Genetic structure and diversity of *Rhododendron arboreum* (Ericaceae) in protected and harvested forests in Northeast India

Hadi Bandali

A Thesis

in

The Department

of

Biology

Presented in Partial Fulfillment of the Requirements

for the degree of Master of Science at

Concordia University

Montreal, Quebec, Canada

August 2008

Your file Votre référence
ISBN: 978-0-494-63253-6
Our file Notre référence
ISBN: 978-0-494-63253-6

ABSTRACT

Genetic structure and diversity of *Rhododendron arboreum* (Ericaceae) in protected and harvested forests in Northeast India

Hadi Bandali

Rhododendron arboreum is an ecologically and economically important forest tree species in high altitude forests. In the Himalayan region of Northeast India, *R. arboreum* is extensively harvested as a source of fuelwood while selected forest stands have been set aside as protected forests by tribal groups resulting in a mosaic of harvested and protected forests. I assessed the genetic structure of *R. arboreum* in these forests using microsatellite markers that were originally developed for *R. metternichii*. Out of eight pairs of primers evaluated, three primer pairs showed reliable amplification and polymorphism in *R. arboreum*. These three primer pairs or loci (RM1D1, RM2D2, and RM9D6) were used in assessing the genetic structure of *R. arboreum* in protected and harvested forests in the Eastern Himalayan region of Northeast India. Leaf samples collected from over 125 trees in five populations of *R. arboreum* representing two protected and three harvested forests, 35 different alleles were detected at three loci, with an average of 11.66 alleles per locus. There were no significant differences in number of alleles between populations. Over all populations, the mean observed and expected heterozygosity values were 0.801 and 0.804, respectively. A total of five private alleles were detected in two of the five populations (Pipraw Jang disturbed and Falokchar Mukto disturbed), suggesting private alleles are confined to harvested populations. The values of population differentiation (Fst) ranged

between 0.015 and 0.057 with a corresponding gene flow (Nm) values of 15.8 and 4.1 respectively suggesting moderate levels of gene flow among populations. Population differentiation showed a significant correlation with the geographic distance.

The high level of heterozygosity suggests a predominantly outcrossing mating system in *R. arboreum*. Although the level of population differentiation is low, the significant correlation between population differentiation (Fst) and geographic distance within a short distance of about 5 km is striking. In order to capture and maintain maximum genetic diversity, representative populations of *R. arboreum* distributed throughout its geographic distribution range should be conserved. These populations may serve as a seed source for regeneration of adjacent harvested forests. Further studies are needed to gain insights into the source of these private alleles as well as average distance of gene flow to determine the optimal distance for establishing protected populations of *R. arboreum*.

Acknowledgements

I would like to express my deep sense of gratitude and sincere appreciation to my supervisor Dr. Selvadurai Dayanandan for giving me this opportunity and his constant encouragement and guidance throughout the course of the study. I remain grateful to him for his tireless effort and help rendered to me in successful completion of the research.

I extend my gratitude to my committee members and external examiner Dr. Patrick Gulick , Dr. William Zerges, and Dr. Paul Widden for their constructive critics and kind support.

I also express my sincere gratitude to Dr. M. L. Khan, Dr. Ashish Paul, and Mr. Jyotishman Deka, Department of Forestry, North Eastern Regional Institute of Science & Technology, Nirjuli, Arunachal Pradesh, India, Dr. Baharul Choudhury, Dr. Atiqur Barbhuiya, Dr. Mona Hamzeh (Concordia University), Dr. Ramesha B.T (Department of Crop Physiology, University of Agricultural Sciences, Bangalore, India) for their valuable guidance during the course of study and compilation of data sets in the thesis. I also wish to thank all my colleagues in the lab Bahar, Margo, Mona, Nathalie and Sandra for their help and friendship.

I wish to express my special thanks to Concordia Centre for Structural and Functional Genomics for laboratory resources. This study was financially supported by NSERC and FQRNT research grants and a Concordia University teaching assistantship.

I feel pleasure to acknowledge my aunt Mrs. Malihe and her family members Mr. Habib, Roksana, Saman, and Saviss for their moral support in Montreal.

Finally my heartfelt appreciation goes to my parents Mr. Davood and Mrs. Najmiyeh for their love, affection, constant encouragement, patience and time to time financial support.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

A- alleles per locus

bp- base pairs

DNA- deoxyribonucleic acid

He- expected heterozygosity

Ho- observed heterozygosity

HWE- Hardy- Weinberg equilibrium

PCR- Polymerase chain reaction

1. Introduction

1.1 Rhododendron

The genus *Rhododendron* is an ecologically and economically important group of plants that occur in high altitude forests. The term '*Rhododendron*' comes from the Greek words 'rhodo' meaning 'rose' and 'dendron' meaning 'tree', in combination 'rosetree' (Hora, 1981). The height of mature individuals of this genus varies from a few centimeters (*R. pumilum*) to over 25 meters (*R. arboreum*) (Tiwari and Chauhan, 2005). *Rhododendrons* may be evergreen, semi deciduous and deciduous shrubs and trees, and many species are known to have adaptations to seasonally cold climatic conditions and highly acidic soil where nutrients are less available to plants. Species of *Rhododendron* grow in wide varieties of habitats including the forest floor, stream sides, marshes, ridges, glades, cliffs, rocks and boulders, open meadows and thickets, and mountain tops. Some of these species grow epiphytically on trees at all levels ranging from the lower trunk to the topmost branches (Hora, 1981).

The genus *Rhododendron* includes over 1200 species and is considered to be the largest genus of the family Ericaceae, with a distribution ranging from Northeast Asia and Eurasia to Western Europe and North America (Rotherham, 1983; Clinton and Vose, 1996). *Rhododendrons* in Asia are widely distributed, and geographically range from the northwestern Himalaya through Nepal, Sikkim, eastern Tibet, Bhutan, Arunachal Pradesh, upper Myanmar, western and central China. More than 90% of the world's natural populations of *Rhododendrons* occur in this region (Singh et al., 2003; Pradhan and Lachungpa, 1990).

Reportedly, 98% of the *Rhododendron* species in India exist in the Himalayan region (Singh et al., 2003). Out of 1200 species known to occur throughout the world, 72 species, 20 subspecies and 19 varieties have been reported from India. The eastern

Himalaya harbors 71 species, as compared to the western Himalayas, with 8 species. The *Rhododendrons* occur between 1200 m (e.g., *R. dalhousiae*, *R. dendricola*) and 6000 m (e.g., *R. nivale*, *R. leptocarpum*), but the majority of the species are found between 2200 and 4000 m. Amongst the Indian species, *R. arboreum* is the most widely distributed and occur from the western to the eastern Himalayan region and other neighboring countries. Southern India houses only one subspecies (*R. arboreum* ssp. *nilagiricum*), in the Nilgiri Hills of the Western Ghats, Tamil Nadu (Mao et al., 2001). *Rhododendron* species are reported to be self- as well as cross-compatible showing the absence of any pre- or post-fertilization barriers (Jain et al., 2000).

Rhododendrons are mostly used by the local inhabitants (Bhutias, Lepchas and Nepalis) of the Indo- Himalayan region. Apart from aesthetic and sacred values, *Rhododendrons* also have medicinal and economic values. The dried flowers of *R. arboreum* are considered to be highly efficacious in treating diarrhea and blood dysentery (Pradhan and Lachungpa, 1990). The wood of *R. arboreum* is used for making '*khukri*' handles (a curved Nepalese knife), pack-saddles, gift-boxes, gunstocks and posts (Pradhan and Lachungpa, 1990).

Among 72 *Rhododendron* species in India, 58 are classified in the endemic, endangered, rare or threatened categories. Most of these species are shrubs and some of them are epiphytic or small trees. *Rhododendrons* growing in high altitude areas are subject to many disturbances due to various natural and anthropogenic factors (Mao et al., 2001; Singh et al., 2003). Natural threats include landslides and forest fire. Heavy snowfall is known to negatively affect the growth of *Rhododendrons* at higher elevations especially in sub alpine and alpine regions. Young trees, saplings and seedlings are the most impacted by avalanches due to low tolerance to the rushing mass of snow (Mao et al., 2001; Singh et al., 2003). Besides natural threats, various anthropogenic factors such as unsustainable extraction for fuelwood, clear felling, logging, and agricultural

practices have collectively intensified the pressure on *Rhododendron* habitat (Mao et al., 2001).

Among the *Rhododendron* species known to occur in the eastern Himalayas, the widely distributed *R. arboreum* is harvested as a source of fuelwood (Figure 1). The widespread harvesting of natural stands of *R. arboreum* could have ecological and genetic consequences with a potential loss of genetic diversity of the species. Genetic diversity is essential for the long-term survival of tree species to enhance adaptive ability and reduce the risk of extinction (Hamrick, 1994). Thus, the information on levels of genetic diversity in natural populations of the target species is invaluable for the formulation of appropriate conservation strategies (Lemes et al., 2003). Although *R. arboreum* is widely harvested in Northeast India (Paul, 2008), and there is a pressing need to assess the conservation status of the species, no information on the genetic structure and diversity of *R. arboreum* in the eastern Himalayas is available. Thus, there is an urgent need to study the levels of genetic diversity of *R. arboreum* in different populations of eastern Himalayas.

DNA markers have been extensively used to assess genetic diversity and gene flow in natural populations of tropical tree species (reviewed in Nybom, 2004 and Mariette et al., 2002). Among the different types of DNA marker, microsatellite markers or simple sequence repeats (SSR's), which are known for their high polymorphism and codominant alleles at a single locus (Bruford and Wayne, 1993; Queller et al., 1993; Ashley and Dow, 1994), have become a popular tool for genetic diversity studies (Dietrich et al., 1992; Weissenbach et al., 1992), and the analysis of paternity and gene flow (Amos et al., 1993; Chase et al., 1996).

1.2 Microsatellite markers for population genetic analysis

Microsatellites, alternatively known as simple sequence repeats (SSRs) or short tandem repeats (STRs), are tandem repeats of sequence units generally less than 5 bp in length, e.g. $(TG)_n$ or $(AAT)_n$ (Bruford and Wayne, 1993). These markers are generally hyper-variable. Their co-dominance and reproducibility make them ideal for genome mapping, as well as for population genetic studies (Dayanandan et al., 1998).

Microsatellite variation results from differences in the number of repeat units. These differences are thought to be caused by errors in DNA replication (Moxon and Willis, 1999; Jarne and Lagoda, 1996); the DNA polymerase "slips" when copying the repeat region, changing the number of repeats (Jarne and Lagoda, 1996). Larger changes in repeat number are thought to be the result of processes such as unequal crossing over (Strand et al., 1993). Such differences are detected on polyacrylamide gels, where DNA fragments with different repeat lengths migrate at different rates based upon their sizes. Consequently, microsatellites are widely used in genetic diversity studies (Rossetto et al., 1999), gene flow and mating systems (Chase et al., 1996) and paternity analysis (Streiff et al., 1999).

One of the advantages of using microsatellite markers in population genetics studies is that this method reveals high level of polymorphisms. A microsatellite locus in soybean (*Glycine max*) is reported to have 26 alleles (Cregan et al., 1995). Furthermore, the ability of the method to differentiate individuals when multiple loci are examined makes the technique very useful for gene-flow experiments, cultivar identification and paternity analyses (Hokanson et al., 1998). Since microsatellites only survey one locus at a time, they are not directly comparable to molecular markers that target multiple regions of the genome simultaneously, such as Amplified Fragment Length polymorphism (AFLPs). Comparisons that include microsatellites should be made with

other single locus, co-dominant markers, such as RFLPs and isozymes. For example, Rossetto et al. (1999) found the observed heterozygosity (H_o) for *Melaleuca alternifolia* microsatellites to be 0.724, much higher than the value for allozymes (H_o=0.154; Butcher et al., 1992). McCouch et al. (1997) compared the number of alleles revealed by RFLPs and microsatellite loci in rice (*Oryza* spp.), and found 2-25 alleles per microsatellite locus compared with 2-4 alleles per RFLP locus, illustrating the high level of polymorphism detected through microsatellite analysis.

Unlike AFLP markers, microsatellites are co-dominant markers, thus heterozygotes can be readily identified. Microsatellite co-dominance will increase the efficiency and accuracy of population genetic measures based on these markers compared with other markers, such as AFLPs and RAPDs. Furthermore, the identity of heterozygotes in the F_1 generation makes gene flow, hybridisation and paternity analyses simpler (Schlotterer and Pemberton, 1994). Since the method is DNA-based, this brings advantages, such as high-throughput and the ability to use dried leaf material. In comparison with allozymes, SSRs are thought to be selectively neutral (Ng and Corlet, 2000).

