DNA BARCODING OF NEPALESE GYMNOSPERMS

M.Sc. thesis
(2020)

Submitted to

CENTRAL DEPARTMENT OF BIOTECHNOLOGY
TRIBHUVAN UNIVERSITY
Kirtipur, Kathmandu, Nepal

For Partial Fulfillment of the Requirement for
M. Sc. Degree in Biotechnology

Submitted by

Ranjeeta Odari

Exam Roll No.: BT 312/072
T. U. Registration No: 5-2-0282-0274-2011

Supervisors

Prof. Tribikaram Bhattarai, PhD
Central Department of Biotechnology
Tribhuvan University Kirtipur, Nepal

Ram Chandra Poudel, PhD
Senior Scientific Officer
Nepal Academy of Science and
Technology (NAST), Khumaltar, Nepal

NEPAL ACADEMY OF SCIENCE & TECHNOLOGY
CENTRAL OFFICE
NAST

Date: 23 January, 2020

RECOMMENDATION

Ms. Ranjeeta Odari, who enrolled in Master of Science in Central Department of Biotechnology, Tribhuvan University, Kirtipur, Nepal has successfully conducted her M.Sc dissertation research work on "**DNA Barcoding of Nepalese Gymnosperms**" in our laboratory for the partial fulfillment of her academic program. The work conducted by Ms. Odari was mutually supervised by me and Prof. Dr. Tribikram Bhattarai from Central Department of Biotechnology, T.U.

I wish her all the best and great success in her research career.

....

Ram Chandra Poudel, Ph.D.
Supervisor
Senior Scientific Officer
Molecular Biotechnology Unit
Nepal Academy of Science and Technology (NAST)
Khumaltar, Lalitpur, Nepal

Address : Khumaltar, Lalitpur, Nepal, GPO Box 3323 Kathmandu, E-mail: Info@nast.gov.np
Telephone: +977-1-5547715, 5547720, 5547721, 5553132, Fax: 977-1-5547713
May 2018

Tribhuvan University
CENTRAL DEPARTMENT OF BIOTECHNOLOGY
Kirtipur, Kathmandu, Nepal

Date: Jan 23, 2020

RECOMMENDATION

This is to certify that **Ms. Ranjeeta Odari** has successfully completed her dissertation work entitled **“DNA Barcoding of Nepalese Gymnosperms”** under our supervision.

This thesis work was performed for the partial fulfillment for award of Master of Science in Biotechnology under the course code BT 621. The result presented here is her original findings. We, hereby, recommend this thesis for final evaluation.

Tribikram Bhattarai
....
Prof. Tribikram Bhattarai, Ph.D
(Supervisor)
Central Department of Biotechnology
Tribhuvan University
Kirtipur, Kathmandu, Nepal

....
Dr. Ram Chandra Poudel
(Supervisor)
Senior Scientific Officer
Nepal Academy of Science and Technology (NAST)
Khumaltar, Lalitpur, Nepal

Tribhuvan University
CENTRAL DEPARTMENT OF BIOTECHNOLOGY
Kirtipur, Kathmandu, Nepal

Date: **Jan 23, 2020**

CERTIFICATE OF EVALUATION

This is to certify that this thesis entitled **"DNA Barcoding of Nepalese Gymnosperms"** presented to evaluation committee by **Ms. Ranjeeta Odari** is found satisfactory for the partial fulfillment of Master of Science in Biotechnology.

Prof. Krishna Das Manandhar, Ph.D
(Head of Department)
Central Department of Biotechnology
Tribhuvan University
Kirtipur, Kathmandu, Nepal

Prof. Tilak Ram Shrestha, PhD
(External Examiner)
Balaju, Kathmandu, Nepal

Tribikram Bhattarai

Prof. Tribikram Bhattarai, Ph.D
(Supervisor)
Central Department of Biotechnology
Tribhuvan University
Kirtipur, Kathmandu, Nepal

Dr. Ram Chandra Poudel
(Supervisor)
Senior Scientific Officer
Nepal Academy of Science and Technology (NAST)
Khumaltar, Lalitpur, Nepal

Prof. Rajani Malla, Ph.D
(Internal Examiner)
Central Department of Biotechnology
Tribhuvan University
Kirtipur, Kathmandu, Nepal

ACKNOWLEDGMENT

Foremost, I would like to express my sincere gratitude to my respected supervisors **Prof. Dr. Tribikram Bhattarai**, Central Department of Biotechnology, T.U. and **Dr. Ram Chandra Poudel**, Senior Scientific Officer, Molecular Biotechnology Unit, Nepal Academy of Science and Technology (NAST) for their continuous support motivation, encouragement, constructive remarks and dynamic supervision throughout the study. I could not have imagined having a better supervisors and mentors for my Master thesis study.

My sincere thanks also goes to **Dr. Sunil Babu Shrestha**, Vice-chancellor of NAST, **Dr. Mahesh Kumar Adhikari**, Secretary of NAST, **Mrs. Jay Shree Sijapati**, Head of Faculty of Science and **Dr. Jyoti Mahrajan**, Coordinator of Molecular Biotechnology Laboratory of NAST for entire financial support in my laboratory work and allowing me to carry out my thesis work at Molecular Biotechnology Laboratory under **Master Thesis Support Program (2074/2075)**. I would like to acknowledge **Dr. Deegendra Khadka**and **Dr. Kishor Pandey**for valuable suggestions and help.

My sincere acknowledgement goes to the **Department of National Parks and Wildlife Conservation (DNPWC)** for permitting me to collect sample from different National Parks of Nepal. Also, my gratitude goes to whole team of **National Herbarium and Plant Laboratories (KATH), Godawari**, Lalitpur for providing me the opportunity of developing the preliminary knowledge about herbaria samples and especially to **Dr. Ganga Dutta Bhatta** for his expert advice and guidelines to carry out this research and also helping me crucially for knowing more about the subject.

I would like to mention my gratitude to **Mr. Deepak Pant** and/of **Central Department of Botany, TU** for assiting me through sample identification. I would also thank Suprim Shrestha, Sapna Shakya, Sarbesh Rijal, Rabin Dahal, Prashant Poudel, Deepa Sharma, Surendra Kumar Subedi for helping me through site identification and sample collection.

My sincere acknowledgement goes to **Prof. Dr. Krishana Das Manandhar**, Head of Department, Central Department of Biotechnology, T.U., Prof. Dr. Rajani Malla, Prof. Dr. Gauri Shankar Manandhar, Prof. Dr. Ganga Prasad Kharel, Prof.Dr.Mohan Kharel, Prof. Dr. Tilak Ram Shrestha, Dr. Pramod Aryal, Dr. Suresh Subedi, Mrs. Jarina Joshi, Mr. Bal Hari Poudel, Mrs. Smita Shrestha, Mrs. Alina Sapkota, Mrs. Mamata Bhattarai and entire teaching faculty members, staffs of Central Department of Biotechnology for their support, encouragement and valuable suggestions.

I would also like to express my gratitude to my entire dear colleagues **7th batch CDBT, T.U.**, seniors, juniors of Central Department of Biotechnology for all those days spent together. You are all amazing people I would never forget your constant support and friendship.

I would also like to thank **University Grant Commission (UGC)**, Sanothimi, Bhaktapur for financial support and awarding **Master's thesis support (2074/2075)** to carry out the field work and preparing the thesis.

Finally, my gratitude goes to entire NAST family and to my laboratory colleagues Ms. Pranita Poudyal, Mr. Prithvi Raj Gurung, Mr. Rupesh Kr. Sah and respected seniors Ms. Medha KC, Ms. Rashmi Thapa, Ms. Barsanti Gautam and Mr. Sandeep Rijal for their cooperation in my thesis work. My work would not have been complete without valuable suggestions and technical support of respected seniors Mr. Pradip Kumar Chaudhary and Mr. Mitesh Shrestha.

My deepest gratitude to my dear friend **Sarbesh Rijal, mymother, brothers and sisters, my friends** and all the concerned **relatives** for their motivation, encouragement, supports and continuous trust to my ability until the successful completion of this research.

…………………………….

Ranjeeta Odari

LIST OF ABBREVIATIONS

µl	Microlitre
AEC	Agro Enterprise Centre
AFLP	Amplified Fragment Length Polymorphism
BLAST	Basic Local Alignment Search Tool
BOLD	Barcode of Life Data System
BP	Base Pairs
CBOL	Consortium for the Barcode of Life
CpDNA	Chloroplast DNA
CTAB	Hexadecyl (Cetyl) Trimethyl Ammonium Bromide
D/W	Double Distilled Water
DNA	Deoxyribonucleic Acid
DNTPs	Deoxynucleotide phosphates
EDTA	Disodium Ethylene Diamine Tetra Acetate
ETBr	Ethidium Bromide
GoN	Government of Nepal
ITS	Internal Transcribed Spacer
*ITS*2	Internal Transcribed Spacer region2
IUCN	International Union for Conservation of Nature
Kb	Kilobase
MatK	maturase K
MEGA	Molecular Evolutionary Genetic Analysis
ml	Mililitre
NAST	Nepal Academy of Science and Technology
NCBI	National Center for Biotechnology Information
NJ	Neighbour Joining

PCR	Polymerase Chain Reaction
rbcL	Ribulose 1, 5-bisphosphate carboxylase/oxygenase Large subunit
TAE	Tris-Acetate-EDTA
TE	Tris-EDTA
TU	Tribhuvan University
UGC	University Grant Commission
WHO	World Health Organization

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

The morphology-based identification of gymnosperms is not always easy and correct which shows the importance of molecular based identification. With an aim to contribute in gene-based identification of extanct gymnosperms targeting to aid morphology-based grouping to species, four different markers *MatK, ITS, Ycf1* and *petN-psbM* were selected and studied to know whether they could be helpful to aid in the identification of 59 samples of gymnosperms collected from Nepal that included 33 different species. The cladogram derived from the marker-based analysis was used to validate phylogenetic relationship between the species.

Efficiencies of *MatK* and *ITS* were analysed for all individuals species studied. *Ycf1* and *petN-psbM* were used to aid the strength of identification and efficiency of discrimination only in the individuals of Pinaceae and Cupressaceae family of gymnosperms respectively as these two families contained higher number of species. *MatK* was concluded efficient to discriminate species of gymnosperms in species level as there was significant difference between intraspecific and interspecific p-distance between the studied species. Further, phylogenetic tree given by *MatK* was the strongest and matched with maximum identity to global cladogram of gymnosperms. Discriminating power of *MatK* was increased with *ITS* by 10% in gymnosperms, with *Ycf1* by 31.3 % in Pinaceae, and with *petN-psbM* by 24 % in Cupressaceaeas. Cladograms were constructed with Neighbor-Joining method of phylogenetic tree construction based on *ITS, MatK, ycf1* (Pinaceae), *petN-psbM* (Cupressaceae) and the combinations *ITS+MatK, ycf1+MatK* (Pinaceae), *petN-psbM+MatK* (Cupressaceae). All of the trees generated could match the universal inference of global phylogeny based on existing clade-relationship. The position of species in *ycf1* tree and *petN-psbM* tree were strengthened as compared to that in *MatK* tree only, showing increased bootstrap values, in Pinceae and Cupressaceae respectively.

Present study on gymnosperms in Nepal has shown that *MatK* and *ITS* can differentiate samples of gymnosperms in species level, even if little amount of sample/DNA is available for examination. *Ycf1* and *petN-psbM* can strengthen the species resolution in Pinaceae and Cupressaceae family respectively.

Keywords: DNA barcoding, *ITS, MatK, Ycf1, petN-psbM*, Phylogeny

CHAPTER 1. **INTRODUCTION**

1.1 Background

Gymnosperms are recognized as an antique plant group for botanical study from historical biogeographic approach (Medina and Vega, 2007). The terminology coined by Theopratus, the gymnosperms (gymno –naked ; sperms- seeds) are the most ancient seed plants (Ohri & Khoshoo, 1986) being the connecting link between pteridophytes and the angiosperms which include about 17 families, 83 genera and 700-800 species among which conifers dominate with about 550-600 species belonging to 60 genera (Tewari *et al.*,2010,). Also defined by Robert Brown, as the plants with ovules not enclosed in ovary, they are one of two sub groups of seed-forming 'vascular plants', meaning that they contain a tubular tissue network that transports water and nutrients, and provides mechanical support (Sederoff, 2013). Their ovule forms on a leaf- like structure or on a scale or megasporophyll or on the apex of a shoot and the seeds at maturity are enclosed by fused cones scales or bracts, which (in some members) are fleshy causing fruiting structures (cones) frequenctly confused with berries (as in Junipers) (Mill & Chase, 2011). Unique features of gymnosperms include presence of transfusion tissues and resinal canals in the leaves with great diversity of forms or arrangements: small or large, simple or compound, and also in shape: pinnate as in Cycads, acicular as in conifers or fan-shaped as in species of *Ginkgo* (a living fossil) (Bhatnagar and Moitra, 2004).

Being arresting group of plants for the zealous researchers, the Gymnosperms have some of their earlier species already become extinct and some are preserved as fossils (Tiwari and Chauhan, 2006). The researches till date, upon the group, have dealt with morphology, taxonomy, phylogeny (phylogram and cladogram) and to some extent their physiology.

1.2 Gymnosperms for community

Besides the craze of most of the species of gymnosperms for ornamental assets, timbers and fuel, some ethnobotanical studies have shown how people of particular culture and region make use of autochthonous gymnosperms (Awale *et al.*, 2011 , Motomura *et al.*, 2007). Apart from their religious value as incenses (Junipers), medicinal functions such as treating coughs, hayfever, bronchitis, asthma, hepatic diseases, skin diseases, ulcers, rheumatism and syphilis, chest affection, catarrh of the bladder, urinary disorders, kidney stones, pulmonary affections, haemoptysis and their role as an antidote of snake venom, tannins, coloring matter, antithelmintic, astringent, anti-cancer, anti-induration, diuretic, carminative, stimulant, an anaesthetic drug 'ephedrine' (*Ephedra*) are well known among the local people (Rokaya *et al.*, 2014,Hussain *et al.*, 2006). In Ayurvedic medicine, the leaves of the conifer *Podocarpus neriifolius* (mount peak/ thitmin/ Gunsi) has been used to treat rheumatism and painful joints (Gautam *et al.*, 2005). Significant numbers of citations are available that supports the role of this gymnosperm in the therapeutic use against neurodegenerative conditions like dementia, ischaemia and gastro-realted stress (Shah, *et al.*, 2003). Dropping down the stress-related elevations of hormones, relaxing the cramps and relieving hypertension and cerebral problems by

Ginkgo extracts is well- understood (Nishida & Satoh, 2003). Gymnosperms show ranges of ecological significance from antioxidant properties and use as herbal medicine to toxic reactions. Phytochemical study of Nepalese *Podocarpus* also shows significant biological role of the species and potential for exploitation in herbal health (Gautam *et al.*, 2005). The most important and dragging matter of research in modern ethnobotany is the use of Taxol, an anticancer compound which is found in bark and leaves of *Taxus.* This aspect has also heightened the threat of the species to extinction (Poudel *et al.*, 2013) in many areas.

The significance of gymnosperms extends beyond just holding membership of world forest population. Gymnosperms like *Picea* are well known for controlling nitrate pollution and many other studied urban-cultivated species are found tolerant to dust, smoke and harmful aerosols (Norby *et al.*, 1989).

1.3 Morphology, phenotypic identification and problems

Morphology is basic and foundation approach for sample collection at site and primary description of a plant species. At the same time, there is constant need of experts for an accurate identification of the species as well as the determination of the origin-based relationship. The lack of experienced experts often doesnot allow this to be satisfied and demands a supplementary technique such as DNA barcoding (Shneyer & Rodionov, 2019).

Not only this, although there are enough phenotypic aids that clearly discriminates most of the gymnosperms, some species have minimal key-morphology such as species of *Ephedra*. But still, various key points and characters are proposed and well-explained by the botanist to identify gymnosperms. *Ephedra gerardiana* with straight micropylar tube is discriminated from *E. pachyclada* having curved and twisted micropylar tube (Ohba *et al.*, 2008). Nomenclature of family Cupressaceae was done on the basis of critical morphological analysis addressing different characteristic variations within the species (Laubenfelschad *et al.*, 2012). On the basis of examination of herbarium, monographic study and several revisions , the expertise could correct the pre-existing identification and revise the nomenclature in *Cycas* and *Podocarpus* species (Doyle, 1998). *Pinus* and other genus of Pinaceae, *Podocarpus Gnetum, Ginkgo, Taxus, Cryptomeria, Araucaria* and *Juniperus* have the species with comparatively enough morphological keys for identification than *Ephedra*. However, the families with very closely related species like Cupressaceae have been problematic from the phase of sample identification to genetic analysis in the aspect of difficulties in markers sort out that are strong enough to discriminate the species (Adams & Chaudhary, 2009).

In order to discriminate species which are hard to be identified, to make the identification of species accurate and to reveal potential new species, including cryptic ones, DNA barcoding is wanted and thus, chosen more often in the identification of gymnosperms. Further, most morphology-based identification is not always applicable. In cases where there is very small amount of biological material available for the examination (a leaf, bark or a twig), for example, in forensic laboratories where extremely low quantity of sample (most of which is highly degraded, most of the time) has to be processed (Pereira et al., 2008). Thus, DNA barcoding, which uses the amplification of a very small segment of DNA to give enough information about a species, is taken here as a supplementary technique for proper species-label to aid morphology-based identification in gymnosperms of Nepal.

1.4 DNA barcoding

The constraint of morphology-based identifications and lack of availabilty of taxo-expertises push the investigators to find new approach that helps labelling of unknown samples with correct identity (Hebert *et al.*, 2003). As the metaphor suggests, DNA barcoding identifies the plant species as similar as an eleven-digit Universal Product Code describes the unique product individually (Neigel *et al.*, 2007). DNA provides more information than proteins because of the degeneracy of the genetic code and the presence of large non-coding streches (Pereira *et al.*, 2008).

The most attractive supplement to morphological identification of plants is the DNA barcoding (Wilkinson *et al.*, 2017). A “ big science” program as regarded earlier, DNA barcoding is also considered as a renaissance of taxonomy (X. Li *et al.*, 2014). DNA barcodes are applied in ecological forensics, identification of traded materials, identifying plants in the shortage of taxonomic expertise and helping species discovery in some plant groups (Hollingsworth *et al.*, 2011). It has been used successfully in a variety of biological applications such as discovery of new or cryptic species, in biodiversity and assessments, in the detection of invasive species, in the identification of medicinal plants in mixtures and in ecological investigations (Gao, 2016).

Taxon identification using a standardized DNA region, DNA barcoding, entered *ITS* golden age in 2003 and is now further developed through an international initiative. The term DNA barcode was first used in 1993 and now it is used not only by the taxonomists but by the scientists of other fields like forensic science, biotechnology, food industries etc. The use of ‘DNA barcode’ for identification of taxon was first proposed by Paul Hebert of University of Guelph in 2003 (Techen *et al.*, 2014). DNA barcoding sensu stricto is the identification of species using a single standardized DNA fragment which fits the definition of CBOL (Consortium for Barcode Of Life). DNA barcoding sensu lato is the identification of any taxonomical level using any DNA fragment that has not been recognized as the barcode (eg: with single- nucleotide polymorphisms (SNPs) or with insertions/ deletions (indels). The identification of populations within a given species, based on either on diagnostic markers or on the differences in marker frequencies among population, refers to assignments tests and should not be considered as DNA barcoding (Valentini *et al.*, 2008). Identification by DNA barcoding involves the comparative matching between unknown specimen to one or more sequences that have been accurately identified by other means (Neigel *et al.*, 2007).

DNA barcode incorporates a standardized short sequence of 400–800 bp that is easily generated and characterized for all species on the planet. DNA barcoding has allowed users to efficiently recognize known species and hastened the discovery of species yet to be known. DNA barcoding uses huge collection of online databases as standard to which DNA sequence from unknown sample of field is compared (Genes *et al.*, 2008). Criteria-wise, a barcode should (i) contain significant species-level genetic variability and divergence, (ii) have conserved flanking sites for developing universal PCR primers for wide taxonomic application and (iii) have a short sequence length so as to facilitate current capabilities of DNA extraction and amplification. Among these criteria, successful universal amplification is said to be a primary for selecting a barcode (Genes *et al.*, 2008). An ideal marker short enough to allow amplification of the degraded DNA, however, is difficult to achieve. Above mentioned features provide the DNA barcoding with rapid, accurate and automatable species identifications (Taberlet *et al.*, 2007). DNA barcoding which involves selection of minimum standard loci that can be efficiently sequenced and can resolve large and diverse samples to provide comparable data for easy interpretation and better resolution as per the principles of barcoding standarisation, minimalism and scalability (Hollingsworth *et al.*, 2011). The DNA barcoding of 23 East Asian *Amentotaxus* with rbcL, *MatK*, trnH-psbA, trnl-F and *ITS* markers has pointed some threated species and their endangered status. The identification through the barcoding has critical role in cross border plant trade (Gao *et al.*,2016).

World gymnosperms, mostly conifers, demand identification through DNA barcodes but are getting significantly less consideration and attention (Armenise *et al.*, 2012). Other DNA based ways to identify and know phylogeny of gymnosperms besides common chloroplast and nuclear markers (Vargas *et al.*, 2014 ,Germano & Klein, 1999 ,Krupkin *et al.*, 1996 ,Olmstead & Palmer, 1994 ,Hsiang & Huang, 2002, Adams *et al.* , 2003 ,Ran *et al.*, 2010 , Flory, 1936) are:

1) Microsatellite markers (SSR)
2) Single nucleotide polymorphism (SNP)
3) Amplified fragment length polymorphism (AFLP)
4) Chloroplast Restriction site analysis
5) Restriction mapping
6) Random amplification of polymorphic DNA (RAPD)
7) mitochondrial DNA analysis
8) Chromosomes numbers

When morphology-based bio-systematics of any group of plants has been so trouble-some, the utilization of DNA sequencing data to discriminate between problematic taxa like *Ephedra* and to infer phylogenies becomes unavoidable choice (Faried, El-banhawy, & Elqahtani, 2018).

1.5 DNA barcoding limitations

While meeting the primary goals of species recognition and contribution to the complex taxonomy, the application of DNA barcoding has been misleaded by lack of proper selection of markers as well as low rates of successful identification of the species (Armenise *et al.*, 2012). The attempt to search a single barcode for plants has been understood to be a major driving factor towards plants' DNA barcoding. Unlike other organism group, plants' DNA barcoding is not easy as it does not simply rely on only DNA marker/locus (Taylor & Harris, 2012). The fact that no gene can serve as an ideal barcode has been a weak spot in DNA barcoding. Experts suggest the utility of barcoding should tilt more towards determining intraspecific and interspecific variation rather than marking the sequence matches as perfect and imperfect (Nielsen & Matz, 2006).

Apart from the utility of DNA barcoding for recognition, identification and assigning an unknown sample to a known species aiding conservation of biodiversity, DNA barcoding is also taken for discovery of new species detecting previously unsampled species as distinct, the later concept of which is major criticism subject and considered as a raw approach to declare new species on the basis of single DNA sequence. DNA barcoding is expected to be used as an additional approach with morphology, anatomy and other informative features to aid plant identification rather than an alternative technique. Its ultility as DNA taxonomy is also not easily accepted as aimed purpose of DNA barcoding as revising the samples. It is welcomed that DNA barcoding can set up hypothesis for new species discovery rather than standing as a core basis for new species declaration demanding the need of changed and relevant utilities of DNA barcoding (Moritz & Cicero, 2004, Taylor & Harris, 2012 ,Meyer & Paulay, 2005). Considering it as a replacement of normal taxonomy would be a worst idea that will leave much of the meaningful details untouched (Will, Mishler, & Wheeler, 2005).

Perfect sequence identity within species is termed as impractical by some statisticians and suggestions is forwarded to consider population genetics in order to fully get the utility of DNA barcoding. No doubt, it can be a useful tool to assign unknown specimen to a known predefined group, DNA barcoding is not expected to be in first place as a method of identifying the species (Nielsen & Matz, 2006).

Rather than being a cause for decline in alpha taxonomy (a discipline of detecting, describing, and classifying new species, as well as revising the classification of previously described species) as pointed by the critics, DNA barocding aids taxonomic investigation by revealing cryptic species and solve the synonymy; an effort-consuming task of taxonomy (Hebert & Gregory, 2005).

1.6 Globally selected molecular markers for DNA barcoding

The DNA markers used in algae, animals and fungi have low levels of variability to be used for land plants and so do the plastid phylogenetic markers for land plants (Chase *et al.*, 2005). For plants, standard barcode loci of *MatK*, *ITS* and rbcL gene are commonly used for barcoding whereas for animal's cytochrome oxidase 1 (CO1) gene is used. After several tests of the various loci and intergenic spacers (proposed barcodes), the plastid gene *MatK* increases universality and has become one of the favorite (Genes *et al.*, 2008). Several broad screenings of gene regions in the plant genome made three plastid markers (rbcL, *MatK*, and trnH-psbA) and one nuclear (*ITS*) gene regions the standard barcode of choice in most investigations for plants.

Table 1.1 Characteristics of plant barcode markers (Hollingsworth *et al.*, 2011)

Marker	Genomic source	Type	Amplicon length range (bases)
nr*ITS*	Nuclear	Transcribed spacer and 5.8S gene	407-1630
nr*ITS*2	Plastid	Transcribed spacer	157-670
MatK	Plastid	Protein coding	862-910
rbcL	Plastid	Protein coding	654-654
trnH-psbA	Plastid	Inter-genic spacer	226-934
rpoB	Plastid	Protein coding	536-590
rpoC1	Plastid	Protein coding	610-622
trnL-F	Plastid	Intron and inter-genic spacer	201-2114
trnL (P6)	Plastid	Intron	51-135

According to Chase *et al.*, 2005, due to low amount of variation in plastid DNA, three regions are necessary for plant identification and thus proposed two combinations, viz., rpoC1, rpoB and *MatK* and rpoC1, *MatK* and psbA-trnH as markers for land plants. Table 1.1 shows the various globally used DNA markers (one nuclear and other chloroplast markers) and their detailed information for PCR amplification. The CBOL Plant working Group (2009) pooled data across laboratories and evaluated the seven candidate plastid loci (atpF-atpH spacer, *MatK*gene, rbcL gene, rpoB gene, rpoc1 gene, psbk-psbI spacer and trnH-psbA spacer). They recommended the two-locus combination of rbcL+ *MatK* as the plant barcode based on the straight-forward recovery of the rbcL region and the discriminatory power of the *MatK* region (Fazekas *et al.*, 2008).

Rbcl is used in phylogenetic investigations with over 50000 sequences available in Genbank. Due to its limited discriminatory power, *rbcL* marker is not suitable at species level (Li *et al.*, 2014). Various reports are available of *rbcL* being used in to distinguish angiosperms but not commonly used in gymnosperms. Gymnospermic *rbcl* that codes the larger subunit of Rubisco has undergone the environmental-CO_2-dependent mutations over the period of evolution. This suggested DNA-driven adaptational strategies in gymnosperms physiology to withstand and survive the surrounding changes (Sen *et al.*, 2011). The slow evolving rate of *rbcL* is ususally not effective to be considered as sole DNA barcode for plants (Olmstead & Palmer, 1994).

The studies in molecular-based identification of gymnosperms suggests the potential role of marker *MatK*and *ITS* distinctly. Further, due to presence of relatively more number of species in a single family: Pinaceae (7 species) and Cupressaceae (10 species), it is estimated that the aid of additional marker is required. *Ycf1* and *petN-psbM* were two markers strongly suggested by the researchers in gymnosperms for accurate identification of species in family Pinaceae and Cupressaceae, respectively. Hence, the set of markers *MatK, ITS, Ycf1* and *petN-psbM* were chosen for the supplementary molecular based identification of gymnosperms in Nepal as referred to Cheng et al., 2016; Li et al., 2011; Dong et al., 2015; Adams *et al.*, 2009 (Details are justified in Literature Review).

1.7 Gymnosperms in Nepal

In context of Nepal, where highly stratified life-forms represent rich diversity of the species, Ecosystem, Species and Genetic diversity cover the zone of biodiversity. The genetic diversity is likely to be high in the country due to different ecosystems, physiographic zones and large population of some species.

Table 1.2 Families of Gymnosperms (Gymnosperms Database; Chamberlain, 1935; Govil, 2007, Singh, 2006)

Subclass	**Order**	**Family**
Cycadidae	Cycadales	Cycadaceae Stangeriaceae Zamiaceae
Ginkgoidae	Ginkgoales	Ginkgoaceae
Gnetidae	Welwitschiales	Welwitschiaceae
	Gnetales	Gnetaceae
	Ephedrales	Ephedraceae
Pinidae	Pinales	Pinaceae
	Araucariales	Araucariaceae Podocarpaceae
	Cupressales	Cupressaceae Taxaceae

Table 1.2 shows the occurence of 12 families of gymnosperms with their respective subclasses according to the standard classification. In Nepal, gymnosperms are represented by 24 wild species (Shrestha et al., 2018, Sharma et al., 2008 , Storrs and Storrs, 1990,Bulletin: Department of Medicinal Plants, 1997) and study of herbarium collection in National Herbarium, Godawari has also confirmed it (Figure shown on Annex). Table1.3 shows the total species covered under this study where 22 species arewild and 13 species are cultivated. Species are arranged according to the classification proposed by(Biswas and Mohri, 1997, Neale and Wheeler,2019).

Table 1.3 List of species of gymnosperms collected for the study

Family	**Wild**	**Cultivated**
Cycadaceae	*Cycas pectinata*	*Cycas revoluta*
Ginkgoaceae		*Ginkgo biloba*
Gnetaceae	*Gnetum montanum*	
Ephedraceae	*Ephedra gerardiana*	
	Ephedra intermedia	
Pinaceae	*Abies spectabilis*	
	Abies pindrow	
	Abies densa	
	Larix himalaica	

	Larix griffithiana	
	Pinus wallichiana	
	Pinus roxburghii	
	Tsuga dumosa	
	Picea smithiana	
	Cedrus deodara	
Araucariaceae		*Araucaria columnaris*
		Araucaria bidwilli
		Araucaria angustifolia
Podocarpaceae	*Podocarpus nerifolius*	
Cupressaceae	*Juniperus indica*	*Juniperus horizontalis*
	Juniperus squamata	*Thuja orientalis*
	Juniperus recurva	*Cupressus macrocarpa*
	Cupressus torulosa	*Cupressus sampevirens*
		Cupressus lusitanica
Taxodiaceae		*Cryptomeria japonica*
Taxaceae	*Taxus wallichiana*	*Taxus media*
	Taxus contorta	*Cephalotaxus sp.*
	Taxus mairei	

The extant gymnosperms of Nepal are possible to be endangered due to different human activities believing on the fact that biodiversity in Nepal is under threat due to high pressure by population growth as reported by (Joshi, 2015;Poudel *et al.*, 2012;Gotame, 2008). In Nepal, many researches are focused upon the uses and values of gymnosperms. On of the very importance study is pollutant resistance property of *Thuja orientalis, Cedrus deodara, Pinus roxburghii* and *Araucaria bidwillii* in Kathmandu (Hamal and Chhetri, 2017). A new platform for studying the genetics of *Taxus wallichiana* has been opened by discovering informative marker regions in chloroplast and nucleus (Poudel *et al.*, 2012). Antibacterial and cytotoxic activity of *Juniperus indica* extracts towards some human pathogenic bacteria added to the potential of gymnosperms' study for therapeutic use and drugs development (Mahajan *et al.*,2012).

1.8 Conservation Genetics and Contributions of DNA barcoding

Gymnosperms are not safe from the extinction threat. About 40 % of species are at high risk of extinction, being overall plant species extinction risk about 21 %. Phylogeny and other DNA base analysis if included in conservation planning, this could possibly make remarkable impact for keeping the gymnosperms around us for more years to come (Forest *et al.*, 2018).

Simply adding a contribution to conserve extant species, somehow, by the genetic studies, the area 'conservation genetics' aims more to avoid extinction of the species (due to loss of genetic diversity) maintaining adaptive evolutionary potential and limiting inbredding in naturally outcrossing species (Neale & Wheeler, 2019). While the biodiversity hotspots throughout the globe are at risk, *in situ* protection and *ex situ* preservation are helping to advance the plant conservation. However, additional approaches are still in demand. Identification of the priority species and integrated conservation measures should be taken under consideration utilizing accessible tools including the *in vitro* methods (Reed *et al.*, 2011).

Providing useful information for different areas in conservation, DNA barcodes facilitates every stakeholders involved: rapid identification for researchers, easy classification for taxonomists and scientific evidences and ideas for policymakers to determine the appropriate scales for conservation (Krishnamurthy & Francis, 2012) .

1.8.1 Community level

Providing knowledge about the existence, prevalence, distribution, benefits and threats of gymnosperms can make people realize not only the beauty of nature but also to contribute a little in their management and protection. This kind of benefits can be explored by scientific research in each possible sector involving the gymnosperms. In context of Nepal, the natural resources utilization is not fully understood at community level. So far, they tend to conserve the species around them but due to extensive exploitation for urbanization and commercialization, the keeping of trees is still a challenge. DNA barcoding of gymnosperms approaches to identify gymnosperms and can help prevent our species to illegalize, add beauty to our nature and use the gymnosperms of aforementioned advantages for our purpose to make ourselves self-reliant in many ways.

1.8.2 Scientific sectors

Interesting variations of gymnosperms from morphology to genetic level has attracted many of the botanists towards knowing more about them. DNA barcoding and DNA - based analysis creates a valid database which can be accessed in coming days for critical studies of gymnosperms from anti-cancer properites of *Taxus* to the development of pollutant-resistant features reported from *Cedrus deodara, Araucaria bidwillii, Thuja orientiales* and *Pinus roxburghii* (Hamal and Chhetri, 2017). As DNA based analaysis involves significant tools and application, it would be an exposure for new learners in Nepal to learn more about DNA barcoding of not only gymnosperms but in general flora as well.

