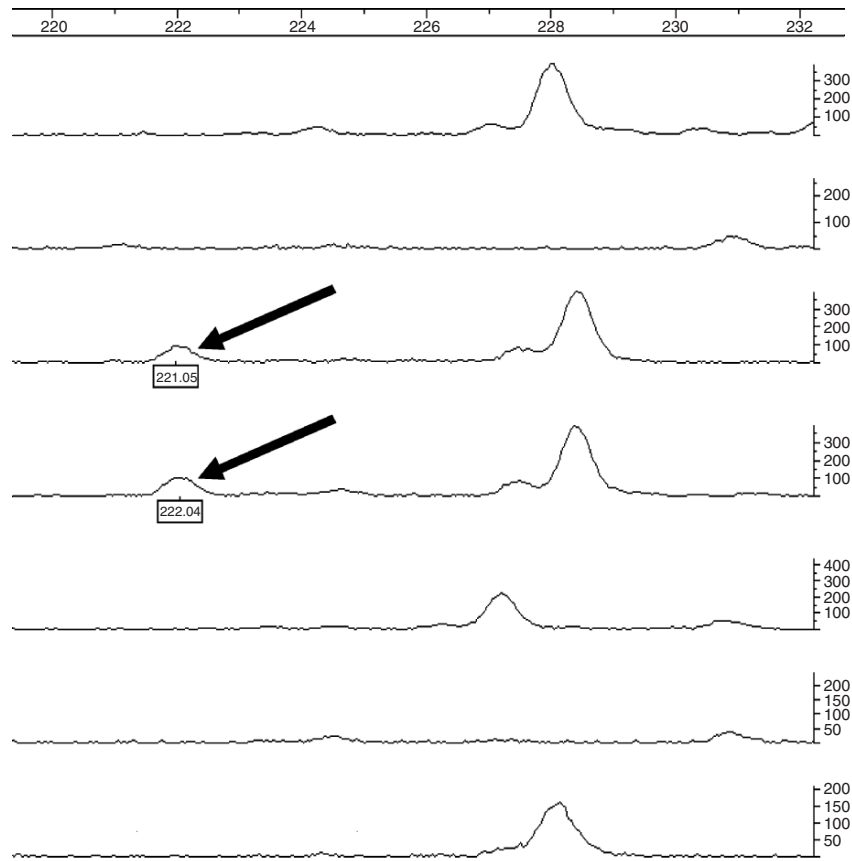
information on the function of a particular DNA fragment or its mode of inheritance is available; hence, RAPDs and AFLPs are “anonymous” markers.

The RAPD technique is a simple molecular tool to assess genetic variation within and among populations which has been widely used for tropical forest plants such as *Gliricidia* spp. (Chalmers et al. 1992), *Cedrela odorata* (Gillies et al. 1997), *Caesalpinia echinata* (Cardoso et al. 1998), *Prunus africana* (Dawson and Powell 1999), and many others. The reproducibility of the RAPD technique has been under dispute (Rabouam et al. 1999).

The AFLP method is more demanding with regard to laboratory equipment and experience, but allows more DNA fragments to be investigated from a single PCR reaction, and shows higher reproducibility in comparison with RAPDs. The potential of the AFLP technique to assess genetic variation patterns in tropical trees was recently reviewed by Kremer et al. (2005). AFLP studies to assess genetic variation within tropical tree species were conducted, for example, for *Moringa oleifera* (Muluvi et al. 1999), *Calycophyllum spruceanum* (Russel et al. 1999), and *Acer skutchii* (Lara-Gomez et al. 2005).