1.3 Genetic studies of the genus *Rhododendron*

There are a few studies that have reported genetic diversity estimates of *Rhododendron* species from different regions of the world. These studies have shown different patterns of genetic structure and genetic differentiation at different spatial scales. For example, Ng and Corlet (2000) through allozyme analysis detected high genetic diversity and allelic richness (H_E=0.20 to 0.39, allelic richness ranges from 2.4 to 4.1) over six species of *Rhododendron* in different provinces of China. The genetic differentiation values (F_{ST}) ranged from 0.056 to 0.393 and showed a significant

difference between species. Three species, *Rhodondendron championiae*, *R. hongkongense*, *R. simiarum*, showed high genetic differentiation (F_{ST} and F_{PT}) and distinct geographical patterns, while the other three species, *R. moulmainense*, *R. farrerae* and *R. simsii,* had low differentiation and little or no geographical structure.

In another study by Pornon et al. (2000), AFLP analysis of two populations of *R. ferrugineum* indicated loss of genetic diversity and suggested that thinning could have occurred in the past. A few studies have indicated low levels of gene flow among populations of different *Rhododendron* species. For instance, Kameyama et al. (2001) assessed the patterns and levels of geneflow in *R. metternichii* using microsatellite markers, and showed that long distance gene flow is very low and could be due to the flowering phenology and topographical and vegetational heterogeneity within populations.

1.4 Study objectives

The objective of the present study is to assess the genetic structure of *Rhododendron arboreum* in harvested and protected forests in Arunachal Pradesh in Northeast India. The specific objectives are 1) to estimate the amount of genetic diversity within populations, 2) to measure the amount of genetic differentiation among populations, and 3) to analyze the relationship between geographical distance and the degree of population differentiation.

2. Materials and Methods

2.1 Study sites

The study sites are located in the Tawang district of Arunachal Pradesh in India (Figure 2).

2.1.1 Arunachal Pradesh

The state of Arunachal Pradesh is located in the northeast region of India. It is situated between 26°28′ to 29°30′ N latitude and 91°30′ to 97°30′ E longitude covering an area of 83,743 sq. km. The total forest cover of the state is 67,777 sq. km covering 80.93 % of the total geographical area out of which very dense forest covers 14,411 sq. km, moderately dense forest 37,977 sq. km and open forest with 15,389 sq. km (FSI, 2005). The state is surrounded by Bhutan in the west, China (Tibet) in the north, Myanmar in the east and Nagaland and Assam in the south. The vegetation of Arunachal Pradesh has been classified into six broad forest types, namely tropical, subtropical, temperate, subalpine, alpine and secondary forests (Kaul and Haridasan 1987). The state is divided into three physiographical zones viz. the flood plains, the foothills and the Greater Himalayas. It also has three climatic zones, the hot and humid subtropical area in the foothills, a cooler micro-thermal zone in the Lesser Himalayas and the alpine zone in the Greater Himalayas.

2.1.2 Tawang district

The Tawang district is situated between 27°25′ to 27°52′ N latitude and 91°16′ to 91°59′ E longitude. It is surrounded by Tibet in the north, Bhutan in the south– west and the Sela range in the east separates it from West Kameng district. The Tawang district spans over 2,172 sq. km and constitutes 2.59% of the total geographical area of the

state. The total forest cover of the district is 1,218 sq. km i.e., 56.08% of the total geographical area out of which very dense forest comprises of 78 sq. km, moderately dense forest 721 sq. km and open forest covers about 419 sq. km (FSI, 2005). The region is drained by two rivers, Tawang Chu and Nyamjang Chu, and their tributaries, which converges into Bhutan. The district comprises three administrative sub-divisions (Jang, Tawang, and Lumla), and nine administrative circles (Jang, Mukto, Thingbu, Tawang, Bongkhar, Kitpi, Lumla, Dudunghar and Zemithang). The Monpa tribe is the predominant community inhabiting Tawang district except for Shyo village, which is dominated by a community of Tibetain origin. They follow Buddhism and belong to the Mongoloid heritage (Anonymous, 2005). The total population of the district is 38,924 with population density of 18 persons per sq. km, marginally exceeding the average for the state (13 persons per sq. km) (Anonymous, 2005). Agriculture and animal husbandry are the primary occupations in the region. Two thirds of the district is located in a very high mountainous region. Most of the villages are concentrated in the basins of the two river valleys. The subtropical, temperate and alpine forests cover the mountains between 1100 m and 6000 m above sea level (Anonymous, 2005).

2.1.3 Topography

The topography of Tawang district is mountainous with snow covered Himalayan ranges from 3350 m to 6700 m with bare mountain, mostly uninhabited. While the high altitude mountainous belt ranges from 1800 m to 3350 m with plateaus and narrow valleys and a sparsely distributed population. The district is drained by two main rivers, the Tawangchu and the Nyamyanjchu and the majority of the villages are concentrated in these river valleys. The soils of the district are (i) rocky and loamy skeletal textured with soil depth shallow to medium and (ii) sandy skeletal, loamy, fine loamy textured with soil depth medium to deep (http://tawang.nic.in).

2.1.4 Geology

The eastern Himalayas of Arunachal Pradesh is geographically divided into three zones extending from south to north *viz.* the Sub-Himalayas, the Lesser Himalayas and the Greater Himalayas. Geologically, the Sub-Himalayan zone consists of Neogene mollassic sediments (Siwalik) whereas the Lesser Himalayas consist of Upper Proterozoic and Lower Palaeogene self sediments (Bomdila Group, Buxa-Miri Formations). The Greater Himalayas have been characterized by para and other metamorphites with acid intermediate igneous intrusions from the Precambrian to Tertiary ages (Sela Group, Siang Group, Lumla Formations etc.). These zones are directly or indirectly controlled by the distinct structural features known as MBF, MCT, Lohit and Mishmi thrusts. The terrains of the study sites are hilly with very steep slopes and with rocky mountains. These rocks of the Himalayan type are mainly shales, schists and conglomerates. They belong to the Sela Group and Bomdila Group of the Greater and Lesser Himalayas. The eruptic rocks in Arunachal Himalaya are represented by extensive occurrence of Abor Volcanics, which occur interbedded both with Buxa Permian Gondwanas and Rajmahal traps (Pascoe, 1950; Anonymous, 1976).

2.1.5 Climate

The study sites are located in cool temperate regions of the Tawang district of Arunachal Pradesh. The seasons can be broadly divided into: winter (December to February), when air temperature decreases resulting in a very cold climate with snowfall, pre-monsoon (March to May) with little rain, monsoon (June to September) with high rainfall and post-monsoon (October to November) with little or no rain. The area receives a high amount of rainfall and maximum rainfall occurs during June-August. In Pakchung, Paipraw and Falockchar in the Tawang district, the average monthly rainfall ranges

between 4 mm to 500 mm and mean monthly temperature range between -2.5 °C to 25.5 °C. Relative humidity in these localities vary from a minimum of 47% to a maximum of 80%.

2.1.6 Description of each study population

The present study was carried out in *Rhododendron* forest stands at the Pakchung, Paipraw and Falockchar localities under the Mukto circle of Tawang district, Arunachal Pradesh (Figure 2, Table 1). The geographic distances among populations are given in Table 2. The harvesting of trees for fuelwood in selected areas in these forests is done annually during the months of October to December. Age of population is considered approximately 200-300 years. Samples were collected randomly from 25-30 individuals separated by a minimum distance of 50 m or more. Trees from the disturbed areas were residual trees. The study populations were classified as disturbed (harvested) or protected, based on the level of harvesting of *R. arboreum.* Thus, the disturbance index was calculated as DI= Basal area of cut stumps/ Total basal area (residual trees + cut stumps) x 100. The disturbance index in sites PJD, PMD and FMD were 89%, 86% and 89% respectively (Paul, 2008). The level of disturbance of the Paipraw Jang Disturbed (PJD) and Falockchar Mukto Disturbed (FMD) stands are almost similar. The Pakchung Mukto Disturbed (PMD) stand showed slightly less disturbance as compared to the other disturbed stands. There were no cut stumps in the Paipraw Jang Protected (PJP) and Pakchung Mukto Protected (PMP) stands. Thus the degree of disturbance in these two populations was considered to be zero. The Pakchung Mukto Disturbed (PMD) population is about 3-4 hectares in size, mountainous and lacks human settlement. There is no agricultural activity in and around the area, but regular cattle grazing was observed. This site was considered to be disturbed due to the

excessive harvesting pressure (logging of trees for fuelwood). The Pakchung Mukto Disturbed (PMD) and Pakchung Mukto Protected (PMP) populations are separated by a footpath. The distance between these two populations is about 200 m. The density of *Rhododendron arboreum* in PMD is 132 ha^{-1} and it is the dominant tree species in the area. The site was covered by associated trees, shrubs, and herb species. The associated trees are *Acer pectinatum*, *Lyonia ovalifolia, Rhododendron arboreum,* and *Rhododendron grande.* The associated shrubs include *Daphne papyracea, Desmodium elegans, Eurya acuminata, Viburnum cylindricum, Viburnum mullaha,* and *Luculia gratissima.* The herbs species include *Ainsliaea aptera, Anemone rivularis, Bergenia ciliata, Bidens pilosa, Dryopteris* sp., *Fragaria nubicola, Globba multiflora, Imperata cylindrica, Pilea scripta, Rubus calycinus, Swertia chirata, Symplocos theifolia, Selinum* sp., *Potentilla nepalensis,* and *Anaphalis adnata* (Table 3).

The Pakchung Mukto Protected (PMP) site has an area of 4-5 hectares. The area is mountainous and there is no nearby human settlement. There is also no agricultural activity in and around the area. Sparse grazing was observed and harvesting was totally absent. The density of *R. arboreum* in PMP is 1082 ha^{-1} and it is the dominant tree species in the area. The area includes various trees, shrubs, and herbaceous species. The tree species contains *Acer pectinatum*, *Lyonia ovalifolia, Rhododendron arboreum, Rhododendron grande, Michelia* sp., *Quercus* sp., and *Pyrus pashia.* The associated shrubs include *Daphne papyracea, Desmodium elegans, Eurya acuminata, Gaultheria fragrantissima, Indigofera dosua, Pyrus expansa Viburnum cylindricum, Viburnum mullaha*, and *Luculia gratissima*. The herbaceous species are *Ainsliaea aptera, Anaphalis adnata, Bergenia ciliata, Bidens pilosa, Dryopteris* sp., *Fragaria vesca, Globba multiflora, Imperata cylindrica, Pilea scripta, Potentilla polyphylla, Swertia chirata, Symplocos theifolia, Viola canescens, Selinum* sp., *Potentilla nepalensis, Hypericum elodeoides,* and *Anaphalis adnata* (Table 3).

The Pipraw Jang Disturbed site (PJD) has an area of 4-5 hectares within the mountainous area and has no human settlement. There is no agricultural activity in and around the area but regular grazing was observed. The site was considered disturbed due to the excessive harvesting pressure (logging of trees for fuelwood). The Pipraw Jang Disturbed (PJD) and Pipraw Jang Protected (PJP) populations are separated by a road. The distance between both sites is 76 m. The density of *Rhododendron arboreum*, the dominant tree species in the area, is 92 individuals ha^{-1}. The site includes different tree, shrub, and herb species. The tree species include *Acer pectinatum, Betula alnoides, Lyonia ovalifolia, Michelia* sp., *Rhododendron arboreum,* and *Rhododendron barbatum*. The associated shrubs include *Aconogonon molle, Berberis wallichiana, Daphne papyracea, Desmodium elegans, Eurya acuminata, Indigofera dosua, Luculia gratissima, Pyrus expansa, Rubus hypargyrus, Viburnum cylindricum,* and *Luculia gratissima*. The herbs species include *Ainsliaea aptera, Anemone polyanthes, Bergenia ciliata, Bidens pilosa, Crawfurdia speciosa, Dryopteris* sp., *Fragaria daltoniana, Fragaria vesca, Globba multiflora, Imperata cylindrica, Pilea scripta, Rubus calycinus, Swertia chirata, Symplocos theifolia, Viola canescens,* and *Rubus nepalensis* (Table 3).