1.8.3 National heritage

Open border is open challenge for us to keep Nepalese heritage untouched. Further, the recent warning of UNESCO to Government of Nepal not to commercialize Chitwan National Park can be understood in many ways. If not able to conserve the national park *in situ* without any human and industrial disturbances, UNESCO may remove the sites like Sagarmatha National Park (high altitude region and rich in gymnosperms) from World Heritage List. Though, government initiations are contributing directly and indirectly towards forest-keeping and utilization, further strong policies are still wanting. In this scenario, knowing our own species and an attempt towards conserving them can make a lot difference.

As Nepal is known for its forests and the gymnosperms are the key forest-makers. Expansion of horizons of gymnosperms database can aid knowing more about them and their current status of conservation and global threats which helps making and implementation of new policies. Through the method of DNA barcoding, the inspection of illegal timber smuggling could be targeted as an forest conservation approach for identifying a particular gymnosperm (Subedi *et al.*, 2013). A need of authentication tag to claim the carted off timber is fulfilled by DNA barcoding which gives threatened species a good story (Nithaniyal *et al.*, 2014). Thus, Nepal is in hunger of such an apposite technique where there are open borders and chances of illegal timber export.

1.9 Rationale

The main purpose of DNA barcoding in gymnosperms of Nepal is to establish its Reference Library that can substantiate the identification and taxonomic classification of gymnosperms and to test the morphology-based identification adding to previously suggested morphological, ecological data and molecular information based on Nepalese gymnosperms (Poudel*et al.*, 2012 ,Poudel *et al.*, 2013 , Gautam *et al.*, 2005 , Adams & Chaudhary, 2009). Extant gymnosperms need to be studied thoroughly in order to conserve and use them sustainably. Few species are already threatened nationally and globally (Gao *et al.*, 2016). Forensic identification of gymnosperms could be adopted in Nepal for authentic identification following the approach suggested by Liu *et al.*, 2018 with development of DNA barcodes for gymnosperm species.

In Context of Nepal, gymnosperms need to be studied in detail and the existing information of this group is not sufficient as rich as the botanical diversity present in Nepal. Also, there is no previous molecular phylogeny and taxonomical study that covers wide range of gymnosperms species growing in Nepal. The aims of this study are documentation, understanding and preservation of gymnosperms along with possible new species discovery by DNA sequencing method.

1.10 Research objectives

1.10.1 General objectives

1. DNA barcoding of wild and cultivated gymnosperms growing in Nepal and preparation of its DNA reference library.

1.10.2 Specific objectives

1. Study of PCR and sequencing efficiency of markers *ITS*, *MatK*, *ycf1* and *petN-psbM* in identifying Nepalese gymnosperms.

2. Study of species level discrimination power of the respective markers in studied Nepalese species.

3. Phylogenetic inference of studied species.

4. Create DNA reference library of gymnosperms of Nepal.

1.11 Research hypothesis

The study, through the application of DNA barcoding, would be helpful to develop easy procedure in differentiating species of gymnosperms present all over Nepal.

Alternatively, the resolving power of the selected barcodes could not be enough to support the taxonomic identification and phylogeny of the gymnosperms.

CHAPTER 2. **LITERATURE REVIEW**

Researches till date reveal much information about the significance of DNA barcoding identification and taxonomy of gymnosperms showing how crucial is the role of marked-based information in knowing more about gymnosperms and dealing with the difficulties of species with minimum morphological differences.

2.1 Molecular markers for gymno-taxon identification

Nuclear region (*ITS*) and chloroplast regions (*MatK* and rbcL) could provide basis for study of regional geographic groups of *Ephedra* along Europe-Mediterranean, American and Asian regions (Huang, Giannasi, and Price, 2005). Information of *ITS* region in Junipers is considered to be more useful for interspecific and intergeneric differences (Adams, Schwarzbach, and Pandey, 2003). Ephedra, a gymnosperm with very less morphological variances, was studied based on one chloroplast marker and ITS region which clearly authenticated the identification of particular species in the study (Faried *et al.*, 2018).

A region of chloroplast DNA, *MatK*, was found to be more evolved and bear more variable regions in *Pinus* family that aids in species discrimination using *MatK* as a barcode (Wang *et al.*, 1999). Cambodian Gnetales have been discerned by DNA barcoding using only chloroplast region (*MatK*) and discovered the inter-species relations of *Gnetum* and species of *Ephedra* (Kim and Won, 2016).

An approach to suggest the strong DNA barcode for *Taxus* recommended trnL- trnF marker robust enough to distinguish the *Taxus* at species label and trnL-trnF + *ITS* as best combination for other plants (Liu *et al.*, 2018). A strategy to determine universal barcodes for gymnosperms has purposed that Cycadales can be differentiated enough at species level by the *ITS* (nuclear) + trnH-psbA (chloroplast) as DNA marker. Understood as most promising maker, *ITS* shows enough variability and presents strong base for gymnosperm barcoding highlighting that nuclear markers comparatively show more reliabilty than the chloroplast markers (Sass *et al.*, 2007). Nuclear *ITS* stood as promising DNA marker than chloroplast regions after a test in *Cycas* which is always taken as important genus to test proposed DNA barcoding regions making ITS a basis to be selected as potential marker with variations inspite of having sequencing difficulties (Sass, Little, Stevenson, & Specht, 2007). Discrimination of *Juniperus communis* upto its fifteen varities has been studied by analysing nuclear DNA (*ITS*) and some chloroplast regions (*MatK* and non-coding regions) which provides the strength of applied markers in identification and taxonomy (Adams and Schwarzbach, 2012). Nuclear DNA and chloroplast DNA has made the understanding of *Juniperus* species easy with the discrimination of varities within the species possible providing the comparison on the basis of differences in DNA regions of the species belonging to two geographical regions (Adams *et al.*, 2013). Three different species of *Picea* has been discriminated by single nucleotide polymorphisms in *ITS* and chloroplast regions (*MatK* and non-coding chloroplast regions) (Germano and Klein, 1999). The chloroplast marker *ycf1* (an open reading frame of unknown function) in section of *Pinus; Trifoliae* provided 46.9% alignment success containing highest variable regions 74.9 % overreaching the potential of other markers *MatK* , rbcL and common non-coding plastid markers while analysing

49 species of *Pinus* over North America. Although not considered as a universal DNA barcode, *ycf1* is claimed to be more effective in identification of closely related species inside *Pinus* family (Hernández-León *et al.*, 2013). Based on primer-tested DNA barcoding of Chinese gymnosperms in 2013, it is said that *ITS2* region (segment of a whole ITS region) has been proved the best and the most universal DNA barcode for gymnosperms (Li *et al.*, 2013).

2.2 Errors in DNA barcoding

Possible errors in interpretation of sequence data in DNA barcoding includes false positive (assignment of genetically distinct individual within a species as cryptic) and false negative (failure in discrimination of an indiviual as a unique species based on absence of sequence variation) throwing the misleading conclusions regarding the utility of DNA barcoding approach for species identification and taxonomy (Wiemers & Fiedler, 2007).

As it is unrealistic assumption that a representative of each species on Earth can be newly collected, such efforts would have to be global in scale, and there are many political as well as practical difficulties that make such efforts impractical. Therefore, reliance upon already collected material (often with highly degraded DNA) has been viewed as crucial. This is one of the reasons why plastid and mitochondrial DNA have been viewed as the most appropriate regions to sequence: these highly repeated genomes are the most likely to survive reasonably intact within herbarium and museum specimens (Chase *et al.*, 2005).

2.3 Plastid genome for marker selection

The importance of plastid genome sequences is obvious in DNA barcoding as most of the plant barcodes are chloroplast targeted. Due to small genome, paternal inheritance, slow to rapid substitution rates, high copy number per cell, primer availability and ease of amplification, the chloroplast region has always been a spot of choice of barcode (Neale and Wheeler, 2019). Chloroplast DNA (cpDNA) molecules are ideal target across the conifers and other gymnosperms’ DNA barcoding because of the highly conserved gene order, the absence of recombination and low levels of nucleotide substitution. The high copy feature of plastid genome, meaning that genomic DNA are rich in plastids, makes the process easier than the low-copy nuclear genes for sequencing (Hollingsworth *et al.*, 2016). The further study aids given by chloroplast genome are phylogeography, introgression studies, phylogentic studies and multiple taxonomic levels (Neale and Wheeler, 2019). Chloroplast contains both highly conserved genes that are fundamental to plant life and more variable regions which are informative over broad time scales. Chloroplast DNA sequences data are versatile tool for plant identification or barcoding and establishing genetic relationships. Close and distant evolutionary distances in plants has been understood using chloroplast DNA and no single locus has been identified that can distinguish between all plant species (Nock *et al.*, 2011).

2.3.1 Chloroplast markers

2.3.1.1 *MatK* (maturase K gene)

Single marker *MatK* is proved to be universal enough in DNA barcoding of gymnosperms if the correct and specific primers are designed throwing out the pre-existing low confidence with *MatK* (Li *et al.*, 2011). Fig 2.1 shows the location of *MatK* region in whole chloroplast of *Podocarpus* between *trnK-UUU* and *psbA* region.The region *MatK* is taken as highly applicable marker among plant families due to its high evolutionary rate, suitable length and obvious inter-specific divergence as well as a low transition/ transversion rate (Li *et al.*, 2014). One of the ways to improve the amplification and sequencing of *MatK* from land plants is to use clade-specific primers (gymnosperms specific primers). In this study, gymnosperm-specific *MatK* primers are used following Li *et al.*, 2011.

2.3.1.2 *Ycf1* region

A less than 1 Kb fragment, *ycf1* gene, is highly conserved and more variable than *MatK* and *rbcl* in seed plants giving satisfactory performance (Fig 2.2) as a DNA barcode (Dong *et al.*, 2015). Fig2.1 shows the location of *ycf1*region in whole chloroplast of *Podocarpus* between *rpa 15* and *psb8* region.With the most unique sequence as claimed, *ycf1* gene can be used in critical molecular analysis and phylogenetic analysis. Though it lacks the potential of being universal barcode, it is understood to be very useful in species identification where a sample is already known to be of Pinaceae (Hernández-León *et al.*, 2013).

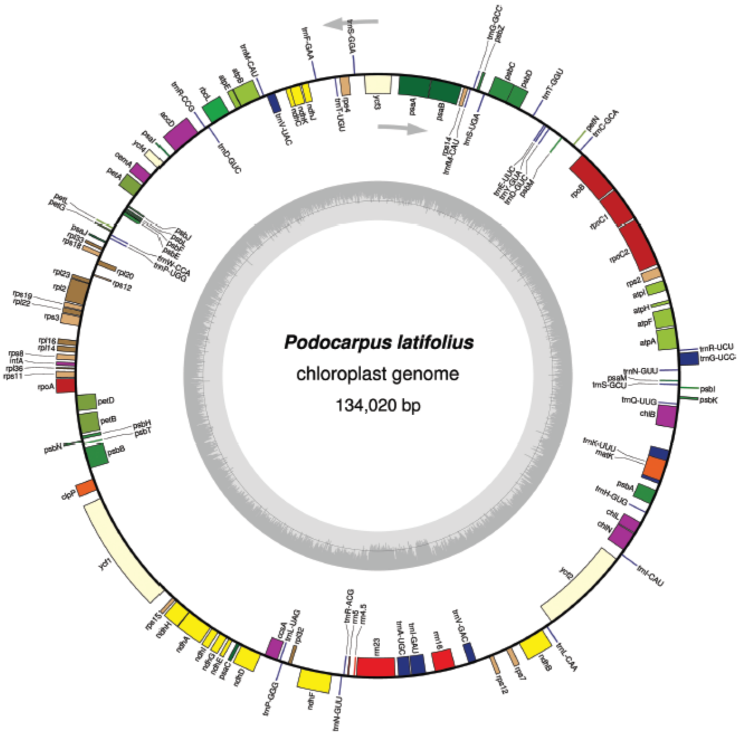

Figure 2.1 Schematic diagram of chloroplast genome of *Podocarpus latifolius* showing *MatK*, trnV, ycf and pet psbM markers used in this study (Saina *et al.*, 2019).

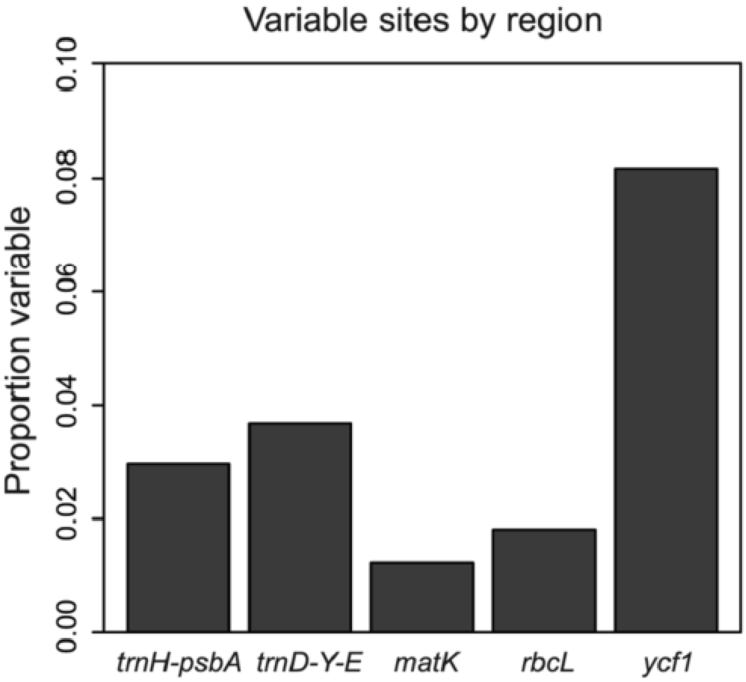

Figure 2.2 A graph showing strength of *ycf1* as a marker over other markers in variability in the study of Pinus -family (Hernández-León *et al.*, 2013).

2.3.1.3 PetN – psbM region

The region between PETN gene (coding a subunit of cytochrome b6-f complex) and psbM (coding photosystem II central protein) has been very strong informative site in case of Junipers (Cupressaceae family). Fig 2.1 shows the location of *petN-psbM* region in whole chloroplast of *Podocarpus* between *trnc-GCA* and *trnD-GUC* region.The region *petN-psbM* was analysed to verify a new variety *Juniperus recurva* var. *uncinata* and section *Sabina* of genus *Juniperus* (Adams *et al.*, 2009), (Robert P Adams & Schwarzbach, 2013). Study of chloroplast DNA variation of *Juniperus* (Robert P Adams & Kauffmann, 2010), report of a new genus *Hesperocyparis* (Adams *et al.*, 2009), speciation and taxonomic study of *Juniperus* (Adams*et al.*, 2010), taxonomy and sequence analysis of *Juniperus* based on chloroplast region (Adams and Schwarzbach, 2012), sequence analysis of Mediterranean *Juniperus*(Adams *et al.*, 2013) and discovery of new variey of Juniper *Juniperus recurva* var. *uncinata* from Jomsom, Nepal (Adams *et al.*, 2009), all of these studies on Cupressaceae family are based on unique informative nature of *petN-psbM* region that is considered to be important marker in species identification.

2.4 Nuclear markers

The Internal Transcribed Spacer (*ITS*) region between 18S - 5.8S - 26S ribosomal RNA cistron of nuclear DNA, is a powerful marker in phylogeny at the species level because of its high level of divergence. The higher discriminatory power of *ITS* over plastid regions at low taxonomic level has been widely studied and suggested as an important plant barcode. The China Plant BOL Group has recently supported *ITS* to be used as a core plant barcode being *ITS* more powerful than the plastid markers (Li *et al.*, 2014). Fig 2.3 shows different pairs of primers designed by Cheng *et al.*, 2016 to amplify various regions of ribosomal DNA that could be used as universal barcode for different groups of plants.

ITS, despite of its high differentiation power, is limited in utility because of the incomplete lineage and the presence of divergent paralogous and pseudogenes which may create a problem in amplification and sequencing (Gao et. al, 2016). Considerably large size of *ITS* region, often, is understood along with the heterogenic nature and variation of this marker gene has been an interesting concern in barcoding (Gernandt *et al.*, 2001). In addition, the chance of endophytic fungal DNA contamination is there. Few reports mention the difficulty in amplifying the *ITS* region (Hollingsworth *et al.*, 2011). The amplification and sequencing problems of whole *ITS* region can be compensated by targeting only *ITS*2 region but has also has some drawbacks as in some plants cases *ITS*2 is unable to provide resolution within the species (Li *et al.*, 2014). The use of *ITS*2 marker is beneficial as it delivers easy post PCR and sequencing procedures. Further, *ITS*2 region is more length conserved than the whole *ITS* region (Hollingsworth *et al.*, 2011). Though molecular data, yet, have not been proved sufficient to properly explain inter relationship between seed plants, the combination of few molecular markers can provide proximate phylogenic information (Bow *et al.*, 2000). The taxonomy and phylogeny study of Egyptian *Ephedra*is studiedthrough the method of DNA barcoding with markers trnH/psbA and *ITS* (Faried *et al.*, 2018).

Recently in 2019, DNA barcoding efficiency of *ITS* is still believed and many studies have been done based on this nuclear marker (Wu *et al.*, 2019,Shu-maw Chaw & Zharkikh, 1996). *ITS* is observed to be amplified easily due to high copy number of ribosomal RNA genes, highly variable, have low evolutionary pressure and non-functional sequences and very important feature its biparental inheritance. The information of *ITS* region has also been extensively used to study molecular phylogeny at generic and intrageneric (species and sub species) levels (Neale and Wheeler, 2019).

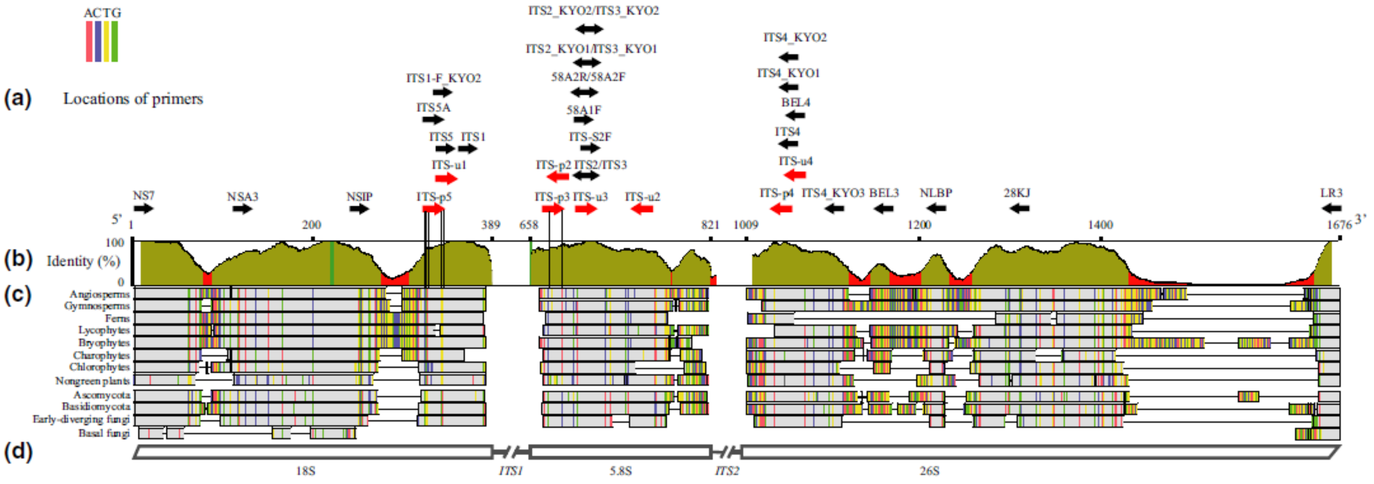

Figure 2.3 Schematic diagram of Ribosome coding nuclear DNA in different groups showing *ITS* region and site of primer binding to amplify different region (Cheng *et al.*, 2016)

2.5 DNA-based phylogeny and cladogram

One of the important purposes of studying gymnosperms genetics is phylogeny which is defined as the evolutionary history of species relationships (as a phylogenetic tree). A typical molecular phylogenetics involves a primary decision for analysis in relation to the target group such as ‘family’, the assembly of representative taxa, the acquisition of sequence information and the construction of phylogenetic tree by using optimality criteria (Hajibabaei *et al.*, 2007). Recently, phylogeny is taken as a useful tool to establish and understand relationship between the species and species groups rather than finding the results of their evolutionary history , an updated utility upon which debates are ongoing regarding phylogeny and its application (Farjon, 2007).

The use of DNA barcode fragments to construct, interpret and understand phylogenetic tree has reached a major research heights in recent years (Wu *et al.*, 2019). Still, since no evolutionary model is fully able to reflect its complexicity, the absolute information cannot be obtained even by the most adequate method. The selection of an appropriate model for phylogenetic tree construction is, thus, not concluded to be mandatory (Sarvananda *et al*., 2016). While phylogenetic trees help us understand evolutionary relationships among taxa, the cladorams provide only the topological information that means showing the interrelationship between the species under study. The cladogram shows how the branching events thrived to give a nested pattern of relationships of present day species (Zachos, 2015).

Chloroplast regions have been commonly used in Junipers to understand phylogenetic relationship within the species. Nuclear DNA and chloroplast DNA have been common basis to describe the phylogeny in *Juniperus* species constructing trees based only on nuclear DNA (*ITS*) whereas combined *ITS* and chloroplast data sets provides the congruency. Interrelation of species of Pinaceae and Conifers, Cycads, Ginkgo, Gnetales and their phylogeny has been studied based on chloroplast and nuclear markers (Bowe *et al.*, 2000).

To obey the strict principles of phylogeny is to establish absolute links of extant species with the extinct species showing a clear ancestor-descendents relationship over time which is practically very impractical to achieve in most of the studies. The choice then shifts towards creating a cladogram, a backbone of phylogenetic tree, that gives synchronous representation of evolutionary history showing the edge group (terminals of a tree) descending from a common ancestor and sharing relationships as sister groups (Podani, 2009).

2.6 Review on techniques and instrumentations of concern

2.6.1 Herbarium and Sample Preparation

Rapid dehydration for herbarium samples rather than slow dessication has been understood as the best way to preserve plant samples so that the DNA be left less damaged over time. The DNA quality of herbarium specimen does not meet satisfaction for further studies in DNA sequences and is worse as compared to that of fresh tissue due to collection and preservation strategies (Ribeiro & Lovato, 2007). The specimen related factors: specimen preparation method, poisoning, freezing and age of samples and the taxon related factors: leaf texture, tissue type and leaf biochemistry fully determine the success of DNA extraction from herbarium specimens (Särkinen *et al.*, 2012).

The silica gel method of drying DNA sample has been proved the most effective for high yield and quality of the DNA (Staats *et al.*, 2011). The PCR success of chloroplast regions from silica preserved specimen is absolute (100 %) than from samples prepared with air-based and alcohol-based drying (Särkinen *et al.*, 2012).

2.6.2 DNA extraction

Plants being rich in different kind of biochemical composition with their individuality, rely upon a single method for DNA extraction is not suitable (Ribeiro & Lovato, 2007). Dried plants in herbarium are valuable source of information about genetics and phylogeny of gymnosperms but the extraction of DNA from then and post processing has been a major challenge as the DNA in herbarium samples are highly degraded (Staats *et al.*, 2011). The most popular CTAB (Cetyltrimethylammonium bromide) method ensures the removal of polyphenolics and polysachharides requiring additional RNA removing approaches (Jobes et al., 1995). CTAB, a nonionic detergent properly precipitates nucleic acids and acidic polysaccharides from low ionic strength solutions keeping the proteins and neutral polysaccharides in solution. The feature of CTAB not precipitating nucleic acids and forming complexes with proteins makes it useful for purification of nucleic acid from organisms which produce large quantities of polysaccharides such as plants (Tan & Yiap, 2009).

2.6.3 DNA sequencing and PCR principle

Sequences-generating PCR duplicates DNA using the basic elements that are used in the natural DNA replication processes. The kinetics of PCR can be figured in three phases: the early cycles during which the primers search for the DNA template for their complementary binding, the mid cycles where the actual copies are made by exponential rate. The reagents used in the reactions should be accurately in optimal concentration and the components like Mg ions, Taq polymerase need to increased or decreased in concentration changing the number of cycles and playing with annealing temperature according to the problems arised in the amplification. $Mgcl_{2}$, for example, needs to be decreased to get rid of non-specific bands, along with increasing the annealing temperature and reducing the number of cycles (McPherson and Moller, 2000). The combination of PCR and fluorescence-based analysis method is reliable method for today's sequencing projects (Koepf, Salser, & Hunkaplller, 1989). As a technology to automate template preparation prior to DNA sequencing, the PCR incubation cycle includes denaturation of the template at 94° C followed by the hybridization with deoxyribonucleotide primers (52 ° C to 55 °C) and extension of the hybrid by the action of DNA polymerase I (Taq polymerase).

PCR protocol plays significant role in amplification- derived errors in the sequencing data which needs to be taken care and requires certain optimization efforts(Wilkinson *et al.*, 2017). A research in Brazil on optimization of PCR to avoid non-specific amplification of *ITS* region from nuclear DNA de Miranda *et al.*, 2010 states the addition of dimethylsuphoxide (DMSO) from 1-10 % and bovine serum albumin in the PCR reaction mixture.

Table 2.1 shows various modes of sequencing used globally according to the type of sample and advantage of the technique required. The mode of sequencing used in the study is 'Cycle sequencing' which increases the sensitivity of the DNA sequencing and permits the use of very small amounts of DNA starting material (that means minimum or no amplification allowed). The technique embraces the ddNTPs terminators during the thermal cycle reaction (PCR) simultaneously with the linear amplification using a single primer in a reaction. The cyclic PCR parameters can be different favoring the efficient ddNTPs incorporation (McPherson and Moller, 2000). The main principle follows DNA sequencing by chain termination method (Sanger's sequencing), considered as 'gold standard' method of DNA sequencinginvolves the inhibitory action of 2',3'-dideoxyribonucleotide triphosphate (ddNTPs) on DNA polymerase I. As the ddNTPs contains no 3'-hydroxyl group, the elongation of the DNA strand towards 3' end is, thus, terminated.

When a primer, template, fluorescent-labeled ddNTPs and dNTPs are incubated in PCR along with DNA polymerase, the extension of the hybridized primer occurs resulting the fragments of DNA each having same 5' terminals and different 3'- fluoro-labelled dideoxyribonucleotide residues (any one of either ddATPs, ddGTPs, ddCTPs or ddTTPs). The completed sequence reaction creates fragments which need no manual gel electrophoresis instead are directly loaded into single lane capillary gel in electrophoretic chamber of the analyser where the separation occurs on the basis of their molecular size (length of different fragments). The capillary gel electrophoretic separation uses thin-coated capillaries that contain non-polymerized gel matrices and laser detection system. The movement of polymer in and out of the capillaries in slab gel with the loading and running of samples makes the fragment detection. The separated fragments are optically analysed on the basis of fluorescence- based DNA sequence analyser. As the method links dideoxynucleotide triphosphates (ddNTPs) with the fluorescent moiety for the synthesis of all four reactions in a single vial, each terminated DNA fragment is, now, fluorescence-tagged respective to four different bases. After the excitation of fluorescent moiety of the DNA fragment with the laser beam present at the end of the gel column/ capillary in which the sample (with mixed DNA fragments terminating with different bases) is run according to the fragment length, the information is displayed and stored in computer system as four different colors corresponding to four different bases (Carrilho, Kist, & Franc, 2002, Sanger et al., 1977).

Table 2.1 Sequencing techniques and their specifications (Stranneheim & Lundeberg, 2012).

		Sequencing principle	**Sample preparation**	**Detection**	**Read length (bp)**	**Sample through put**
First generation sequencing	Sanger sequencing	Enzymatic chain termination	Cloning/ PCR	Optical	900	Low
Consensus sequencing	454	Pyro sequencing	Emulsion PCR	Optical	700	Medium
	Illumina/ Solexa	Reversible dye terminator SBS	Bridge amplification	Optical	100	High
	SOLiD	Sequencing by ligation	Emulsion PCR	Optical	75	High
	Complete Genomics	Sequencing by ligation	Restriction/ Circularization/ RCA	Optical	35	High
	Ion Torrent	Ions-sensitive SBS	Emulsion PCR	Electronic	200	Medium
Single-molecule sequencing	Helicos	Reversible single-dye terminator SBS	No amplification	Optical	35	Medium
	Pacific Bio sciences	ZMW sequencing	No amplification	Optical	>2000	Medium
	Nanopore sequencing	Ionic current shift	NA	Electronic/Optical	NA	NA

*RCA : Rolling circle amplification , SBS: Sequencing by synthesis, NA : not available

The downstream steps of sequence data analysis, handling and storage have been major challenges in DNA barcoding as well. The choice of methodological and coherent assembly and mapping algorithms is determining factor for accurate management of the massive amount of data (Stranneheim & Lundeberg, 2012).

2.6.4 Sequence analysis

A successful completion of DNA sequencing techniques leaves a set of data from which useful information is to be extracted to conclude the remarks of study. The conversion of obtained biological data into meaningful information is provided to us by computational methods with statistical models with sound mathematical framework for analysing bio-data holding crucial place in biological data processing (Yoon, 2009).The sequence analysis starts with knowing the absolute nitrogen bases in a DNA sequencing and then analysing other aspects of it. Similarities and differences in the sequences can be inferred by doing the pairwise and multiple sequence alignment, looking after the informative sites (conserved and variable sites), finding the differences in length of sequences, similarity search tools, pairwise distance between the sequences, relating the sequences to observe the phylogenetic tree topology (Stranneheim & Lundeberg, 2012 ,Wilkinson *et al.*, 2017 ,Pereira *et al.*, 2008).

Various tools used for sequence analysis are given below:

2.6.4.1 Bioedit and Sequencher

Bioedit vs 7.2.5 serves as most popular program used in sequence related analysis. It has sequence aligment feature which is used in this study. It is used to save the alignment file in different usable formats (fasta, mega, genebank) (Hall, 2011).

Sequencher vs 4.1.4 has facility to let import sequences in many formats, shows chromatogram and is usually exploited to generate consensus sequences. It displays chromatograms in multiple columns and allows easy visualization and editing (Nishimura & Rating, 2000).

2.6.4.2 BLAST

Basic Local Alignment Search Tool (BLAST) is used for similarity search of all the retrieved sequences using Nucleotide-BLAST in NCBI. The maximum score, query cover and percentage identity were observed and reported for comparistion of BLAST identification with other approaches in the analysis- distance based and tree-based.

2.6.4.3 TaxonDNA: Species identifier

The application of TaxonDNA Species identifier vs 1.8 is used to derive information given by various parameter such as Pairwise Explorer, Inter and Intraspecific distance, Species summary, etc. (Tahir *et al.*, 2018).

2.6.4.4 MEGA and Neighbor- Joining method

Molecular Evolutionary Genetic Analysis (MEGA7) is used to generate phylogenetic tree, find informative sites, nucleotide composition, etc. The Neighbor-Joining method is used to derive topology of tree using the loaded sequences and also provides branch length if wanted. The parameter with ‘bootstrapping’ involves picking a random subset of the data and re-runnig the phylogenetic analysis, and the value shown in branch split is the percentage of bootstrap replicates in which the node showed up. Thus, 100 means that the node is well-supported: it showed up in all bootstrap replicates. We consider and expect usually more than 50 % node value in order to validate our data.The distance on the basis of which the bootsrap Neighbour-Joining tree is constructed is the p-distance which is the proportion of nucleotide sites and which two sequences are different (Saitou & Nei, 1987).

CHAPTER 3. **MATERIALS AND METHODS**

The study started with collection of gymnosperms from sites of various alitudes in Nepal. Figure 3.1 shows the overall design of this study. All the samples and data obtained were deposited in Molecular Biotechnology Unit, NAST, Lalitpur, Nepal.

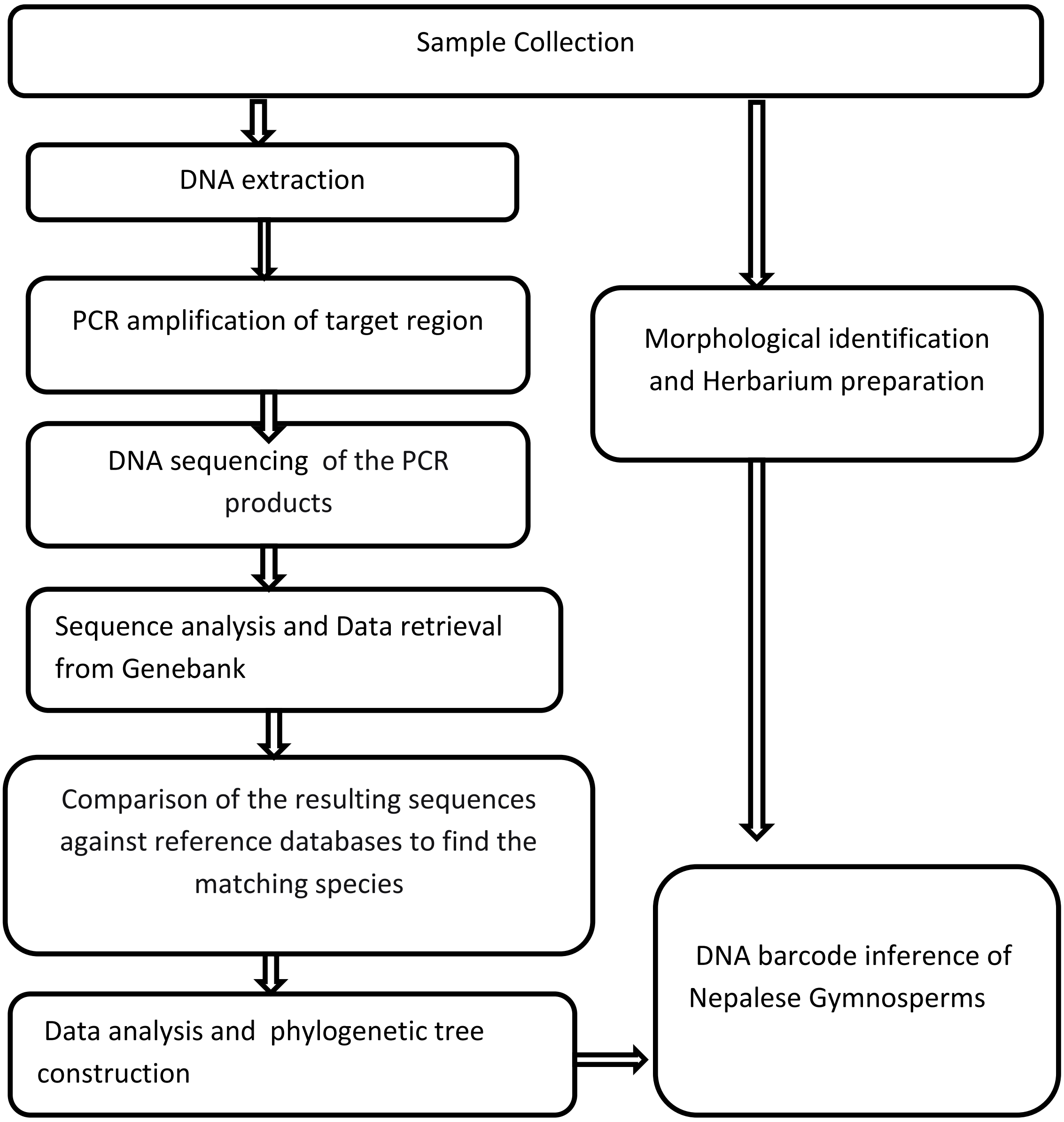

Figure 3.1 An overall study design in flowchart

3.1 Study of herbarium specimens of gymnosperms

Herbarium specimens of gymnosperms deposited in National Herbarium and Plant Laboratories (KATH), Godawari were examined for morphology-based identification study. Altogether, more than 460 herbarium specimens present in KATH were studied with an aim to gain key information about their morphological differences in order to ease the sample collection later on. The altitudinal variation of species was noted and regional distribution was also studied.