### **Restriction Fragment Length Polymorphism**

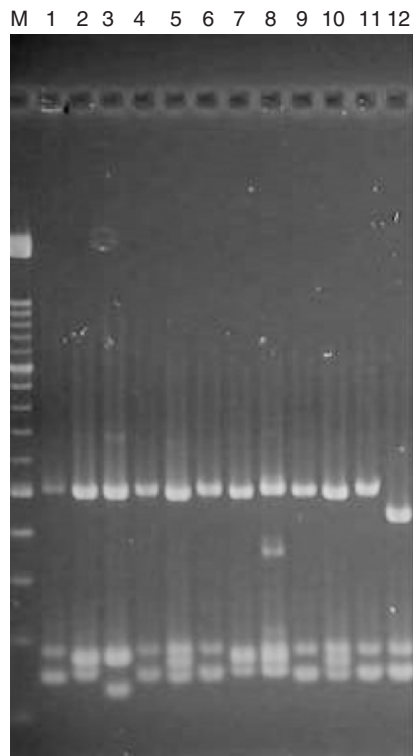
A simple method to observe variation among DNA fragments from different plants is to digest the fragments into smaller pieces by restriction enzymes. Restriction enzymes cut DNA at a particular sequence. These recognition sequences are usually between four and six base pairs in length. The restriction fragment length polymorphism (RFLP) technique is based on the use of restriction enzymes to observe genetic variation. For example, nuclear RFLPs



**Fig. 2.4.** Variation at amplified fragment length polymorphisms in seven *Shorea parvifolia* trees after separation in an automated sequencer (ABI 3100). Only fragments in the range 220–232 bp are shown. The fragment with a size of 220 bp (*arrows*) is “diagnostic” for the population Berau (Borneo), since it was not observed in six other populations. (From Cao, unpublished)

were developed for *Acacia mangium* to assess genetic diversity and differentiation among populations (Butcher et al. 1998) and with the objective to incorporate marker-based approaches in breeding programs (Butcher et al. 2000; Butcher 2004).

More frequently, the RFLP techniques is combined with a PCR step and is used to assess variation of cpDNA. For example, Tsumura et al. (1996) and Indrioko et al. (2006) (Fig. 2.5) analyzed phylogenetic relationships among Southeast Asian dipterocarps using the PCR-RFLP technique (Example 8.1), and Cavers et al. (2003) observed strong genetic differentiation among five cpDNA haplotypes of *Cedrela odorata* in Central America (Example 3.6).



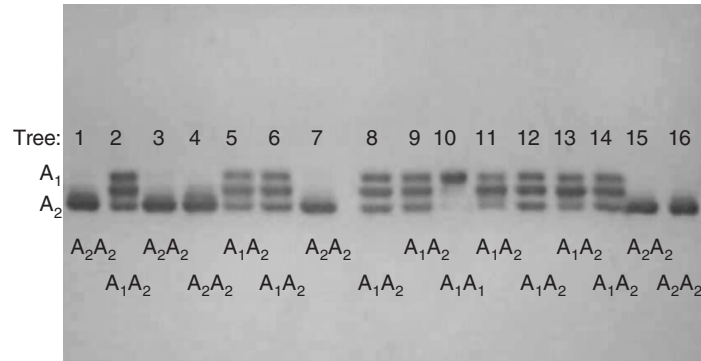
**Fig. 2.5.** Polymerase chain reaction restriction fragment length polymorphisms of Indonesian Dipterocarpaceae after amplification with the primer pair *TrnL–TrnF* and restriction with *TaqI*. *M* marker (size standard), 1 *Hopea celebica*, 2 *Vatica rassak*, 3 *V. pauciflora*, 4 *H. bancana*, 5 *Dipterocarpus oblongifolius*, 6 *S. javanica*, 7 *V. granulata*, 8 *Upuna borneensis*, 9 *S. mecistopteryx*, 10 *D. rigidus*, 11 *S. mecistopteryx*, 12 *H. grifithii*. (From Indrioko, unpublished)

### 2.3.1.2

#### Biochemical Markers – Isozymes

The sequence of nucleotides of the DNA is converted to a sequence of amino acids of polypeptides by the processes of transcription and translation. Enzymes are the most important group of polypeptides catalyzing all kinds of biochemical reactions in the metabolism; thus, there is a close relation between enzymes and controlling genetic information according to the “one gene – one polypeptide hypothesis.” Isozymes are enzymes with similar or even identical functions. Electrophoresis of isozymes followed by biochemical staining is a simple way to observe differences with regard to their electric load and/or spatial structure (Rothe 1994; Fig. 2.6). It is often possible to directly relate the variation observed by isozyme electrophoresis to genetic variation (Bergmann and Hattemer 1998). A formal inheritance study (Sect. 2.3.2) is strongly recommended prior to the use of isozymes as genetic markers.