The Pipraw Jang Protected (PJP) includes approximately 5-6 hectares. The area is mountainous and there is no human settlement and no agricultural activity in and around the area. Sparse grazing was observed and harvesting is totally absent. The density of *Rhododendron arboreum,* which is the dominant tree species in the area, is 1422 ha^{-1}. The tree species in the area includes *Acer pectinatum, Acer caudatum, Betula alnoides, Castanopsis tribuloides, Lyonia ovalifolia, Michelia* sp., *Rhododendron arboreum,* and *Rhododendron barbatum, Quercus* sp. The shrubs are *Aconogonon molle, Agapetes incurvata, Berberis wallichiana, Corylopsis himalayana, Daphne papyracea, Desmodium elegans, Eurya acuminata, Gaultheria fragrantissima, Indigofera dosua, Luculia gratissima, Pyrus expansa, Rubus hypargyrus, Viburnum cylindricum,*

and *Luculia gratissima.* The associated herbs species include *Ainsliaea aptera, Anemone polyanthes, Arisaema speciosum, Bergenia ciliata, Bidens pilosa, Crawfurdia speciosa, Dryopteris* sp., *Fragaria vesca, Globba multiflora, Imperata cylindrica, Pilea scripta, Rubus calycinus, Swertia chirata, Symplocos theifolia, Viola canescens,* and *Rubus nepalensis* (Table 3).

The Falokchar Mukto Disturbed (FMD) site consists of 3-4 hectares in the mountainous area and has no human settlement. There is no agricultural activity in and around the area and regular grazing was observed. The site was considered to be disturbed due to the excessive harvesting pressure (logging of trees for fuelwood). The density of *Rhododendron arboreum*, which is the dominant tree species in the area, is 112 ha^{-1}. The site includes other associated tree, shrub and herb species. The tree species are *Betula utilis, Lyonia ovalifolia, Rhododendron arboreum,* and *Rhododendron barbatum.* The shrubs species are *Aster albescens, Berberis wallichiana, Daphne papyracea, Eurya acuminata, Spiraea canescens, Viburnum cylindricum,* and *Luculia gratissima*. The associated herbs are *Ainsliaea aptera, Anaphalis adnata, Anemone rivularis, Arisaema speciosum, Bergenia ciliata, Bidens pilosa, Corydalis geraniifolia, Dryopteris* sp., *Eupatorium* sp., *Fragaria nubicola, Imperata cylindrica, Leucas lanata, Pilea scripta, Potentilla polyphylla, Viola canescens, Rubus nepalensis, Potentilla nepalensis, Hypericum elodeoides,* and *Anaphalis adnata* (Table 3).

In summary, *R. arboreum* is the dominant tree in the study area. In disturbed sites, *R. arboreum* has been harvested on a large scale for fuelwood. The distance between populations ranged from about 76 m to over 5000 m (Table 2). The sites were separated by mountains, agricultural land, human settlements, roads and streams.

Table 1. Protection status, number of individuals analyzed, latitude, longitude and elevation of five *R. arboreum* populations.

Population	Population ID	Protection status	No. of individuals analyzed	Latitude (N)	Longitude (E)	Elevation (m)
Pakchung Mukto	PMD	Disturbed	24	27° 31′31.0″	91° 55′18.1″	2601
Pakchung Mukto	PMP	Protected	23	27° 31′25.5″	91° 55′20.7″	2642
Pipraw Jang	PJD	Disturbed	29	27° 33′09.9″	91° 57′50.2″	2922
Pipraw Jang	PJP	Protected	24	27° 33′08.9″	91° 57′47.3″	2949
Falokchar Mukto	FMD	Disturbed	25	27°32′35.0″	91°56′42.0″	3153

Table 2. Pairwise distance between *R. arboreum* populations studied.

Population ID	Aerial distance (m)
PJD-FMD	2140
PJD-PJP	76
PJP-FMD	2050
PMD-FMD	3020
PMD-PJD	5170
PMD-PJP	5090
PMP-FMD	3000
PMP-PJD	5200
PMP-PJP	5130
PMP-PMD	200

Abbreviations of population names are given in Table 1.

Table 3. Associate plant species and density of *R. arboreum* observed in five study sites in the eastern Himalayas, India. Names of study populations are given in Table 1 (Paul, 2008).

Site	Pipraw Jang Protected (PJP)	Pipraw Jang Disturbed (PJD)	Pakchung Mukto Protected (PMP)	Pakchung Mukto Disturbed (PMD)	Falokchar Mukto Disturbed (FMD)
Trees	*Acer pectinatum*, *Acer caudatum*, *Betula alnoides*, *Castanopsis tribuloides*, *Lyonia ovalifolia*, *Michelia* sp., *Rhododendron arboreum*, *Rhododendron barbatum*, *Quercus* sp.	*Acer pectinatum*, *Betula alnoides*, *Lyonia ovalifolia*, , *Michelia* sp., *Rhododendron arboreum*, *Rhododendron barbatum*	*Acer pectinatum* , *Lyonia ovalifolia*, *Rhododendron arboreum*, *Rhododendron grande*, *Michelia* sp., *Quercus* sp., *Pyrus pashia*	*Acer pectinatum*, *Lyonia ovalifolia*, *Rhododendron arboreum*, *Rhododendron grande*,	*Betula utilis*, *Lyonia ovalifolia*, *Rhododendron arboreum*, *Rhododendron barbatum*
Shrubs	*Aconogonon molle*, *Agapetes incurvata*, *Berberis wallichiana*, *Corylopsis himalayana*, *Daphne papyracea*, *Desmodium elegans*, *Eurya acuminata*, *Gaultheria fragrantissima*, *Indigofera dosua*,	*Aconogonon molle*, *Berberis wallichiana*, *Daphne papyracea*, *Desmodium elegans*, *Eurya acuminata*, *Indigofera dosua*, *Luculia gratissima*, *Pyrus expansa*, *Rubus hypargyrus*, *Viburnum*	*Daphne papyracea*, *Desmodium elegans*, *Eurya acuminata*, *Gaultheria fragrantissima*, *Indigofera dosua*, *Pyrus expansa Viburnum cylindricum*, *Viburnum mullaha*, *Luculia gratissima*,	*Daphne papyracea*, *Desmodium elegans*, *Eurya acuminata*, *Viburnum cylindricum*, *Viburnum mullaha*, *Luculia gratissima*	*Aster albescens*, *Berberis wallichiana*, *Daphne papyracea*, *Eurya acuminata*, *Spiraea canescens*, *Viburnum*

	Luculia gratissima, Pyrus expansa, Rubus hypargyrus, Viburnum cylindricum, Luculia gratissima,	*cylindricum, Luculia gratissima,*			*cylindricum, Luculia gratissima,*
Herbs	*Ainsliaea aptera, Anemone polyanthes, Arisaema speciosum, Bergenia ciliata, Bidens pilosa, Crawfurdia speciosa, Dryopteris* sp., *Fragaria vesca, Globba multiflora, Imperata cylindrica, Pilea scripta, Rubus calycinus, Swertia chirata, Symplocos theifolia, Viola canescens, Rubus nepalensis,*	*Ainsliaea aptera, Anemone polyanthes, Bergenia ciliata, Bidens pilosa, Crawfurdia speciosa, Dryopteris* sp., *Fragaria daltoniana, Fragaria vesca, Globba multiflora, Imperata cylindrica, Pilea scripta, Rubus calycinus, Swertia chirata, Symplocos theifolia, Viola canescens, Rubus nepalensis,*	*Ainsliaea aptera, Anaphalis adnata, Bergenia ciliata, Bidens pilosa, Dryopteris* sp., *Fragaria vesca, Globba multiflora, Imperata cylindrica, Pilea scripta, Potentilla polyphylla, Swertia chirata, Symplocos theifolia, Viola canescens, Selinum* sp., *Potentilla nepalensis, Hypericum elodeoides, Anaphalis adnata,*	*Ainsliaea aptera, Anemone rivularis, Bergenia ciliata, Bidens pilosa, Dryopteris* sp., *Fragaria nubicola, Globba multiflora, Imperata cylindrica, Pilea scripta, Rubus calycinus, Swertia chirata, Symplocos theifolia, Selinum* sp., *Potentilla nepalensis, Anaphalis adnata,*	*Ainsliaea aptera, Anaphalis adnata, Anemone rivularis, Arisaema speciosum, Bergenia ciliata, Bidens pilosa, Corydalis geraniifolia, Dryopteris* sp., *Eupatorium* sp., *Fragaria nubicola, Imperata cylindrica, Leucas lanata, Pilea scripta,*

					Potentilla polyphylla, Viola canescens, Rubus nepalensis, Potentilla nepalensis, Hypericum elodeoides, Anaphalis adnata,
Density of *Rhododendron arboreum*	1422 ha^{-1}	92 ha^{-1}	1082 ha^{-1}	132 ha^{-1}	112 ha^{-1}

2.1.7 Soil type of the study area

Rhododendrons are known to thrive in acidic soils, and growth is hindered in soils with a pH of 7.0 or higher (Kaisheva, 2006). The soil of Tawang district is rocky and loamy textured with a shallow to medium soil depth. A detailed analysis of soil composition in each site (Paul 2008) is given in Table 4. The pH, organic C and total N content vary depending upon the depth. The soil pH in all study sites was acidic. Extreme acidity (pH 3.24) is recorded in the Paipraw Jang Disturbed (PJD) site while very strong acidity is reported in the Pakchung Mukto Protected (PMP) (pH 3.45) and Paipraw Jang Protected (PJP) (pH 3.51) sites. Soil organic carbon is 9.55% and 7.40% in the Falockchar Mukto Disturbed (FMD) and Paipraw Jang Protected (PJP) sites respectively, which is higher than the Pakchung Mukto Disturbed (PMD) (6.52%), Paipraw Jang Disturbed (PJD) (6.45%) and Pakchung Mukto Protected (PMP) (6.45%) sites. Both soil organic carbon (SOC) and total Kjeldahl nitrogen (TKN) of soil decreased with increased soil depth (Table 4).

Table 4. Chemical properties of soil in five study locations (Paul 2008).

Parameters/ Depth	0-10 cm	10-20 cm	20-30 cm
PJP (Paipraw Jang Protected)			
pH	3.51 ± 0.00	3.65 ± 0.00	3.96 ± 0.01
Soil Organic Carbon (%)	7.40 ± 0.05	6.65 ± 0.05	4.20 ± 0.04
Total Kjeldahl Nitrogen (%)	0.56 ± 0.04	0.33 ± 0.02	0.23 ± 0.05
PJD (Paipraw Jang Disturbed)			
pH	3.24 ± 0.01	3.35 ± 0.00	3.61 ± 0.01
Soil Organic Carbon (%)	6.45 ± 0.04	4.73 ± 0.17	4.35 ± 0.04
Total Kjeldahl Nitrogen (%)	0.28 ± 0.00	0.21 ± 0.00	0.19 ± 0.02
PMP (Pakchung Mukto Protected)			
pH	3.45 ± 0.01	3.60 ± 0.01	3.84 ± 0.00
Soil Organic Carbon (%)	6.45 ± 0.05	5.82 ±0.87	5.01 ± 0.05
Total Kjeldahl Nitrogen (%)	0.48 ± 0.02	0.39 ± 0.00	0.31 ± 0.02
PMD (Pakchung Mukto Disturbed)			
pH	3.57 ± 0.01	3.69 ± 0.01	3.84 ± 0.00
Soil Organic Carbon (%)	6.52 ± 0.17	5.13 ± 0.15	4.71 ± 0.26
Total Kjeldahl Nitrogen (%)	0.43 ± 0.02	0.36 ± 0.02	0.25 ± 0.01
FMD (Falockchar Mukto Disturbed)			
pH	3.58 ± 0.02	3.62 ±0.01	3.95 ± 0.01
Soil Organic Carbon (%)	9.55 ± 0.10	8.70 ± 0.15	7.72 ± 0.31
Total Kjeldahl Nitrogen (%)	0.63 ± 0.00	0.40 ± 0.02	0.35 ± 0.04