3.2 Sample collection

Altogether 59 individual samples were collected from different regions of Nepal as shown in Figure 4.1. Total collection of 59 samples comprises 35 different species with their duplicates. Of 35 unique species, 22 species were native to Nepal and 13 were cultivated.

During the sample collection, a 20-30 cm twig with leaves (and fruit/cone if present) was taken. From the same twig, few clean and young leaves were separated, and put into paper bag (tea bag) and kept into silica gel box. The branch was pressed into a newspaper fold and brought into lab in a herbarium press for drying. In the lab, the silica gel crystals in the box of leaves sample were replaced by the fresh ones and the samples were stored for DNA extraction. Drying of the herbarium specimens was performed through regular change of newspapers. The newspaper of the herbarium sets were changed frequently (at least three times a week).

Driedleaves samples for DNA extraction of *Abies densa* and *Larix griffithiana* was obtained from herbarium specimens of respective species deposited in National Herbarium and Plant Laboratories (KATH), Godawari.

3.3 Morphological identification and species description

Specimens collected from different parts of the country were identified using standard literature (Byng, 2015; Govil, 2007; Singh, 2006; Bhatnagar and Moitra, 2004; Chamberlain, 1935; Storrs and Storrs, 1990; Sharma et al., 2008, Bulletin Nepal: Department of Medicinal Plants 1997; Shrestha et al., 2018) tallied with herbarium specimens and also acquired help from plant taxonomist. Species description was done through careful observation of herbarium and their diagonostic characters were measured and noted.

3.4 DNA Extraction

DNA extraction from the silica gel preserved leaves was carried out following three standard techniques as per time constraints, resources availability and nature of the sample.

In all of the techniques, DNA extraction was started by grinding of material either manually or by using tissue-lyser technique (FastDNA Spin Kit). After the sample is grinded, the cell wall/membrane and nuclear membrane was ruptured using technique-specific detergent buffer (Eg: CTAB in CTAB method and commercial reagents in the supplied kit) to release chloroplast as well as nuclear DNA. The released mixture of cellular mass with DNA was treated with proteinases, RNases to minimize protein and RNA contaminations and DNA precipitated using isopropanol. After precipitation followed by centrifugation, the settled DNA pellet was dried and dissolved in nuclease free water or TE buffer for further use.

3.4.1 Modified CTAB DNA extraction

CTAB Extraction Buffer is a widely-used reagent to isolate DNA from plant tissues. Polysaccharides and polyphenols are problematic contaminants associated with DNA isolated from plants. CTAB Extraction Buffer effectively eliminates polysaccharides and polyphenols by employing the cationic detergent CTAB (hexadecyltrimethylammonium bromide or cetyltrimethylammounium bromide), and the polyphenol binding agent, Polyvinylpyrrolidone. The components of CTAB buffer and preparation are given in Annex.

Following Doyle, 1991, the modified CTAB method uses the lysis power of CTAB detergent enhanced by preheating for absolute results in DNA extraction. It provides easy homogenization, proper incubation and effective optimization for good quality DNA extraction. The working concentration of CTAB buffer was preheated in 50ml falcon tube at 60-65˚C in a water bath. About 50-100mg silica gel preserved plant leaves (gymnosperm needles, stem or buds) was taken in sterile chilled mortar and pestle and grounded with liquid nitrogen. One ml CTAB buffer was added to resuspend the paste and the suspension was transferred to 2ml microcentrifuge tube. The microcentrifuge tubes with suspension was incubated at 60-65 ˚C in a shaking water bath (35-50rpm) for 30-60 minutes with mixing by inversion every 10 minutes. After cooling the tubes at room temperature, 800 µl chloroform isoamylalcohol (24:1) was added to each tube and mixed by gentle inversion for 10minutes. The tubes were centrifuged at 9000 rpm for 10minutes at room tempertature (RT). The upper aqueous layer (the supernatent) was carefully transferred to a fresh 2ml tube and repeated with chlorform isoamyl alcohol treatment. The supernatent now is carefully transferred to 1.5ml microcentrifuge tube and isopropanol equal to 2l3 volume of supernatent was added, mixed well by inversion and incubated at -20 ˚C for 45minutes -1 hour. The tubes were centrifuged at 10,000 rpm for 10 minutes at RT. The supernatent was discarded and the pellet was washed with 70% ethanol (300-500 µl) and the tubes were centrifuged at 10,000rpm for 1-2mins. This ethanol washing step was repeated and finally the ethanol was then completely pipetted out and the pellet was washed with absolute ethanol. The supernatent was completely pipetted out and the tubes were completely dried in microcentrifuge tube heat block. The extracted DNA in the form of pellet attached in the bottom of tubes were hydrated with 50- 100 µl TE buffer and stored at -20 ˚C for further use.

For samples whose DNA pellet was turbid or with visible debrisor visible plaques, further purification was done using Qiagen Spin Column. The impure DNA solution was passed through the spin-column, incubated for 5 minutes at room temprature and centrufuged at 8000 rpm for a minute. The column was washed with 70 % ethanol and centrifuged for a minute and then air dried properly. After complete drying of the spin column, the

DNA bound in the column was eluted with warm TE buffer (50-60 ˚C) and then stored at -20 ˚C for further use.

3.4.2 FastDNA Spin Kit DNA extraction

CTAB buffer is effective for removing polysaccharides and polyphenol contaminants. Most CTAB protocols simply precipitate DNA which is followed by chloroform extraction. However, by incorporating a silica spin columns into the protocol, higher purity DNA can be isolated. Spin columns have a glass fiber filter that captures DNA so that it can be more stringently washed. DNA is easily eluted with water of TE buffer. This added step results in cleaner DNA than extractions without spin columns.

FastDNA Spin Kit DNA isolation was carried out starting with 100-200mg sample which was added to Lysing Matrix A tube. 800 µl CLS-VF and 200 µl of PPS reagents were added to the sample in Lysing Matrix A tube. The tubes were homogenized in the FastPrep Instrument for 40 seconds at a speed setting of 6.0 The tubes were centrifuged at 14,000 g for 5-10mins to pellet debris. The supernatent was transferred to a 2ml microcentrifuge tube and equal volume of Binding Matrix was added. The suspension was mixed by inversion and incubated at RT for 5 minutes with gentle agitation (mixed by gentle inversion). Half of the suspension (approximately 800 µl) was transferred to a SPIN Filter and the SPIN Filter was centrifuged at 14, 000 g for a minute. The catch tube was emptied and the second half volume of suspension was passed through the same SPIN Filter. The SPIN Filter was centrifuged at 14,000 g for a minute and the catch tube was emptied. Prepared SEWS-M of volume 500 µl was added to the SPIN Filter and gently resuspended with the tips of the pipette. The Filter set was cenrtifuged at 14,000 g for a minute and the catch tube now is discarded and replaced. The Filter set was centrifuged at 14,000 rpm for a minute for complete removal of residual ethanol. The catch tube was replaced with 1.5 ml microcentrifuge tube. The Binding Matrix above the SPIN Filter was resuspended with 100 µl of DES and incubated at 55 ˚C water bath (heat block) after prope capping of the tube. The tubes were centrifuged at 14,000 g for a minute for DNA elution. The SPIN was discarded and the DNA in the microcentrifuge tube was stored at -20 ˚C for further use.

3.4.3 DNeasy Plant Mini Kit DNA extraction

Less than 100mg sample was disrupted using liquid nitrogen in sterile Mortar and pestle. The sample was crushed in 200 µl AP1 buffer for making suspension and the suspension was poured into 2ml microcentrifuge tube. 200 µl of AP1 buffer was added followed by addition of 4µl of RNase A and vortexed before incubation at 65 ˚C for 10 minutes. Buffer P3 of volume 130 µl was added, mixed well and incubated for 5 minutes on ice. The lysate was centrifuged at 14,000 rpm for 5 minutes. The lysate (excluding pellet) was pipetted into a QIAShredder spin column placed in a 2ml collection tube and centrifuged at 14,000 rpm for 2 minutes. The flow- through was transferred into a new tube without disturbing the pellet if present. About 1.5 volumes of Buffer AW1 was added and mixed by pipetting. Among the total volume, 650 µl of the mixture was transferred into a DNeasy Mini spin column placed in a 2ml collection tube and centrifuged at 8000 rpm for a minute followed by the repetition with transferring remaining volumes of the mixture. The spin column was placed into a new 2 ml collection tube for and 500 µl

Buffer AW2 was added. The spin was centrifuged at 14, 000 rpm for 2 minutes. The spin column was transferred to a new 1.5 ml microcentrifuge tube and 100 µl Buffer AE was added for elution. The spin with elution buffer was incubated for 5 minutes at room temperature and centrifuged at 8000 rpm for 1 minute. The elution was repeated for better recovery of DNA.

3.5 Marker and primer selection

Following the global and gymnosperms specific literatures on molecular characterization, the marker of chloroplast region MatK, ycf1 and petN-psbM(reference: Table 3.1) and single marker of nuclear region ITS were selected for the study. Table 3.1 shows the details of all the markers used under the study. Among them, ITSprimer was universal marker for gymnosperms, MatK was specific marker designed only for gymnospermic MatK region (amplifying a segment inside MatK region), ycf1 was used only for Pinaceae family and petN-psbM was Cupressaceae family- specific.

Table 3.1 Primer sets for PCR amplification

Region	**Direction**	**Primer sequence 5'-3**	**Reference**	**Remarks**
ITS	P5_forward	CCTTATCAYTTAGAGGAAGGAG	Cheng *et al.*, 2016	specific for plants
	U4_reverse	RGTTTCTTTTCCTCCGCTTA	Cheng *et al.*, 2016	universal primer
MatK	*MatK* _forward	ATY GYR CTT TTA TGT TTA CAR GC	(Li *et al.*, 2011)	Gymnosperm-specific
	MatK _reverse	TCA YCC GGA RAT TTT GGT TCG	(Li *et al.*, 2011)	Gymnosperm-specific
	MatK _eph_forward	TCATTCAGAGCTGTTAGTTAG	(Li *et al.*, 2011)	*Ephedra*-Specific
	MatK _eph_reverse	ATCGTACTTTTATGCTTACAGGC	(Li *et al.*, 2011)	*Ephedra*-Specific
Ycf1	*Ycf1*_forward	TGAAAGCTCTAAGCAATGGATCYCC	(Dong *et al.*, 2015)	Pinaceae-specific
	*Ycf1*_reverse	ATACGACCAATATTTTTRGCTATTAT	(Dong *et al.*, 2015)	Pinaceae specific
PetN-psbM	petN5_forward	AAC GAA GCG AAA ATC AAT CA	(Adams, *et al.*, 2009)	Cupressaceae-specific
	psbM_reverse	AAA GAG AGG GAT TCG TAT GGA	(Adams, *et al.*, 2009)	Cupressaceae-specific

3.6 Polymerase Chain reaction (PCR) amplification and PCR success rate

After the selection of appropriate markers and reviewing the PCR conditions, the PCR reaction mixture for amplification of markers *ITS, MatK, ycf1* and p*etN-psbM* were carried out. All the necessary optimization was carried out according to the requirement such as alteration of number of cycles, annealing temperatures and components for the successful PCR. All the PCR reaction were carried out in PCR tubes, prepared in ice and then incubated in PCR machine for an average of 3 hrs.

The reaction mixtures for amplification of *ITS* region and *MatK* region of gymnosperms from the DNA templates with all the basic components and BSA, if required, were prepared. For amplification of *ycf1* region, DNA samples of Pinaceae family only were taken as template and PCR amplification was carried out as the barcode was Pinaceae family specific. For amplification of *petN-psbM* region, DNA samples of Cupressaceae family only were taken as template and PCR amplification was carried out as the barcode was family specific. These two family-specific markers were used in order to strengthen the identification and to enhance resolution as the families were species rich with 10 species out of 59 belonging to Pinaceae and 9 species to Cupressaceae (references in Table 3.1).

A PCR reaction mixture was prepared by mixing Master Mix (PCR buffer, dNTPs, Taq polymerase, Magnesium chloride) to final concentration of 1X, forward and reverse primers to that of 0.4 pmoles, bovine serum albumin (BSA) to that of 0.1 µg/µl (if required), Magnesium Chloride to that of 1mM (if required) and Taq polymerase to that of 0.25 U/10 µL (if required). The final volume of reaction mixture (10 µl) is fulfilled by nuclease free water. The details of every reaction mixtures are shown in the appendices.

After preparation of reaction mixture, the PCR tubes were vortexed for 2 secs, short-spined and then incubated in PCR machine (thermocycler) for targeted amplification. PCR incubation condition varied according to the marker and details for *ITS*, *MatK*, *ycf1* and *PetN-psbM* regions respectively are presented in respective tables in appendices. PCR amplifications were carried out in thermal cycler (Applied Biosystems, USA).

3.7 Agarose gel electrophoresis

After PCR incubation, the success of PCR amplification was checked by the method of Agarose gel electrophoresis.After the electrophoresis, the target DNA-Etbr complex in the agarose gel is illuminated on the exoposure to UV-light and compared to standard DNA ladder, thereby, confirming the presence of a target band.

Agarose gel (1%) was prepared in freshly prepared 1X TAE buffer in a conical flask by dissolving the agarose granules completely with the help of an oven (medium heat for 2minutes). After a little cooling, EtBr (ethidium bromide) was added to palm bearable hot agarose gel to meet the final concentration (0.5 µg/ml). The gel after properly mixing with EtBr was poured into gel casting tray with comb and allowed to set for about 20-30 minutes. The comb was carefully taken out from the solidified gel and the gel was put into electrophoretic chamber containing freshly prepared 1X TAE buffer. 1X TAE buffer was added until complete sinking of the gel into the buffer. PCR product (2.5-3 µl) was mixed with loading dye (to meet the 1X concentration of loading dye) and carefully loaded into the wells of the solidifed gel avoiding the air bubbles. Standard DNA marker (Generuler Plus 100bp ladder) was also loaded to correspond the band size. The electrophoresis was started to run in 100volts for 30 minutes and the gel was taken out to visualize under UV rays in a UV- transilluminator and also in Syngene (Syngene

Bio-imaging, UK) gel documentation instrument. The handling of EtBr and the gel was peformed using gloves for self protection and to avoid the possible cross contamination, hazards of EtBr.

PCR efficiency was calcualted by sorting out the successful PCR confirmation for all the markers used ITS, MatK , ycf1 and petN-psbM.

3.8 Purification of the PCR products

In order to avoid the unwanted components other than the amplicon for the further processing of the target amplicon, the PCR reaction mixture with target PCR product was purified using EXOSap- IT ^{TM} (Product No. 75001/75002; Applied Biosystems, USA).

PCR products were purified followingEXOSap kit protocol. To 5µl PCR product, 2 µl of EXOSap mix is added and mixed well by pipetting. The mixture was incubated at 37˚C for 10 minutes followed by incubation at 80 ˚C for 10 minutes (both in a PCR machine). Purified PCR product was tested by loading 1µl of the each product into 1% agarose gel and electrophoresis was performed. The confirmation of presence of purified product was done and the difference in the purity before and after the purification was also analysed. The rest of the purified product was preserved for downstreaming applications (DNA sequencing).

3.9 DNA sequencing

3.9.1 Cycle sequencing

DNA sequencing reaction involves the PCR incubation of amplified DNA (to be sequenced) with different reagents and using only a single primer at once; either forward or a reverse. DNA sequencing was performed by using high quality single band PCR purified product after gel visualization under UV-light. The sequencing reactions were prepared on 96-well plate by following BigDye ^{TM} Terminator v3.1 Cycle Sequencing Kit.

Primers, Ready Reaction mix (fluorescent-labeled ddNTPs, dNTPs, DNA polymerase), PCR buffer, PCR products and BSA were completely thawed and kept on ice after short-spining. BSA was added to enhance the effectiveness of the reaction. Among the components used, the Ready Reaction mixture being light-sensitive was handled very carefully with the minimum exposure into the light as possible (covering with aluminium foil, turning off the chamber light, etc). The sequencing reaction mixtures were prepared separately for each of the respective markers of respective direction (forward or reverse) according to Table 3.2. After preparation of mixture on 96-well plate, the plate was properly sealed with MicroAmp ^{TM} Clear Adhesive film (Applied Biosystems, USA). The plate was vortexed for 2-3 secs and short-spinned in a swinging bucket centrifuge (ST 8R, Thermo Scientific, USA). The plate was then immediately put in PCR machine for incubation with appropriate sequencing reaction condition (Table 3.3).

Table 3.2 DNA sequencing reaction mixture

Components	**Volume in µl**	**Final concentration**
Ready Reaction mix (2.5X)	0.5	
PCR buffer (5X)	1.75	1X
Primer (forward or reverse)	1.5	1.5 pm/µl (µM)
PCR product	1	not quantified
BSA (1mg/ml)	0.025	2.5 µg/ml
NFW	5.225	
Total	10	

Table 3.3 DNA Sequencing reaction condition

Steps	Initial Denaturation	Denaturation	Annealing	Extension	Final Extension	Hold
Cycles	1	35			1	1
Temperature	94 ˚C	94 ˚C	52 ˚C	72 ˚C	72 ˚C	4 ˚C
Time	3mins	30sec	30sec	45sec	10mins	Infinite

3.10 Purification of the synthesized DNA fragments

To avoid the possible interference of used buffers, primers and other utilized reagents, the DNA fragments synthesized after the PCR incubation were further purified prior to the sequence-reading electrophoresis. Purification of *ITS*, *MatK*, *ycf1* and *petN-psbM* markers was carried out following one of two different purification methods.

3.10.1 Purification by Beads

Purification of *ITS*, *MatK*, *ycf1* and *petN-psbM* markers was done by using BigDye Terminator ^{TM} kit (Applied Biosystems, USA) to remove the salts, unincorporated dye terminators and dNTPs. The reaction mixture for purification was prepared according to Table 3.4 using SAM solution and BigDye terminator solution in proportionate volume.

Table 3.4 Reaction mixture of SAM/BigDye Terminator beads

Reaction volume	Plate type	SAM Solution	Bigdye XTerminator Solution
10 µl	96-well	45 µl	10 µl
20 µl	96-well	90 µl	20 µl

Starting with the sequencing reaction plate, the 96-well plate was short spinned in centrifuge. The SAM solution and BigDye terminator solution were brought to room temperature and kept on ice. The required reaction volumes were separated according to the Table 3.4. The adhesive film was carefully removed from the 96-well plate. To each well, the volume of SAM Solution was added followed by the addition of BigDye

XTerminator solution specified as above using pipette tips of orifice 1mm or greater. The plate was then sealed with adhesive seal cover and the contents of the plate were mixed thoroughly by vortexing the plate for 30 minutes. The plate was then centrifuged at 1000 xg for 2 minutes.

3.10.2 Sequencing purification by EDTA precipitation

A master mix I of 10 µl Mili-Q and 2 µl of 125 mM EDTA per reaction was prepared. Twelve µl master mix I was added to each 10 µl of sequencing reaction in PCR strips with attached caps. The PCR strips were capped properlyand the reagents were mixed by vortexing. A master mix II of 2 µl of 3M Sodium acetate pH 4.6 and 50 µl of absolute ethanol per reaction was prepared and 52 µl of the master mix II was added to each reaction. The PCR strips were capped properly and the reagents weremixed by vortexing. The capped-strips were incubated at room temperature for 15 minutes and centrifuged at 3000 g for 30 minutes at room temperature.The supernatant was decanted by inverting the plate on the tissue paper. The plate was invert-spinned at 180 g to remove residual supernatant. After that 100 µl of freshly prepared 70 % ethanol was added to each reaction of the plate and the capped strips were spinned at 3000 g for 5 minutes at room temperature. The ethanol was drained out (washed) by inverting the PCR strips immediately over layers of UV-sterile tissue paper after opening the caps. The washing was done with absolute care and was repeated for removal of residual ethanol. Finally 10 µl of Hi-Di formamide was added to each of the reactions in the plate and spinned. Denaturation at 95 °C for 5 minutes was performed in a thermo cycler machine and the plate was promptly chilled (snap-chilling) on ice.

3.11 Sequencing: Capillary electrophoresis

The plate with the purified reactions was then forwarded for analysis on Genetic Analyser (Sequencer). The plate was loaded on the deck of the sequencer, locked and linked to the 3500XL Genetic Analyser (Applied Biosystems, USA) for capillary electrophoresis using BigDye XTerminator run module. Applied Biosystems Gene Analyser ABC 3500XL is based on Sanger's dideoxy chain termination method. The base allocation and chromatographic inference for the markers *ITS*, *MatK, ycf1* and *petN-psbM* was given by the analyser which were further assessed to generate information.

3.12 Sequence editing and Alignment

The Sequence file were imported to Sequencher vs 4.1.4 to check the sequence peaks to confirm the quality of the sequences and the ambigious peaks were excluded for single strand. Whereas for the double strands (with both forward and reverse sequences), the consensus sequences (the contigs) were prepared.

To prepare the contigs, sequences (forward and reverse primer based) obtained through the Genetic analyser were imported to the project of Sequencher vs 4.1.4 and contig of overlapping fragments was created by assembling the sequences keeping the assembly parameters: minimum match percentage 70-85 and minimum overlap percentage 20. The sorted sequence was exported as Consensus as taken for further analysis. Based on the base call of a nitrogen base corresponding to forward and reverse sequence, the assignment of bases, deletion of bases, deletion of gap, assignment of gaps, replacement of nitrogen bases were performed in order to obtain a reliable and valid sequence.

Consensus sequences were then imported to Bioedit vs 7.2.5 for multiple sequence alignment. The primer regions (forward and reverse) were excluded from the multiple sequence alignment comparing with the primer sequences and the fasta file of alignment was saved for further analysis. The individuals sequences based on the alignment were also saved for all the markers ITS, MatK , ycf1 and petN-psbM for further anlaysis.

3.13 NCBI BLAST

Similarity search was performed by using NCBI-nucleotide BLAST. The method takes the top hit and percentage identity in order to give the identity score from query-reference comparision. For this method, all the top hits within 10 points deviation down of the maximum score were considered: if the maximum score (-10 points) included only a single species then a species level identification was assigned; if the maximum score (-10 points) included multiple species in the same genus then a genus level identification was assigned; if the maximum score (-10 points) included multiple species in different genera in the same family then a family level identification was assigned. Meaning that, maximum identity and one species within 1 % deviation was assigned species level confidence; with more than one species was assigned genus level confidence and so on.

The consensus sequence thus obtained was entered in NCBI nucleotide-BLAST to observe the matches of samples (queries) with the available specimens (subjects). Considering the maximum score, identity and the E-value , the samples were checked for the match of sequence based species identification with the phenotypic based sample labelling.

Morphologically identified gymnosperms species were compared with informative DNA sequences and the matches were examined. Respective sequences of the barcodes were entered into ‘NCBI nucleotide BLAST’ and observed for the possible mis-matches, query coverage, matching score and identical results. Further, the species identification is compared with BLAST method (percentage identity), distance method (interspecific value larger than the intraspecific one) and the tree method (branch bootstrap value more than 50%). Species discrimination is considered as successful if minimum uncorrected interspecific p-distance is greater than maximum intraspecific p-distance between the species.

The ratio of successfully identified species to all the sample species was calculated as the proportion of the species that were discriminated.

3.14 Nucleotide composition

The nucleotide compositions of all the marker-sequences corresponding to the respective markers of the samples were obtained by ‘Model’ application on MEGA. The percentage composition of A,T,G and C were known and further analysed to find the average G+C content in the genes in gymnosperms used as markers.

3.15 Genetic distance calculation

The result of multiple sequence alignment was observed in order to detect the variations. The sequence were back tracked and checked if there was actually a variation in the base or it was false-positive detection as a result of sequencing errors. The variable hotspots were noted and variable region for each gene *MatK*, *ITS*, *ycf1* and *petN-psbM* were reported.The sequences from different regions from chloroplast and nuclear markers' region were aligned by CLUSTALW (in Bioeditvs 7.2.5). The aligned fasta file for each marker (locus) was imported to MEGA v. 7.0 software and processed to estimate parameters like number of constant sites (C), variable sites (V), parsimony informative sites (Pi) and singleton sites (S).

Based on p-distance, the genetic differences between the species (inter-specific distances) and within the duplicates of species that were available (intra-specific distance) were calculated using Distance tool in Species Identifier of TaxonDNA vs 1.8 software. The interspecific distance is expected to be more than intraspecific one to indicate the distinguishing feature of species. If the interspecific distance between two species was greater than the intraspecific distance, the unique species identification was considered successful. For those species whose duplicates (intra- specific distance) was not available, average intraspecific distance within all the samples was taken for comparision.The percentage of successfully identified species to all the sample species was calculated as the proportion of the species that were discriminated.

3.16 Phylogenetic tree construction and distance estimation

The phylogenetic trees were constructed for markers *ITS*, *MatK*, *ycf1* and *petN-psbM* using neighbour-joining method (NJ) in MEGA 7.0 software program using p-distance as a genetic measure. Bootstrap testing of 1000 replicates was performed to estimate the confidence level of the topology of the consensus tree. Node support was estimated based on the following scale: 50-74 % (weak bootstrap support) and 75-100 % (strong bootstrap support) following Felsenstein and Kishino, 2012; Eizirik *et al.*, 2001; Manning *et al.*, 2014. The efficiency of species resolution was considered successful only when all the single clades have at least ≥ 50 % bootstrap value. In addition, the sampled species was considered successful when discriminate individuals of species formed a monophyletic group in the phylogenetic tree (Ford *et al.*, 2009).

The aligned sequences in MEGA-acceptable format (fasta format) were imported to MEGA7 and continued with 'Analyze'option. The tree was contructed on 'Phylogeny' application withNeighbor-Joining method with number of bootstrap 1000 as changed parameter. The tree was visualized in different formats as available in MEGA7 and one of it is saved for interpretation and analysis.

3.17 Assignment of species to samples

An identification of each specimens collected in the study has followed two different complementary methods aforementioned. The morphological method included tallying the collected herbarium voucher with the standard herbarium specimens of gymnosperms deposited at National Herbarium and Plant Laboratories (KATH), Godawari, Lalitpur and with the help of standard gymnosperms literatures. The second aiding method was the molecular or DNA-based method using four different markers starting with the DNA extraction, PCR and finally to the sequencing of the markers and then relying upon the BLAST method, Distance method and Tree method in the analysis of the obtained sequences.

3.18 Discrimination power of the markers

The efficiency of *ITS*, *MatK*, *ycf1* and *petN-psbM* in species identification was carried out using three analytical methods. Species discrimination power of each marker along with their possible combination was calculated based on the ratio of species discriminated by BLAST method with features available online on BLAST https://blast.ncbi.nlm.nih.gov/Blast.cgi (Altschup *et al.*, 1990). The second method was distance method (TaxonDNA method) (Meier *et al.*, 2006). The third method was phylogeny (tree method) using MEGA v 7.0 (Tamura *et al.*, 1994). The possible combinations of the markers were *ITS+MatK*, *ycf1 +MatK* and *petN-psbM+MatK*.

CHAPTER 4. **RESULTS**

4.1 Study of herbarium specimens

A total of 460 specimens of gymnosperms deposited in National Herbarium and Plant Laboraories (KATH), Godawari have been studied. The detail of specimens is given in Annex 1. Distribution map of gymnosperms based on the herbarium specimens of KATH is shown in Fig 4.1.

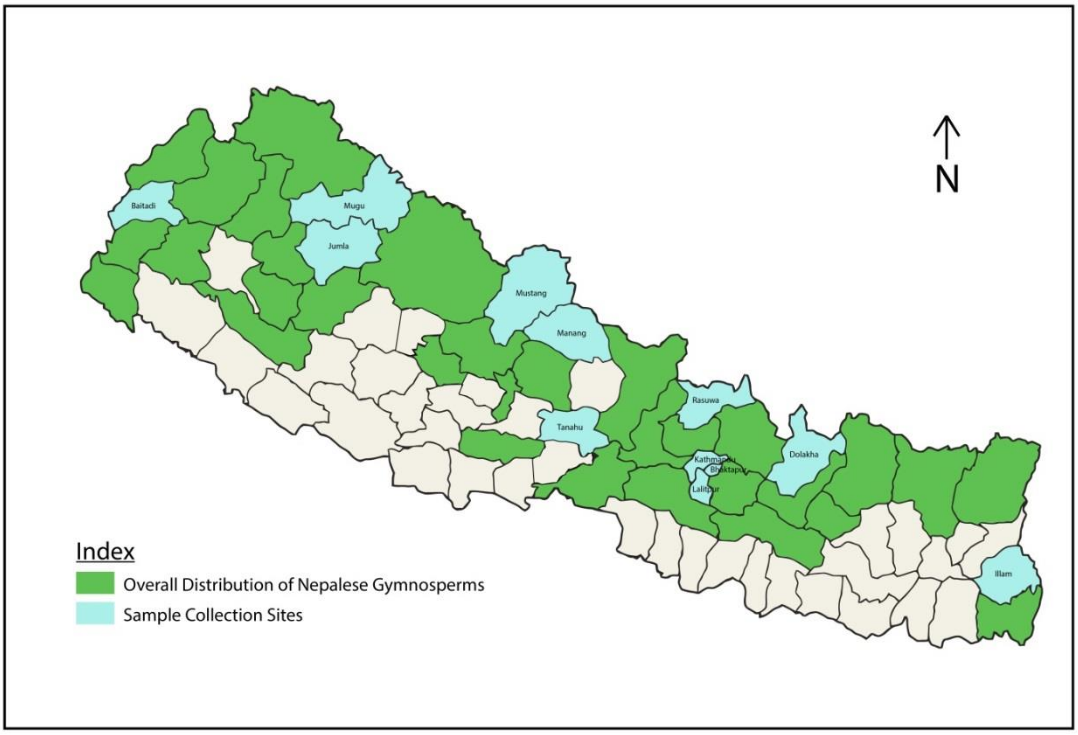

Figure 4.1 Distribution of gymnosperms in Nepal based on the herbarium specimens of KATH.

Overall distribution of Nepalese Gymnosperms: district-wise national distribution studied from the herbarium collection-catalogue of National Herbarium and Plant Laboratories, Godawari. Sample collection sites: districts from where 59 individuals of gymnosperms were collected.

4.2 Study area and sample collection

Three different study sites representing ecological zones of west, central and east Nepal were selected for the collection of indigenous species of gymnosperms. Cultivated species of gymnosperms were collected mainly from Kathmandu valley. Contributing altogether 47 samples, Central Nepal was major sample collection area while 10 samples are representing western Nepal and 2 samples were taken from eastern Nepal. Langtang National Park (Rasuwa) was site for collection of 9 different species and was the richest site among the study sites in terms of species richness. Kathmandu valley was the second most visited field as 14 cultivated species of gymnosperms were possible to be collected The GPS location of most of the sample collection sites was reported. At least altitudes of all of the sampling sites were estimated for distribution study and herbarium preparation. Interaction with the local people provided some useful information about the occurrence, exploitation and significance of the species for local ecology and livelihood of the people. The samples were morphologically studied to allocate the species name which was finally confirmed after analysis.

The complete set of sampling comprises a total of 33 different species among which 13 species were cultivated. Herbarium specimens of collected samples were prepared following standard technique. The herbarium samples corresponding to each of the species were prepared as mentioned in materials and methods. Leaves samples were properly stored back in silica gel after DNA extraction with periodic replacement of silica crystals. All the herbarium specimens and silica samples prepared during this study are stored in Plant Specimen curation unit of Molecular Biotechnology laboratory, NAST.

Habit wise all the studied gymnosperms were either trees (Pinaceae, Cycadaceae, Araucariaceae, Taxaceae, Cupressaceae) or shrubs (Ephedraceae, Cupressaceae) except *Gnetum montanum* which is a woody climber. The number of species analysed and their habit is shown in Figure 4.2.

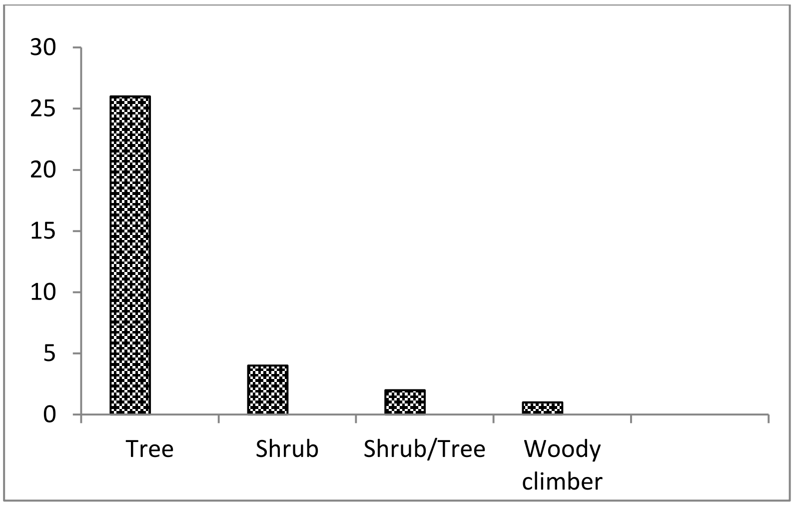

Figure 4.2 Graph showing habit of gymnosperms collected and analysed.