Isozyme gene loci became the first widely used markers to assess patterns of genetic variation of tropical forest species at single gene loci. The ease of isozyme



**Fig. 2.6.** Zymogram showing variation of 16 *Dalbergia sissoo* trees from Nepal at the enzyme system alcohol dehydrogenase. The variation is controlled by a gene locus *ADH-A*. The inferred genotypes of the respective trees are indicated. (From Pandey, unpublished)

inventories and the comparatively low costs have made them useful tools for population genetics studies, although they were often substituted by various types of molecular DNA markers during the last decade.

### 2.3.2

#### The Gene As a Unit of Heredity

Genetics as the science of heredity dates back to the experiments of the Austrian monk Gregor Mendel (1866), if not earlier. The definition of a gene as a unit of heredity was introduced in the early twentieth century, long before the role of DNA as the material basis of genetic information was recognized. A gene in this sense is identified by the observation of segregation within progenies of particular parents.

The basic idea for the identification of a gene as a unit of heredity has remained unchanged since the experiments Mendel (1866). He investigated segregation ratios in progenies after controlled pollination at simple morphological traits such as the color of petals in peas.

If a plant is heterozygous at a controlling gene locus, it will transmit only one of the two different alleles to a particular progeny (the terms heterozygosity and allele are described later). Thus, segregation is expected in progenies if at least one parent was heterozygous at a controlling gene locus. It is possible to compare the observed segregation in a sample of progenies after controlled pollination with an expectation based on the assumption of a simple “Mendelian” inheritance of the trait. The controlling gene or the gene locus is identified, if the



differences between the observed and expected values are not significant (Hattermer 1991). Suitable statistical tests are, for example, described by Sokal and Rohlf (1998, p. 686 ff.;  $\chi^2$  test or goodness-of-fit test, G test).

The formal identification of a genetic marker locus by an observation of segregation is possible only for environmentally stable, genetic traits. It is also advisable for the identification of gene loci based on the observation of biochemical or molecular markers. For example, Moran and Bell (1983) observed segregation at the isozyme system *GOT* in progenies of *Eucalyptus regnans* after controlled pollination and found no evidence to reject the hypothesis of the genetic control of the variation by a single gene locus.

Forest trees, in particular those of the tropics, are extraordinarily difficult to cross owing to their large size, short viabilities of pollen, unknown reproductive biology, and other obstacles. Thus, the identification of a gene locus by inheritance studies is preferentially conducted on the basis of alternative approaches such as the observation of segregation in the haploid megagametophyte (the “endosperm”) of gymnosperms (Bergmann 1974) or the observation of segregation in the progenies of putatively heterozygous seed trees after open pollination (Gillet and Hattermer 1989). The first method has been used to identify enzyme gene loci in *Pinus merkusii* from Thailand (Changtragoon and Finkeldey 1995b); the latter approach was used to clarify the inheritance of isozymes in *Pterocarpus indicus* from the Philippines (Finkeldey et al. 1998).

---

### 2.3.3

#### The Mode of Inheritance

The analysis of variation patterns at environmentally stable, genetic traits (Chap. 3) requires an understanding of the transmission of the trait from parents to progenies and from one generation to the next. Two considerations are of particular importance in this context: the transmission of genetic information from both parents or from a single parent only, and the impact of a single or two alleles at a gene locus on the observed trait expression (the phenotype).