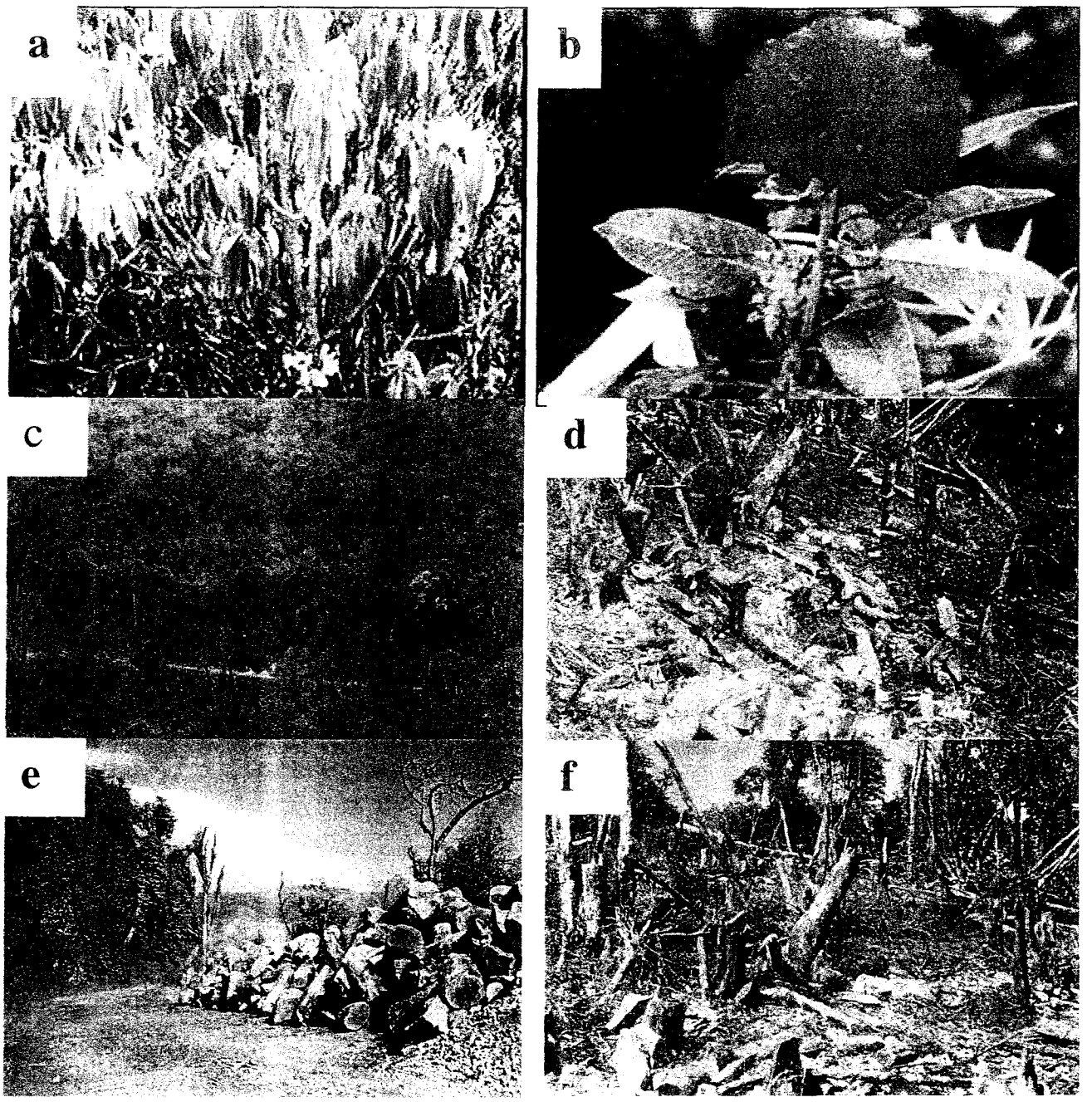

Figure 1. *R. arboreum* tree with flowers (a,b), distance view of *Rhododendron* forest (c), harvesting of *R. arboreum* (d), pile of *R. arboreum* logs on roadside (e) and a forest after harvesting *R. arboreum* (f).

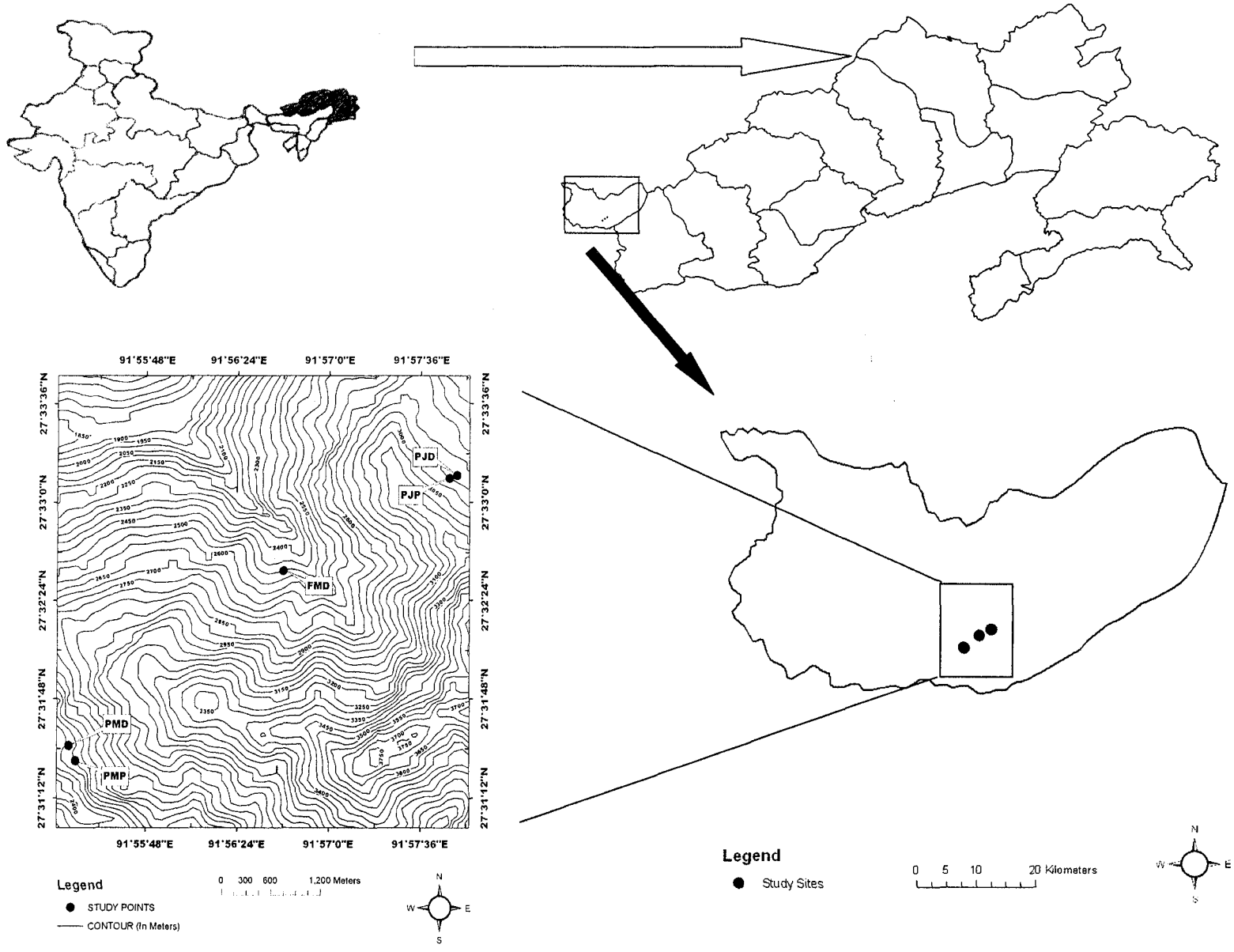

Figure 2. Location of Arunachal Pradesh, Tawang district and study populations of *R. arboreum* in the Eastern Himalayan region.

2.2 Sample collection

Leaf samples from 25 - 30 individual trees from each study site (Table 1) were collected, dried and transported to the laboratory for DNA extraction.

2.3 Isolation of genomic DNA from *R. arboreum*

Total genomic DNA was extracted from the leaves following the protocol of Dayanandan et al. (1997) (Appendix 1, Protocol 1). Five µl of genomic DNA were loaded on to the ethidium bromide (0.33µg/ml) premixed 1% agarose gel to evaluate the approximate quality and concentration. The gel was observed under a UV Trans illuminator (Appendix 1, Protocol 2).

2.4 Standardization of PCR amplification with SSR primers

PCR amplifications were performed with a thermal cycler (GeneAmp PCR System 9600, ABI). Eight microsatellite primer pairs (Naito et al., 1998; Kameyama et al., 2002) (RM1D1, RM1D5, RM1D12, RM2D2, RM3D4, RM7D9, RM9D1, and RM9D6) were used for amplifying genomic DNA of *R. arboreum* (Table 5). PCR amplification was carried out in a 25 μl reaction mixture, which included 5 pmol of each primer, 250 μM dNTPs, 10 mM Tris-HCl (pH 8.3), 1.5 mM $MgCl_2$, 0.5 U Taq DNA polymerase and approximately 25 to 50 ng of template DNA using a GeneAmp PCR (System 9600, ABI). The forward primer of each primer pair was labeled with either FAM, HEX, or TET fluorescent dye (Operon Technologies, USA). PCR cycles were as follows: 3 minutes at 94°C for initial denaturation followed by 35 cycles with 40 seconds at 94°C for cycling denaturation, 30 seconds at 50°C for annealing temperature, 1 minute at 72°C for elongation, and final elongation for 4 minutes at 72°C (Appendix 1, Protocol 3). Table 5 shows the optimal annealing temperature of each primer.

Table 5. Primer sequence, length and estimated annealing temperature of eight microsatellite primers from *R. metternichii* (Naito et al., 1998; Kameyama et al., 2002) tested for PCR amplification against *R. arboreum* genomic DNA.

Primer ID	Sequence 5' - 3'	Length of primer (bp)	Annealing Temperature (°C)
RM1D1F	ATCTGGAGGCCATTGGTAGT	20	60.4
RM1D1R	TATTGGGTCCGATGACAGAC	20	60.4
RM1D5F	CAAATACTCCTCTCAGTTAC	20	56.3
RM1D5R	TTAAACATGGTGATTGTGTC	20	54.3
RM1D12F	GCACAAGCCCGAGTAGAAAG	20	62.5
RM1D12R	GGTGGTAACCCCTGAAATAG	20	60.4
RM2D2F	ATGTGTTTCGTTGCTACTGT	20	56.3
RM2D2R	ATGGTTGGTTTGTTTTCCTA	20	54.3
RM3D4F	CTCCCAACAAACAAATCCAT	20	56.3
RM3D4R	CACCGAACGAAGACACTCAG	20	62.5
RM7D9F	TTTACTGGCTGAAAATACAC	20	54.3
RM7D9R	TAAGGTGATGAAATATGAGA	20	52.2
RM9D1F	TGACCAAGTGCGACCTAATC	20	60.4
RM9D1R	TAACTCCTAACAACCAAAAC	20	54.3
RM9D6F	CTCGCCTCCCAAAAGCAAT	19	60.2
RM9D6R	CGTGTCCTCACCCCCGTAAC	20	66.6

2.5 SSR genotyping and scoring the genetic data

The multilocus genotype of each DNA sample was scored by using 3 pairs of microsatellite primers (RM1D1, RM2D2, and RM9D6). The size of the PCR products was determined by automated fluorescent scanning detection with an ABI 310 genetic analyzer and Genescan analysis software (Applied Biosystems, Foster city, California). The genotyping procedure was followed by mixing 10 μl of Hi-Di formamide (deionized formamide) with 0.1 μl of Tamara GeneScan-500 size standard (Applied Biosystems). To the mixture, aliquoted into separate tubes, the following amounts of PCR product were added; 1 μl of undiluted, 1 μl of 1:5, 1:10 and 1:20 dilution of PCR product. The mixture was subjected to denaturation for 2 minutes at 95 °C and immediately placed on ice for a minimum of 5 minutes. Finally, the mixture was loaded onto an ABI 310 genetic analyzer (Applied Biosystems, Foster city, California) for capillary electrophoresis for 30 minutes to detect SSR polymorphism across *R. arboreum* samples. The dilution factor that yielded unsaturated clean peaks was used for genotyping individuals. The electrophorograms were analyzed with the GeneScan version 3.7 application software (Applied Biosystems) following the manufacturer's instructions.

2.6 Data analysis

Allele frequencies and expected heterozygosity (He) in each population at each locus were calculated using Popgene v 3.2 (Yeh et al., 1997). The observed heterozygosity (Ho) was calculated directly from observed genotypes. The effective number of alleles (*Ne*) was estimated using the formula $Ne = 1/\Sigma xi^2$, where, *xi* is the frequency of the *i*th allele for each locus (Crow and Kimura, 1964). Genetic structure was assessed by comparison of allele frequency distributions and the estimation of pairwise F_{ST} values, where F stands for *F*-statistics and S for sub- and T total- population

(Reynolds et al., 1983; Slatkin, 1995), for all possible population combinations using Genetic Data Analysis (GDA) version 1.0 software package (Lewis and Zaykin, 2001). F_{ST} (F-statistics) was estimated by estimating Wright (1951) hierachical F coefficients by the method of Weir and Cockerham (1984) as implemented in the program Genetic Data Analysis (GDA) version 1.0 software package (Appendix 3, Test 2). Gene flow (Nm) was calculated based on the estimates of F_{ST} (Slatkin and Barton, 1989).