Botanical family wise the largest number of species; 21 samples belonged to the species-rich family Pinaceae followed by 17 to Cupressaceae, 5 samples to Taxaceae, 4 to Ephedraceae, 3 to Araucariaceae, 2 samples to Cycadaceae, 2 to Taxodiaceae and one each to Gnetaceae, Ginkgoaceae and Podocarpaceae. Studied species of gymnosperms and their families are shown in Figure 4.3. Cupressaceae family contributed largest number of samples.

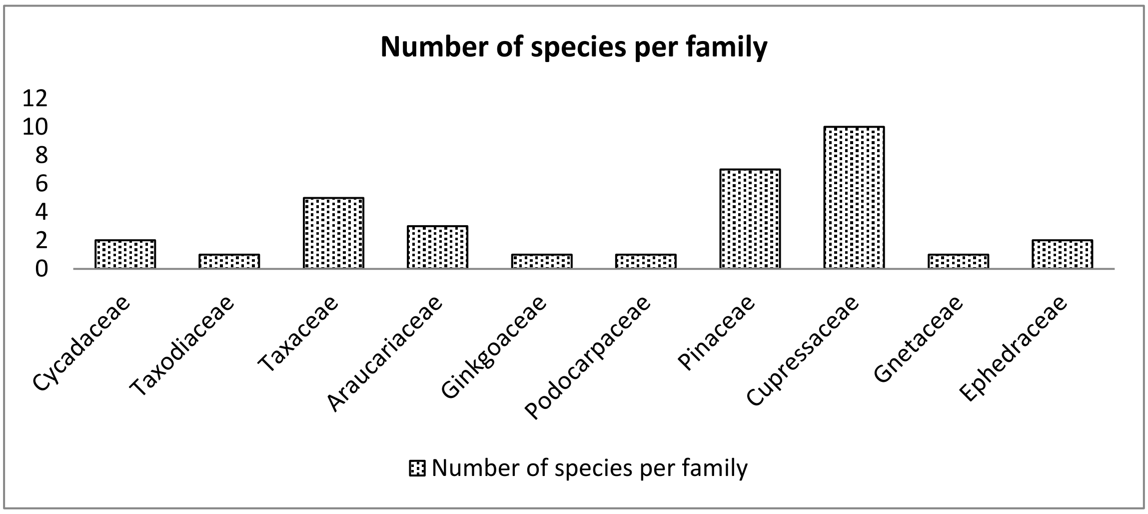

Figure 4.3 A bar graph showing composition of different species (without duplicates) of respective families in overall sampling

4.3 Distribution of Gymnosperms

Gymnosperms are found in different altitudinal ranges as low as 100 m in the tropics to 4400 m in alpine region. As the gymnosperms are available in diverse altitudinal range, their altitude variation was studied, maximum and minimum altitudes were known and the floating graph was plottedas shown in Fig 4.4. The altitude variation was studied with herbaria- collection catalogue of National Herbarium and Plant Laboratory, Godawari. Altitude of sample collection lies within the aforementioned range of altitude for all the species.

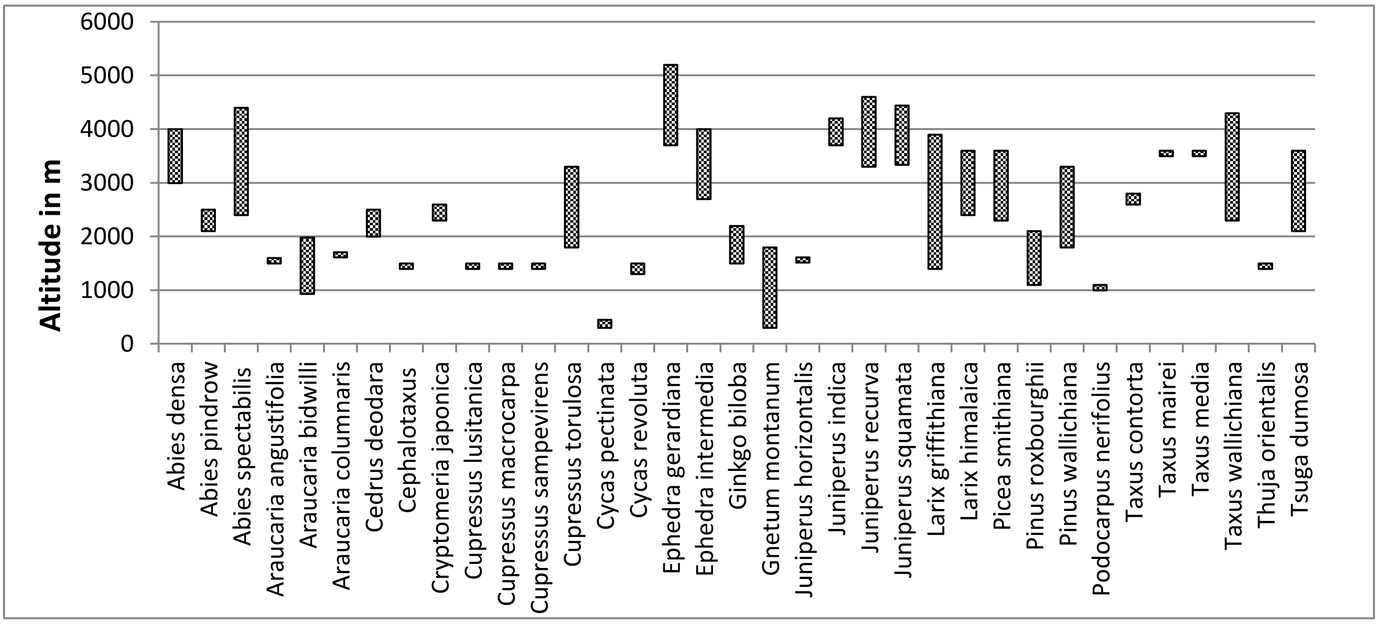

Figure 4.4 A floating-bar graph showing altitude variation of species under study.

4.4 Species description

Gymnosperm specimens collected during the field trip were examined and their characteristics morphological characters were studied. Descriptive features of each species were listed as in the following paragraphs. Four species of *Ephedra* could be differentiated only at genus level in the field and labeled as *Ephedra* on the field, except those, all other individuals were identified right on the field. *Cephalotaxus* (cultivated) was also difficult to be differentiated at species level during collection.

Family: Araucariaceae
Araucaria angustifolia (Bert.)Ktze.
Distribution: Brazil, Argentina and Paraguay.
Collection area: Lalitpur, Nepal.
Evergreen, monoecious trees 25–35(–52) m tall, with a straight trunk and horizontal branches, the crown becoming flat-topped with age. Leaves sharply acute, needle-like, lanceolate, midrib visible, usually in pairs, dark green or glaucous, thick. Male cones and female cones present, seeds light brown.

Araucaria bidwilli Hook.
Distribution: Australia.
Collection area: Kathmandu, Nepal.
Evergreen, monoecious trees to 50 m tall; branchlets dense, pendulous, green, glabrous. Leaves radially spreading, bright green, lanceolate, ovate-lanceolate, or triangular-ovate, flattened, slighty incurved, hard, thick, leathery, lacking a midvein but with numerous,

parallel, thin veins, stomatal lines abaxial; leaves of same year dimorphic: those of young trees and vegetative branchlets longer and more loosely arranged than those of old trees and cone-bearing branchlets, apex acuminate or acute.

Araucaria columnaris (J.R. Forst.) Hooker
Distribution: New Caledonia (Europe)
Collection area: Lalitpur, Nepal.
Evergreen, monoecious, narrowly conical tree to 60 m. tall. Trunk clear, then later with numerous short shoots. Juvenile leaves needle-like, imbricate, lanceolate, apex incurved. Adult leaves scale-like imbricate, triangular. Male cone and female cone present.

Family: Cupressaceae
Cupressus torulosa D. Don.
Distribution: Himalaya (Kashmir to C. Nepal).
Collection area: Jumla and Mustang, Nepal.
Evergreen, monoecious trees to 45 m tall; branchlets loosely or densely arranged, terete, ultimate ones pendulous or irregularly spreading, slender. Leaves closely arranged, scalelike, triangular, thick with closely adpressed tips so that the brances are no rough to touch, closely imbricated on the youngest twigs, flat or slightly gibbous and with a rounded, sometimes inconspicuous central abaxial gland, apex usually slightly obtuse.

Cupressus lusitanica Miller
Distribution: Belize, Guatemala, Honduras, El Salvador and Mexico.
Collection area: Lalitpur, Nepal.
Evergreen, monoecious trees 25-30 m tall. Shoots quadrangular, pendulous, forming flattened foliage sprays. Irregularly arranged sprays of foliage.Foliage blue-green, four-ranked, ovate, closely pressed, usually with long, pointed apex.Cones globose.

Cupressus macrocarpa Hartweg ex Gordon
Distribution: USA
Collection area: Lalitpur, Nepal.
Evergreen, monoecious, trees to 25 m; crown generally broadly spreading, especially on exposed headlands, fairly sparse, often composed of few major limbs from near ground, more upright in sheltered locations. Bark rough, fibrous. Branchlets decussate. Leaves not glaucous. Seed cones oblong, grayish brown, not glaucous; scales 4-6 pairs, smooth.

Cupressus sempervirens L.
Distribution: Greece, Turkey, Crete, North Iran, Lebanon, and Syria; and Cyprus. In North Africa;Tunisia and North Libya.
Collection area: Lalitpur, Nepal.
Evergreen, monoecious, Tree 20-30 m in height. Trunk straight.Shoots radiating in all directions, round or quadrangular.Leaves scale-like, decussate, small, ovate, obtuse, dark green, with a dorsal gland in the shape of longitudinal furrow. Flowers appear early in spring. Cones on short stalk, glossy, brown to gray, pendulous, globose to elliptic, 2-3 cm long.

Juniperus horizontalis Moench
Distribution: Canada, France, USA.
Collection area: Lalitpur, Nepal.
Evergreen, dioecious shrubs. Branches creeping; branchlets erect, leaves green but turning reddish purple in winter, abaxial gland elliptic, conspicuous, exudate absent, margins entire, scalelike, apex rounded or obtuse to acute and apiculate, spreading. Seed cones generally with curved peduncles, globose to ovoid.

Juniperus indica Bertoloni
Distribution: Himalaya; Karakoram.Bhutan, Nepal, W. China.
Collection area: Rasuwa, Manang and Mugu, Nepal
Evergreen, dioecious tree 20 m high, rarely small trees, with hard trunk, ultimate branchlets densely arranged. Leaves both scale like and needle like; needle like leaves usually present on young branches, apex acuminate; scale like leaves decussate, linear, apex obtuse. Fruit a fleshy berry, one seeded, brown then shining blue.

Juniperus recurva Buch.-Ham.
Distribution: Chitral, Himalaya (Kashmir to Bhutan), Assam, Burma, W. China.
Collection area: Rasuwa, Nepal
Evergreen, monoecious or rarely dioecious Shrubs or trees; bark light grayish brown or brown; crown conical or broadly pyramidal; branches ascending in apical part of plant and spreading toward base; branchlets pendulous, curved. Leaves in whorls of 3, loosely appressed, greenish white or slightly glaucousadaxially, all needlelike, nearly straight, slightly incurved, concave adaxially, base decurrent, convex with longitudinal grooves at base abaxially, apex sharply pointed. Leaf base decurrent or lanceolate, in threes and pungent. The under surface of the leaf furrowed from the base upto halfway: no green midrib above.

Juniperus squamata Buch.-Ham.
Distribution: Afghanistan, Himalaya, N. Burma, W. China.
Collection area: Rasuwa, Nepal.
Evergreen, monoecious shrubs, erect or procumbent, or small trees to 12 m tall; branches ascending or horizontally spreading; branchlets densely arranged, straight or curved, usually short, not angled. Leaves in whorls of 3, spreading or ascending, sometimes nearly appressed, needlelike, straight or slightly curved, slightly concave, with white stomatal bands adaxially, obtusely ridged with longitudinal, thin groove on ridge or at base abaxially, base decurrent, apex acute or acuminate.

Thuja orientalis L.
Distribution: Korea, E. Russia, and China
Collection area: Lalitpur, Nepal.
Evergreen, monoecious trees to 20 m tall. Leaves decussate. Scale-like yellowish-green leaves in flattened sprays in youth, becoming darker green with age, base decurrent, 1-3 mm long, apex bluntly pointed, dimorphic along branchlets. Seed cones terminal, solitary Foliage may bronze up in winter.

Family: Cycadaceae

Cycas pectinata Griff.

Distribution: Himalaya (Nepal, Sikkim), Assam (Khasia), Manipur, Burma, Yunnan.

Collection area: Lalitpur, Nepal.

Tall evergreen dioecioustree 2-6 m with crown of leaves at the apex of trunk. Trunk glabrous.Leaves 40–80 in number and 1-2 m long. Margins of the leaflets are flat. Leaflets deeply pectinate. Megasporophyll with 4-6 ovules, ovules yellowish orange.

Cycas revolute Thunb.

Distribution: Southern Japan.

Collection area: Kathmandu, Nepal.

Tall evergreen dioecious tree with crown of leaves at the apex of trunk. Base and sometimes distal part with numerous adventitious lateral branches. Leaves are 40–100 or more, 1-pinnate; petiole subtetragonal in cross section. Margins of the leaflets are revolute or recurved. Megasporophyll with 2-6 ovules, bright reddish ovules.

Family: Ephedraceae

Ephedra gerardiana Wallich ex C.A. Meyer

Distribution: Afghanistan, Himalaya (Karakoram to Bhutan).

Collection area: Rasuwa and Mustang, Nepal.

Evergreen, dioecious, Shrubs small, to 15 cm, stems erect, ; woody stems buried in soil, parts above ground scabrous, rarely almost smooth; branchlets directed upward, dark green, later brown, short, stout, usually with 1–3 nodes, longitudinally furrowed, often scabridulous. Leaves opposite, 2–3 mm, connate for aprroximately 2/3 their length.

Ephedra intermedia Schrenk & Meyer.

Distribution: Afghanistan, Kashmir,Nepal.

Collection area: Mustang, Nepal.

Evergreen, dioecious subshrubs or shrubs to 1 m, densely branched, sometimes with creeping stem producing single, erect, green primary branches; branchlets yellowish or bluish green, often pruinose. Leaves in whorls of 3 or opposite, connate for at least 2/3 their length.

Family: Ginkgoaceae

Ginkgo biloba L.

Distribution: China.

Collection area: Lalitpur, Nepal.

Tree deciduous, dioecious, upto 40 m tall; trunk to 4 m d.b.h.; bark light gray or grayish brown, longitudinally fissured especially on old trees; crown conical initially, finally broadly ovoid; long branchlets pale brownish yellow initially, finally gray, internodes, short branchlets blackish gray, with dense, irregularly elliptic leaf scars; winter buds yellowish brown, ovate. Leaves with petiole; blade pale green.

Family: Gentaceae

Gnetum montanum Markgr.

Distribution: Himalaya (Nepal to Bhutan), Assam, Burma, Thailand, Indo-China, W. & S. China.

Collection area: Tanahu, Nepal.

Vines evergreen, dioecious, upto more than 10 m tall; branchlets orbicular or compressed orbicular in cross section, smooth, sometimes wrinkled longitudinally. Leaf blade usually oblong, sometimes oblong-lanceolate or elliptic, leathery or nearly so, lateral veins each side, base rounded or broadly cuneate, apex obtuse to acute. Male inflorescences lax, each collar with (20–)25–45 flowers plus 10–15 sterile female flowers.

Family: Pinaceae

Abies densa Griff.

Distribution: E. Nepal to Assam, and S.E. Tibet.

Collection area: Sankhuwasabha, Nepal

Trees evergreen, monoecious, up to 60 m tall; bark scaly; branchlets initially grayish yellow or light brown, turning yellowish gray or grayish brown in 2nd or 3rd year; winter buds resinous. Leaves radially spreading or ascending on upper side of branchlets, pectinately arranged in 2 lateral sets on lower side, linear.

Abies pindrow (Royle ex D. Don) Royle

Distribution: Western Himalaya, in Afghanistan, Pakistan, India, and Nepal.

Collection area: Jumla and Mugu, Nepal.

Trees evergreen, monoecious,upto 60 m tall.Twigs stiff, pinkish to pale brown, soon fading to pale gray, glabrous, lightly ridged, with circular leaf scars. Leaves spirally arranged but appearing roughly 2-ranked on shade foliage, the upper leaves somewhat shorter than the lower, twisted at base, flattened, grooved and dark green above. Seed cones lateral, erect, on short peduncles, cylindrical with obtuse apex.

Abies spectabilis D. Don

Distribution: Himalaya (Chitral to C. Nepal).

Collection area: Rasuwa and Illam, Nepal.

Trees evergreen, monoecious,upto 50 m tall; trunk to more than 1.5 m d.b.h.; bark rough, scaly; branchlets yellowish gray, brown, or reddish brown, furrowed, pubescent or glabrous, densely leafy; winter buds globose or ovoid, resinous. Leaves ascending on upper side of cone-bearing branchlets, pectinately arranged in 2 lateral sets on young and vegetative branchlets, bright green adaxially, linear, stomatal lines in 2 white bands abaxially, resin canals 2, marginal, apex emarginate or 2-cleft.

Cedrus deodara Roxb.ex D. Don.

Distribution: Afghanistan, Himalaya (Kashmir to W. Nepal).

Collection area: Jumla and Bhaktapur, Nepal.

Trees evergreen, monoecious,upto 60 m tall; bark dark gray, cracking into irregular scales; branches horizontal, slightly tilted or slightly pendulous; long branchlets usually pendulous, pale grayish yellow and densely pubescent with some white powder in 1st year, thereafter grayish; winter bud scales curved outward at base. Leaves radially spreading on long branchlets, in apparent fascicles of 15–20 on short branchlets, initiallypale green, aging dark green, linear but broader distally, triangular in cross section, 2.5–5 cm × 1–1.5 mm, hard, stomatal lines 2 or 3 adaxially and 4–6 abaxially, apex acuminate.

Larix griffithiana Carriere; Traite Gen.
Distribution: Himalaya (Nepal, Sikkim, Bhutan), Assam
Trees deciduous, monoecious,upto over 20 m tall; bark gray-brown or dark brown, deeply fissured longitudinally; long branchlets initially reddish brown, light brown, or yellowish brown; short branchlets 6–8 mm in diam., nearly smooth, with remnants of bud scales and rings of revolute scales bases; winter buds ovoid-globose or globose, not resinous. Leaves keeled abaxially and toward base adaxially.

Larix himalaica Cheng & L. K. Fu.
Distribution: C. Nepal, S. Tibet.
Collection area: Rasuwa, Nepal.
Tree deciduous, monoecious. Branchlets pendulous; long branchlets light yellow or brownish yellow, bright in 2nd year, finally yellowish gray; short branchlets, glabrous, with rings of reflexed bud scales; winter buds subglobose or conical-globose. Leaves linearflat or keeled toward base adaxially, keeled abaxially.

Picea smithiana (Wall) Boiss.
Distribution: Afghanistan, Himalaya (Kashmir to C. Nepal).
Collection area: Mugu and Manang, Nepal.
Tree evergreen, monoecious, up to 60 m tall; bark pale brown, breaking into irregular plates; crown conical; branchlets pendulous, pale brown or pale gray when young, glabrous; winter buds reddish brown, conical or ovoid, scales slightly open, rarely appressed at base of branchlets. Leaves spreading radially, directed obliquely forward, quadrangular-linear, slender, curved, quadrangular or subquadrangular in cross section, stomatal lines 2–5 along each surface, apex acute or acuminate.

Pinus roxburghii Sarg.
Distribution: Afghanistan, Himalaya (Kashmir to Bhutan).
Collection area: Rasuwa, Dolakha, Baitadi and Kathmandu, Nepal.
Tree evergreen, monoecious,upto 55 m tall; trunk to over 1 m d.b.h.; bark dark red-brown, thick, deeply and longitudinally fissured, scaly; winter buds brown, small, ovoid, not resinous. Needles 3 per bundle, slender, flabellate-triangular in cross section.

Pinus wallichiana A.B Jacks.
Distribution: Afghanistan, Himalaya (Kashmir to Nepal).
Collection area: Rasuwa, Mugu, Manang and Dolakha, Nepal.
Tree evergreen, monoecious,upto 70 m tall; bark dark gray-brown, minutely scaly and flaking; crown broadly pyramidal; 1st-year branchlets green (drying red-brown), shiny, faintly whitish bloomed, glabrous; winter buds red-brown, cylindric-obovoid or cylindric-conical, slightly resinous. Needles 5 per bundle, pendulous, slender, triangular in cross section.

Tsuga dumosa D. Don.
Distribution: Himalaya (Kumaun to Bhutan), N. Burma.
Collection area: Rasuwa, Mugu and Manang, Nepal.
Tree evergreen, monoecious,upto 40 m tall; bark brownish gray or gray-brown, thick, longitudinally fissured; crown pyramidal; branchlets initially yellowish or reddish brown,

turning light brown or dark gray in 2nd or 3rd year, ridged and grooved, brown lanate. Leaves pectinately arranged, linear, rarely narrowly linear-lanceolate, grooved adaxially.

Family: Podocarpaceae

Podocarpus nerifolius D. Don.
Distribution: Himalaya (Nepal, Sikkim), Assam, Burma, China, Taiwan, Malaysia, New Guinea, Fiji.
Collection area: Kathmandu, Nepal.
Tree evergreen, dioecious, 12- 25 m tall with whorled branches. Bark grayish brown, thin and fibrous. Linear leathery leaves nearly sessile. Leaves on young trees being longer than those on older trees; trunk usually to 5 cm d.b.h.; bark grayish brown, thin, fibrous, peeling off in longitudinal flakes; branches spreading or ascending. Leaves foliage bud scales erect, triangular, 1–1.5 mm wide, apex acute. Leaf blade lanceolate, usually slightly curved, leathery, midvein raised adaxially, flat or slightly raised abaxially, base cuneate into short petiole, apex long acuminate; juvenile leaves wider, with obtuse, mucronate apex. Leaves much longer, 10-17.8 cm and are linear.

Family: Taxaceae

Cephalotaxus Siebold and Zuccarini ex Endlicher
Distribution:Korea, China, Japan, Burma, Laos, Vietnam and India.
Collection area: Kathmandu, Nepal.
Evergreen dioecious, rarely monoecious tree or shrub with branches opposite or whorled. Buds ovate, covered with numerous persistent imbricate scales. Leaves spirally arranged on terminal branchlets, needle-like, usually pectinate. Cones are drupe-like, with a leathery, fleshy outer covering, green to reddish.

Taxus contorta Griff.
Distribution: South West Tibet (Jilong region); central and W Nepal; India: Himachal Pradesh, Jammu & Kashmir, Uttar Pradesh; N Pakistan; Afghanistan.
Collection area: Manang, Nepal.
Evergreen, dioeciousshrubs, or trees up to 20 m tall with numerous branches forming a rounded or pyramidal crown. Twigs alternate, slender, round, finely grooved along leaf bases, green turning yellow-brown or gray. Leaves spirally inserted but arranged in a V-formation, linear, slightly twisted at the short-petiolate base,Seed cones axillary, solitary.

Taxus mairei (Lemee & Lev.) S.Y. Hu ex T.S. Liu
Distribution: Taiwan and China.
Collection area: Kavrepalanchowk, Nepal.
Evergreen, dioecious shrubs, or trees to 30 m tall, twigs alternate, slender, round, finely grooved along decurrent leaf bases, green turning orange- or purple-brown. Leaves 2-ranked, spreading at nearly right angles to the shoot.Seed cones axillary, solitary or in pairs.

Taxus media Rehder
Hybrid of *Taxus baccata* and *T. cuspidata*created by T.D. Hatfield.
Collection area: Makwanpur, Nepal.

Evergreen, dioecious medium-sized tree or large shrub, depending on cultivar can also be dwarfed, needle like leaves, two-ranked, blunt bud scales, dark green color on upperside, light to medium green on underside, needles can be straight or curved, leaves apex is pointed.

Taxus wallichiana Zucc.
Distribution: East Asiaand Malesia; Afghanistan, Himalaya (Kashmir to Bhutan), Assam, N. Burma, Indo-China, W. China, Malaysia
Collection area: Kavrepalanchowk, Nepal.
Evergreen, dioecious trees or shrubs to 30 m tall, leafy branchlets, leaves linear to lanceolate, gradually tapered distally, usually falcate, apex gradually acuminate or abruptly tapered and indistinctly mucronate; winter bud scales early deciduous or some persistent at base of branchletsLeaves subsessile or with petiole.The seed cone is highly modified, berry-like.

Family: Taxodiaceae
Cryptomeria japonica (L.F.) D. Don.
Distribution: Japan
Collection area: Illam and Bhaktapur, Nepal.
Evergreen, monoecious trees upto 40 m tall; trunk to at least 2 m d.b.h.; bark reddish brown, fibrous, peeling off in strips; crown pyramidal; main branches whorled, horizontally spreading or slightly pendulous; branchlets usuallypendulous. Leaves spirally arranged, persisting 4-5 years, awl shaped, sometimes slightly twisted, pointing forward, cone globular and solitary.

4.5 DNA extraction

DNA, from leaves samples of 59 gymnosperms, except *Larix* (stem), extracted from CTAB and performed with agarose gel electrophoresis method were visualized under UV illumination. DNA bands of 12 samples in agarose gel electrophoresis are shown in Fig 4.5. DNA extraction of two herbaria samples *Abies densa* and *Larix griffithiana* (not shown in figure) was problematic showing no visible bands (not enough DNA).

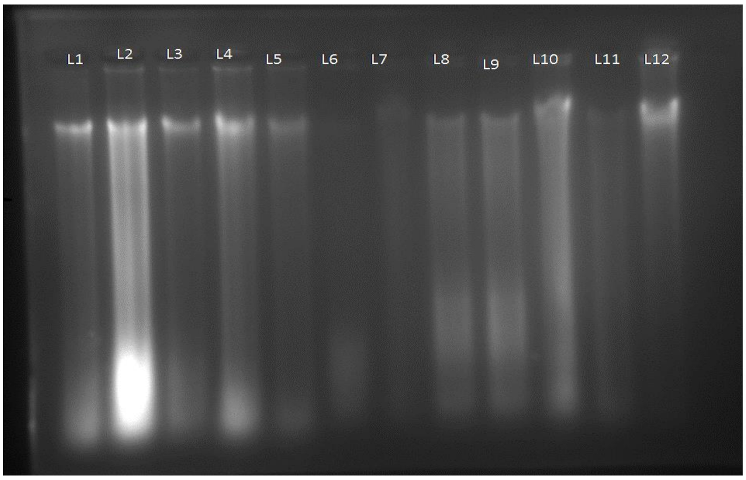

Figure 4.5 Plant DNA from silica gel preserved samples extracted by CTAB method on 1% agarose, L: lanes. Sample in L6 shows very faint band of DNA in 1 % gel but has shown PCR success.

4.6 Amplification of barcode markers

The success of PCR amplification reactions were observed performing agarose gel electrophoresis. The amplified bands of DNA fragments in gel were observed under UV-illumination. The illumination or visible bands corresponds to the positive results. Negative control (reaction mixture with no template) and positive standard DNA (DNA ladder or marker) were used to compare size and validate the presence of target amplicons. Representative diagram of agarose gel electrophoresis of PCR amplification is shown in Fig 4.6 and Fig 4.7 showing confirmation of amplification of *ITS*, *MatK* , *ycf1* and *petN-psbM* regions.

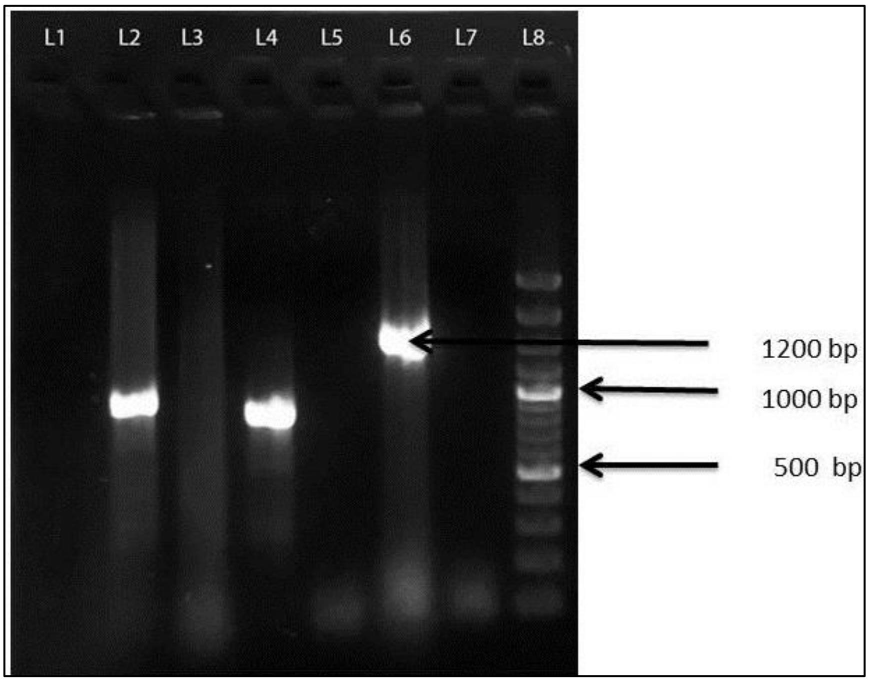

Figure 4.6 Visualization of PCR amplified DNA bands on agarose gel showing PCR products of ycf1 gene from chloroplast of the samples. L6: amplified product; L8: DNA ladder Gene ruler 100 bp Plus, L7: Negative Control

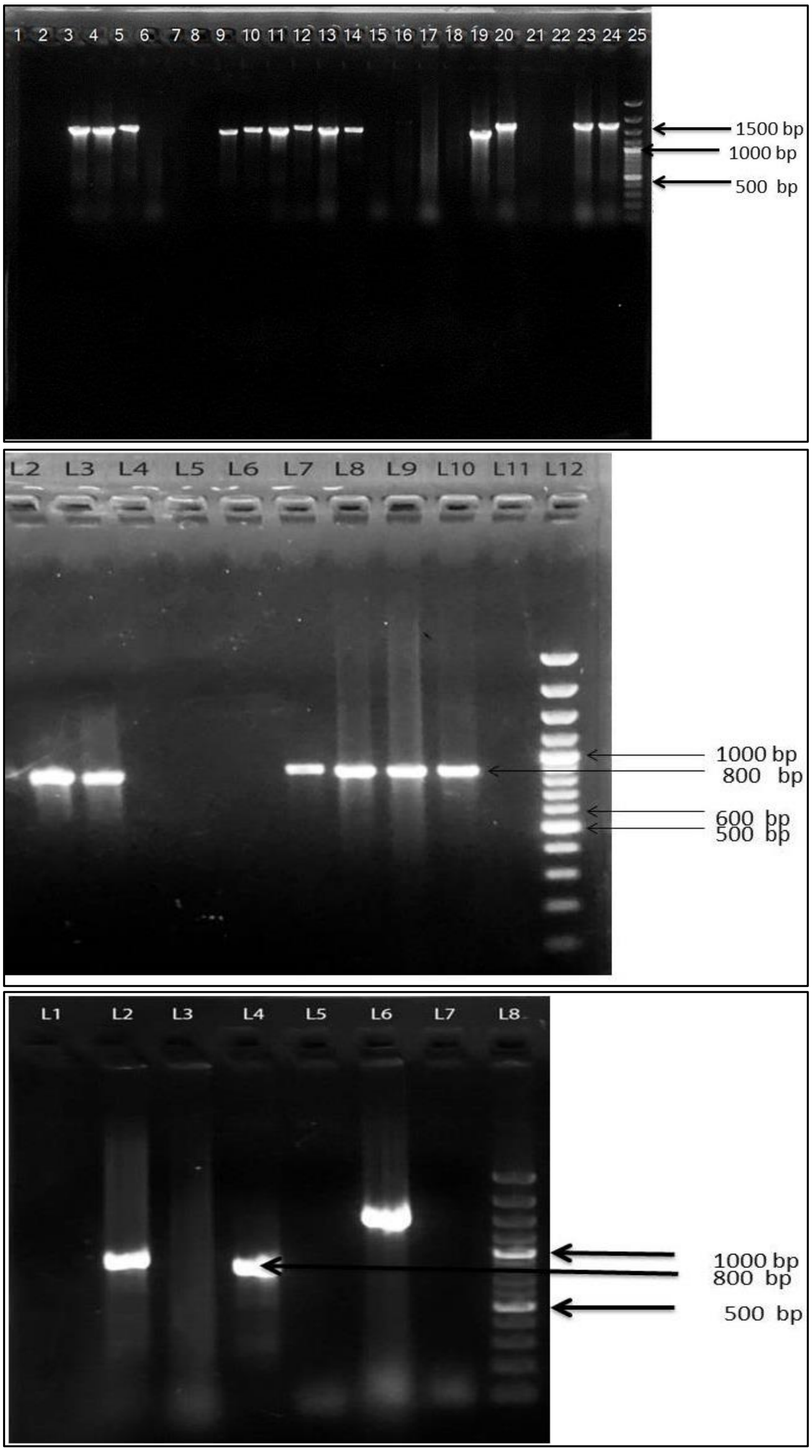

Figure 4.7 Visualization of PCR amplified DNA bands on agarose gel. (A) PCR amplification of ITS region from nuclear DNA of gymnosperms on 1% agarose gel. Lanes 1-24: PCR products, lane 25: DNA ladder- Gene ruler 100 bp plus; Lanes 1,2,6,7,8, 15, 18 and 21: No amplification, Lane 22: Negative Control (B) PCR amplification of *Matk*gene; L2, L3, L7, L8, L9, L10 showing positive amplification, L12: DNA ladder Gene ruler 100 bp Plus, Lane 11: Negative Control (C) PCR amplification of pet-psbM region of Junipers (Cupressacae) family; L2, L4 showing positive amplification, L8: DNA ladder Gene ruler 100 bp Plus, Lane 7: Negative Control.