Most of the DNA in a plant cell is located in the nucleus. Nuclear DNA (nDNA) is inherited by both parents. For diploid organisms, one set of chromosomes is transmitted by each of the two parents to their common progenies; thus, nDNA is typically **biparentally** inherited. In addition, DNA of plants is located in two types of plastids: mitochondria (mitochondrial DNA, mtDNA) and chloroplasts (cpDNA). DNA in plastids, both mitochondria and chloroplasts, is usually transmitted by a single parent to its progenies (uniparental inheritance). mtDNA is mainly inherited from the female or seed parent only

(**maternal** inheritance). The same holds for cpDNA in angiosperms. However, cpDNA of gymnosperms is typically **paternally** inherited, i.e., only the genetic information of the pollen parent is transferred to a progeny. cpDNA was investigated in most studies on genetic variation of tropical forest plants based on uniparentally inherited markers (Example 3.6). Maternally inherited mtDNA was studied in several animal species such as Central American frogs (Crawford 2003) and the giant panda in China (Lu et al. 2001).

The different types of genetic information at any biparentally inherited gene locus are called **alleles**. Since two sets of chromosomes are inherited (one from the seed parent, the other from the pollen parent), each nuclear gene is represented in two copies in each progeny of a diploid species. Polyploids are not discussed in this introduction. A tree is **homozygous** (e.g.,  $A_1A_1$ ) at a particular locus A, if two identical alleles (both  $A_1$ ) were inherited from its parents. A **heterozygous** tree (e.g.,  $A_1A_2$ ) received two different alleles ( $A_1$  from the seed parent and  $A_2$  from the pollen parent or vice versa) from its parents.

An allele is **dominant**, if its possession in a single copy is sufficient to express a particular phenotype. If  $A_1$  is dominant, the phenotypic expressions of trees with the genotypes  $A_1A_1$ ,  $A_1A_2$ ,  $A_1A_3$ , etc. are identical. The other alleles ( $A_2$  and  $A_3$  in the example) are **recessive**. If it is possible to distinguish all heterozygotes (e.g.,  $A_1A_2$ ) from all homozygotes ( $A_1A_1$  and  $A_2A_2$ ), the alleles are defined as **codominant**. A simple observation of genetic structures (Sect. 2.4) is only possible at codominant marker loci.

#### 2.3.4

##### Definition and Classification of Gene Markers

On the basis of the previous considerations, it is possible to define a gene marker in different contexts:

- A **gene marker** (synonyms gene locus and marker locus) **in a wide sense** is an environmentally stable trait; thus, the variation is determined by genetic factors only.
- A **gene marker in a narrow sense** is an environmentally stable, biparentally inherited, codominant trait.

Although many different traits may be regarded as gene markers in a wide sense, only the molecular and biochemical markers briefly described in Sect. 2.3.1 have considerable practical importance to assess genetic variation patterns of tropical forest species. A rough classification of important, currently available marker types with regard to their mode of inheritance and their variability is presented in Table 2.1.

**Table 2.1.** Classification of genetic markers in a broad sense commonly used in studies on genetic variation of tropical forest species with regard to their mode of inheritance and their variability

Variation	Uniparental	Anonymous <sup>a</sup>	Codominant <sup>b</sup>
High	Chloroplast simple sequence repeats; sequences of chloroplast DNA	Amplified fragment length polymorphisms	Microsatellites (simple sequence repeats)
Moderate to low	Polymerase chain reaction restriction fragment length polymorphisms of chloroplast DNA and mitochondrial DNA	Random amplified polymorphic DNAs	Isozymes

<sup>a</sup>Mode of inheritance usually unknown; interpretation as dominant markers<sup>b</sup>Gene markers in a narrow sense

## 2.4 Genetic Structures Within Populations

The final result from a laboratory study to assess the genetic constitution of a plant at a particular biochemical or molecular gene marker system is a specific pattern, which can be used:

- To infer a genotype in case of gene markers in a narrow sense (e.g., isozymes, SSRs)
- To identify a haplotype in the case of uniparentally inherited markers (e.g., PCR-RFLPs of cpDNA)
- To assign a genetic fingerprint usually consisting of many anonymous marker loci to the plant (e.g., AFLPs)

Genetic structures are frequency distributions of such genetic types in populations. The most important frequency distributions to assess levels of genetic variation within and among populations are allelic and genotypic structures.