The mean observed and expected heterozygosity values among populations were compared using ANOVA. The relationship between pair-wise geographical distance and F_{ST} were analyzed by calculating the correlation coefficient between pair-wise geographical distance and linearized Fst values. The values of linearized Fst were calculated as Fst/(1-Fst). These analyses were done in a Microsoft Excel based programme called poptools v 2.6.2 (Hood, 2004).

Nei's standard genetic distance (Nei, 1972) was calculated using the GDA software. A dendrogram based on genetic distances was constructed using the Neighbor Joining method as implemented in PHYLIP ver. 3.68 (Felsenstein, 2005). The relationship between genetic distance and geographical distance was investigated by calculating the correlation coefficient as well as constructing a graph between these two variables.

3. Results

3.1 Cross-amplification of microsatellite markers of *R. metternichii* in *R. arboreum*

Since there were no microsatellite markers available for *R. arboreum* in the literature, microsatellite markers that were developed for *R. metternichii* were used with *R. arboreum* samples using the cross-amplification approach. Eight pairs of primers (RM1D1, RM1D5, RM1D12, RM2D2, RM3D4, RM7D9, RM9D1, and RM9D6) were evaluated for the success of cross-amplification against *R. arboreum* and five pairs (RM1D1, RM2D2, RM3D4, RM9D1, and RM9D6) produced clear PCR amplification products detectable on agarose gels. Those five primers were labeled with FAM, HEX and TET fluorescent dyes and used for genotyping of *R. arboreum* samples. Out of five, three (RM1D1, RM2D2, and RM9D6) showed polymorphism in ABI 310 Genetic analyzer and these three pairs of primers were further used in assessing the genetic diversity of *R. arboreum* in protected and disturbed forests in the eastern Himalayan region of India. All three primers (RM1D1, RM2D2, and RM9D6) produced amplification products that are comparable to the size reported for *R. metternichii*, indicating site specific amplification of similar microsatellite regions in *R. arboreum* samples.

3.2 Within population genetic diversity of *R. arboreum*

3.2.1 Observed and effective number of alleles:

Among more than 125 samples analyzed from 5 different populations of *R. arboreum*, 35 different alleles were detected at three loci, with an average of 11.66 alleles per locus. Locus 9D6 showed 15 alleles over all populations, followed by 11 at

locus 1D1 and 9 alleles at locus 2D2. Similarly, in each population, the population PJD (Pipraw Jang disturbed) exhibited a maximum of 15 alleles at locus 9D6 with a mean of 10.33 alleles per locus, followed by 8.66, 8, 7.33 and 7 in FMD (Falokchar Mukto disturbed), PMP (Pakchung Mukto protected), PMD (Pakchung Mukto disturbed) and PJP (Pipraw Jang disturbed protected) populations respectively (Table 6). Populations did not differ significantly in the number of alleles observed over all loci as confirmed by ANOVA test ($F=0.545$, $P=0.707$).

In each population, the mean effective number of alleles was compared with the observed number of alleles in that population (Table 6). The effective number of alleles is the minimum number of alleles required to maintain the current level of heterozygosity in the population if all allele frequencies were equal. Most populations showed a higher number of observed alleles than the effective number of alleles. However, the differences were not statistically significant (student t test, $P>0.05$). Similarly, there was no significant difference in the effective number of alleles between populations (ANOVA: $F=0.383$; $P=0.816$).

Table 6. Observed heterozygosity (Ho), expected heterozygosity (He), observed number of alleles (Na) and effective number of alleles (Ne) observed in five populations of *R. arboreum* in the eastern Himalayas, India. Names of study populations are given in Table 1.

Population ID	Locus	Ho	He	Na	Ne
PMD	1D1	0.79	0.73	6	3.62
	2D2	0.52	0.54	6	2.15
	9D6	0.95	0.87	10	6.73
	Average	**0.75 ± 0.21**	**0.71 ± 0.16**	**7.33 ± 2.30**	**4.17 ± 2.33**
PMP	1D1	0.69	0.65	6	2.75
	2D2	0.55	0.61	6	2.50
	9D6	0.92	0.88	12	7.43
	Average	**0.72 ± 0.18**	**0.71 ± 0.14**	**8 ± 3.46**	**4.23 ± 2.78**
PJD	1D1	0.72	0.83	9	5.72
	2D2	0.72	0.71	7	3.36
	9D6	0.89	0.90	15	9.33
	Average	**0.78 ± 0.09**	**0.82 ± 0.09**	**10.33 ± 4.16**	**6.13 ± 3.0**
	1D1	0.79	0.75	4	3.90
	2D2	0.92	0.78	7	4.31
	9D6	0.92	0.86	10	6.72
PJP	**Average**	**0.87 ± 0.07**	**0.80 ± 0.05**	**7 ± 3**	**4.97 ± 1.92**
	1D1	0.88	0.81	8	5.00
	2D2	0.92	0.78	7	4.35
	9D6	0.83	0.82	11	5.18
FMD	**Average**	**0.87± 0.04**	**0.8 ± 0.02**	**8.66 ± 2.08**	**4.84 ± 0.83**

3.2.2 Heterozygosity: Over all populations of *R. arboreum* the observed heterozygosity (Ho) was highest (0.90) at locus 9D6 followed by 0.77 and 0.72 at 1D1 and 2D2 loci respectively (Table 7). Across populations, PJP (0.87) and FMD (0.87) had higher Ho values compared to PJD (0.78), PMD (0.75) and PMP (0.72) (Table 6). However, populations did not differ significantly in their Ho values (ANOVA: F=0.787, P=0.559).

Table 7. Observed heterozygosity (Ho), expected heterozygosity (He), observed number of alleles (Na) and effective number of alleles (Ne) of each locus in five populations of *R. arboreum* in the eastern Himalayas, India. Names of study populations are given in Table 1.

Locus	**Ho**	**He**	**Na**	**Ne**
1D1	0.776	0.776	11	4.414
2D2	0.725	0.724	9	3.594
9D6	0.904	0.912	15	10.953
Average ± SD	**0.801 ± 0.092**	**0.804 ± 0.097**	**11.666 ± 3.055**	**6.320 ± 4.032**

3.2.3 Private alleles: In total, five private alleles were detected in two of the five populations, PJD and FMD. Three alleles 256, 270 and 274 at locus 1D1 and one allele, 184, at locus 9D6 were found in the population PJD. In the FMD population, one private allele (276) was detected at locus 1D1.

3.3 Among population genetic diversity in *R. arboreum* populations in the eastern Himalayas

3.3.1 Allele Frequency: The frequency distribution of alleles varied over populations. A few alleles were rare and found in some populations with a frequency as low as 0.02. Similarly few alleles were restricted to only one or two populations, again with a low frequency. Alleles unique to a given population were classified as private alleles in those populations (Table 8).

3.3.2 Population differentiation (F_{ST}) and gene flow (Nm): Pair wise population differentiation (F_{ST}) was estimated by analyzing between population differences (Hudson et al., 1992). The mean F_{ST} and Nm over all populations and at all the three loci were 0.052 and 4.53 respectively. In pair wise population F_{ST} analysis, populations PMD and PMP exhibited the lowest F_{ST} of 0.015 with the corresponding highest Nm value of 15.8 (Table 9). The F_{ST} value between populations PJP and PJD was 0.019 suggesting a moderate level of gene flow (Nm=12.6). In contrast to the previous pairs of populations, the value of F_{ST} was highest (0.057) between populations PMD and PJP as well as between populations PJP and PMP indicating a relatively low level of gene flow (Nm=4.1).

Table 8. The frequency of alleles observed at three microsatellite loci in five populations of *R. arboreum* in the eastern Himalayas, India. Names of study populations are given in Table 1

Locus ID	Allele size (bp)	Allele Frequency of populations				
		PMD	PMP	PJD	PJP	FMD
1D1	254	0.166	0.217	0.137	0.229	0.160
	256			0.069		
	260	0.416	0.543	0.310	0.312	0.340
	262	0.229	0.108	0.137	0.208	0.140
	264	0.020	0.021			0.020
	266			0.086		0.080
	268	0.020	0.087	0.155	0.250	0.160
	270			0.034		
	272	0.145	0.021	0.051		0.020
	274			0.017		
	276					0.080
2D2	116				0.020	0.040
	120	0.120	0.037	0.155	0.300	0.080
	122	0.660	0.592	0.413	0.240	0.360
	124	0.080	0.092	0.310	0.260	0.260
	126	0.040	0.037	0.017	0.120	0.080
	128	0.060	0.185	0.069		0.120
	132			0.017	0.040	
	136		0.055		0.020	0.060
	138	0.040		0.017		

9D6	178		0.037	0.089	0.040	
	180	0.023	0.111	0.035		0.062
	182	0.142	0.148	0.107		0.020
	184			0.035		
	190	0.095	0.111	0.125	0.140	0.354
	192	0.023	0.055	0.017	0.080	0.083
	194	0.119		0.035	0.020	0.104
	196	0.261	0.259	0.107	0.020	0.020
	198	0.023	0.092	0.053	0.040	0.187
	200		0.055	0.214	0.180	0.020
	202			0.053	0.240	
	204	0.142	0.037	0.053	0.120	0.041
	206	0.071	0.037	0.017	0.120	0.020
	208		0.018	0.017		
	210	0.095	0.037	0.035		0.083

Table 9. Pairwise population differentiation (F_{ST}: below diagonal) and pair wise gene flow estimates (Nm: above diagonal) of five populations of *R. arboreum* in the eastern Himalayas. Names of study populations are given in Table 1.

Population ID	**PMD**	**PMP**	**PJD**	**PJP**	**FMD**
PMD	-	15.8	8.0	4.1	5.9
PMP	0.0156	-	8.7	4.1	7.1
PJD	0.0305	0.0279	-	12.6	12.9
PJP	0.0572	0.0571	0.0195	-	7.9
FMD	0.0409	0.0341	0.0191	0.0307	-

3.3.3 Geographic distance and F_{ST}: The analysis of the correlation between the geographic distances of the populations and their respective pairwise linearized F_{ST} values showed a significant association between F_{ST} and the geographic distance of the populations (Figure 3; r^2 = 0.4917, P=0.023), suggesting isolation by distance between the populations of *R. arboreum* in the eastern Himalayas.

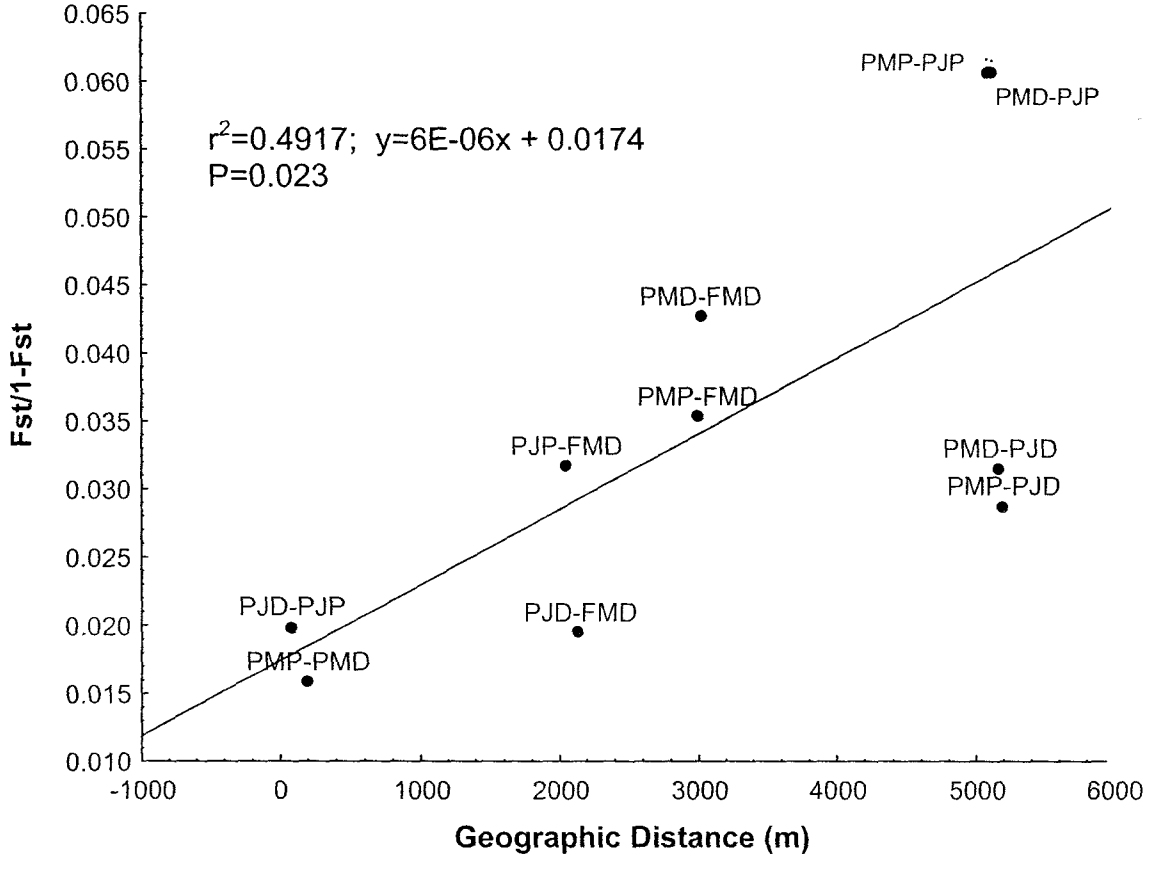

Figure 3. Correlation between linearized Fst and geographic distance between five populations of *R. arboreum* in the eastern Himalayan region. Names of study populations are given in Table 1.