4.7 PCR clean up

Succesfully amplified PCR products were purified to avoid the chemical interference in sequencing by SAP-EXO purification kit. After the reaction, the samples were loaded in 1% agarose gel to visualize the improved quality of DNA bands and also to compare the tentative intensity of DNA band per microliter of sample.

4.8 PCR and Sequencing efficiency

PCR and sequencing efficiency of four different markers were checked and result is shown in Fig 4.8. The PCR efficiency of *ITS* marker was 80% and involved most difficulties in optimization to get the target product. The rest of 20 % samples could not be studied for their molecular information due to unsuccessful PCR. The PCR success rate of *MatK*, *ycf1* and *PetN-psbM* regions, on the other hand, was absolute (100%) each. The Pinaceae family specific marker *ycf1* was the most efficient one with absolute PCR efficiency. Sequencing efficiency was absolute only for *ycf1* marker (100%). *MatK*marker was the one with second highest sequencing efficiency of 96.61 %, followed by *petN-psbM* marker 94.2 %. *ITS* marker has the lowest sequencing efficiency of 76.59 %.

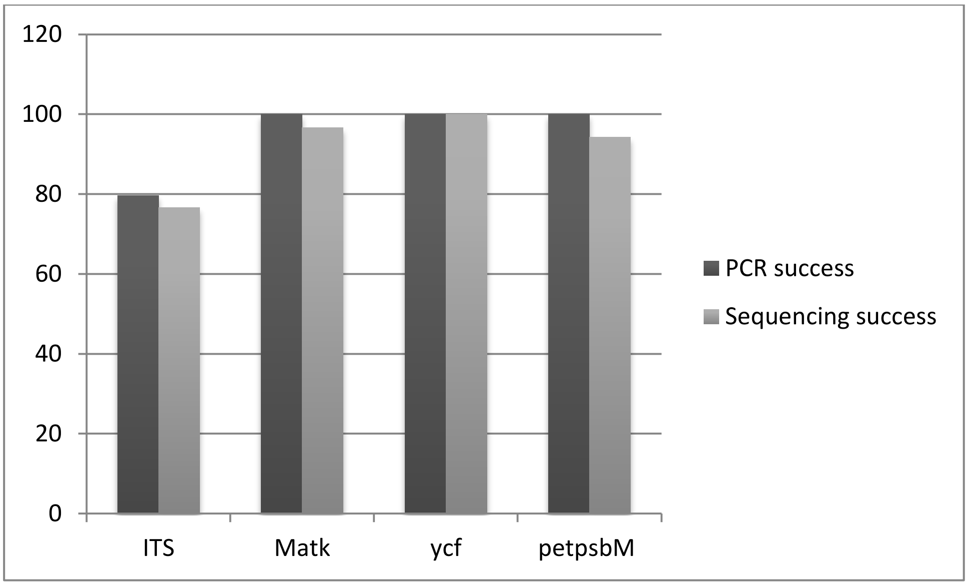

Figure 4.8 Sequencing and PCR efficiency of five different markers used to analyse the gymnosperms. Markers *ycf1* and *petN-psbM* have shown excellent PCR and sequencing success

4.9 Length of markers

Variation in the length of markers was analysed for every sample and the results were reported as in Table 4.1, Table 4.2, Table 4.3 and Table 4.4. Some of the regions with length too smaller than the fellow sequences were affected by abortion during sequencing.

Table 4.1 Length of *ITS* region amplified by *ITS*p5/u4 primers

Species	Lab code	Length of *ITS* (bp)	With primer (bp)	Remarks
Abies pindrow	GW11	1672	1752	F/R
Abies spectabilis	GR5	1688	1728	F/R
Araucaria angustifolia	GM18	1049	1188	F/R
Araucaria columnaris	GM16	1099	1138	F/R
Cryptomeria japonica	GM15	1065	1145	F/R
Cupressus macrocarpa	CG2	1144	1162	F/R
Cupressus sampervirens	CG3	1110	1132	F/R
Cupressus torulosa	GM14	1149	1174	F/R
Cupressus torulosa	MG3	1149	1175	R
Cycas revoluta	GL11	1046	1056	F/R
Ephedra gerardiana	GR1	1126	1156	F/R
Ephedra intermedia	MG2	1562	1595	F/R
Ephedra intermedia	MG1	1562	1593	R
Ginkgo biloba	GK14	1205	1264	F/R
Gnetum montanum	GT6	1202	1241	F/R
Juniperus horizontalis	GM19	1003	1146	F/R
Juniperus indica	GR6	960	1003	R
Juniperus indica	GW19	960	1009	R
Juniperus indica	GW18	960	1057	R
Juniperus indica	GM6	960	1176	R
Juniperus indica	GM7	960	1166	R
Juniperus recurva	GR8	1109	1152	F/R
Juniperus squamata	GR9	1123	1182	F/R
Larix himalaica	GR7	1795	1819	F/R
Podocarpus nerifolius	GK16	1289	1311	F/R
Taxus contorta	GT2	1166	1166	F/R
Taxus mairei	GT3	1166	1166	F/R
Taxus wallichiana	GT1	1167	1167	F/R
Thuja orientalis	CG1	1083	1113	R
Tsuga dumosa	GR3	1749	1796	F/R

Note: F-Forward sequence, R- Reverse requence, F/R-contig of forward and reverse sequence

Table 4.2 Length of *MatK* region amplified with Gym_*MatK* F1/R1 primers

Species	Length of *MatK* (bp)	With primer (bp)	Remarks
Abies pindrow_GM2	816	845	F/R
Abies pindrow_GW11	816	886	F
Abies pindrow_GW20	816	850	F
Abies spectabilis_GR5	818	888	F/R

Araucaria angustifolia_GM18	825	893	F/R
Araucaria bidwilli _GM17	822	889	F/R
Araucaria columnaris_GM16	820	897	F/R
Cedrus deodara_GM11	817	882	F/R
Cedrus deodara_GM13	817	884	F
Cedrus deodara_GW13	817	849	F
Cephalotaxus spp_GT5	822	888	F/R
Cryptomeria japonica_GM15	820	888	F/R
Cryptomeria japonica_MG6	820	845	F
Cupressus macrocarpa_CG2	602	623	F/R
Cupressus sempervirens_CG3	814	874	F/R
Cupressus torulosa_GM14	812	862	F/R
Cupressus torulosa_MG3	812	846	F
Cycas pectinata_GL10	832	900	F/R
Cycas revoluta_GL11	835	907	F/R
Ephedra gerardiana_GR1	830	857	F/R
Ephedra gerardiana_MG7	830	854	F/R
Ephedra intermedia_MG1	1139	1164	F
Ephedra intermedia_MG2	1139	1156	F
Ginkgo biloba_GK14	817	880	F/R
Gnetum montanum _GT6	809	855	F/R
Cupressus lusitanica_GW12	806	843	F/R
Juniperus horizontalis_GM19	817	883	F/R
Juniperus indica_GR6	809	887	F/R
Juniperus indica_GW18	809	844	F
Juniperus indica_GW19	809	842	F
Juniperus indica_GW22	809	843	F
Juniperus indica_MG4	809	845	F
Juniperus recurva_GR8	815	878	F/R
Juniperus squamata_GR9	812	828	F/R
Larix himalaica_GR7	810	892	F/R
Picea smithiana_GM1	816	879	F/R
Picea smithiana_GW15	816	850	F
Pinus roxburghii_GM5	818	838	F/R
Pinus roxburghii_GM8	818	847	F
Pinus roxburghii_GM10	818	846	F
Pinus roxburghii_GR4	818	888	F
Pinus wallichiana_GM4	816	847	F/R
Pinus wallichiana_GM9	816	848	F
Pinus wallichiana_GR2	816	895	F
Pinus wallichiana_GW21	816	847	F
Pinus wallichiana_MG5	816	847	F
Podocarpus nerifolius_GK16	828	899	F/R
Taxus contorta_GT2	823	886	F/R

Taxus mairei_GT3	822	887	F/R
Taxus media_GT4	824	898	F/R
Taxus wallichiana_GT1	824	887	F/R
Tsuga dumosa_GM3	818	846	F/R
Tsuga dumosa_GR3	818	884	F
Tsuga dumosa_GW17	818	851	F

Note: F-Forward sequence, R- Reverse requence, F/R-contig of forward and reverse sequence

Table 4.3 Length of *ycf1* region of Pinaceae family amplified by *ycf1*g_F/R primers

Lab code	**Species**	**Length of *ycf1* (bp)**	**With primer (bp)**	**Remarks**
GW11	***Abies pindrow***	1111	1181	F/R
GM2	***Abies pindrow***	1111	1122	F
GW20	***Abies pindrow***	1111	1140	F
GR5	***Abies spectabilis***	1373	1416	F/R
GM11	***Cedrus deodara***	1130	1136	F/R
GW13	***Cedrus deodara***	1130	1192	F
GM13	***Cedrus deodara***	1130	1136	F
GR7	***Larix himalaica***	1390	1417	F/R
GW15	***Picea smithiana***	1410	1428	F/R
GM1	***Picea smithiana***	1410	1436	F
GR4	***Pinus roxburghii***	1091	1114	F/R
GM5	***Pinus roxburghii***	1091	1146	F
GM8	***Pinus roxburghii***	1091	1140	F
GM10	***Pinus roxburghii***	1091	1150	F
GR2	***Pinus wallichiana***	1115	1196	F/R
GW21	***Pinus wallichiana***	1115	1184	F
GM4	***Pinus wallichiana***	1115	1165	F
MG5	***Pinus wallichiana***	1115	1065	F
GR3	***Tsuga dumosa***	1457	1477	F/R
GW17	***Tsuga dumosa***	1457	1485	F
GM3	***Tsuga dumosa***	1457	1493	F

Note: F-Forward sequence, R- Reverse requence, F/R-contig of forward and reverse sequence

Table 4.4 Length of *petN-psbM* region of Cupressaceae family amplified by petNF/psbMR primers

Lab code	**Species**	**Length of petN- psbM region (bp)**	**With Primer (bp)**	**Remarks**
CG2	***Cupressus macrocarpa***	791	824	F/R
GM14	***Cupressus torulosa***	792	822	F/R
GW12	***Cupressus lusitanica***	788	815	F/R
sCG4	***Cupressus lusitanica***	788	818	F
GM19	***Juniperus horizontalis***	856	912	F/R
GR6	***Juniperus indica***	872	821	F/R
GW19	***Juniperus indica***	872	901	F/R
GW18	***Juniperus indica***	872	915	F/R
GW22	***Juniperus indica***	872	906	F/R
GM6	***Juniperus indica***	872	911	F/R
MG4	***Juniperus indica***	872	906	F/R
GR8	***Juniperus recurva***	865	908	F/R
GR9	***Juniperus squamata***	788	828	F/R
CG1	***Thuja orientalis***	593 *	*640	F
CG3	***Cupressus sempervirens***	801	839	F/R
MG3	***Cupressus torulosa***	792	828	F/R

Note: F-Forward sequence, R- Reverse requence, F/R-contig of forward and reverse sequence, * signifies length of Forward sequence.

Table 4.5 Summarized List of species with their markers length

Species	**MatK (bp)**	**ITS (bp)**	**ycf1(bp)**	**petN- psbM (bp)**
Abies pindrow	816	1672	1111	NA
Abies spectabilis	818	1688	1373	NA
Araucaria angustifolia	825	1049	NA	NA
Araucaria bidwilli	822	NA	NA	NA
Araucaria columnaris	820	1099		NA
Cedrus deodara	817	NA	1130	NA
Cephalotaxus spp	822	NA	NA	NA
Cryptomeria japonica	820	1065	NA	NA
Cupressus macrocarpa	602	1144	NA	791
Cupressus sempervirens	814	1110	NA	801
Cupressus torulosa	812	1149	NA	792
Cycas pectinata	812	NA	NA	NA
Cycas revoluta	832	1046	NA	NA
Ephedra gerardiana	835	1126	NA	NA
Ephedra intermedia	1139	1562	NA	NA
Ginkgo biloba	1139	1205	NA	NA
Gnetum montanum	817	1202	NA	NA
Cupressus lusitanica	809	NA	NA	788
Juniperus horizontalis	806	1003	NA	856
Juniperus indica	817	960	NA	872
Juniperus recurva	809	1109	NA	865
Juniperus squamata	815	1123	NA	788
Larix himalaica	812	1795	1390	NA
Picea smithiana	810	NA	1410	NA
Pinus roxburghii	818	NA	1091	NA
Pinus wallichiana	818	NA	1115	NA
Podocarpus nerifolius	816	1289	NA	NA
Taxus contorta	828	1166	NA	NA
Taxus mairei	823	1166	NA	NA
Taxus media	822	NA	NA	NA
Taxus wallichiana	824	1167	NA	NA
Tsuga dumosa	824	1749	1457	NA
Thuja orientalis	NA	NA	*NA*	593 *

NA: Not applicable, *Sequence aborted

4.10 Nucleotide composition of each marker

Nucleotide composition provided the idea about the compostion of A, T, G and C in terms of their percentage of total length.

***ITS* region**: The nucleotide compostions of *MatK* region in gymnosperms studied are shown in Table 4.5. Average compostion of A was 17.5 %, T was 21.1%, G 31.6 % and C was29.7 % and the G+C content was calculated using this value.

Table 4.6 Nucleotide compostion of *ITS* region of ribosomal DNA in gymnosperms amplified with *ITS* p5F and U4R primers

Species	**T(U)**	**C**	**A**	**G**	**Total**
Abies_pindrow_GW11	24.8	28.6	14.6	32.0	1102.0
Abies_spectabilis_GR5	25.4	28.1	14.4	32.2	1107.0
Araucaria_angustifolia_GM18	19.2	33.4	13.1	34.3	1047.0
Araucaria_columnaris_GM16	18.5	35.1	13.1	33.2	1119.0
Cryptomeria_japonica_GM15	19.4	31.1	18.7	30.8	855.0
Cupressus_macrocarpa_CG2	19.8	29.7	19.4	31.1	954.0
Cupressus_sempevirens_CG3	19.1	29.6	19.3	31.9	931.0
Cupressus_torulosa__MG3	20.2	28.3	20.2	31.4	947.0
Cupressus_torulosa_GM14	20.6	29.7	18.8	31.0	927.0
Cycas_revoluta_GL11	16.2	33.3	21.2	29.3	1108.0
Ephedra_gerardiana_GR1	20.6	28.2	18.3	33.0	1119.0
Ephedra_intermedia_MG1	24.2	24.6	21.1	30.2	745.0
Ephedra_intermedia_MG2	20.8	28.3	18.2	32.7	1121.0
Ginkgo_biloba_GK14	19.7	32.6	14.0	33.6	1079.0
Gnetum_montanum_GT6	20.7	26.3	16.4	36.6	1003.0
Juniperus_horizontalis_GM19	18.6	30.5	19.6	31.2	957.0
Juniperus_indica_GM6	22.1	30.2	18.6	29.1	959.0
Juniperus_indica_GM7	20.0	29.6	19.1	31.3	957.0
Juniperus_indica_GR6	20.9	30.1	18.9	30.0	952.0
Juniperus_indica_GW18	18.9	32.0	18.7	30.4	957.0
Juniperus_indica_GW19	20.6	30.0	18.7	30.8	958.0
Larix_himalaica_GR7	27.3	26.1	16.5	30.0	1149.0
Tsuga_dumosa_GR3	27.6	26.4	14.8	31.2	1131.0
Podocarpus_nerifolius_GK16	18.7	30.6	21.6	29.1	1163.0
Taxus_contorta_GT2	23.9	29.3	15.9	30.9	996.0
Taxus_mairei_GT3	23.4	30.1	15.8	30.7	996.0
Taxus_wallichiana_GT1	24.1	29.2	15.7	31.0	997.0
Thuja_orientalis_CG1	18.8	29.1	18.4	33.7	946.0
Juniperus_recurva_GR8	17.1	31.0	18.3	33.7	410.0
Juniperus_squamata_GR9	17.3	31.5	17.3	34.0	400.0
Avg.	21.1	29.7	17.5	31.6	969.7

MatK region: The nucleotide compostions of *MatK* region in gymnosperms studied are shown in Table 4.6.The average value of A contained in overall *MatK* region was 35 %, T 31.3 %, G 18.6 % and C 16.3 %. The G+C value was calculated and reported.

Table 4.7 Nucleotide compostion *MatK* barcode region in gymnosperms amplified with *MatK* Gym F1 and *MatK* Gym R1 primers

Species	**T(U)**	**C**	**A**	**G**	**Total**
Abies_pindrow_GM2	30.3	16.7	34.4	18.6	816.0
Abies_pindrow_GW11	30.1	16.6	34.5	18.7	817.0
Abies_pindrow_GW20	30.0	16.7	34.3	19.0	816.0
Abies_spectabilis_GR5	30.3	16.5	34.4	18.8	816.0
Araucaria_angustifolia_GM18	29.8	17.7	35.0	17.5	823.0
Araucaria_bidwilli__GM17	29.8	17.9	34.8	17.6	823.0

Araucaria_columnaris_GM16	29.6	17.4	35.4	17.6	823.0
Cedrus_deodara_GM11	28.8	18.0	34.4	18.8	816.0
Cedrus_deodara_GM13	28.8	17.9	34.5	18.8	817.0
Cedrus_deodara_GW13	28.8	18.0	34.4	18.8	817.0
Cephalotaxus_spp_GT5	36.2	17.3	28.5	17.9	820.0
Cryptomeria_japonica_GM15	29.6	16.5	36.8	17.1	820.0
Cryptomeria_japonica_MG6	29.9	16.5	36.6	17.1	820.0
Cupressus_macrocarpa_CG2	31.0	16.1	37.0	15.9	646.0
Cupressus_sempervirens_CG3_	30.1	15.6	38.4	15.9	813.0
Cupressus_torulosa_GM14	29.9	15.6	38.2	16.3	815.0
Cupressus_torulosa_MG3	30.1	15.6	38.1	16.2	816.0
Cycas_pectinata_GL10	29.3	17.4	34.9	18.5	834.0
Cycas_revoluta_GL11	29.1	17.6	34.7	18.6	834.0
Ginkgo_biloba_GK14_A	35.8	18.4	26.4	19.4	815.0
Gnetum_montanum__GT6	39.5	13.0	29.0	18.5	794.0
Cupessus_lusitanica_CG4	30.6	15.4	37.7	16.3	811.0
Juniperus_horizontalis_GM19	30.1	16.3	37.7	16.0	815.0
Juniperus_indica_GR6	30.1	15.6	38.2	16.1	814.0
Juniperus_indica_GW18	30.2	16.1	37.6	16.1	814.0
Juniperus_indica_GW19	30.2	16.0	37.6	16.2	814.0
Juniperus_indica_GW22_	30.2	16.0	37.5	16.3	815.0
Juniperus_indica_MG4	29.7	15.7	38.3	16.3	815.0
Juniperus_recurva_GR8	29.9	15.8	38.0	16.2	815.0
Larix_himalaica_GR7	28.4	17.5	35.2	18.9	818.0
Picea_smithiana_GM1	29.0	17.0	36.2	17.8	818.0
Picea_smithiana_GW15	29.0	17.1	35.9	18.0	817.0
Pinus_roxburghii_GM5	28.7	16.7	35.3	19.3	819.0
Pinus_roxburghii_GM8	28.7	17.0	35.1	19.2	818.0
Pinus_roxburghii_GM10	28.7	16.7	35.3	19.2	818.0
Pinus_roxburghii_GR4	28.9	16.7	35.3	19.1	818.0
Pinus_wallichiana_GM4	29.1	16.1	35.7	19.1	818.0
Pinus_wallichiana_GM9	29.1	16.4	35.5	19.0	819.0
Pinus_wallichiana_GR2	29.0	16.1	35.9	18.9	818.0
Pinus_wallichiana_GW21	29.1	16.0	35.8	19.2	819.0
Pinus_wallichiana_MG5	29.1	15.9	35.8	19.3	819.0
Podocarpus_nerifolius_GK16	29.0	17.1	34.3	19.6	823.0
Taxus_contorta_GT2	36.2	16.7	29.3	17.7	818.0
Taxus_mairei_GT3	36.2	16.8	29.4	17.6	817.0
Taxus_media_GT4	36.2	16.7	29.5	17.6	818.0
Taxus_wallichiana_GT1	36.4	16.7	29.3	17.6	819.0
Tsuga_dumosa_GM3	30.0	17.2	34.8	18.0	818.0
Tsuga_dumosa_GR3	30.0	17.2	35.0	17.8	818.0
Tsuga_dumosa_GW17	30.0	17.3	34.8	18.0	817.0
Juniperus_squamata_GR9	30.3	15.7	38.0	16.0	813.0
Cupressus_lusitanica_GW12	30.4	15.4	37.6	16.5	816.0
Ephedra_gerardiana_GR1	41.9	13.0	31.1	13.9	797.0
Ephedra_gerardiana_MG7	41.9	12.9	31.2	13.9	797.0
Ephedra_intermedia_MG1	38.5	13.7	33.4	14.3	1122.0
Ephedra_intermedia_MG2	38.2	14.3	33.3	14.1	1122.0
Avg.	31.3	16.3	35.0	17.4	824.1

***Ycf1* region**: This region was specific to Pinaceae family and the nucleotide compostions of *ycf1* region in Pinaceae family are shown in Table 4.7. The average value of A contained in overall *MatK* region was 31.7 %, T 29.7 %, G 19.9 % and C 18.7 %. The G+C value was calculated and reported.

Table 4.8 Nucleotide compostion *ycf1* barcode region in Pinaceae family of gymnosperms amplified with *ycf1*g F and *ycf1*g R primers

Species	T(U)	C	A	G	Total
Abies_pindrow_GM2	29.2	18.5	31.7	20.6	1054.0
Abies_pindrow_GW11	29.3	18.5	31.8	20.5	1070.0
Abies_pindrow_GW20	29.3	18.5	31.5	20.6	1062.0
Abies_spectabilis_GR5	29.3	18.5	31.7	20.4	1071.0
Cedrus_deodara_GM11	30.0	19.0	31.0	20.0	1044.0
Cedrus_deodara_GM13	30.0	19.0	31.1	19.9	1038.0
Cedrus_deodara_GW13	30.0	18.9	31.2	19.9	1048.0
Larix_himalaica_GR7	30.1	17.9	32.4	19.6	1078.0
Picea_smithiana_GM1	29.3	18.3	32.3	20.1	1083.0
Picea_smithiana_GW15	29.3	17.9	32.9	19.9	1085.0
Pinus_roxburghii_GM5	29.9	20.8	30.2	19.1	977.0
Pinus_roxburghii_GM8	31.4	20.0	29.9	18.7	934.0
Pinus_roxburghii_GM10	30.3	20.6	30.2	18.9	977.0
Pinus_roxburghii_GR4	30.0	20.6	30.2	19.2	984.0
Pinus_wallichiana_GM4	29.3	17.9	32.4	20.4	1042.0
Pinus_wallichiana_GR2	29.4	17.5	32.6	20.5	1055.0
Pinus_wallichiana_GW21	29.4	17.9	32.5	20.2	1046.0
Pinus_wallichiana_MG5	29.7	17.6	32.6	20.0	1033.0
Tsuga_dumosa_GM3	29.7	18.8	32.2	19.4	1079.0
Tsuga_dumosa_GR3	29.6	18.7	32.4	19.3	1085.0
Tsuga_dumosa_GW17	29.4	18.8	32.5	19.3	1072.0
Avg.	29.7	18.7	31.7	19.9	1043.7

***petN-psbM* region:** This region was specific to Cupressaceae family and the nucleotide compostions of *petN-psbM* region in Cupressaceae family are shown in Table 4.8. The average value of A contained in overall *MatK* region was 38.4 %, T 34.3 %, G 13.9 % and C 13.4 %. The G+C value was calculated and reported.

Table 4.9 Nucleotide compostion petN5- psbM barcode region in Cupressaceae family of gymnosperms amplified with petN5 F and psbM R primers

Species	T(U)	C	A	G	Total
Cupressus_macrocarpa__CG2	34.4	13.3	38.5	13.8	799.0
Cupressus_sempervirens_CG3	33.7	13.3	38.9	14.0	809.0
Cupressus_torulosa_GM14_	34.3	13.5	38.5	13.7	802.0
Cupressus_torulosa_MG3	34.3	13.5	38.4	13.8	802.0
Cupressus_lusitanica_CG4	34.2	13.3	38.9	13.6	799.0
Cupressus_lusitanica_GW12	34.1	13.3	38.9	13.8	800.0
Juniperus_horzontalis_GM19	33.9	13.3	39.6	13.3	844.0
Juniperus_indica_GM6	32.7	13.1	39.9	14.2	852.0
Juniperus_indica_GR6	33.0	14.6	38.6	13.8	828.0

Juniperus_indica_GW18	32.7	13.1	39.9	14.2	852.0
Juniperus_indica_GW19	32.7	13.1	39.9	14.2	852.0
Juniperus_indica_GW22	32.7	13.1	39.9	14.2	852.0
Juniperus_indica_MG4	32.7	13.1	39.9	14.2	852.0
Juniperus_recurva_GR8	39.4	13.6	33.2	13.7	873.0
Juniperus_squamata_GR9	39.0	14.0	32.5	14.5	798.0
Thuja_orientalis_CG1	34.9	13.8	38.4	12.9	581.0
Avg.	34.3	13.4	38.4	13.9	811.3

4.11 Features of markers

All of four markers were anlaysed for their G+C content, conserved sites, variable sites, parsimony informative sites, singleton sites and then comparative inference was extracted as shown in Table 4.9. Among them, *ITS* has the highest G+C value and the highest number of variable sites with respect to its length.The variation of *ITS* could not be considered better than *MatK* as it is resultant increment due to the presence of pseudogenes and homolgous copies that has created the ambiguous readings in most of the sequences. Due to this, information for identification of sample with BLAST, tree and distance method given by *ITS*, is considered, here, and supportive marker to strengthen the power of *MatK*. *MatK* was, thus,taken primary in terms of sequence informations, conserved and variable sites ratio and discrimination power of samples at species level.

Table 4.10 The characteristics of DNA markers used in the study (Analysed with MEGA)

DNA markers	**Maximum total aligned length**	**G+C content %**	**Conserved sites (C)**	**Variable sites (V)**	**Parsimony informative sites (Pi)**	**Singelton sites (S)**
MatK	824	33.7	342	828	703	52
ITS	1274	61.2	112	1129	1038	204
ycf1	1043	38.6	715	435	346	88
petN-psbM	811	27.3	250	666	439	227

4.12 Sequence alignment and variation

The multiple sequence alignment file carried out in Bioedit vs 7.2.5 was imported to TaxonDNA SpeciesIdentifier software and then interspecific and intraspecific pairwise distances were calculated using ‘Pairwise explorer’ function.

4.12.1 *MatK* distance

The average interspecific distance was 1.5 %, the value frequent upto 10 % and intraspecific distance was 0.5 %, the value frequent upto less than 10 %. Being average intraspecific distance less than the interspecific one as shown in Fig 4.9, this marker is considered powerful to distinguish gymnosperms samples upto the species level.

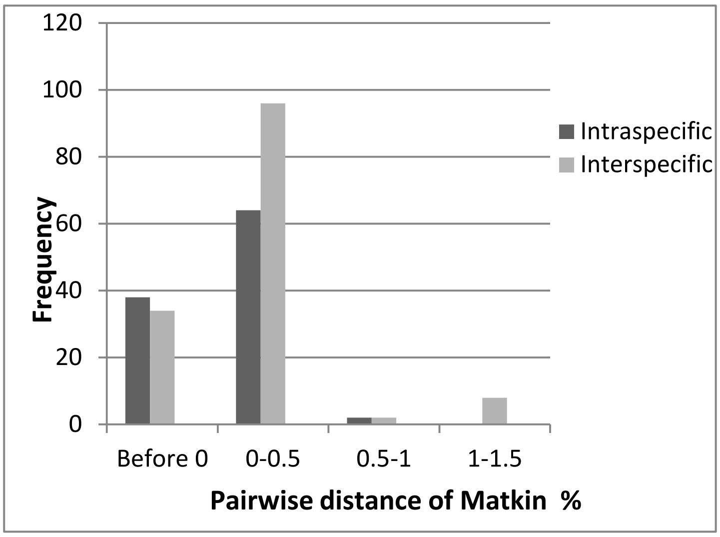

Figure 4.9 Pairwise distance of *Matk*region showing intraspecific and interspecific distances in % calculated by Sequence Identifier

4.12.2 *ITS* distance

The average interspecific distance was 19 %, the value frequent upto 8 % and intraspecific distance was 4.5 %, the value frequent upto 2 %. Being average intraspecific distance less than the interspecific one as shown in Fig 4.10, this marker is considered powerful to distinguish gymnosperms samples upto the species level.

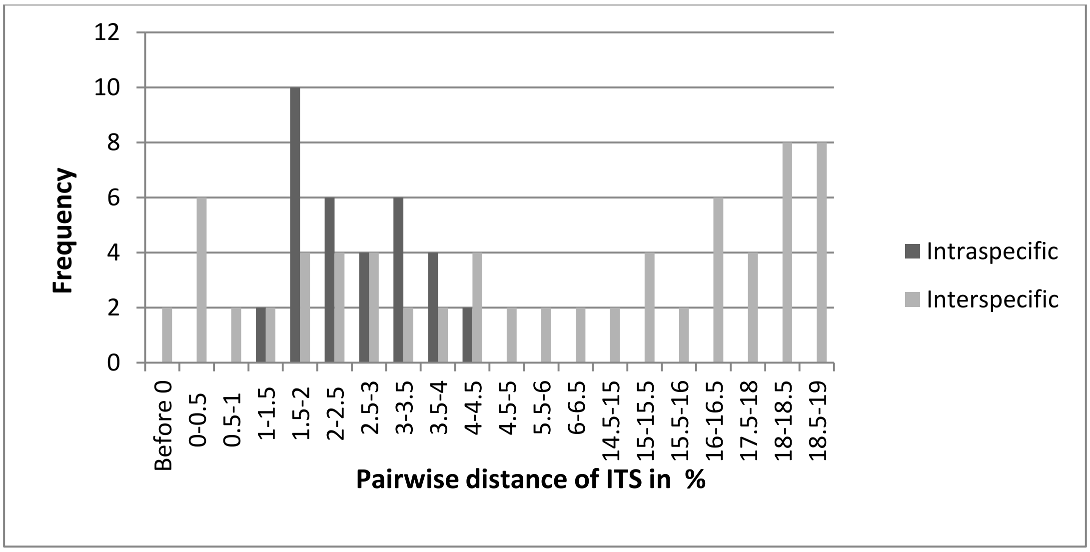

Figure 4.10 Pairwise distance of *ITS* region showing intraspecific and interspecific distances in % calculated by Sequence Identifier

4.12.3 *MatK* + *ITS* distance

The average interspecific distance of *MatK* with *ITS* was increased to 16.5 %, the value frequent upto 2 % and intraspecific distance was 3 %, the value frequent upto 8 %. Being average intraspecific distance less than the interspecific one as shown in Fig 4.11, this combination of markers (*MatK* + *ITS*) is considered powerful to distinguish gymnosperms samples upto the species level.

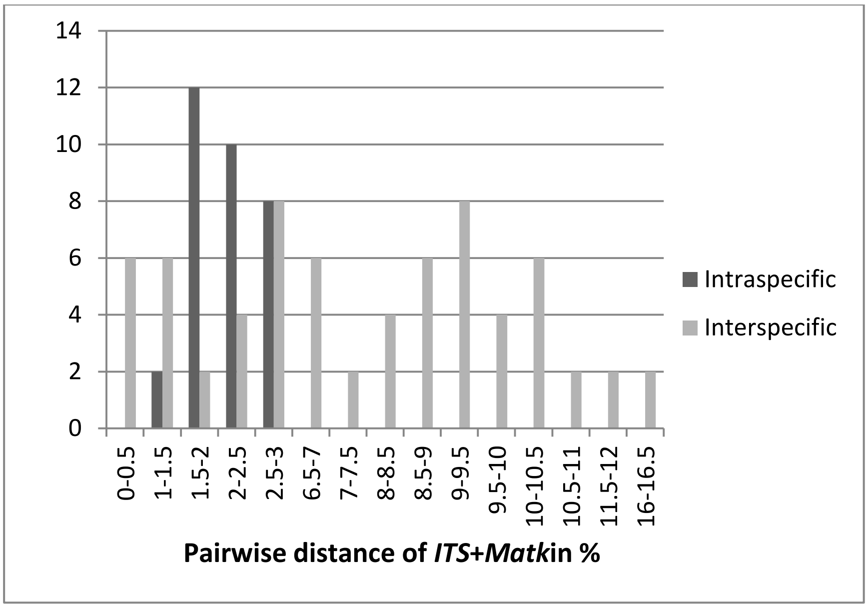

Figure 4.11Pairwise distance of *ITS* +*MatK* region showing intraspecific and interspecific distances in % calculated by Sequence Identifier

4.12.4 *Ycf1* distance

This could measure the distance between the species of Pinaceae family only. The average interspecific distance was 21 %, the value frequent upto less than5 % and intraspecific distance was 6.5 %, the value frequent upto 7 %. Being average intraspecific distance less than the interspecific one as shown in Fig 4.12, this marker is considered powerful to distinguish gymnosperms samples upto the species level.