### 2.4.1 Allelic and Genotypic Structures

Genetic variation is assessed by an investigation of all plants or a (random) sample of plants from a population. At a gene marker in a narrow sense, the genotypes of these plants are known, and it is possible to calculate the (relative) frequency of a particular genotype  $A_iA_j$  as

$$P_{ij} = N_{ij}/N, \quad (2.1)$$

where  $N_{ij}$  is the (absolute) frequency of genotype  $A_iA_j$  in the population (or sample) and  $N$  is the population size (or sample size).

The frequency vector of all (relative) frequencies of genotypes is defined as the **genotypic structure** of the population at the respective marker gene locus. It obviously holds that  $\sum_i \sum_j P_{ij} = 1$ .

Each genotype of a diploid species consists of two alleles; thus, the total number of alleles in a population (or a sample) is twice the number of plants investigated, and the frequency of a particular allele  $A_i$  can be calculated as

$$p_i = N_i/2N, \quad (2.2)$$

where  $N_i$  is (absolute) frequency of allele  $A_i$  in the population (or sample) and  $N$  is the population size (or sample size). The respective allele  $A_i$  is counted twice in case of homozygous plants  $A_iA_i$ .

The frequency vector of all (relative) frequencies of alleles is defined as the **allelic structure** of the population at the respective marker gene locus. Again, it holds that  $\sum p_i = 1$ .

If the genotypic structure of a population is known, it is easily possible to compute the allelic structure at the respective gene locus. The frequency of allele  $A_i$  ( $p_i$ ) can be computed by calculating the sum of the frequency of the corresponding homozygote ( $A_iA_i$ ) and half of the frequencies of all heterozygotes where the respective allele occurs ( $A_iA_j, A_iA_k, \dots$ ). However, the calculation of genotypic structures based on allele frequencies requires additional information or assumptions (e.g., Sect. 6.1.2).

**Example 2.1:** Genetic Structures at an Isozyme Gene Locus in *Dalbergia sissoo*

The observed variation at the isozyme gene locus *ADH-A* is illustrated for 16 *Dalbergia sissoo* trees in Fig. 2.6. Three different genotypes were observed. Six trees (nos. 1, 3, 4, 7, 15, and 16) show the homozygote genotype  $A_2A_2$ , tree 10 is the only tree with genotype  $A_1A_1$ , and the remaining nine trees exhibit the heterozygous genotype  $A_1A_2$ . Thus, the **genotypic structure** of the sample is as follows:

$$\begin{aligned} P_{11} &= 1/16 = 0.0625, \\ P_{12} &= 9/16 = 0.5625, \\ P_{22} &= 6/16 = 0.3750. \end{aligned}$$

Eleven of the 32 observed alleles (each tree has two alleles) are  $A_1$ , the remaining 21 alleles are of type  $A_2$ . The **allelic structure** may also be computed as follows:

$$\begin{aligned} p_1 &= P_{11} + 0.5P_{12} = 0.0625 + 0.5625/2 = 11/32 = 0.34375, \\ p_2 &= P_{22} + 0.5P_{12} = 0.375 + 0.5625/2 = 21/32 = 0.65625. \end{aligned}$$

### 2.4.2

#### Variation at Uniparentally Inherited Markers

The uniparental inheritance of cpDNA and mtDNA implies the absence of recombination among genes. Since cpDNA and mtDNA are usually transmitted without any changes in maternal or, in the case of cpDNA of gymnosperms, paternal lineages, it is possible to interpret variation patterns observed at different loci together and to assign particular cpDNA or mtDNA **haplotypes** to different plants. The interpretation of genetic diversity within and among populations (Chap. 3) is conventionally based on these haplotypes rather than on variation at particular loci.