3.3.4 Genetic distance: Nei's (1978) standard genetic distances between all pairs of populations ranged from 0.077 between PMD and PMP to 0.426 between PMD and PJP (Table 10). Pairwise genetic distances between populations showed a positive but statistically non-significant correlation (r^2=0.3452, P=0.0695) with the geographical distances (Figure 4). Nei's standard genetic distances revealed three clusters corresponding to geographical location of populations. One of the clusters (cluster I) contained the PJP and PJD populations, the other cluster (Cluster-II) included PMD and PMP populations and the remaining branch contained the FMD population (Figure 5).

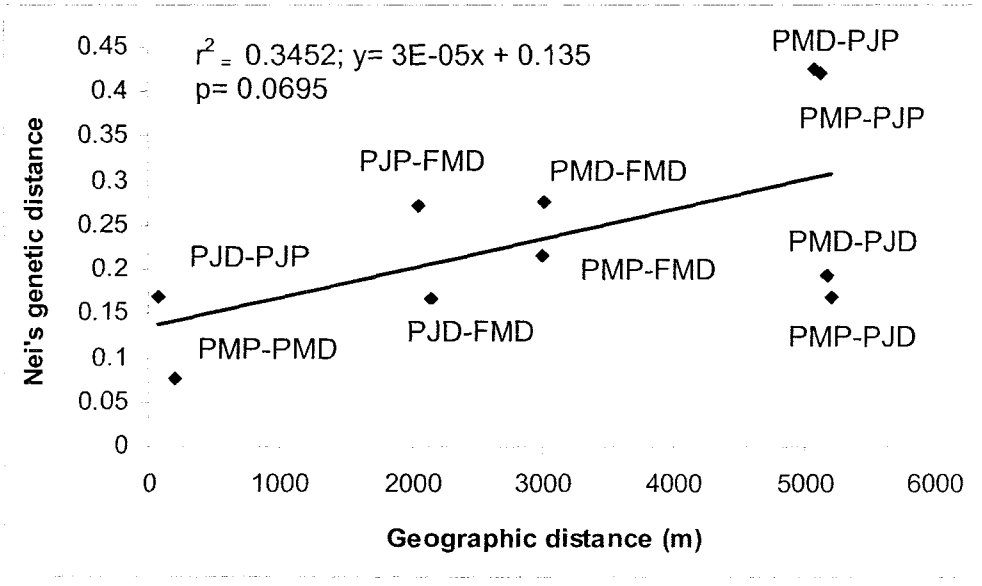

Figure 4. Correlation between pairwise Nei's genetic distance and geographic distance between five populations of *R. arboreum* in the eastern Himalayan region. Names of study populations are given in Table 1.

Table 10. Nei's genetic identity (above diagonal) and genetic distance (below diagonal) of five populations of *R. arboreum* in the eastern Himalayas. Names of study populations are given in Table 1.

Population ID	**FMD**	**PMD**	**PJP**	**PJD**	**PMP**
FMD	-	0.757	0.761	0.845	0.805
PMD	0.277	-	0.652	0.824	0.925
PJP	0.272	0.426	-	0.844	0.656
PJD	0.167	0.193	0.169	-	0.843
PMP	0.216	0.077	0.420	0.169	-

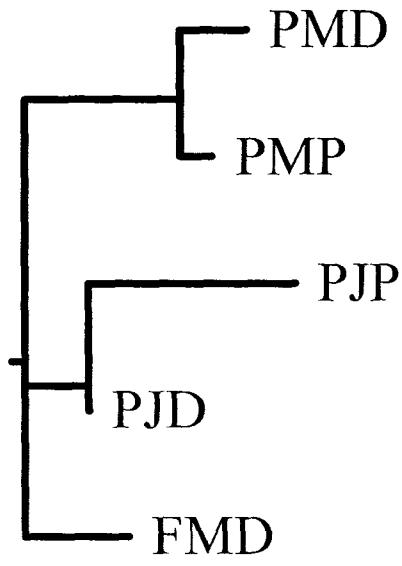

Figure 5. Neighbour joining dendrogram of Nei's genetic distances between populations of *R. arboreum*. PMD = Pakchung Mukto-harvested; PMP = Pakchung Mukto – protected; PJD = Pipraw Jang- harvested; PJP = Pipraw Jang – protected; FMD = Falokchar Mukto- harvested.

4. Discussion

The cross-amplification of microsatellite markers that were originally developed for *R. metternichii* with *R. arboreum* samples indicates that flanking regions of microsatellite repeats are conserved across congeneric species. The three loci with successful cross amplification (1D1, 2D2 and 9D6) showed high levels of polymorphism. These three loci consisted of CT/GA repeats. Reportedly CT/GA repeats are one of the most common dinucleotide repeats found in plants (Morgante and Olivieri, 1993; Akagi et al., 1997; Steinkellner et al., 1997).

This cross-utility of microsatellite markers suggests that these three markers are likely to be of wider application for population genetic studies in the species *R. arboreum* as well as other *Rhododendron* species to assess genetic diversity. This is in agreement with many other studies in which the cross utility of microsatellite markers has been demonstrated to address paternity, geneflow and mating systems in different taxa of plants (Dayanandan et al., 1997; Terauchi and Konuma, 1994; Chase et al. 1996a; Chase et al., 1996b). The cross utility of microsatellite markers has also been demonstrated for several animal groups (Moore et al., 1991; Schlotterer et al. 1991; Levine et al., 1995).

4.1 Within population genetic diversity in *R. arboreum*

The genetic diversity within populations of *R. arboreum* is similar to that reported for related species of *Rhododendron* with similar habitat and life history traits (Kameyama et al., 2001; Jain et al., 2000; Singh et al., 2003). For example, the mean observed heterozygosity (Ho=0.801) of *R. arboreum* was comparable to the observed heterozygosity (Ho) values of *R. metternichii*, a shrub restricted to mountain regions of

Japan (Kameyama et al., 2001). Similarly a recent study by Jain et al. (2000) reported high level of genetic diversity among different populations of *R. arboreum* in the foothills of the Himalayas. However, the present study is the first to show levels of genetic variability at a microsatellite locus for *R. arboreum* from the eastern Himalayas.

In the present study, at the population level, the observed heterozygosity (Ho) values of *R. arboreum* are high and similar to the expected heterozygosity (He) values. Jain et al. (2000) also reported a high level of observed heterozygosity for *R. arboreum*. The high level of observed heterozygosity could be attributable to predominant outcrossing in the species (Jain et al., 2000). This species is thought to be predominantly pollinated by insects, which supports high outcrossing rates as in other insect pollinated plants (Hamrick and Godt, 1996). In the present study, the maximum number of alleles (15) was found in the locus 9D6 followed by 1D1 (11 alleles) and 2D2 (9 alleles). However, in *R. metternichii* the locus 9D6 showed a low number of alleles (6 to 8) and other loci 2D2 and 1D1 showed 10 alleles each (Naito et al., 1998 and Kameyama et al., 2001). Populations did not differ significantly in the mean number of alleles observed. However the disturbed populations FMD (8.66) and PJD (10.33) had a relatively higher number of alleles compared to their corresponding protected populations. This is primarily due to the occurrence of private alleles in these populations. Although private alleles in a population is a measure of genetic distinctiveness, these results have to be interpreted with caution as they are dependent on the sample size. In the present study, private alleles were detected in two out of five populations, four in PJD, and one in FMD. This result reveals that private (unique) alleles are concentrated in populations that are disturbed. The distribution of alleles and the presence of private alleles demonstrates the importance of protecting such distinct and disturbed populations from further degradation.

Another measure reflecting the number of alleles is the effective number of alleles. The effective number of alleles is less influenced by rare alleles. In the present study, at all populations and all loci, the mean number of observed alleles is relatively more than the effective number of alleles in all populations, reflecting unequal frequencies of alleles at a given locus. For example, in population PJD the number of alleles observed (Na) is 10 as compared to the effective number of alleles (6) required to maintain a heterozygosity of 0.78.

In summary, within a population, diversity estimates for *R. arboreum* show that the observed heterozygosity levels are comparable to those reported for *R. metternichii*, a short tree distributed in southwest Japan (Kameyema et al., 2001). The Ho estimates are much higher as compared to *R. farrerae* and *R. simsii*; which are widespread and abundant in South China (Ng and Corlett, 2000). Moreover, Hamrick (1994) reported genetic diversity estimates of tropical and temperate species: *R. arboreum,* a temperate species maintains a higher level of genetic variation compared to tropical tree species.

4.2 Among population genetic diversity in *R. arboreum*

F_{ST} is a measure of genetic differentiation or evolutionary divergence among populations across the geographical distribution range of a species. The mean F_{ST} value of 0.052 observed over populations of *R. arboreum* indicates a low level of genetic differentiation, suggesting that 5.2% of the variation is partitioned among populations and the remaining variation occurs within populations, a trend typical of tree species (Hamrick et al., 1992).

In pairwise population comparisons, the F_{ST} ranged from 0.015 to 0.057 and populations PMD and PMP exhibited the lowest F_{ST} of 0.015 (Table 9). Population pairs PMD and PJP as well as PMP and PJP showed the highest F_{ST} (0.057). Ng and Corlett

(2000), based on allozyme studies, reported different levels of F_{ST} for common and rare and endemic species of *Rhododendron* in mountain regions of China. The F_{ST} value for *R. arboreum* is comparable to that of *R. simsii*, (F_{ST}=0.056), a widely distributed tree species in mountainous grassland regions of China (Ng and Corlett, 2000).

Although the level of population differentiation is low, the statistically significant correlation between population differentiation and geographic distance within a relatively short distance (5 km) is striking. These results suggest that, despite high levels of gene flow between the populations, genetic structuring of *R. arboreum* populations in the eastern Himalayas is correlated with geographic distance. The following explanations are consistent with the available ecological data on the species. First, in temperate mountainous environments such as the eastern Himalayas, movement of pollen by pollinators for a long geographic distance could be influenced by low temperature, strong winds and short growth season leadings to geographical structuring of pollen movement between populations. Kudo (1993) reported a significant reduction in fruit set in *R. auruem*, a species restricted to mountains in Japan suggesting pollinator limitation due to extreme environmental conditions. *Rhododendrons* are known to be pollinated by birds (Subramanya and Radhamani 1993), bumblebees and flies. Hirao et al. (2006) reported bumblebees and flies as major pollinators of *R. aureum*. Stout (2007) reported bumblebees as the most common pollinator of *R. ponticum*. Although the pollinators of *R. arboreum* are unknown, they are most likely pollinated by birds and bees, leading to high level of out crossing and low population differentiation. However, relatively more near neighbor pollination events compared to long distance pollination may give rise to geographically structured populations.