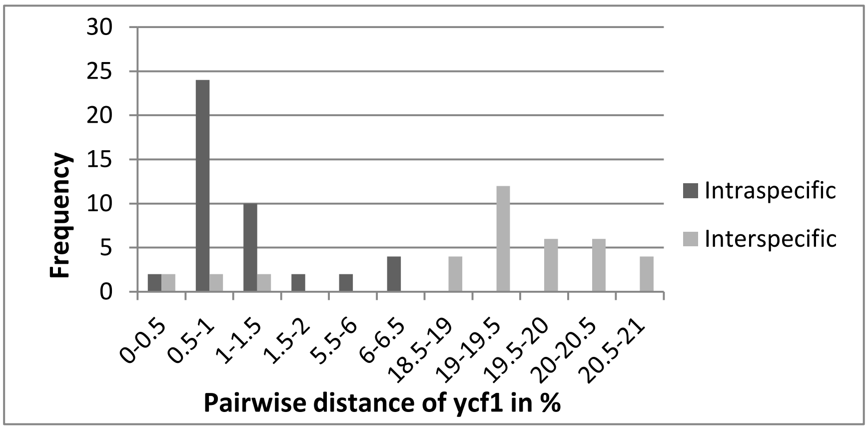

Figure 4.12 Pairwise distance of *ycf1* (Pinacae) region showing intraspecific and interspecific distances in % calculated by SequenceIdentifier

4.12.5 *PetN-psbM* distance

This could measure the distance between the species of Cupressaceae family only. The average interspecific distance was 17.5 %, the value frequent upto 10 % and intraspecific distance was 11.5 %, the value frequent upto less than 5 %. Being average intraspecific distance less than the interspecific one as shown in Fig 4.13, this marker is considered powerful to distinguish gymnosperms samples upto the species level.

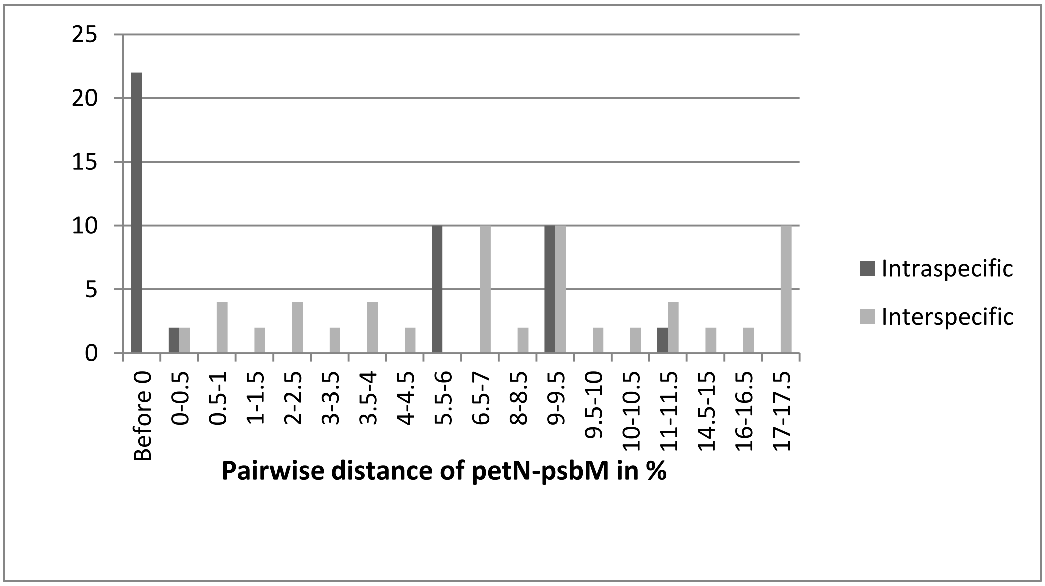

Figure 4.13 Pairwise distance of *petN-psbM* (Cupressaceae) region showing intraspecific and interspecific distances in % calculated by Sequence Identifier

4.12.6 *MatK* + *petN-psbM* distance

This could measure the distance between the species of Cupressaceae family only. The average interspecific distance of *MatK* with *petN-psbM* was increased to 9.5 %, the value frequent upto less than 8 % and intraspecific distance was 5.5 %, the value frequent upto 2 %. Being average intraspecific distance less than the interspecific one as shown in

Fig 4.14, this combination of markers (*MatK* + *petN-psbM*) is considered powerful to distinguish individuals of Cupressaceae family upto the species level.

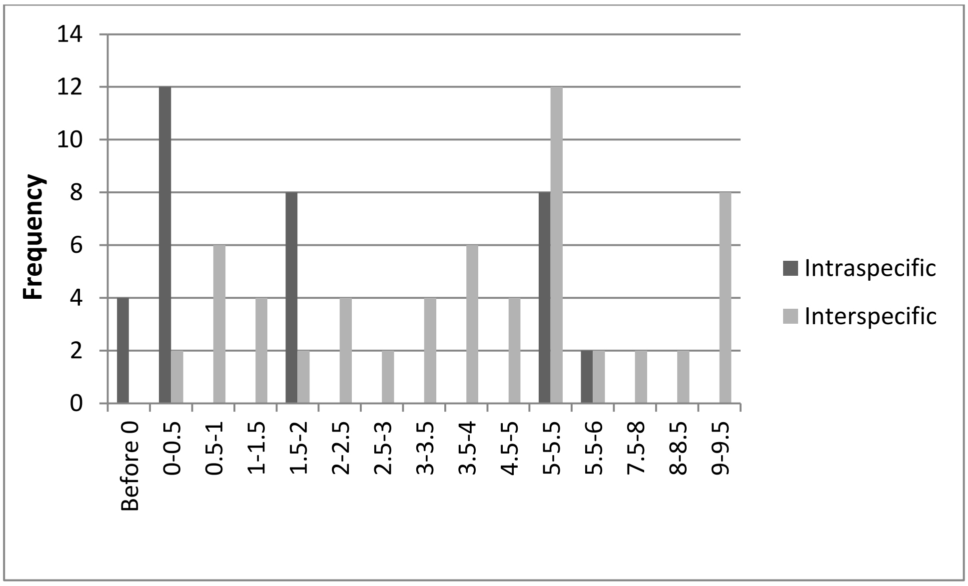

Figure 4.14 Pairwise distance of *MatK* + *petN-psbM* (Cupressaceae) region showing intraspecific and interspecific distances in % calculated by Sequence Identifier

4.12.7 *MatK* + *ycf1* distance

This could measure the distance between the species of Pinaceae family only. The average interspecific distance of *MatK* with *ycf1* was increased 11.5 %, the value frequent upto less than 4% and intraspecific distance was 3.5 %, the value frequent upto 2 %. Being average intraspecific distance less than the interspecific one as shown in Fig 4.15, this combination of markers (*MatK* + *ycf1*) is considered powerful to distinguish individuals of Pinaceae upto the species level.

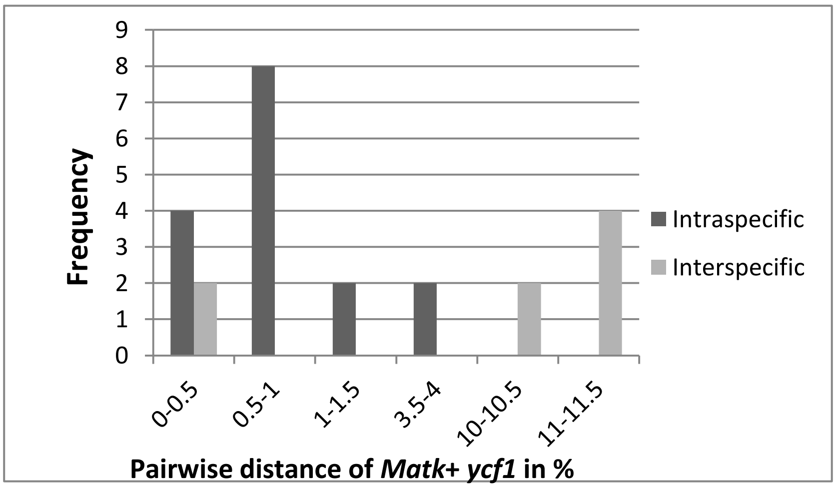

Figure 4.15 Pairwise distance of Matk+ ycf1 (Pinaceae) region showing intraspecific and interspecific distances in % calculated by Sequence Identifier

Details of intraspecific and interspecific distances of *MatK* region in terms of pairwise-distance are shown in Table 4.11 and Table 4.12 respectively.

Table 4.11 Intraspecific variation in terms of pairwise distance of *MatK* in gymnosperms calculated by TaxonDNA-SpeciesIdentifier

Replica 1	**Replica 2**	**Variation**
Abies pindrow GM2	*Abies pindrow GW20*	0.00%
Cedrus deodara GM13	*Cedrus deodara GW13*	0.00%
Cryptomeria japonica GM15	*Cryptomeria japonica MG6*	0.00%
Ephedra gerardiana MG7	*Ephedra gerardiana GR1*	0.00%
Ephedra intermedia MG1	*Ephedra intermedia MG2*	0.00%
Juniperus indica GW19	*Juniperus indica GW18*	0.00%
Juniperus indica GW22	*Juniperus indica MG4*	0.00%
Juniperus indica MG4	*Juniperus indica GM6*	0.00%
Juniperus indica MG4	*Juniperus indica GW22*	0.00%
Pinus roxburghii GM10	*Pinus roxburghii GM8*	0.00%
Pinus roxburghii GM10	*Pinus roxburghii GR4*	0.00%
Pinus roxburghii GM8	*Pinus roxburghii GM10*	0.00%
Pinus roxburghii GM8	*Pinus roxburghii GR4*	0.00%
Pinus roxburghii GR4	*Pinus roxburghii GM10*	0.00%
Pinus roxburghii GR4	*Pinus roxburghii GM8*	0.00%
Pinus wallichiana GM9	*Pinus wallichiana GW21*	0.00%
Pinus wallichiana GM9	*Pinus wallichiana MG5*	0.00%
Pinus wallichiana GW21	*Pinus wallichiana GM9*	0.00%
Pinus wallichiana GW21	*Pinus wallichiana MG5*	0.00%
Pinus wallichiana MG5	*Pinus wallichiana GM9*	0.00%
Pinus wallichiana MG5	*Pinus wallichiana GW21*	0.00%
Tsuga dumosa GM3	*Tsuga dumosa GR3*	0.00%
Tsuga dumosa GR3	*Tsuga dumosa GM3*	0.00%
Abies pindrow GM2	*Abies pindrow GW11*	0.12%
Abies pindrow GW11	*Abies pindrow GM2*	0.12%
Abies pindrow GW11	*Abies pindrow GW20*	0.12%
Abies pindrow GW20	*Abies pindrow GW11*	0.12%
Cedrus deodara GM11	*Cedrus deodara GM13*	0.12%
Cedrus deodara GM11	*Cedrus deodara GW13*	0.12%
Cedrus deodara GM13	*Cedrus deodara GM11*	0.12%
Cedrus deodara GW13	*Cedrus deodara GM11*	0.12%
Cupressus torulosa GM14	*Cupressus torulosa MG3*	0.12%
Cupressus torulosa MG3	*Cupressus torulosa GM14*	0.12%
Juniperus indica GM6	*Juniperus indica GR6*	0.12%
Juniperus indica GR6	*Juniperus indica GW22*	0.12%
Juniperus indica GR6	*Juniperus indica MG4*	0.12%
Juniperus indica GW18	*Juniperus indica GW22*	0.12%
Juniperus indica GW18	*Juniperus indica MG4*	0.12%

Juniperus indica GW19	*Juniperus indica GW22*	0.12%
Juniperus indica GW19	*Juniperus indica MG4*	0.12%
Juniperus indica GW22	*Juniperus indica GR6*	0.12%
Juniperus indica GW22	*Juniperus indica GW18*	0.12%
Juniperus indica GW22	*Juniperus indica GW19*	0.12%
Juniperus indica MG4	*Juniperus indica GR6*	0.12%
Juniperus indica MG4	*Juniperus indica GW18*	0.12%
Juniperus indica MG4	*Juniperus indica GW19*	0.12%
Picea smithiana GM1	*Picea smithiana GW15*	0.12%
Picea smithiana GW15	*Picea smithiana GM1*	0.12%
Pinus roxburghii GM10	*Pinus roxburghii GM5*	0.12%
Pinus roxburghii GM5	*Pinus roxburghii GM10*	0.12%
Pinus roxburghii GM5	*Pinus roxburghii GM8*	0.12%
Pinus roxburghii GM5	*Pinus roxburghii GR4*	0.12%
Pinus roxburghii GM8	*Pinus roxburghii GM5*	0.12%
Pinus roxburghii GR4	*Pinus roxburghii GM5*	0.12%
Pinus wallichiana GM4	*Pinus wallichiana GM9*	0.12%
Pinus wallichiana GM4	*Pinus wallichiana GW21*	0.12%
Pinus wallichiana GM4	*Pinus wallichiana MG5*	0.12%
Pinus wallichiana GM9	*Pinus wallichiana GM4*	0.12%
Pinus wallichiana GM9	*Pinus wallichiana GR2*	0.12%
Pinus wallichiana GR2	*Pinus wallichiana GM9*	0.12%
Pinus wallichiana GR2	*Pinus wallichiana GW21*	0.12%
Pinus wallichiana GR2	*Pinus wallichiana MG5*	0.12%
Pinus wallichiana GW21	*Pinus wallichiana GM4*	0.12%
Pinus wallichiana GW21	*Pinus wallichiana GR2*	0.12%
Pinus wallichiana MG5	*Pinus wallichiana GM4*	0.12%
Pinus wallichiana MG5	*Pinus wallichiana GR2*	0.12%
Tsuga dumosa GM3	*Tsuga dumosa GW17*	0.12%
Tsuga dumosa GR3	*Tsuga dumosa GW17*	0.12%
Tsuga dumosa GW17	*Tsuga dumosa GM3*	0.12%
Tsuga dumosa GW17	*Tsuga dumosa GR3*	0.12%
Juniperus indica GR6	*Juniperus indica GW18*	0.23%
Juniperus indica GR6	*Juniperus indica GW19*	0.23%
Juniperus indica GW18	*Juniperus indica GR6*	0.23%
Juniperus indica GW19	*Juniperus indica GR6*	0.23%
Pinus wallichiana GM4	*Pinus wallichiana GR2*	0.23%
Pinus wallichiana GR2	*Pinus wallichiana GM4*	0.23%

Table 4.12 Interspecific variation of *MatK* in terms of pairwise distance in gymnosperms calculated by TaxonDNA-SpeciesIdentifier

Species 1	**Species 2**	**Variation**
Abies pindrow GW11	*Abies spectabilis GR5*	0.00%
Araucaria angustifolia GM18	*Araucaria columnaris GM16*	0.00%
Araucaria bidwilli GM17	*Araucaria angustifolia GM18*	0.00%
Araucaria columnaris GM16	*Araucaria bidwilli GM17*	0.00%
Cupressus sempervirens CG3	*Cupressus macrocarpa CG2*	0.00%
Cupressus torulosa CG5	*Cupressus macrocarpa CG2*	0.00%
Cycas pectinata GL10	*Cycas revoluta GL11*	0.00%
Juniperus horizontalis GM19	*Juniperus recurva GR8*	0.00%
Juniperus indica GW22	*Juniperus horizontalis GM19*	0.00%
Juniperus indica GW22	*Juniperus recurva GR8*	0.00%
Juniperus indica MG4	*Juniperus horizontalis GM19*	0.00%
Juniperus indica MG4	*Juniperus recurva GR8*	0.00%
Juniperus recurva GR8	*Juniperus indica GW22*	0.00%
Juniperus recurva GR8	*Juniperus indica MG4*	0.00%
Taxus contorta GT2	*Taxus media GT4*	0.00%
Abies pindrow GM2	*Abies spectabilis GR5*	0.12%
Abies pindrow GW20	*Abies spectabilis GR5*	0.12%
Abies spectabilis GR5	*Abies pindrow GM2*	0.12%
Abies spectabilis GR5	*Abies pindrow GW20*	0.12%
Juniperus horizontalis GM19	*Juniperus indica GR6*	0.12%
Juniperus horizontalis GM19	*Juniperus indica GW18*	0.12%
Juniperus horizontalis GM19	*Juniperus indica GW19*	0.12%
Juniperus indica GR6	*Juniperus horizontalis GM19*	0.12%
Juniperus indica GR6	*Juniperus recurva GR8*	0.12%
Juniperus indica GR6	*Juniperus squamata GR9*	0.12%
Juniperus indica GW18	*Juniperus horizontalis GM19*	0.12%
Juniperus indica GW18	*Juniperus recurva GR8*	0.12%
Juniperus indica GW19	*Juniperus horizontalis GM19*	0.12%
Juniperus indica GW19	*Juniperus recurva GR8*	0.12%
Juniperus recurva GR8	*Juniperus indica GR6*	0.12%
Juniperus recurva GR8	*Juniperus indica GW18*	0.12%
Juniperus recurva GR8	*Juniperus indica GW19*	0.12%
Juniperus squamata GR9	*Juniperus indica GR6*	0.12%
Pinus wallichiana GM9	*Pinus roxburghii GM10*	0.12%
Pinus wallichiana GM9	*Pinus roxburghii GM8*	0.12%
Pinus wallichiana GM9	*Pinus roxburghii GR4*	0.12%
Pinus wallichiana GW21	*Pinus roxburghii GM10*	0.12%
Pinus wallichiana GW21	*Pinus roxburghii GM8*	0.12%
Pinus wallichiana GW21	*Pinus roxburghii GR4*	0.12%

Pinus wallichiana MG5	*Pinus roxburghii GM10*	0.12%
Pinus wallichiana MG5	*Pinus roxburghii GM8*	0.12%
Pinus wallichiana MG5	*Pinus roxburghii GR4*	0.12%
Taxus mairei GT3	*Taxus contorta GT2*	0.12%
Taxus mairei GT3	*Taxus media GT4*	0.12%
Taxus media GT4	*Taxus mairei GT3*	0.12%
Taxus media GT4	*Taxus wallichiana GT1*	0.12%
Taxus wallichiana GT1	*Taxus contorta GT2*	0.12%
Taxus wallichiana GT1	*Taxus media GT4*	0.12%
Juniperus squamata GR9	*Juniperus horizontalis GM19*	0.23%
Juniperus squamata GR9	*Juniperus indica GW22*	0.23%
Juniperus squamata GR9	*Juniperus indica MG4*	0.23%
Juniperus squamata GR9	*Juniperus recurva GR8*	0.23%
Pinus roxburghii GM5	*Pinus wallichiana GW21*	0.23%
Pinus roxburghii GM5	*Pinus wallichiana MG5*	0.23%
Pinus roxburghii GM8	*Pinus wallichiana GM4*	0.23%
Pinus roxburghii GM8	*Pinus wallichiana GR2*	0.23%
Pinus roxburghii GR4	*Pinus wallichiana GM4*	0.23%
Pinus roxburghii GR4	*Pinus wallichiana GR2*	0.23%
Pinus wallichiana GM4	*Pinus roxburghii GM10*	0.23%
Pinus wallichiana GM4	*Pinus roxburghii GM8*	0.23%
Pinus wallichiana GM4	*Pinus roxburghii GR4*	0.23%
Pinus wallichiana GM9	*Pinus roxburghii GM5*	0.23%
Pinus wallichiana GR2	*Pinus roxburghii GM10*	0.23%
Pinus wallichiana GR2	*Pinus roxburghii GM8*	0.23%
Pinus wallichiana GR2	*Pinus roxburghii GR4*	0.23%
Pinus wallichiana GW21	*Pinus roxburghii GM5*	0.23%
Pinus wallichiana MG5	*Pinus roxburghii GM5*	0.23%
Taxus mairei GT3	*Taxus wallichiana GT1*	0.23%
Taxus wallichiana GT1	*Taxus mairei GT3*	0.23%
Cupressus macrocarpa CG2	*Cupressus torulosa MG3*	0.30%
Cupressus torulosa MG3	*Cupressus macrocarpa CG2*	0.30%
Cupressus sempervirens CG3	*Cupressus torulosa MG3*	0.35%
Cupressus torulosa MG3	*Cupressus sempervirens CG3*	0.35%
Juniperus indica GW18	*Juniperus squamata GR9*	0.35%
Juniperus indica GW19	*Juniperus squamata GR9*	0.35%
Juniperus squamata GR9	*Juniperus indica GW18*	0.35%
Juniperus squamata GR9	*Juniperus indica GW19*	0.35%
Pinus roxburghii GM5	*Pinus wallichiana GM4*	0.35%
Pinus roxburghii GM5	*Pinus wallichiana GR2*	0.35%
Pinus wallichiana GM4	*Pinus roxburghii GM5*	0.35%
Pinus wallichiana GR2	*Pinus roxburghii GM5*	0.35%
Cupressus macrocarpa CG2	*Cupressus torulosa GM14*	0.44%
Cupressus torulosa GM14	*Cupressus macrocarpa CG2*	0.44%

Cupressus sempervirens CG3	*Cupressus torulosa GM14*	0.47%
Cupressus torulosa GM14	*Cupressus sempervirens CG3*	0.47%
Ephedra gerardiana GR1	*Ephedra intermedia MG2*	1.05%
Ephedra gerardiana MG7	*Ephedra intermedia MG1*	1.05%
Ephedra intermedia MG1	*Ephedra gerardiana GR1*	1.05%
Ephedra intermedia MG2	*Ephedra gerardiana MG7*	1.05%

4.13 Phylogentic inference: Cladogram

Phylogenetic tree was constructed using four markers viz. *ITS*, *MatK*, *ycf1* and *petN-psbM* (Fig 4.16, Fig 4.17, Fig 4.18 and Fig 4.19 respectively) with three combinations between the markers *MatK* + *ITS*, *MatK* + *ycf1* and *MatK* + *petN-psbM* (Fig 4.20, Fig 4.21 and Fig 4.22 respectively).Altogether, 30 individuals with *ITS* markers were taken for tree construction, followed by 54 individuals with *MatK* marker, 21 individuals with *ycf1* marker and 17 individuals with *petN-psbM* marker. Generally, species generated separate clusters (monophyletic clades) in the tree with the bootstrap value > 50 % were considered to be discriminated successfully.

Though bootstrap value for *ITS* marker-based Neighbor-joining tree was less than 50 % in some individuals, the tree topology was matched with the global tree. The average bootstrap value for *ITS* marker tree was 70.37 %. For *MatK* tree only 5 out of 54 individuals have bootstrap value less than 50 % but the overall tree topology was similar with universal tree and has shown higher frequency absolute species discrimination than *ITS* with 11.11 % individauls showing 100 % bootstrap values. Similarly, for *ycf1* marker-based tree 33.33 % individuals have shown absolute species discrimination with no occurrence of bootstrap value less than 50 %. Similarly, for *petN-psbM* marker-based tree 35.29% individuals have shown absolute species discrimination with 3 individuals showing bootstrap value less than 50 %.

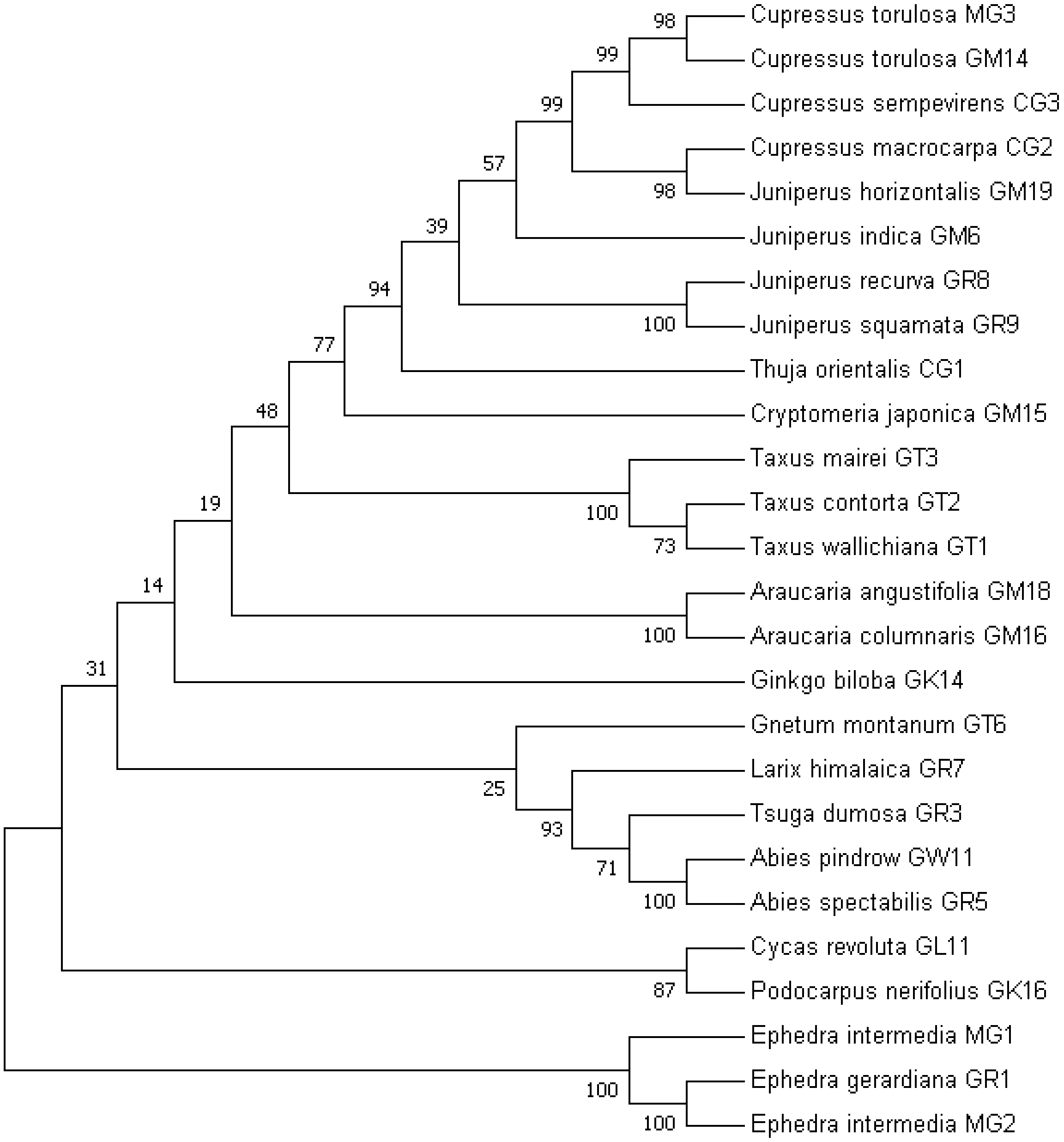

Figure 4.16 A Neighbor-Joining tree on the basis of *ITS* region showing relationships between 30 individuals of gymnosperms.

The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 4.43636212 is shown. The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The evolutionary distances were computed using the p-distance method and are in the units of the number of base differences per site. The analysis involved 26 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 173 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.

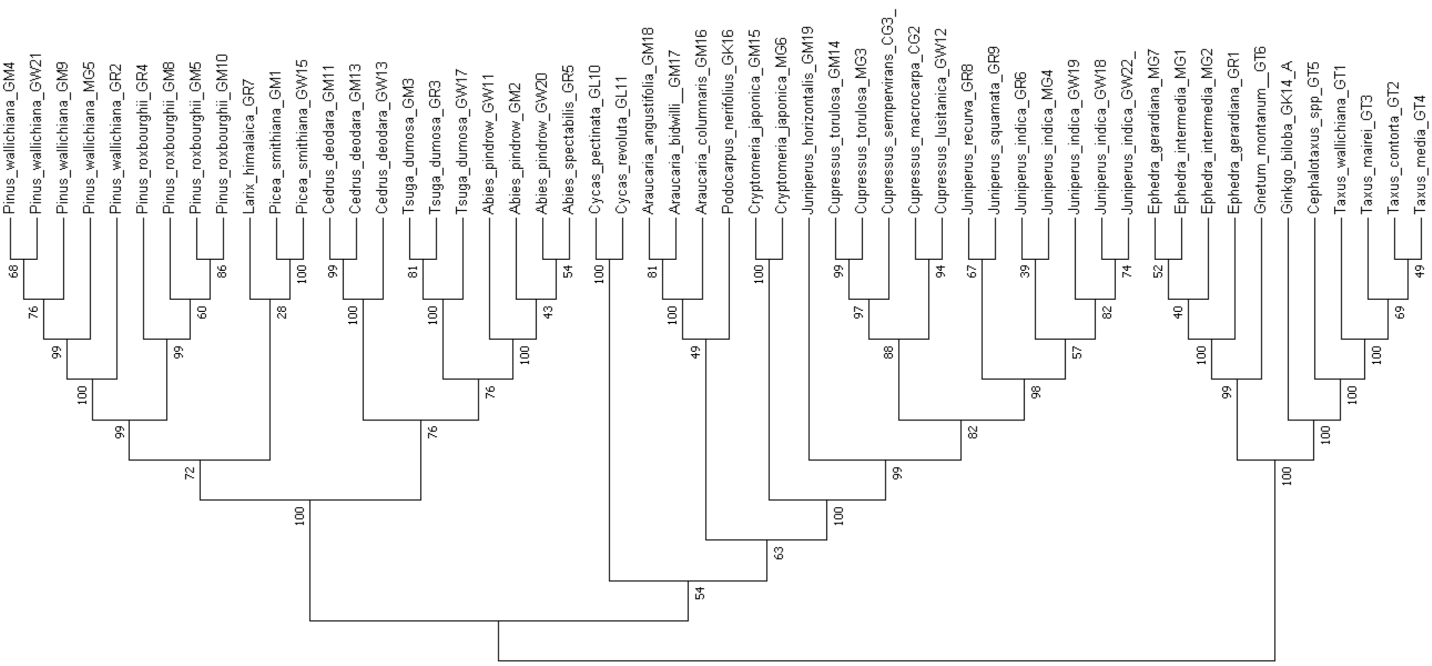

Figure 4.17 A Neighbour- joining tree based on *MatK* barcode showing relationships between 54 individuals of Nepalese gymnosperms

The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 2.08210140 is shown. The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The evolutionary distances were computed using the p-distance method and are in the units of the number of base differences per site. The analysis involved 54 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 571 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.

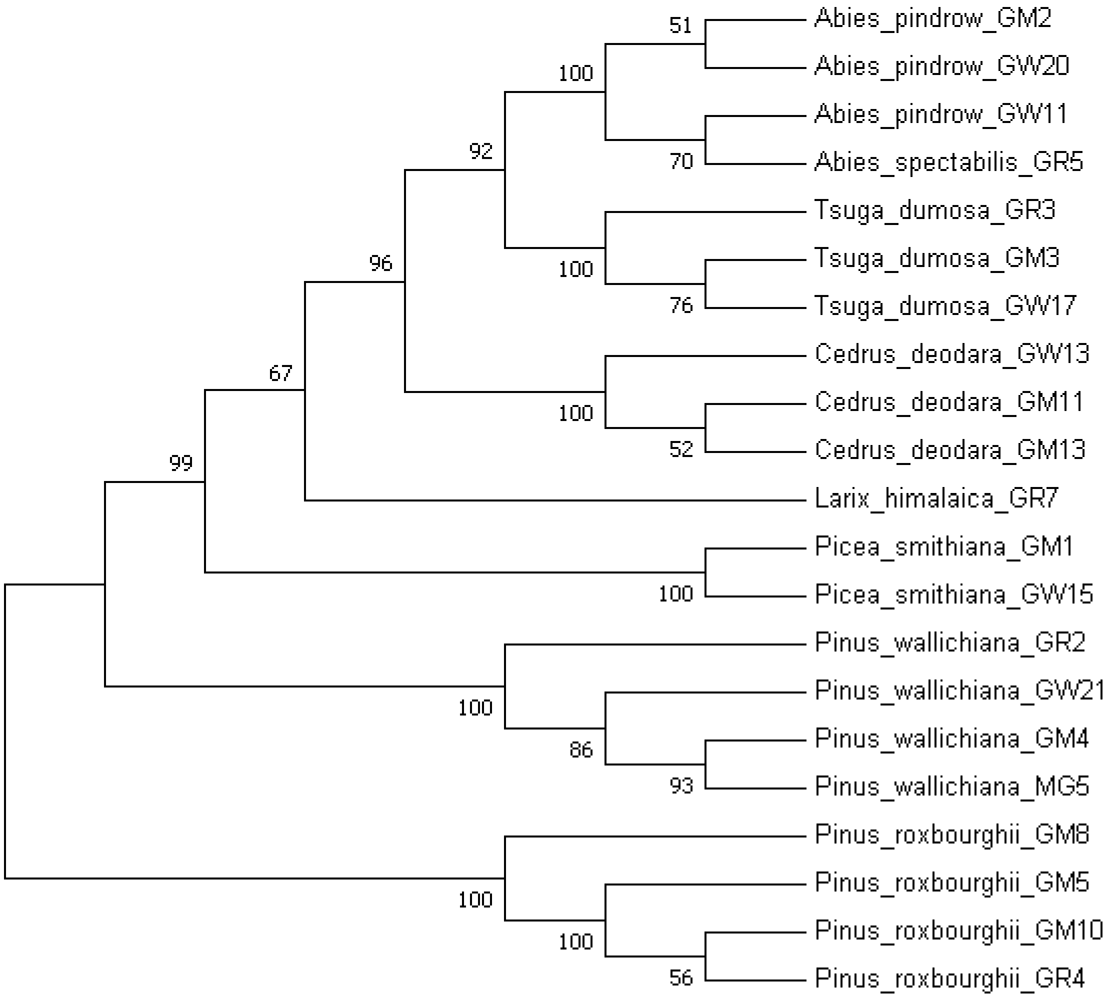

Figure 4.18 A Neighbor-Joining tree based on *ycf1* barcode showing relationships between 21 individuals in Pinaceae family of gymnosperms

The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.52289838 is shown. The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The evolutionary distances were computed using the p-distance method and are in the units of the number of base differences per site. The analysis involved 21 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 826 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.