#### Example 2.2: Diversity of cpDNA Haplotypes in *Dalbergia sissoo*

Six to ten seeds from randomly selected, different seed trees were investigated in each of ten populations of *D. sissoo* in Nepal. The inventory was conducted in five regions, each represented by one natural population and a neighboring plantation. Variation was detected by the PCR-RFLP technique (Sect. 2.3.1) by restriction of the amplified fragment *Trn K1/K2* with the enzymes *AluI* and *RsaI*, and by amplification of the chloroplast microsatellites (cpSSRs) *ccmp6* and *ccmp7* (Weising and Gardner 1999). A total of eight haplotypes with two or three haplotypes in each population were observed (Table 2.2). Haplotypes 1–3 were only observed in natural populations, while haplotypes 4–8 occurred mainly in plantations (Pandey et al. 2004).

**Table 2.2.** Variation of chloroplast DNA haplotypes in five natural populations and five plantations of *Dalbergia sissoo* in Nepal. (Adapted from Pandey et al. 2004)

Region	Population	N	Chloroplast DNA haplotypes							
			1	2	3	4	5	6	7	8
<b>Natural populations</b>										
A	Hetauda N	10	9	1						
B	Shivapur	10	8		2					
C	Hattisar	6	5	1						
D	Godawari	10	4		6					
E	Pipariya	10			9	1				
<b>Plantations</b>										
A	Hetauda P	10					4	6		
B	Surai	10				1			7	2
C	Thakurdwara	10				4			3	3
D	Attaria	10				6			4	
E	Shuklaphanta	10				5			5	

N sample size

---

## 2.5 Evolution and Evolutionary Factors

Today's biodiversity is the result of the biological evolution which started at least three and a half billion years ago on earth. The number of existing species is still unknown, since a large proportion of the species living on our planet have not yet been described. The majority of "unknown" species are insects living in the canopy of tropical forests (Erwin 1988). Current estimates of the global species diversity vary from approximately two million to 30 or even 50 million species (Erwin 1997).

The long history of evolution on earth does not preclude the study of its effects in short periods. Evolution is permanently ongoing and it is possible to observe its effects even in long-living organisms such as forest trees. Furthermore, the majority of evolutionary change is within existing species and does not immediately result in the creation of new taxa.

The main principle of evolution was recognized and described by Charles Darwin (1859): Biological evolution is based on heritable variation within species or populations. Darwin recognized the key role of natural selection in this context. Fundamental observations, which led to the development of the selection theory, were made both by Darwin and Alfred Russel Wallace, who independently developed a similar theory, after extended studies in tropical areas (Lefèvre 1984; Sect. 8.2).

Our advanced knowledge about the process of inheritance allow us to define evolution more precisely today than during Darwin's era. **Evolution is a change of the genetic (allelic or genotypic) structure of a population** at one or several gene loci. Only changes improving the adaptedness of a population to particular environmental conditions were regarded as evolution by Dobzhansky et al. (1977); however, Kimura (1983) proposed including all changes of genetic structures including those at selectively neutral loci into the concept.

Evolutionary factors are the causes of changes of genetic structures (Hedrick 2000). The significance of evolutionary factors for the dynamics of genetic structures of tropical forest plants is the main topic in later chapters of this book; thus, only a brief description is given here:

- **Mutation** is the prerequisite for any genetic variation. Mutations are random changes of genetic information of an organism which might affect a single nucleotide (a spontaneous change of a single nucleotide resulting in a SNP; Sect. 2.3.1), the chromosomal structure (e.g., fissions, translocations, inversions), or the number of genes (e.g., duplications, polyploidy). Most mutations are detrimental or neutral for an organism. Only few advantageous mutations increase the fitness, i.e., the capacity to produce offspring.