The dispersal distance of seeds through wind is inversely related to the geographical distance, which could lead to genetic structuring in wind dispersed species. *Rhododendron* species produce many small seeds, which could be considered as an

adaptation to wind dispersal in alpine conditions. The small size and morphology of *R. arboreum* seeds suggests wind as a primary dispersal agent (Kameyama et al., 2000). Even if seeds reach long distances, their establishment may be limited due to constraints on seed germination and seedling survival, such as limitation of light and nutrients in those habitats (Tiwari and Chauhan, 2006).

4.3 Comparison of genetic diversity in *R. arboreum* populations in protected and harvested forests.

The comparison of genetic diversity parameters between protected and harvested populations of *R. arboreum* showed no discernible pattern, except for the presence of four rare unique (private) alleles confined to the PJD population, and one in the FMD population. However, populations show geographically correlated genetic differentiation, possibly related to limited seed dispersal. This suggests that *R. arboreum* individuals in the disturbed sites may have colonized through seed sources from the adjacent protected forest. There is also a possibility that some individuals in the disturbed forests may have originated from remnants or cut stumps through coppicing after harvesting.

The presence of low frequency private alleles in one of the disturbed sites indicates that some individuals in the disturbed site may have originated through gene flow from outside source populations. The extensive harvesting of *R. arboreum* trees in the disturbed forests leads to canopy opening and soil perturbation, making sites suitable for the germination and establishment *R. arboreum* seeds dispersed through wind from outside source populations.

4.4 Conservation implications

Although the *R. arboreum* populations analyzed in the present study show a relatively low level of population differentiation, there is a significant correlation between geographic distance and population differentiation. In order to capture and maintain maximum genetic diversity, representative populations of *R. arboreum* distributed throughout its geographic distribution range should be conserved. These populations may serve as seed source for regeneration of adjacent harvested forests. However, the presence of private alleles in harvested populations indicates that a substantial amount of genetic diversity is present in harvested populations. Further studies are needed to gain insights into the source of these private alleles as well as average distance of gene flow to determine the optimal distance for establishing protected populations of *R. arboreum.*

In summary, *R. arboreum* possess relatively high levels of genetic diversity within populations and the populations show relatively low, but geographically correlated genetic differentiation. Moreover, a considerable level of genetic diversity is found in harvested forests. Therefore, efforts should be made to conserve representative populations distributed throughout the distribution range. Further studies are needed to assess the gene flow distances to determine the optimal distances for establishing protected populations as a means for effective conservation of *R. arboreum* genetic resources.

Literature Cited

AKAGI, H., Y. YOKOZEKI, A. INAGAKI, AND T. FUJIMURA. 1997. Highly polymorphic microsatellites of rice consist of AT repeats, and a classification of closely related cultivars with these microsatellite loci. *Theoretical and Applied Genetics* 94: 61–67.

AMOS, B., C. SCHLOTTERER, AND D. TAUTZ. 1993. Social structure of pilotwhales revealed by analytical DNA profiling. *Science* 260: 670-672.

ANONYMOUS. 1976. Geological and mineral resources of the states of India. Geological Survey of India Misc. Publication No. 30, Govt. of India.

ANONYMOUS. 2005. Statistical abstract of Arunachal Pradesh. Directorate of Economics and Statistics. Government of Arunachal Pradesh.

ASHLEY, M. V., AND B. D. DOW. 1994. " The use of microsatellite analysis in population biology: background, methods and potential applications" *In* Schierwater, B. B, Streit, G. P. Wagner, and R. DeSalle, [eds.], Molecular ecology and evolution: approaches and applications, 185- 201. Birkhauser, Basel.

BRUFORD, M. AND R. WAYNE. 1993. Microsatellites and their application to population genetic studies. *Current Opinion in Genetics and Development* 3: 939-943.

BUTCHER, P. A., J. C. BELL, AND G. F. MORAN. 1992. Patterns of genetic diversity and nature of the breeding system in *Melaleuca alternifolia* (Myrtaceae). *Australian Journal of Botany* 40: 365–375

CHASE, M., R. KESSELI, AND K. S. BAWA. 1996a. Microsatellite markers for opulation and conservation genetics of tropical trees. *American Journal of Botany* 1: 51-57.

CHASE, M. R., C. MOLLER, R. KESSELI, AND K. S. BAWA. 1996b. Distant gene flow in tropical trees. *Nature* 383: 398-399.

CLINTON, B. D., AND J. M. VOSE. 1996. Effects of *Rhododendron maximum* L. on *Acer rubrum* L. seedling establishment. *Castanea* 61: 38–45.

CREGAN, P. B., A. A. BHAGWAT, M. S. AKKAYA, AND J. RONGWEN. 1995. Microsatellite fingerprinting and mapping in soybean. *Methods in Molecular Cell Biology* 5: 49-61.

CROW, J. F., AND M. KIMURA. 1964. The theory of genetic loads. *Proceedings of the XIth International Congress of Genetics,* Vol. 2, pp. 495–505. Caryologia 6 (supplement).

DAYANANDAN, S., K. S. BAWA, AND R. V. KESSELI. 1997. Conservation of microsatellites among tropical trees (Leguminosae). *American Journal of Botany* 84:1658-1663.

DAYANANDAN, S., O. P. RAJORA, AND K. S. BAWA. 1998. Isolation and characterization of microsatellites in trembling aspen (*Populus tremuloides*). *Theoretical and Applied Genetics* 96: 950-956.

DIETRICH, W., H. KATZ, S. E. LINCOLN, H. S. SHIN, J. FRIEDMAN, N. DRACOPOLI, AND E. S. LANDER. 1992. A genetic map of the mouse suitable for typing intraspecific crosses. *Genetics* 131: 423-447.

FELSENSTEIN, J. 2005. PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

FSI 2005. State of forest report 2005. Forest Survey of India. Ministry of Environment and Forests, Dehradun, India.

HAMRICK, J. L. 1994. Genetic diversity and conservation in tropical forests. *In* Drysdale, R. M., S. E. T. John, and A. C. Yapa, [eds.], *Proceedings of the International*

Symposium on Genetic Conservation and Production of Tropical Forest Tree Seed, pp. 1-9. ASEAN-Canada Forest Tree Seed Centre Project, Muak-Lek, Saraburi, Thailand.

HAMRICK, J. L., AND M. J. W. GODT. 1996. Effects of Life History Traits on Genetic Diversity in Plant Species. *Philosophical Transactions: Biological Sciences* 351: 1291-1298.

HAMRICK, J. L., M. J. W. GODT, AND S. L. SHERMAN-BROYLES. 1992. Factors influencing levels of genetic diversity in woody plant species. *New Forests* 6: 95-124.

HIRAO, A. S., Y. KAMEYAMA, M. OHARA, Y. ISAGI, AND G. KUDO. 2006. Seasonal changes in pollinator activity influence pollen dispersal and seed production of the alpine shrub *Rhododendron aureum* (Ericaceae). *Molecular ecology* 15: 1165-1173.

HOKANSON, S. C., A. K. SZEWC-MCFADDEN, W. F. LAMBOY, AND J. R. MCFERSON. 1998. Microsatellite (SSR) markers reveal genetic identities, genetic diversity and relationships in a Malus x domestica Borkh. *Theoretical and Applied Genetics* 97: 671-683.

HOOD, G. M. 2004. PopTools version 2.6.2 Available on the internet. URL http://www.cse.csiro.au/poptools.

HORA, B. 1981. The Oxford Encyclopedia of Trees of the World, Oxford University Press, Conscent Books, New York.

HUDSON, R. R., M. SLATKIN, AND W. P. MADDISON. 1992. Estimation levels of gene flow from DNA sequence data. Genetics 132: 583-589.

JAIN, A., M. K. PANDIT, S. ELAHI, A. JAIN, A. BHASKAR, AND V. KUMAR 2000. Reproductive behavior and genetic variability in geographically isolated populations of *R. arboreum* (Ericaceae). *Current Science* 79(9): 1377-1381.

JARNE, P., AND P. J. L. LAGODA. 1996. Microsatellites, from molecules to populations and back. *Trends in Ecology and Evolution* 11: 424-429.

KAISHEVA, M. V. 2006. The effect of metals and soil pH on the growth of Rhododendron and other alpine plants in limestone soil. Ph.D. Thesis, College of Science and Engineering, University of Edinburgh, Scotland.

KAMEYAMA, Y., Y. ISAGI, K. NAITO, AND N. NAKAGOSHI. 2000. Microsatellite analysis of pollen flow in *Rhododendron metternichii* var. *hondoense. Ecological Research* 15: 263–269.

KAMEYAMA, Y., Y. ISAGI, AND N. NAKAGOSHI. 2001. Patterns and levels of gene flow in *Rhododendron metternichii* var. *hondoense* revealed by microsatellite analysis. *Molecular Ecology* 10: 205-216.

KAMEYAMA, Y., Y. ISAGI, AND N. NAKAGOSHI. 2002. Relatedness structure in *Rhododendron metternichii* var. hondoense revealed by microsatellite analysis. *Molecular Ecology* 11:519-527.

KAUL, R. N., AND K. HARIDASAN. 1987. Forest types of Arunachal Pradesh - A preliminary study. *Journal of Economic and Taxonomic Botany* 9(2): 379-388.

KUDO, G. 1993. Relationships between Flowering Time and Fruit Set of Entomophilous Alpine Shrub, *Rhododendron aureum* (Ericaceae) Inhabiting Snow Patches., *American Journal of Botany* 80 (11):1300 - 1304.

LEMES, M. R., R. GRIBEL, J. PROCTOR, AND D. GRATTAPAGLIA. 2003. Population genetic structure of mahogany (*Swietenia macrophylla King, Meliaceae*) across the Brazilian Amazon based on variation at microsatellite loci: implications for conservation. *Molecular Ecology* 12: 2845-2883.

LEVINE, I., H. CHENG, J. BACTER, AND J. HILLEL. 1995. Turkey microsatellite DNA loci amplified by chicken specific primers. *Animal Genetics* 26:107–110.

LEWIS, P. O., AND D. ZAYKIN. 2001. Genetic data analysis: computer program for the analysis of allelic data. Version 1.0, Department of Ecology and Evolution, University of Connecticut.

MAO, A. A., K. P. SINGH, AND P. K. HAJRA. 2001. Rhododendrons. In Singh, N. P., and D. K. Singh [eds.], Floristic Diversity and Conservation Strategies in India, 2167–2202. vol. IV. Botanical Survey of India, Kolkata, India.

MARIETTE, S., V. LECORRE, F. AUSTERLITZ, AND A. KREMER. 2002. Sampling within the genome for measuring within-population diversity: trade-offs between markers. *Molecular Ecology* 11: 1145–1156.

MCCOUCH, S. R., X. CHEN, O. PANAUD, S. TEMNYKH, Y. Xu, Y.G. CHO, N. HUANG, T. ISHII, AND M. BLAIR. 1997. Microsatellite marker development, mapping and applications in rice genetics and breeding. *Plant Molecular Biology* 35: 89-99.

MOORE, S. S., L. L. SARGEANT, T. J. KING, J. S. MATTICK, M. GEORGES, AND D. J. S. HETZEL. 1991. The conservation of dinucleotide microsatellites among mammalian genomes allows the use of heterologous PCR primer pairs in closely related species. *Genomics* 10: 654-660.

MORGANTE, M., AND A. M. OLIVIERI. 1993. PCR-amplified microsatellites as markers in plant genetics. *The Plant Journal* 3: 175–182.

MOXON, E. R., AND C. WILLS. 1999. DNA microsatellites: agents of evolution? *Scientific American* 1: 72-77.

NAITO, K., Y. ISAGI, AND N. NAKAGOSHI. 1998. Isolation and characterization of microsatellites of *Rhododendron metternichii* Sieb. Et Zucc. Var. hondoense Nakai. *Molecular Ecology* 7:927-928.

NEI, M. 1972. Genetic distance between populations. *American Naturalist* 106: 283-292.

NEI, M. 1978. Estimation of average heterozygosity and genetic distance from a Small number of individuals. *Genetics* 147: 1943- 1957.

NG, SAI-CHIT, AND R. T. CORLETT. 2000. Genetic variation and structure in six *Rhododendron* species (Ericaceae) with contrasting local distribution patterns in Hong Kong, China. *Molecular Ecology* 9: 959–969.

Nybom, H. 2004. Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology* 13: 1143-1155

PASCOE, E. H. 1950. A manual of the geography of India and Burma. 2 Vols. Geological Survey of India, Govt. of India Publication.

PAUL, A. 2008. Studies on diversity and regeneration ecology of rhododendrons in Arunachal Pradesh. Ph.D. Thesis, Assam University, Silchar, India.