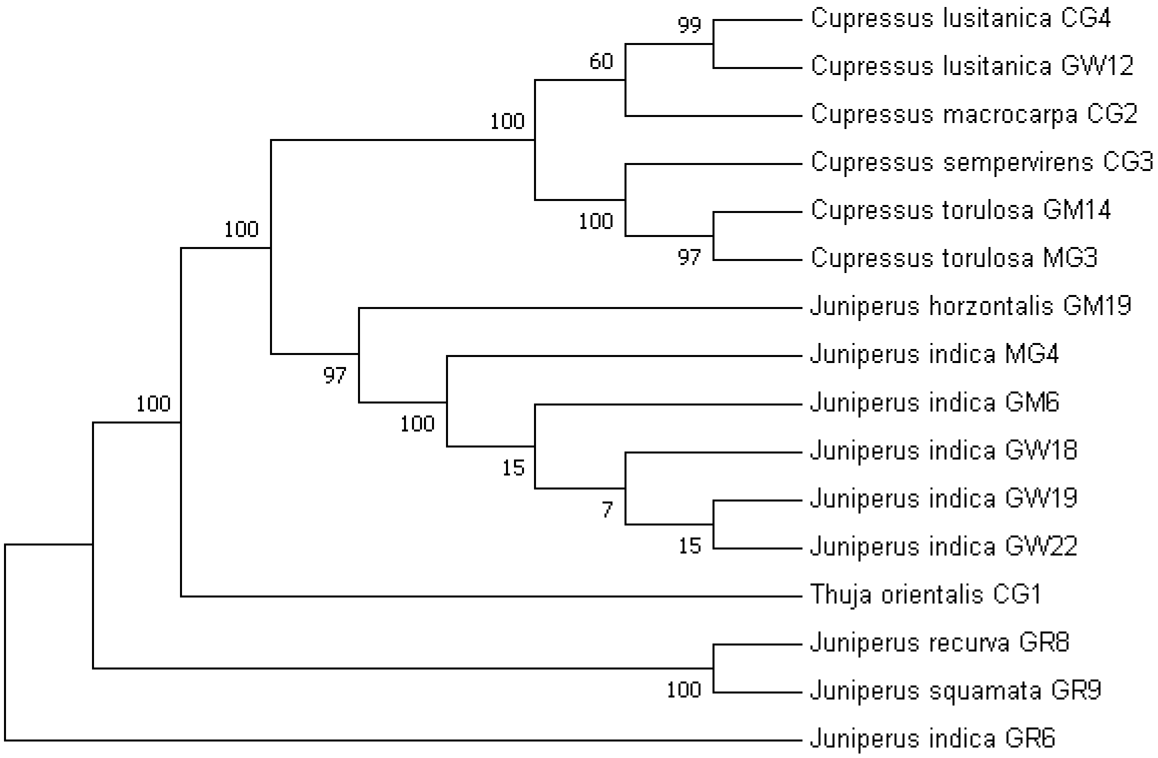

Figure 4.19 A Neighbor-Joining tree based on petN5-psbM region showing relationships between 17 individuals in Cupressaceae family of gymnosperms

The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 1.05636867 is shown. The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The evolutionary distances were computed using the p-distance method and are in the units of the number of base differences per site. The analysis involved 16 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 474 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.

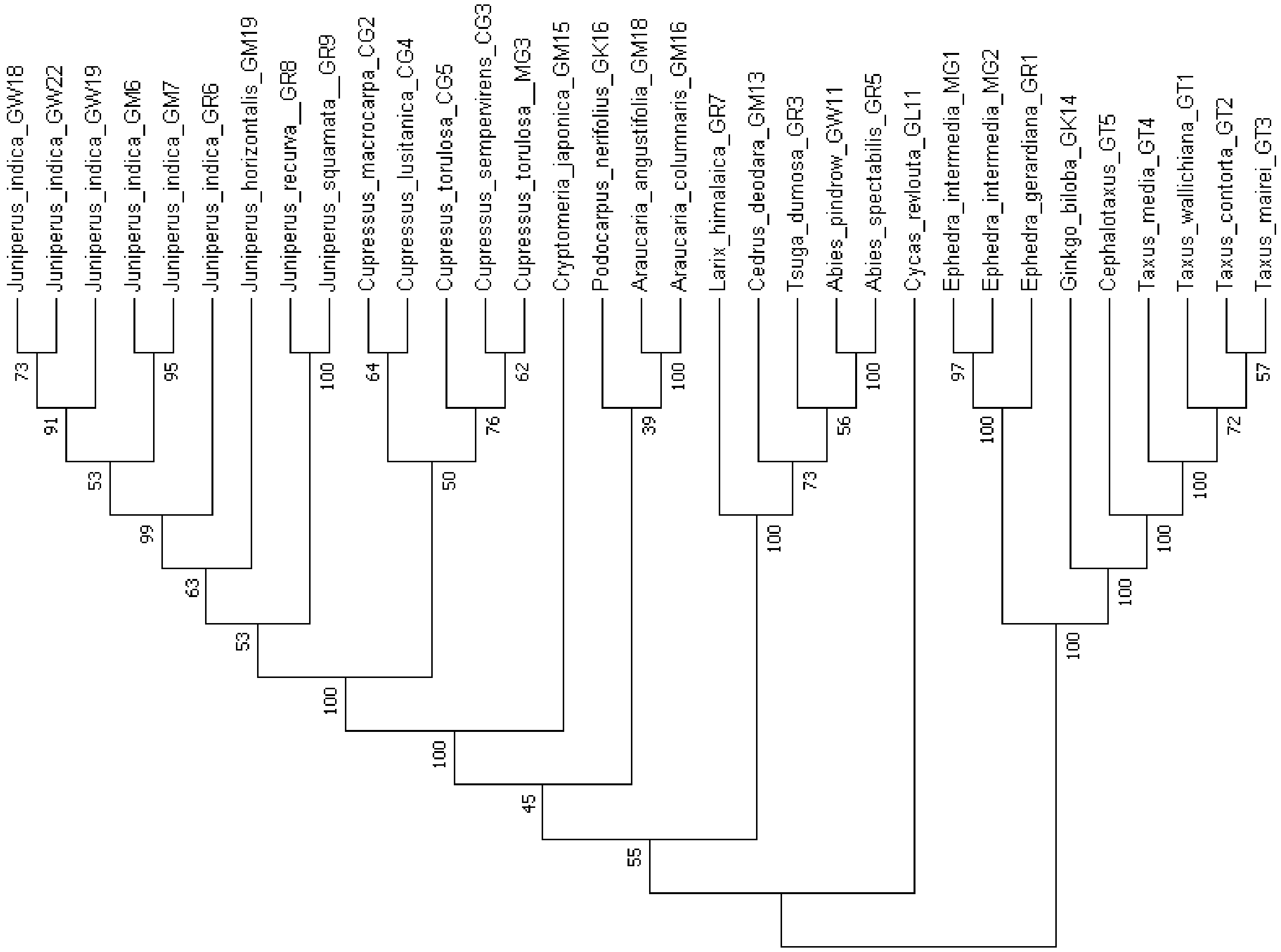

Figure 4.20 A Neighbourhood-Joining tree of *ITS +MatK* showing relationships between 33 individuals of gymnosperms

The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 1.72254148 is shown. The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The evolutionary distances were computed using the p-distance method and are in the unitsof the number of base differences per site. The analysis involved 33 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 602 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.

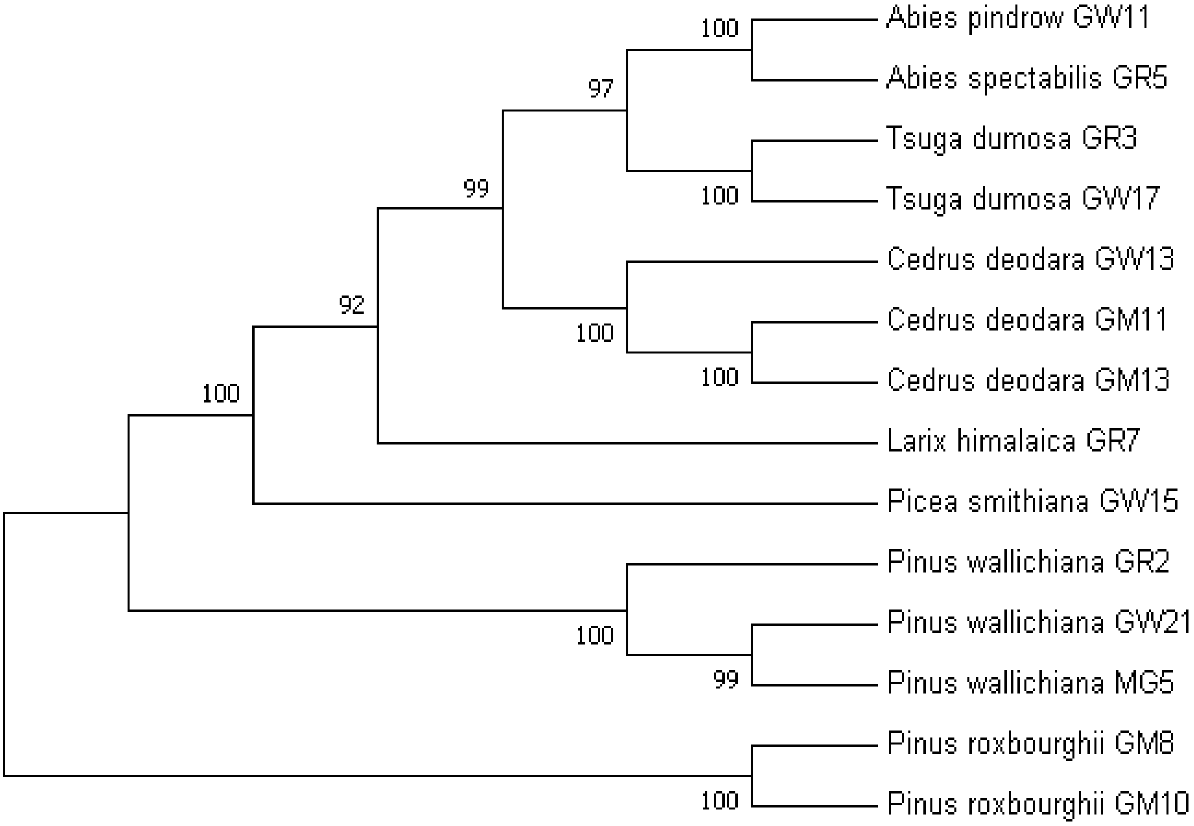

Figure 4.21 A Neighbourhood-Joining tree of *ycf1 +MatK* showing relationships between 14 individuals in Pinacae family of gymnosperms

The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.38084529 is shown. The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The evolutionary distances were computed using the p-distance method and are in the units of the number of base differences per site. The analysis involved 14 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 1664 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.

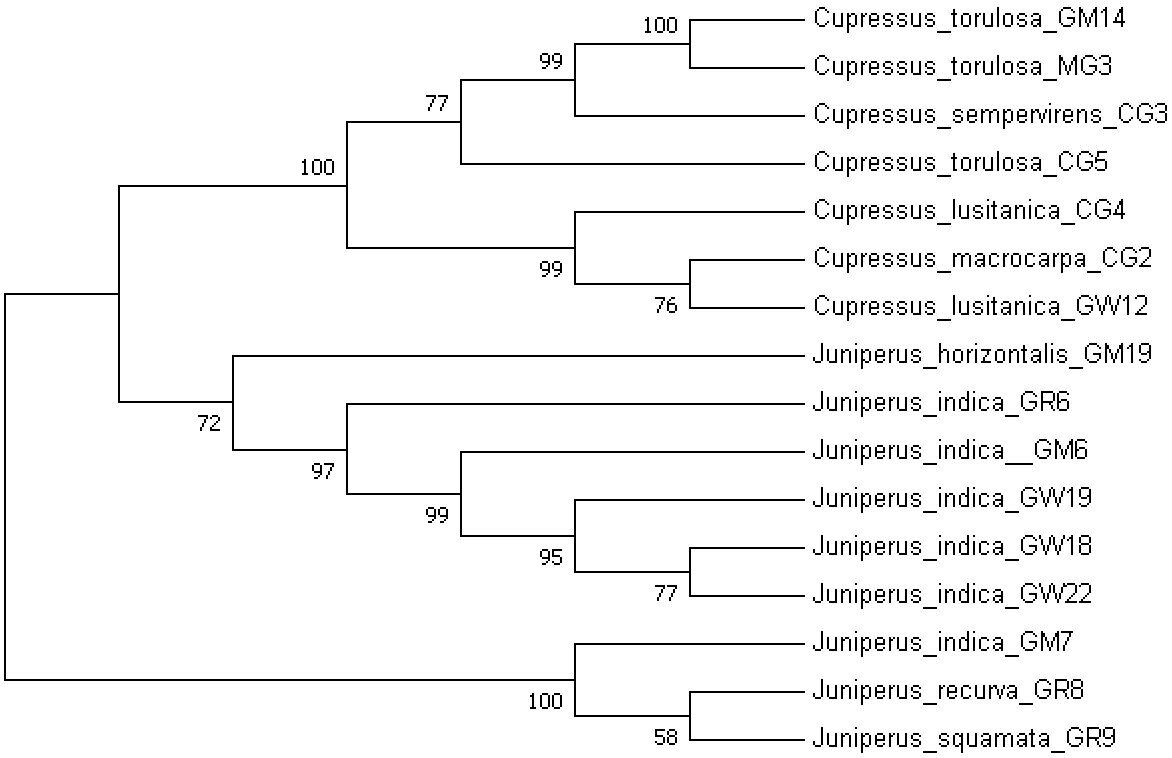

Figure 4.22 A Neighbor-Joining tree of *petN-psbM +MatK* showing relationships between 16 individuals in Cupressaceae family of gymnosperms.

The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.37506665 is shown. The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The evolutionary distances were computed using the p-distance method and are in the units of the number of base differences per site. The analysis involved 16 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 1348 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.

4.14 Species discrimination potential of markers

One nuclear marker (*ITS* region) and three other cpDNA region *MatK*, *ycf1* and *petN-psbM* region were used to understand whether they were strong enough to segregate the studied sample of gymnosperms upto the species level (Figure 4.23). Among the cpDNA markers *MatK* was gymnosperms specific, *ycf1* was Pinaceae family specific and *petN-psbM* was Cupressaceae family specific. The markers alone along with their combined effort were analysed. For the combination, the sequences were simply added in the text file (notepad file) without any space/deletion and then observed for changes in the analytical approaches like BLAST, pairwise-distance and Neighbor-Joining tree topology.

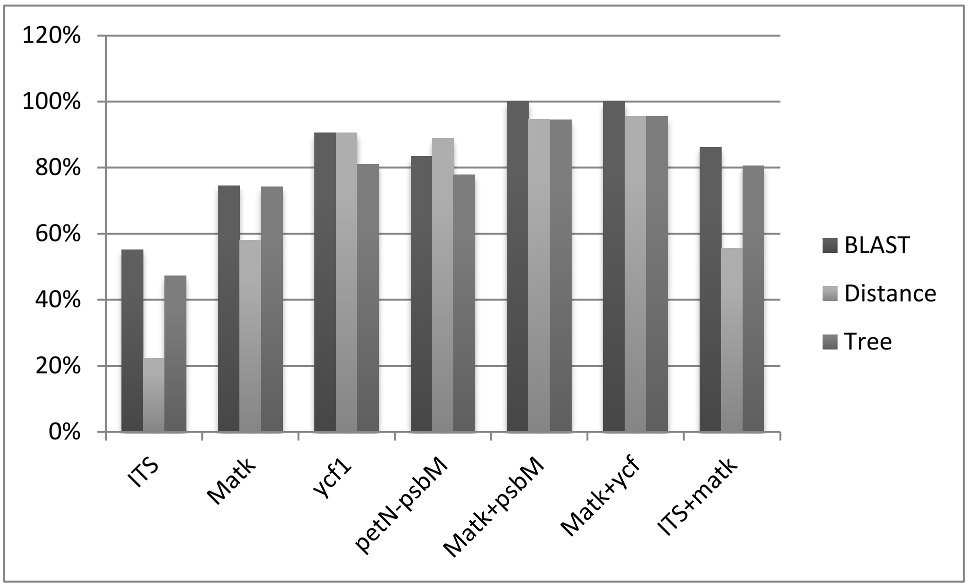

Figure 4.23 Species discrimination attempt of 3 different markers and their possible combinations. *Ycf1* marker is Pinaceae-oriented and *petN-psbM* is Cupressacae-oriented

MatK had the highest species resolving power of average 72.33 %. Though *petN-psbM* and *ycf1* both have stronger differentiation capacity than *MatK*, they were family specific and thus are used only to aid the identification rather preceeding as the strongest marker by themselves. Considering *MatK* as the strongest and primary marker, *ITS* increased the *MatK* discrimination by approximately 10%. With *petN-psbM*, 24 % discrimination power of *MatK* was increased and *ycf1* contributed 31.3 % increment in resolving the species upto the species level.

CHAPTER 5. **DISCUSSION**

5.1 Status of conservation

According to the report of IUCN in 2015, 34% of conifers are in Red List which means that they are threatened in state and on their way to extinction, hence, requiring core attention (IUCN Red List, 2015) in their conservation. Among those conifers, 32.6 % of *Abies*, 70.6% of *Araucaria*, 46.7% of *Cupressus* , 24.5% of *Juniperus* , 41.7% of *Picea,* 15.9% of *Pinus*, 37.1% of *Podocarpus,* 62.5% of *Taxus* and 83.3% of *Amentotaxus* are threatened with vulnerable, endangered and critically endangered status (Neale and Wheeler, 2019). In Nepal, the list of gymnosperms presented by the experts earlier suggests the total of 24 wild species (KATH, Godawari; DPR report, GON) and also additional cultivated species besides those possible to collect in this study. Also, some cultivated species, which were never reported before were supported well by morphology and DNA barcode to be possible new cultivated species in Nepal. *Araucaria angustifolia, Cupressus lusitanica, Cupressus macrocarpa, C. sempervirens, C. tonkinensis* are the species that are reported as new cultivated species.

5.2 Identification challenges

Expert identification was carried out based on intensive study on morphology of gymnosperms of particular family, genus and species. DNA barcoding was taken not to replace the phenotype-based identification but as an accessory scientific evidence to validate the identification. But an expert examination fails whenever a species has no diagnostic characters and key features that can be used to differentiate its name, position and its kind in scientific classification and this is where DNA barcoding is needed (Neigel *et al.*, 2007).

There are examples of different combination of barcodes giving no resolution to the gymnosperms' barcoding and species identification (Armenise *et al.*, 2012). Instead of relying on a single approach, a more natural way to identify and name a plant species would be the combination of morphological (reproductive structure, leaves, phylotaxy, wood structure), anatomical, biochemical and DNA traits to get a correctly identified species that, as a result, comes with sufficient ecological and evolutionary history (Neale and Wheeler, 2019).

5.3 Sampling and preservation

The attempt of discovering possible fields, altitude and habitat of each species was carried out. Each species has a range of altitude and the diagrammatic representation of altitude represents only the altitude of first found species and not the duplicates. The range of altitude for particular species is listed in the appendix.

The preservation using silica gel is chosen and found satisfactory to function as a good preserving chemical. As experimented, the herbarium sample of *Abies densa* and *Larix griffithiana*are concluded not tohave good quality and quantity of DNA products as preserved in silica gel that's why it is assumed that DNA anlaysis from the herbarium sample was not successful (Staats *et al.*, 2011, Ribeiro & Lovato, 2007). As fungal contamination is seen in both of the herbarium samples, it might be due to the chance of endophytic fungal DNA contamination (Hollingsworth *et al.*, 2011).

5.4 DNA extraction and PCR

DNA extraction of gymnosperms was not as easy for all the collected samples. With CTAB extraction method, some samples have shown significant amount of phenolics and colored compound in final DNA diluent solution. Three different approaches of DNA extraction were performed based on availability of Kit and reagents, nature of sample and time constraints. Some samples extracted from CTAB method required further purification passing through DNA-binding Silica Spin column for better PCR results. As the target PCR band was obtained and the product was further purified, there was no significant difference in the sequence quality of products given by DNA that were extracted using three different methods. Rather, the comparision could be made as per the markers choice.

DNA extraction success rate for the gymnosperms studied was 96.61 % because of failure in DNA extraction from two herbarium samples namely *Abies densa* and *Larix griffithiana*. The PCR efficiency of *ITS* marker was 80% and that of *MatK*, *ycf1* and *PetN-psbM* regions, on the other hand, was absolute (100%) each. Sequencing efficiency was absolute only for *ycf1* marker (100%) with*MatK* marker being second highest in sequencing efficiency with 96.61 % success, followed by *petN-psbM* marker 94.2 %. *ITS* marker has the lowest sequencing efficiency of 76.59 %.

The PCR success of amplification of *ycf1* gene with the primers used in this study (*ycf1*g) is studied to be 90.91% as stated in Dong *et al.*, 2015 whereas our study reveals the PCR success of absolute percentage but considering the samples taken only from the Pinaceae family. *MatK* and *petN-psbM* markers both have shownabsolute PCR efficiency.

Herbarium samples of *Abies densa* and *Larix griffithiana* were brought from National Herbarium Laboratory Godawari and DNA extraction was carried out but no successful PCR of the targeted gene could occur which was realized only after sequencing. Sequencing provided DNA of fungus not the gymnosperms. The sample might have been degraded rapidly as were they not preserved in silica gel or could not be enough in concentration to act as template DNA for PCR. The sequencing success rate of herbarium samples for the chloroplast marker *MatK* reported to be only 40-50% as compared to 90% success rate in case of fresh samples caused due to aging and improper preservation creating DNA fragmentation and quality issues leading difficulties in DNA analysis (Korpelainen & Pietiläinen, 2019). With the species identification power 65.09% ± 6 % with no PCR failure, *MatK* is studied to be very strong barcode in a recent study (Wu *et al.*, 2019).

5.5 Sequence analysis and interpretation

Not all sequences of identified species matched exactly with that of databases available. The possible failure of matching a barcode sequence of an unidentified specimen with any database sequence reflects either the absence of sequences from the relevant species in the database or the absence of a particular haplotype (a set of genetic determinant on a single chromosome) for that species (Neigel *et al.*, 2007). A mismatch scenario may be due to an extreme case where a sequence sampled from a sub-population that has no ongoing gene-flow with anyof the populations from which the database sequenceshave sampled and thus does not show any similarity to be considered as correct match (Nielsen & Matz, 2006).The error of a query sequence not matching the subject sequence is intrepretable. The database sequence to which the query is observed to be matched may not always represent true species and the query, gone through some genetic modification, might not be observed as matched with the true sequence of the species to which it actually belongs (Nielsen & Matz, 2006).

The PCR and sequencing difficulties were found tremendously prevalent while working with *ITS* regions. The non-specific bands, smearing, and lack of full length sequencing with ambivalent chromatogram lead to unreliable analysis of *ITS* results. However, only the mapping of Neighbor-Joining tree with the global one is considered a good achievement whereas observations of abundant variable sites were not defended by the experimental scenario behind it. This kind of failure is well-explained by some other literatures too. The presence of paralogous and pseudogenes definitely creates a problem in amplification and sequencing (Gao et. al, 2016, Hollingsworth *et al.*, 2011). The amplification and sequencing problems of whole *ITS* region can be compensated by targeting only *ITS*2 region but has also has some drawbacks as in some plants cases *ITS*2 is unable to provide resolution within the species (Li *et al.*, 2014). The use of *ITS*2 marker is beneficial as it delivers easy post PCR and sequencing procedures. Further, *ITS*2 region is more length conserved than the whole *ITS* region (Hollingsworth *et al.*, 2011). *ITS*2 in this study, however, could not be extracted from whole *ITS* sequences for all of the samples that’s why the overall analysis based on *ITS*2 could not be performed which could have aided to the validity of *ITS* analysis.

ITS region with regard to aforementioned references is a strong barcode and has informative variable sites. But, variable reasons in the sequences that we obtained might have been increased due to quality of sequence generated after troublesome PCR. The same size of PCR bands obtained might have contained mixture of different composition of DNA bands but of same size of which the segregation is so difficult (Gao et. al, 2016).

5.5.1 Variation in length and informative sites

The analysis of *ITS* sequences was tedius as fine sequences of every sample were not obtained but the average length matched with the global finding. The maximum length of amplicon amplified by using *ITS* p5/u4 primer is proposed to be 1134 bp to 2914 bp(Cheng *et al.*, 2016). *ITS* length of 3 individuals were reported to be less than 1000 bp, 17 individuals to be 1000 bp to 1200 bp, 4 individuals to be 1200 bp -1500 bp, 4 individuals to be 1600 bp to 1800 bp.The length of *ITS* region in gymnosperms varies from that of angiosperms and pteridophytes (usually larger) and ranges from size 610 bp to 3100 bp (Maggini *et al.*, 1998). This study shows minimum length of 988 bp to 1688 bp.

The average size of *MatK* region amplified by using Gym_F1/R1 primer is 838. The purposed length of amplicon using this particular primer is 832 (Li *et al.*, 2011).Altogether, 14 species were found with *MatK* amplicon between 811bp and 832 bp, 18 species with MatK amplicon between 834 bp and 863 bp, 1 species of 623 bp and 1 species (*Ephedra intermedia)* of 1142 bp. Length 623 bp of *Cupressus torulosa*is explained as readable length while the overall length was around the normal length. Enough misleading based appeared showing ambivalent chromatogram which could not be interpreted. Although species of *Ephedra* involved different sets of primers, since both the primers were designed from similar positions, it does make the alignment possible with rest of the sequences(Li *et al.*, 2011).

The length of *ycf1* in pinaceae ranges from 1097 bp to 1390 bp as seen in our data of amplicon amplified with primers *ycf1*g F/R which follows the recommendation of previous designed primers from Dong *et al.*, 2015 . The length of *petN-psbM* region ranged from 593 bp to 865 bp (Adams & Schwarzbach, 2013, Adams & Chaudhary, 2009).

The alignment coverage of *ycf1* gene in a study to reveal *Pinus* phylogeny had shown enough variation with 74.9 % informative sites leaving all other markers behind in variation in previous study(Hernández-León *et al.*, 2013) which is contradicted as the parsimony informative sites in *ycf1* in Pinaceae family were found to be 29.7%. The *ycf1* gene has enough discriminatory power to exhibit species resolution 71.87% by showing that *MatK* came with increased resolution and discriminatory rate when combined with *ycf1* barcode (Dong *et al.*, 2015). Here, in the study, ycf was used only for Pinaceae family to aid the better resolution because it has been said that *ycf1* works better for this particular family. *MatK* and *petN-psbM* region have shown respectively 70.4 % and 72.4 % variable sites with 59.7 % and 47.7 % parsimony informative sites.

5.6 Utility of the markers

MatK region could be enough to separate almost all of the gymnosperms in the study 59 indiviudals of 34 species except species of *Abies (*Figure 4.17 and Figure 4.23). The discrimination power of *MatK* between species of particular genus like *Pinus* has opened dimensions for the possible use of this marker in forensics and disguised sample. *MatK* only is enough to discriminate most species of gymnosperms under study expcept minimal variation in *Abies* and Junipers as clear clades of species are shown in phylogenetic tree in Figure 4.17 and they have significant interspecific variation (Figure 4.9). *A. spectabilis* hybridizes freely with *A. pindrow* forming intermediate populations in the altitudinal middle zone of their common distribution" showing their close relatedness (Vidakovic 1991). A recent study on DNA barcoding to identify the closely related species of *Pinus* has found the none of the chloroplast markers (*MatK*, rbcL,*ycf1*,*trnH-psbA, trnL-trnF*) effective enough to discriminate closely related *Pinus* (Celiński *et al.*, 2017) but in our study *MatK* has clearly distinguished two close species of *Pinus*; *P. roxburghii* and *P. wallichiana*.Excellent power of *MatK* is shown in case of *Pinus*showing significant interspecific variation observed in MatK sequence as shown in our results (Figure 4.17 and Table 4.11). Chloroplast genome only (*MatK* and rbcL) could stand as a basis for phylogenetic analysis within the *Pinus* family (Gernandt *et al.*, 2005).

The discrimination power of *MatK* was increased with *ycf1* by 31.3 %, with *ITS* by 10% and with *petN-psbM* by 24 % as observed in our study. With the 100 % PCR success rate, a recent DNA barcoding report shows the barcoding efficiency and species resolution of *ITS* marker 66.48 % ± 5.97 % and species identification power increased to 99.88% ± 8.62 % with the addition of *ITS* marker information (Wu *et al.*, 2019). The *ITS* marker was proved a strong marker in case of gymnosperms too. Incase where chloroplast marker could not give much variation in some species, *ITS* had a significant role in validating the morphological identification. However, our study could not give empirical value for accuracy of *ITS* region although it has shown clearly the genus and species level discrimination.

MatK was found to be strongest marker which if aided by *ycf1* (for Pinaceae) and *petN-psbM* (for Cupressacae) can increase the resolution of identification. *ITS* data was supportive enough but not 100% satisfactory for the data validation may be due to sequencing errors.

5.7 Phylogeny of gymnosperms

The overall glance of the cladogram of this study matches properly with universal tree of global gymnonsperms. Several studies assure and review the identification approach of present day gymnosperms.

The global tree of gymnosperms shows that species of araucariaceae are linked closely with Podocarpaceae more than to Pinaceae (Saina *et al.*, 2019) as seen in our study too. *Cupresssus* forms a monophyletic clade with Junipers. The group of Cupressaceae is sister to the next family Taxodiaceae This clade of two family is then distant to clade of closely related family Taxaceae and Cupressaceae (Neale and Wheeler, 2019) whereas the *MatK* tree clearly separates Cupressaceae and Taxaceae and proves their relation to be paraphyletic arising from different common branch. The species of Taxaceae and *Ephedra-Gnetum* clade are monophyletic according to the tree given by *MatK* sequences showing 100% clade strength. Pinaceae forms closer clade-relationship with Taxodiaceae-Cupressacae than with Cycadaceae and Taxacae-Ephedracae (Gernandt & Liston, 1999, Liston *et al.*, 1999, Willyard *et al.*, 2007).

Elaborating *MatK* tree as an example (Figure 4.17), each node give rise to two sister taxa which by it means the taxa are of same age and none of two is basal or ancestor. The terminal taxon, one at the end of a cladogram or any kind of phylogenetic tree, is never considered to be a basal species (descendent). It doesn't mean rests of the species are evolved after advancement of species of *Taxus* in the tree. It just shows the existing inclination and relationship between existing species and clades.It is to be understood the way that the species is far away from its ancestor group than/is more close to the sister group (Zachos, 2015).

5.7.1 *Ginkgo-Taxus* family and Gnetales

Single species belonging to Ginkgoacae, *Ginkgo biloba* has shown monophyletic relationship with Cephalotaxacae-Taxacae clade in *MatK* tree Figure 4.17. This *Ginkgo-Taxus* clade then shows most proximate properties with Gnetales (Gnetaceae-Ephedracae). Previous study of Biswas and Johri,1997 and Ran *et al.*, 2010links Ginkgo near to Conifers than to Cycads which doesn't coincide with our finding. But, *ITS*tree shows *Ginkgo biloba* close to Taxaceae-Araucariaceae clade Figure 4.16. A conclusion can be drawn that Ginkgo is close clade and monophyletic to Taxaceae family.

The only species of *Gnetum* in Nepal, *G. montanum* is studied with the knowledge of *MatK*facts to be closely related to Ephedraceae family (Rydin, Pedersen, and Friis, 2004) which coincides with results on this study too Figure 4.17. The *ITS* tree on the other hand showed the clade relation of *Gnetum* more close to Pinaceae rather than Ephedraceae Figure 4.16 and comparing the relative branch strength of phylogentic tree, the former marker inference (*MatK*) was considered stronger.

Gnetum and two species of *Ephedra* are found to form a single clade as per study on the basis of *MatK* phylogeny. The study to allocate the lineage of extant gymnosperms carried out in 1999 has separated the family Gnetidae from angiosperms and listed it with gymnosperms analyzing the mitochondrial, plastid and 18S rDNA sequences (Bowe *et al.*, 2000).Phylogeny and evolutionary study of *Ephedra* is an interesting subject among the study of gymnosperms as species of *Ephedra* are most closely related and very similar species difficult to understand their inter-specific resolution (Faried *et al.*, 2018). New insight into the history of a controversial gymnosperm *Gnetum* (Gnetales) approached by the study of nuclear and chloroplast DNA regions has stated that *Gnetum montanum* is more closely related to other Gnetales such as *Ephedra* species than other conifers and cycadales (Hou *et al.*, 2014).

Gnetum debate : (Bowe *et al.*, 2000) has concluded the monophyletic origin of gymnosperms and figured the proximity of Gnetales to gymnosperms instead of angiosperms. *ITS* tree has brough Gnetum close to Pinaceae family agreeing with (Chaw *et al.*, 2000) while *MatK* tree has clearly shown itsnearness towards the Ephedraceae species. This study finds *MatK* cladogram as more reliable comparing the cleaness and accuracy of *MatK* sequences validation as *ITS* involved tedius sequence editing. Ephedraceae species and Gnetum are monophyletic meaning forms a group that has splitted from a common branch.

Being morphologically different from other gymnosperms, the phylogenetic placement of the Gnetales (*Gentum, Ephedra*) is unclear. Gnetales, the most debated group in seed phylogeny, is understood to be linked with conifers as gnetifers where it is linked with Cupressaceae as gnecup relation (gnecup hypothesis), with pines as gnepine relation (gnepine hypothesis) (Goremykin *et al.*, 2011) each of which are discussed differently by different studies few of which co-incide most of them contradict. One clear finding is that *Gnetum* is observed more closely related to conifers and distant to angiosperms (Farjon, 2007) and *Gneutm* is put in family Gnetaceae, close to the other family Ephedraceae (Govil, 2007).

5.7.2 Junipers

On the basis of study of phylogentic tree, three species of Junipers, *Juniperus squamata* and *Juniperus recurva* are found very closely related and develop sisterhood to establish cladistic relation with *Juniperus indica.* Among three species of Junipers, it can be said on the basis of *MatK* based phylogeny that *J. recurva* and *J. squamata* are more undifferentiated in terms of discrimination which was also observed during sample collection(Adams & Schwarzbach, 2013 ,Adams & Chaudhary, 2009). *Cupressus* species are monophyletic with Junipers forming a close clade. This group of Cupressaceae is sister to the next family Taxodiaceae with single genus *Cryptomeria japonica* as family member under study showing 100% clade discrimination. *Cupressus torulosa* has closely related to *Juniperus communis* and than to *Juniperus horizontalis* than to *Juniperus indica, J. recurva* and *J. squamata* (Mao *et al.*, 2010). Among the species studied, *Thujaoreintalis* is most distant one.

5.7.3 Pines and Cycadales

Cycas pectinata and *C. revoluta* form a distinct clade (Wei *et al.*, 2015), (Xiao & Möller, 2015) showing close relationship with Pinacae family as shown by MatK tree presented in this study. The conceptual view based on the several comparision and studies carried out by International Association of Plant Taxonomy has brought the coniferales (Pines) and Cycadales (Cycas group) together to be in origin (Beck, 2013).