PORNON, A., N. ESCARAVAGE, P. THOMAS, AND P. TABERLET. 2000. Dynamics of genotypic structure in clonal *Rhododendron ferrugineum* (Ericaceae) populations. *Molecular Ecology* 9: 1099 - 1111.

PRADHAN, U. C., AND S. T. LACHUNGPA. 1990. Sikkim Himalayan Rhododendrons, Primulaceae Books, Kalimpong.

QUELLER, D. C., J. E. STRASSMANN, AND C. E. HUGHES. 1993. Microsatellites and kinship. *Trends in Ecology and Evolution* 8: 285–288.

REYNOLDS, J., B. S. WEIR, AND C. C. COCKERHAM. 1983. Estimation of the coancestry coefficient: basis for a short-term genetic distance. *Genetics* 105: 767–779.

ROSSETTO, M., R. W. SLADE, P. R. BAVERSTOCK, R. J. HENRY, AND L. S. LEE. 1999. Microsatellite variation and assessment of genetic structure in teatree (*Melaleuca alternifolia*- Myrtaceae). *Molecular Ecology* 8: 633-643.

ROTHERHAM, I. D. 1983. The ecology of *Rhododendron ponticum* L. with special reference to its competitive and invasive capabilities, Ph.D. Thesis, University of Scheffielt, UK.

SCHLOTTERER, C., B. AMOS, AND D. TAUTZ. 1991. Conservation of polymorphic simple sequence loci in cetacean species. *Nature* 354: 63-65.

SCHLOTTERER, C., AND J. PEMBERTON. 1994. The use of microsatellites for genetic analysis of natural populations. *In* Schierwater, B., B. Streit, G. P. Wagner, R. DeSalle [eds.], Molecular Ecology and Evolution: Approaches and Applications, pp. 203-214. Birkhauser Verlag Basel, Switzerland.

SINGH, K. K., S. KUMAR, L. K. RAI, AND A. P. KRISHNA. 2003. Rhododendron conservation in Sikkim Himalaya. *Current Science* 85: 602–606.

SLATKIN, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* 139: 457-462.

SLATKIN, M., AND N. H. BARTON.1989. A comparison of three indirect methods for estimating average levels of gene flow. *Evolution* 43:1349-1368.

STEINKELLNER, H., C. LEXER, E. TURETSCHEK, AND J. GLOSSL. 1997. Conservation of (GA)n microsatellite between *Quercus* species. *Molecular Ecology* 6: 1189–1194.

STOUT, J. C. 2007. Pollination of invasive *Rhododendron ponticum* (Ericaceae) in Ireland. *Apidologie* 38: 198-206.

STRAND, M., T. A. PROLLA, R. M. LISKAY, AND T. D. PETES. 1993. Destabilisation of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. *Nature* 365: 274-276.

STREIFF, R., A. DUCOUSSO, C. LEXER, H, STEINKELLLNER, J, GLOESSL, AND K. KREMER. 1999. Pollen dispersal inferred from paternity analysis in a mixed oak stand of *Quercus robur* L. & *Q. petraea* (Matt.) Liebl. *Molecular Ecology* 8: 831-841.

SUBRAMANYA, S., AND T. R. RADHAMANI. 1993. Pollination by birds and bats. *Current Science* 65(3): 201-209.

TERAUCHI, R., AND A. KONUMA. 1994. Microsatellite polymorphism in *Dioscorea tokoro*, a wild yam species. *Genome* 37:794-801.

TIWARI, O. N., AND U. K. CHAUHAN. 2005. Genus *Rhododendron* status in Sikkim Himalaya: an assessment. *American Rhododendron Society Journal* 59(3): 147.

TIWARI, O. N., AND U. K. CHAUHAN. 2006. Rhododendron conservation in Sikkim Himalaya. *Current Science* 90(4): 532-541.

WEIR, B. S., AND C. C. COCKERHAM. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358 -1370.

WEISSENBACH, J., G. GYAPAY, C. DIB, A. VIGNAL, J. MORISSETTE, P. MILLASSEAU, G. VAYSSEIX, AND M. LATHROP. 1992. A second generation linkage map of the human genome. *Nature* 359: 794-801.

WRIGHT, S. 1951. The genetical structure of populations. *Annals of Eugenetics* 15: 323-354.

YEH, F.C., R. C. YANG, T. B. J. BOYLE, Z. H. YE, AND J. X. MAO 1997. POPGENE, a user friendly shareware for population genetic analysis. Molecular Biology and Biotechnology Centre, University of Alberta, Edmonton, Canada.

Appendix 1: Protocols

Protocol 1: DNA Extraction Procedure.

1- Grind 0.5 g of leaves using mortar and pestle in the presence of liquid nitrogen. Put the leaf powder (0.35- 40 g) in centrifuge tube and add 1 ml Extraction Buffer. Place on ice 45 minutes.

2- Centrifuge the sample(s) at 7000 rpm (3000 G) for 8 minutes at room temperature. Pour off the supernatant. Pellet(s) may be frozen at -80 °C for future extraction, if necessary.

3- Resuspend the pellet with 600 µl of Extraction/Nuclei Lysis (1:1) Buffer. Add RNAse, if required. Add 120 µl of 5% Sarkosyl solution. Mix with a vortex mixer.

4- Incubate at 63 °C for at least 30 minutes in water bath. Preferably vortex the tube(s) about halfway through the heating procedure.

5- In the fume-hood, add 600 µl of Chloroform/Isoamyl Alcohol (24:1) mixture. Mix gently by inverting the tubes 20 to 25 times to form an emulsion. Spin at 7000 rpm (5000 G) for 5 minutes. Transfer the top aqueous phase to a new 15 ml centrifuge tube containing 600 µl of Isopropanol. Gently invert tube(s) and allow the DNA to precipitate out of solution at room temperature for a minimum of 30 minutes.

6- Spin at 14000 rpm (20000 G) for 10 minutes. Decant off the supernatant.

7- Wash the pellet(s) with 750 µl of 70% ethanol. Finger vortex and leave 5 minutes. Centrifuge at 14000 rpm (20000 G) for 10 minutes. Pour off the supernatant and dry in the vacuum overnight.

8- Dissolve the dried pellet(s) in 150 µl of TE Buffer.

Extraction Buffer (500 ml):

Sorbitol (MW 182.2)	31.9 g
Sodium bisulfite	1.9 g
1 M Tris (pH 7.5)	50 ml
0.5 M EDTA	5 ml

Add distilled water to volume.

Nuclei Lysis Buffer (200 ml):

NaCl	23.4 g
CTAB (MW 364.5)	4.0 g
1 M Tris (pH 7.5)	40 ml
0.5 M EDTA	20 ml

Add distilled water to volume. It may be necessary to heat this solution in order to aid in dissolving.

Protocol 2: 1% agarose gel electrophoresis.

For 100 ml gel:

Agarose powder	1 g
5 X TBE (Tris-borate)	20 ml
dH_2O	80 ml

Microwave for about 1 minute to dissolve the agarose. Leave it to cool for 5 minutes down to about 60°C, and add 3.3 µl of ethidium bromide (10g/L) and swirl to mix.

5 X TBE (Tris-Borate):

Tris Base (MW 121.14)	54 g
Boric Acid (MW 61.83)	27.5 g
0.5 M EDTA (PH 8.0)	20 ml

Dissolve the ingredients in final volume of 1L of distilled water.

Protocol 3: PCR thermal cycling program.

Lid preheated to 105 °C

1- Denaturation	94 °C for 2 minutes
	94 °C for 1 minute
2- Annealing	Optimum temperature for 30 seconds
(R= 3 °C/S)	
3- Elongation	72 °C for 1 minute
4- Go to step 2	Repeat 30 34 times
5- Final elongation	72 °C for 4 minutes
6- Cooling and storage	4 °C

Optimum temperatures for each primer pair are written in Table 5.

Protocol 4: Genotyping procedure.

1- Mix and vortex 50 µl of Hi-Di formamide (deionized formamide) with 0.1 µl of Tamara GeneScan-500^{TM} size standard (Applied Biosystems). It depends on the amount of samples. For example, mix 500 µl of Hi-Di formamide with 5 µl of size standard for 45 samples.

2- Add 10 µl from mixture of Hi-Di formamide and size standard to each strip capst tube.

3- Add 1 µl of 1:20 dilution of PCR products to primer pairs.

4- Denature at 94 °C for 2 minutes and place on ice for 5 minutes.

5- Proceed to genotype with the ABI 310 genetic analyzer and Genescan software (Applied Biosystems, Foster City =, California).

Appendix 2: Hardy-Weinberg Equilibrium

The Hardy–Weinberg principle states that the genotype frequencies in a population remain in equilibrium (constant) from generation to generation unless specific disturbing influences include non-random mating; new mutations, selection, random genetic drift and gene flow are introduced. When the hardy–Weinberg assumptions are met these disturbing influences can cause deviations from expectation, but depending which assumption is met, such deviations may or may not be statistically detectable.

The Hardy-Weinberg model enables to compare a population's actual genetic structure over time with the genetic structure we would expect if the populations are in Hardy-Weinberg equilibrium (i.e., not evolving). If genotype frequencies differ from those we would expect under equilibrium, we can assume that one or more of the model's assumptions are being violated, and attempt to determine which one(s).

The Hardy-Weinberg model could calculate allele and genotype frequencies. Because we are dealing with frequencies, both equations must add up to 1. The equation $p + q = 1$ describes allele frequencies for a gene with two alleles. (but the equation can also be modified and used with three or more alleles.) If we know the frequency of one allele (p) we can easily calculate the frequency of the other allele (q) by $1- p = q$.

In a single locus with two alleles A and a with allele frequencies of p and q, respectively, the HWP predicts that the genotypic frequencies for the AA homozygote to be p^2, the Aa heterozygote to be $2pq$ and the other aa homozygote to be q^2. The frequency of the possible diploid combinations of these alleles: AA, Aa, or aa, is expressed as:

$p^2 + 2pq + q^2 = 1$

This form of the equation is expected to cases where three or more alleles are present.

Appendix 3: Test Statistics

Test 1: Expected heterozygosity.

Heterozygosity is of major measures to genetic variation in natural populations. It shows about the structure and even history of a population. High heterozygosity means lots of genetic variability. Under Hardy-Weinberg equilibrium (HWE), if the observed heterozygosity is low it might invoke forces such as inbreeding. If heterozygosity is high, there is a suspicion in an isolate-breaking effect.

The value of heterozygosity measures range from zero (no heterozygosity) to nearly 1.0 (for a system with a large number of equally frequent alleles). In expected heterozygosity (H_E, or gene diversity, D) the simplest way to calculate it for a single locus is as:

$$1 - \sum_{i=1}^{k} P_i^2$$

(Nei,1987)

where p_i is the frequency of the i^{th} of k alleles. If we want the gene diversity over several loci we need double summation and subscripting as follows:

$$1 - \frac{1}{m} \sum_{j=1}^{m} \sum_{i=1}^{k} p_i^2$$

Test 2: F_{ST} (Hierarchical F coefficients)

F_{ST} is a parameter applied to measure population differentiation, (Wright 1951; Cockerham and Weir 1987). It is used based on genetic polymorphism, such as microsatellites. F-statistics (Wright 1951; Nei 1987; Weir 1996) are very commonly used (where T stands for total) quantifies the differences in allele frequencies among populations.

Estimation of F_{ST} for over multiple loci:

Deviation from HWE within subpoplutions: $F_{IS} = (H_S - H_O)/ H_S$

Deviation from HWE in total populations: $F_{IT} = (H_T - H_O)/ H_T$

Genetic differentiation among subpopulations: $F_{ST} = (H_T - H_S)/ H_T$

Based on the above formula

H_O = average observed heterozygosity within a subpopulations over all loci

H_S = average expected heterozygosity within subpopulations over all loci

H_T = the average of the expected heterozygosity over all subpopulations

In F_{ST} the range of the values is from 0 (no genetic differentiation) to 1 (fixation of alternative alleles).

According to Wright:

Little, moderate, high and very high differentiation ranges from 0 – 0.05, 0.05 – 0.15, 0.15 – 0.25, and > 0.25, respectively.

One of the most commonly used of estimators of F coefficients is theta (θ) that corrects error associated with differences in population sizes and is defined as an unbiased estimator of F_{ST} (Weir and Cockerham 1984).

F_{ST} is related to Nm (number of migrants per generation) by the following formula:

$Nm = 1- F_{ST} /4 F_{ST}$

Where N is the local populations size and m is the migrations rate among populations.