Within Pinaceae *Abies*, *Tsuga* and *Cedrus* (78 %) forms a clade distinctly unique than the *Pinus* and *Picea-Larix* clade (65 %). *Abies-Tsuga-Cedrus* and *Pinus-Picea* clade shows 100% discrimination.Both MatK and ycf1 markers support this universality of pinaceae tree with ycf1 marker adding power to the strength of bootstrap tree. *Abies i*s understood to be more close sister-clade of *Cedrus* and *Tsuga* whereas *Picea* is more closely related to *Larix* and *Pinus*, all of these species being in a very close relation to each other within the Pinaceae on the basis of chloroplast DNA analysis by (Lin *et al.*, 2010). The phylogeny of gymnosperms has been studied to find the relationships between four genera *Larix, Pinus* and *Tsuga*.

The phylogeny of species of *Pinus* is illustrated clearly single marker *MatK* and its combination with *ycf1* marker as *ITS* from *Pinus* species could not be identified. Figure 4.17, Figure 4.18 and Figure 4.21 show that species of *Pinus* are sister clade for sure and the next close members to this genus are *Picea* and *Larix* which form closest sister to *Pinus* than other species in the family. *Pinus roxburghii* and *Pinus wallichiana* have been associated closely (Wang *et al.*, 1999). *Abies pindrow, Abies densa* and *Abies spectabilis* are reported to be related sister species comprising a group with species separated from close common point with respect to chloroplast sequences (Xiang *et al.*, 2015).

For the species of Pinaceae, *MatK* marker is enough to show distinction between the species except for the species of genus *Abies* which shows no differences between *A. pindrow* and *A. spectabilis.* The distinction between the Pinus family was strongly supported and aided by the marker *ycf1* showing strong species identificiation success increased by 31.3 %.

CHAPTER 6. **CONCLUSION**

This study helped to know whether the used barcode markers were enough to discriminate the gymnosperms found in Nepal at different levels: family, genus and species. Both universal markers (*ITS* and *MatK*) were able to discriminate all the samples at genus level and most of the samples were identified upto species level. Family specific markers (*ycf1* and *petN-psbM*) aided to discrimination of closely related species within same genus in Pinaceae and Cupressacae respectively. Except the species of *Abies*, all other species identified morphologically in the field were matched with the DNA based information. But species of *Abies* are rich in key morphological variations to be distinguished by phenotypic method.

ITS, although was considered strong by the literature, in our study the absolute strength of *ITS* could not be observed due to sequencing limitations and ambiguity in sequences. *MatK* was concluded efficient to discriminate species of gymnosperms in species level as there was significant difference between intraspecific and interspecific p-distance between the studied species. Further, phylogenetic tree given by *MatK* was the strongest and matched with maximum identity to global cladogram of gymnosperms. Discriminating power of *MatK* was increased with *ITS* by 10% in gymnosperms, with *Ycf1* by 31.3 % in Pinaceae, and with *petN-psbM* by 24 % in Cupressaceae as observed in our study. Species of *Abies* could not be distinguished by *MatK* showing no significant difference in the value of interspecific and intraspecific variation. Cladogram was constructed with Neighbor-Joining method of phylogenetic tree construction based on *ITS, MatK, ycf1* (Pinaceae)*, petN-psbM* (Cupressaceae) and the combinations *ITS+MatK, ycf1+MatK* (Pinaceae)*, petN-psbM+MatK* (Cupressaceae). All of the trees generated could match the universal inference of global phylogeny based on existing clade-relationship except for position of three species (two species of *Ephedra*, and *Gnetum biloba*) *in ITS*-based tree. The position of species in *ycf1* tree and *petN-psbM* tree were strengthened as compared to that in *MatK* tree only showing increased percentage branch strength in Pinceae and Cupressaceae respectively.

Standing on the fact of simple morphometric differences and the values of genetic distances based on limited markers, evidences are not enough to describe and split the species of gymnosperms (Zachos, 2015). Our present study also matches with this conclusion. The recommendation on this basis for proper identification is to use gymnosperms specific *ITS* primers. This study used plants specific *ITS* primer and it is assumed that the use of gymnosperms-specific primers could give the better results. The multiple sets of sample (wide-range sampling) could also increase the reliability of the result. With review of literatures, it is also suggested to consider anatomical and biochemical characters beside external morphology. If the right markers are selected based on comparative and quantitative analysis, DNA barcoding has great potential for enhancing ecological and evolutionary investigations. Specific markers may be needed for identification of different genera, as in *Ephedra,* specific MatK primer used in this study following Li *et al.*, 2011.

The results of this study showed the efficiency of nuclear marker *(ITS)* and chloroplast markers (*MatK, ycf1* and *petN-psbM)* to identify gymnosperms of Nepal. The results could be used as reference to identify species of gymnosperms in the mixture of samples. The outcome of such finding can be helpful in controlling the cross-border timber transportation of gymnosperms. It is expected that the analysis of gymnosperms on the basis of molecular markers ultimately helps conserving the gymnosperms of Nepal. It has aided in basic information of gymnosperms at molecular level and has directed the information to be used in species level identification adding evidences to claim the regionality of the gymnosperms in Nepal.

REFERENCES

Adams, R P, Boratynski, A., Arista, M., Schwarzbach, A. E., Leschner, H., Liber, Z., … Manolis, A. (2013). Analysis of Juniperus phoenicea from throughout its range in the Mediterranean using DNA sequence data from nrDNA and petN-psbM: The case for the recognition of J. turbinata Guss. *Phytologia*, *95*(2), 202–209. https://doi.org/citeulike-article-id:13699754

Adams, Robert P., Bartel, J. A., & Price, R. A. (2009). A new genus, Hesperocyparis, for the cypresses of the Western Hemisphere. *Phytologia*, *91*(1), 160–185.

Adams, Robert P., Schwarzbach, A. E., & Pandey, R. N. (2003). The concordance of terpenoid, ISSR and RAPD markers, and ITS sequence data sets among genotypes: An example from Juniperus. *Biochemical Systematics and Ecology*, *31*(4), 375–387. https://doi.org/10.1016/S0305-1978(02)00157-6

Adams, Robert P, Bartel, J. A., & Price, R. A. (2009). A new genus, *Hesperocyparis*, for the cypresses of the western hemisphere (cupressaceae). Phytologia (April 2009) 91(1), *91*(April), 160–185.

Adams, Robert P, & Chaudhary, R. P. (2009). *Juniperus recurva* var. *uncinata*, the hooked branchlet juniper, a new variety from nepal.Phytologia (December 2009) 91(3) 361 361–382.

Adams, Robert P, & Kauffmann, M. E. (2010). VARIATION IN nrDNA AND cpDNA OF JUNIPERUS CALIFORNICA, J. GRANDIS, J. OCCIDENTALIS AND J. OSTEOSPERMA (CUPRESSACEAE). *Phytologia*, *92*(August), 266–276.

Adams, Robert P, & Schwarzbach, A. E. (2013). Phytologia (May 2013) 95(2), *95*(May), 2012–2014.

Altschup, S. F., Gish, W., Pennsylvania, T., & Park, U. (1990). Basic Local Alignment Search Tool 2Department of Computer Science, 403–410.

Armenise, L., Simeone, M. C., Piredda, R., & Schirone, B. (2012). Validation of DNA barcoding as an efficient tool for taxon identification and detection of species diversity in Italian conifers. *European Journal of Forest Research*, *131*(5), 1337–1353. https://doi.org/10.1007/s10342-012-0602-0

Beck, C. B. (2013). On the Origin of gymnosperms. TAXON.VOL XV. No9, *15*(9), 337–339.

Biswash, C. & Mohri, B.M (1977). The Gymnosperms.Springer-Verlag Berlin Heidelberg.ISBN 978-3-662-13164-0.

Bowe, L. M., Coat, G., & dePamphilis, C. W. (2000). Phylogeny of seed plants based on all three genomic compartments: Extant gymnosperms are monophyletic and

Gnetales' closest relatives are conifers. *Proceedings of the National Academy of Sciences*, *97*(8), 4092–4097. https://doi.org/10.1073/pnas.97.8.4092

Carrilho, E., Kist, T. B. L., & Franc, L. T. C. (2002). A review of DNA sequencing techniques. *Quarterly Reviews of Biophysics*, *2*, 169–200. https://doi.org/10.1017/S0033583502003797

Celiński, K., Kijak, H., Wojnicka-Półtorak, A., Buczkowska-Chmielewska, K., Sokołowska, J., & Chudzińska, E. (2017). Effectiveness of the DNA barcoding approach for closely related conifers discrimination: A case study of the Pinus mugo complex. *Comptes Rendus - Biologies*, *340*(6–7), 339–348. https://doi.org/10.1016/j.crvi.2017.06.002

Chase, M. W., Salamin, N., Wilkinson, M., Dunwell, J. M., Kesanakurthi, R. P., Haidar, N., & Savolainen, V. (2005). Land plants and DNA barcodes: Short-term and long-term goals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1462), 1889–1895. https://doi.org/10.1098/rstb.2005.1720

Chaw, Shu-maw, & Zharkikh, A. (1996). Molecular Phylogeny of Extant Gymnosperms Analysis of Nuclear 18s rRNA Sequences and Seed Plant Evolution :

Chaw, Shu-miaw, Parkinson, C. L., Cheng, Y., Vincent, T. M., & Palmer, J. D. (2000). Seed plant phylogeny inferred from all three plant genomes : Monophyly of extant gymnosperms and origin of Gnetales from conifers.

Cheng, T., Xu, C., Lei, L., Li, C., Zhang, Y., & Zhou, S. (2016). Barcoding the kingdom Plantae: New PCR primers for ITS regions of plants with improved universality and specificity. *Molecular Ecology Resources*, *16*(1), 138–149. https://doi.org/10.1111/1755-0998.12438

de Miranda, V. F. O., Martins, V. G., Furlan, A., & Bacci, M. (2010). Plant or fungal sequences? an alternative optimized PCR protocol to avoid ITS (nrDNA) Misamplification. *Brazilian Archives of Biology and Technology*, *53*(1), 141–152. https://doi.org/10.1590/S1516-89132010000100018

Don, D. (2011). Isolation and characterization of sterols and an aliphatic alcohol from Tsuga ISOLATION AND CHARACTERIZATION OF STEROLS AND AN ALIPHATIC ALCOHOL FROM TSUGA DUMOSA, (April 2016), 10–12. https://doi.org/10.3126/sw.v9i9.5511

Dong, W., Xu, C., Li, C., Sun, J., Zuo, Y., Shi, S., … Zhou, S. (2015). ycf1, the most promising plastid DNA barcode of land plants. *Scientific Reports*, *5*, 8348. https://doi.org/10.1038/srep08348

Eizirik, E., Murphy, W. J., & Brien, S. J. O. (2001). Molecular Dating and Biogeography of the Early Placental Mammal Radiation, 212–219.

Faried, A., El-banhawy, A., & Elqahtani, M. (2018). Taxonomic , DNA Barcoding and Phylogenetic Reassessment of the Egyptian Ephedra L . (Ephedraceae), *17*, 1–13.

Farjon, A. (2007). In defence of a conifer taxonomy which recognises evolution [1]. *Taxon*, *56*(3), 639–641. https://doi.org/10.2307/25065847

Fazekas, A. J., Burgess, K. S., Kesanakurti, P. R., Graham, S. W., Newmaster, S. G., Husband, B. C., … Barrett, S. C. H. (2008). Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. *PLoS ONE*, *3*(7). https://doi.org/10.1371/journal.pone.0002802

Fir, D., Menziesii, P., Adams, R. P., Hunter, G., Ave, E. P., Angeles, P., … Callahan, F. (2012). Phytologia (August 2012) 94(2) 199, *94*(August), 25–30.

Ford, C. S., Ayres, K. L., Toomey, N., Haider, N., Stahl, J. V. A. N. A., Kelly, L. J., … Wilkinson, M. J. (2009). Selection of candidate coding DNA barcoding regions for use on land plants, 1–11.

Forest, F., Moat, J., Baloch, E., Brummitt, N. A., Bachman, S. P., Ickert-bond, S., … Buerki, S. (2018). Gymnosperms on the EDGE, (March), 1–11. https://doi.org/10.1038/s41598-018-24365-4

Gao, L., Ahrends, A., Baral, S. R., Liu, J., Poudel, R. C., Mo, M., … Li, D. (2012). Using Morphological , Molecular and Climatic Data to Delimitate Yews along the Hindu Kush-Himalaya and Adjacent Regions, *7*(10). https://doi.org/10.1371/journal.pone.0046873

Gao, L. M., Li, Y., Phan, L. K., Yan, L. J., Thomas, P., Phan, L. K., … Li, D. Z. (2017). DNA barcoding of East Asian Amentotaxus (Taxaceae): Potential new species and implications for conservation. *Journal of Systematics and Evolution*, *55*(1), 16–24. https://doi.org/10.1111/jse.12207

Garden, T. B., Museum, N. S., & Products, F. (2007). Variable Wood Formation and Adaptation to the Alpine Environment of Ephedra pachyclada (Gnetales : Ephedraceae) in the Mustang District , Western Nepal, 315–324. https://doi.org/10.1093/aob/mcm111

Gautam, L. N., Awale, S., Kalauni, S. K., Shrestha, K., & Gewali, M. B. (2005). PHYTOCHEMICAL AND BIOLOGICAL STUDIES ON PODOCARPUS NERIIFOLIUS D . DON : A HIMALAYAN CONIFER OF NEPAL, *3*(3), 22–25.

Genes, D. N. A., Kress, W. J., & Erickson, D. L. (2008). DNA barcodes: Genes, genomics, and bioinformatics, *105*(8), 2761–2762.

Germano, J., & Klein, A. S. (1999). Species-specific nuclear and chloroplast single nucleotide polymorphisms to distinguish Picea glauca, P. mariana and P. rubens. *Theoretical and Applied Genetics*, *99*(1–2), 37–49. https://doi.org/10.1007/s001220051206

Gernandt, D. S., Liston, A., & Pin, D. (2001). Variation in the nrDNA ITS of Pinus Subsection Cembroides : Implications for Molecular Systematic Studies of Pine Species Complexes, *21*(3), 449–467. https://doi.org/10.1006/mpev.2001.1026

Gotame, B. (2008). Economics of Biodiversity Conservation in Nepal, (Fao 2001), 55–61.

Government of Nepal, Department of Plant Resources (1997). FLORA OF LANTANG AND CROSS SECTION VEGETATION SURVEY.

Hajibabaei, M., Singer, G. A. C., Hebert, P. D. N., & Hickey, D. A. (2007). DNA barcoding : how it complements taxonomy , molecular phylogenetics and population genetics, *23*(4). https://doi.org/10.1016/j.tig.2007.02.001

Hall, A. T. (2011). BioEdit : An important software for molecular biology, *2*(June), 60–61.

Harvard, S., June, N., & Doyle, F. (2015). Harvard Papers in Botany, *3*(1), 101–106.

Hebert, P. D. N., Cywinska, A., Ball, S. L., & Jeremy, R. (2003). Biological identifications through DNA barcodes, (July 2002), 313–321. https://doi.org/10.1098/rspb.2002.2218

Hebert, P. D. N., & Gregory, T. R. (2005). The Promise of DNA Barcoding for Taxonomy. *Systematic Biology*, *54*(5), 852–859. https://doi.org/10.1080/10635150500354886

Hernández-León, S., Gernandt, D. S., Pérez de la Rosa, J. A., & Jardón-Barbolla, L. (2013). Phylogenetic Relationships and Species Delimitation in Pinus Section Trifoliae Inferrred from Plastid DNA. *PLoS ONE*, *8*(7), 1–14. https://doi.org/10.1371/journal.pone.0070501

Hollingsworth, P. M., Graham, S. W., & Little, D. P. (2011). Choosing and using a plant DNA barcode. *PLoS ONE*, *6*(5). https://doi.org/10.1371/journal.pone.0019254

Hollingsworth, P. M., Li, D., Bank, M. Van Der, Twyford, A. D., Botanic, R., Edinburgh, G., … Hollingsworth, P. M. (2016). Telling plant species apart with DNA : from barcodes to genomes.

Hsiang, T., & Huang, J. (2002). The use of RAPD markers to distinguish among juniper and cedar cultivars. *Canadian Journal of Botany*, *78*(5), 655–659. https://doi.org/10.1139/cjb-78-5-655

Hussain, M., Shah, G. M., & Khan, M. A. (2006). Traditional Medicinal and Economic uses of Gymnosperms of Kaghan Valley ,.

Joshi, A. R. (2015). Strategy for Conserving Biodiversity in Nepal, *25*(3), 25–32.

Koepf, S. M., Salser, W., & Hunkaplller, M. W. (1989). Automated DNA Sequencing Methods Polymerase Chain Reaction reagents tubes, *2201*, 2196–2201.

Korpelainen, H., & Pietiläinen, M. (2019). The effects of sample age and taxonomic origin on the success rate of DNA barcoding when using herbarium material. *Plant Systematics and Evolution*, *305*(4), 319–324. https://doi.org/10.1007/s00606-019-01568-4

Krishnamurthy, P. K., & Francis, R. A. (2012). A critical review on the utility of DNA barcoding in biodiversity conservation, 1901–1919. https://doi.org/10.1007/s10531-012-0306-2

Krupkin, A. B., Liston, A., & Strauss, S. H. (1996). Phylogenetic analysis of the hard pines (Pinus subgenus Pinus, Pinaceae) from chloroplast DNA restriction site analysis. *American Journal of Botany*, *83*(4), 489–498. https://doi.org/10.2307/2446218

Kumar, S., Tamura, K., & Nei, M. (1994). MEGA: Molecular Evolutionary Genetics Analysis software for microcomputers, *10*(2), 189–191.

Laubenfelschad, D. J. De, Griffith, M. P., Laubenfels, D. J. De, Husby, C. E., Griffith, M. P., & Road, O. C. (2012). Further Nomenclatural Action for the Cypresses (Cupressaceae) Further Nomenclatural Action for the Cypresses (Cupressaceae), *22*(1), 8–15. https://doi.org/10.3417/2010056

Li, X., Yang, Y., Henry, R. J., Rossetto, M., Wang, Y., & Chen, S. (2014). Plant DNA barcoding : from gene to genome. https://doi.org/10.1111/brv.12104

Li, Y., Gao, L. M., Poudel, R. C., Li, D. Z., & Forrest, A. (2011). High universality of matK primers for barcoding gymnosperms. *Journal of Systematics and Evolution*, *49*(3), 169–175. https://doi.org/10.1111/j.1759-6831.2011.00128.x

Li, Y., Wu, D., & Gao, L.-M. (2013). Highly Universal DNA Barcoding Primers of\rITS2 for Gymnosperms. *Plant Diversity and Resources*, *35*(6), 751–760.

Liu, J., Milne, R. I., Möller, M., Zhu, G. F., Ye, L. J., Luo, Y. H., … Gao, L. M. (2018). Integrating a comprehensive DNA barcode reference library with a global map of yews (Taxus L.) for forensic identification. *Molecular Ecology Resources*, *18*(5), 1115–1131. https://doi.org/10.1111/1755-0998.12903

Mahajan, B., Shrestha, T.M& Gyawali, R. (2012). NTIBACTERIAL AND CYTOTOXIC ACTIVITY OF JUNIPERUS INDICA BERTOL FROM NEPALESE HIMALAYA . International Journal of Pharmaceutical Sciences and Research.Vol. 3(4): 1104-1107

Manning, J., Boatwright, J. S., Daru, B. H., Maurin, O., Van, M., Bank, D.,&Bank, M. Van Der. (2014). A Molecular Phylogeny and Generic Classification of Asphodelaceae Subfamily Alooideae : A Final Resolution of the Prickly Issue of Polyphyly in the Alooids ? A Molecular Phylogeny and Generic Classification of Asphodelaceae subfamily Alooideae : A Final R, *39*(1), 55–74. https://doi.org/10.1600/036364414X678044

McPherson, M.J. & Moller, S.G. (2000). PCR The basics.BIOS Scientific Publishers.ISBN 1859960170.

Medina, R.C, & Vega, I.L. (2007). Species richness , endemism and conservation of Mexican gymnosperms, 1803–1821. https://doi.org/10.1007/s10531-006-9072-3

Meier, R., Shiyang, K., Vaidya, G., & Ng, P. K. L. (2006). DNA Barcoding and Taxonomy in Diptera : A Tale of High Intraspecific Variability and Low Identification Success, *55*(5), 715–728.

Meyer, C. P., & Paulay, G. (2005). DNA barcoding: Error rates based on comprehensive sampling. *PLoS Biology*, *3*(12), 1–10. https://doi.org/10.1371/journal.pbio.0030422

Mill, R. R., & Chase, M. W. (2011). A new classification and linear sequence of extant gymnosperms, *70*, 55–70.

Moritz, C., & Cicero, C. (2004). DNA Barcoding: Promise and Pitfalls. *PLoS Biology*, *2*(10), e354. https://doi.org/10.1371/journal.pbio.0020354

Neale, D. B., & Wheeler, N. C. (2019). *The conifers: Genomes, variation and evolution*. *The Conifers: Genomes, Variation and Evolution* (pp. 1–590). Springer International Publishing. https://doi.org/10.1007/978-3-319-46807-5

Neigel, J., Domingo, A., & Stake, J. (2007). DNA barcoding as a tool for coral reef conservation. *Coral Reefs*, *26*(3), 487–499. https://doi.org/10.1007/s00338-007-0248-4

Nielsen, R., & Matz, M. (2006). Statistical Approaches for DNA Barcoding. *Systematic Biology*, *55*(1), 162–169. https://doi.org/10.1080/10635150500431239

Nishida, S., & Satoh, H. (2003). Mechanisms for the vasodilations induced by Ginkgo biloba extract and its main constituent , bilobalide , in rat aorta, *72*, 2659–2667. https://doi.org/10.1016/S0024-3205(03)00177-2

Nishimura, D., & Rating, O. (2000). Sequencher 3.1.1, *1.Biotech Software & Internet Report*.Volume 1, Numbers 1 .

Nithaniyal, S., Newmaster, S. G., Ragupathy, S., Krishnamoorthy, D., Vassou, S. L., & Parani, M. (2014). DNA barcode authentication of wood samples of threatened and commercial timber trees within the tropical dry evergreen forest of india. *PLoS ONE*, *9*(9). https://doi.org/10.1371/journal.pone.0107669

Norby, R.J, Weerasuriya,Y & Hanson, P.J (1989). Induction of nitrate reductase activity in red spruce needles by NO2 and HNO3 vapor.Vol 19

Nock, C. J., Waters, D. L. E., Edwards, M. A., Bowen, S. G., Rice, N., Cordeiro, M., & Henry, R. J. (2011). Chloroplast genome sequences from total DNA for plant identification, 328–333. https://doi.org/10.1111/j.1467-7652.2010.00558.x

Ohri, D., & Khoshoo, T. N. (1986). Genome size in gymnosperms. Plant Systematics and Evolution.*153*, 119–132.

Olmstead, R. G., & Palmer, J. D. (1994). Chloroplast DNA systematics: A review of methods and data analysis. *American Journal of Botany*, *81*(9), 1205–1224. https://doi.org/10.2307/2445483

Pereira, F., Carneiro, J., & Amorim, A. (2008). Identification of Species with DNA-Based Technology : Current Progress and Challenges, 187–199.

Podani, J. (2009). Taxonomy versus evolution. *Taxon*, *58*(4), 1049–1053. https://doi.org/10.1002/tax.584001

Poudel, R. C., & Forrest, A. (2011). High universality of matK primers for barcoding gymnosperms, *49*(3), 169–175. https://doi.org/10.1111/j.1759-6831.2011.00128.x

Poudel, R. C., Gao, L., Michael, M., & Baral, S. R. (2013). Yews (Taxus) along the Hindu Kush-Himalayan region : Exploring the ethnopharmacological relevance among communities of Mongol and Caucasian origins, *147*, 190–203. https://doi.org/10.1016/j.jep.2013.02.031

Ran, J. H., Gao, H., & Wang, X. Q. (2010a). Fast evolution of the retroprocessed mitochondrial rps3 gene in Conifer II and further evidence for the phylogeny of gymnosperms. *Molecular Phylogenetics and Evolution*, *54*(1), 136–149. https://doi.org/10.1016/j.ympev.2009.09.011

Ran, J. H., Gao, H., & Wang, X. Q. (2010b). Fast evolution of the retroprocessed mitochondrial rps3 gene in Conifer II and further evidence for the phylogeny of gymnosperms. *Molecular Phylogenetics and Evolution*, *54*(1), 136–149. https://doi.org/10.1016/j.ympev.2009.09.011

Reed, B. M., Sarasan, V., Kane, M., Bunn, E., & Pence, V. C. (2011). Biodiversity conservation and conservation biotechnology tools, 1–4. https://doi.org/10.1007/s11627-010-9337-0

Ribeiro, R. A., & Lovato, M. B. (2007). Comparative analysis of different DNA extraction protocols in fresh and herbarium specimens of the genus Dalbergia. *Genetics and Molecular Research*, *6*(1), 173–187.

Rokaya, M. B., Uprety, Y., Poudel, R. C., Timsina, B., Münzbergová, Z., Asselin, H., … Sigdel, S. R. (2014). Traditional uses of medicinal plants in gastrointestinal disorders in Nepal. *Journal of Ethnopharmacology*, *158*, 221–229. https://doi.org/10.1016/j.jep.2014.10.014

Ruddy, K. J., & Partridge, A. H. (2009). Adherence with adjuvant hormonal therapy for breast cancer. *Annals of Oncology*, *20*(3), 401–402. https://doi.org/10.1093/annonc/mdp039

Saina, J. K., Li, Z.-Z., Gichira, A. W., Avoga, S., Wang, Q.-F., & Kuo, L. (2019). The complete plastome of real yellow wood (Podocarpus latifolius): gene organization and comparison with related species. *Holzforschung*, *0*(0). https://doi.org/10.1515/hf-2018-0155

Saitou, N., & Nei, M. (1987). The Neighbor-joining Method : A New Method for Reconstructing Phylogenetic Trees ', *4*(4), 406–425.

Särkinen, T., Staats, M., Richardson, J. E., Cowan, R. S., & Bakker, F. T. (2012). How to Open the Treasure Chest? Optimising DNA Extraction from Herbarium Specimens. *PLoS ONE*, *7*(8), 1–9. https://doi.org/10.1371/journal.pone.0043808

Sarvananda, L; Abed, Samer; Rohini, Jijy; Sathyamurthy, B. (2016). Molecular identification of the medicinal plant Justicia gendarussa using. *European Journal of Pharmaceutical and Medical Research*, *3*(1), 259–266. https://doi.org/10.26479/2019.0501.55

Sass, C., Little, D. P., Stevenson, D. W., & Specht, C. D. (2007). DNA Barcoding in the Cycadales : Testing the Potential of Proposed Barcoding Markers for Species Identification of Cycads, (11), 1–9. https://doi.org/10.1371/journal.pone.0001154

Sederoff, R. (2013). A spruce sequence. V O L 497 | N A T U R E. Macmillan Publishers Limited. 569

Sen, L., Fares, M. A., Liang, B., Gao, L., Wang, B., Wang, T., & Su, Y. (2011). Molecular evolution of rbcL in three gymnosperm families : identifying adaptive and coevolutionary patterns, 1–19. https://doi.org/10.1186/1745-6150-6-29

Shah, Z. A., Sharma, P., & Vohora, S. B. (2003). G inkgo biloba normalises stress-elevated alterations in brain catecholamines , serotonin and plasma corticosterone levels, *13*, 321–325.

Shneyer, V. S., & Rodionov, A. V. (2019). Plant DNA Barcodes, *9*(4), 295–300. https://doi.org/10.1134/S207908641904008X

Staats, M., Cuenca, A., Richardson, J. E., Ginkel, R. V. van, Petersen, G., Seberg, O., & Bakker, F. T. (2011). DNA damage in plant herbarium tissue. *PLoS ONE*, *6*(12). https://doi.org/10.1371/journal.pone.0028448

Stranneheim, H., & Lundeberg, J. (2012). Stepping stones in DNA sequencing. *Biotechnology Journal*, *7*(9), 1063–1073. https://doi.org/10.1002/biot.201200153

Subedi, A., Kunwar, B., Choi, Y., Dai, Y., Andel, T. Van, Chaudhary, R. P., … Gravendeel, B. (2013). Collection and trade of wild-harvested orchids in Nepal, 1–10.

Taberlet, P., Coissac, E., Gielly, L., Miquel, C., Brochmann, C., Valentini, A., … Willerslev, E. (2007). Power and limitations of the chloroplast trn L (UAA) intron for plant DNA barcoding, *35*(3). https://doi.org/10.1093/nar/gkl938

Tahir, A., Hussain, F., Ahmed, N., & Ghorbani, A. (2018). Assessing universality of DNA barcoding in geographically isolated selected desert medicinal species of Fabaceae and Poaceae, 1–16. https://doi.org/10.7717/peerj.4499

Tan, S. C., & Yiap, B. C. (2009). DNA, RNA, and Protein Extraction: The Past and The Present. *Journal of Biomedicine and Biotechnology*, *2009*, 1–10. https://doi.org/10.1155/2009/574398

Taxonomy, P. (2012). Plant DNA Isolation : A Method to Efficiently Remove Polyphenolics , Polysaccharides , and RNA Author (s): David V . Jobes , David L . Hurley and Leonard B . Thien Reviewed work (s): Published by : International Association for Plant Taxonomy (IAPT), *44*(3), 379–386.

Taylor, H. R., & Harris, W. E. (2012). An emergent science on the brink of irrelevance: A review of the past 8years of DNA barcoding. *Molecular Ecology Resources*, *12*(3), 377–388. https://doi.org/10.1111/j.1755-0998.2012.03119.x

Techen, N., Parveen, I., Pan, Z., & Khan, I. A. (2014). DNA barcoding of medicinal plant material for identification. *Current Opinion in Biotechnology*, *25*, 103–110. https://doi.org/10.1016/j.copbio.2013.09.010

Tewari, S. Chauhan, D.K. (2006).Gymnospermous Biodiversity of Himalayas, their economic importance and conservation. TAXON. BOT. Vol. 30. No. 4.

Tewari, L.M., Kumar, S. & Upreti, K. (2010). Assessmentofgymnosperms diversity of Uttarakhand, Biodiversity Potentials of Himalays. 265-278

Url, S., & Felsenstein, J. (2012). Society of Systematic Biologists Is There Something Wrong with the Bootstrap on Phylogenies ? A Reply to Hillis and Bull Author (s): Joseph Felsenstein and Hirohisa Kishino Reviewed work (s):, *42*(2), 193–200.

Valentini, A., Taberlet, P., & Decollato, V. S. G. (2008). DNA barcoding for ecologists, (December). https://doi.org/10.1016/j.tree.2008.09.011

Valley, K. (2017). AIR POLLUTION TOLERANCE INDEX OF SOME SELECTED GYMNOSPERM SPECIES ALONG THE ROAD SIDE OF, 13–19.

Vargas, P., Caujapé-Castells, J., Nogales, M., Jaén-Molina, R., & Rumeu, B. (2014). Phylogeography and genetic structure of the threatened Canarian Juniperus cedrus (Cupressaceae) . *Botanical Journal of the Linnean Society*, *175*(3), 376–394. https://doi.org/10.1111/boj.12172

Wiemers, M., & Fiedler, K. (2007). Does the DNA barcoding gap exist? - A case study in blue butterflies (Lepidoptera: Lycaenidae). *Frontiers in Zoology*, *4*(September 2004), 1–16. https://doi.org/10.1186/1742-9994-4-8

Wilkinson, M. J., Szabo, C., Ford, C. S., Yarom, Y., Croxford, A. E., Camp, A., & Gooding, P. (2017). Replacing Sanger with Next Generation Sequencing to improve coverage and quality of reference DNA barcodes for plants. *Scientific Reports*, *7*(March), 1–11. https://doi.org/10.1038/srep46040

Will, K. W., Mishler, B. D., & Wheeler, Q. D. (2005). The Perils of DNA Barcoding and the Need for Integrative Taxonomy. *Systematic Biology*, *54*(5), 844–851. https://doi.org/10.1080/10635150500354878

Wu, F., Li, M., Liao, B., Shi, X., & Xu, Y. (2019). DNA Barcoding Analysis and Phylogenetic Relation of Mangroves in Guangdong Province , China, 1–13. https://doi.org/10.3390/f10010056

Yoon, B. (2009). Hidden Markov Models and their Applications in Biological Sequence Analysis, 402–415.

Zachos, F. E. (2015). Ac ce p d cr t. *Mammalian Biology*. https://doi.org/10.1016/j.mambio.2015.10.002

Zachos, F. E. (2015b). Ac ce p d cr t. *Mammalian Biology*. https://doi.org/10.1016/j.mambio.2015.10.002

APPENDICES

Photographs of sample collection and Habit study

Photographs of herbarium preparation
















Sampling details of single set of species under study

Table 0.1 Sampling information of the species studied

Species	Collection code	Lab Code	Locality	Latitude/Longitude	Altitude (m)	District	Collected by
Abies spectabilis	GymNepal 1	GR5	Above Cholangpati		3700	Rasuwa	Ranjeeta Odari and group
Juniperus recurva	GymNepal 2	GR8	Below Lauri Binayak	28.0924° N, 85.3806° E	3850	Rasuwa	Ranjeeta Odari and group
Pinus wallichiana	GymNepal 3	GR2	Cholanpati to Bomba		3200	Rasuwa	Ranjeeta Odari and group
Pinus roxburghii	GymNepal 4	GR4	Above Syafrubesi	28.163522°N 85.336325°E	1400	Rasuwa	Ranjeeta Odari and group
Tsuga dumosa	GymNepal 5	GR3	Bomba to Thulo Syafru	28.1433° N, 85.3611° E	2800	Rasuwa	Ranjeeta Odari and group
Juniperus indica	GymNepal 6	GR6	Kyanjin Ri	28° 13' 29'' N, 85° 35' 22'' E	4200	Rasuwa	Mitesh Shrestha, Suprim Shrestha, Sarbesh Rijal
Juniperus squamata	GymNepal 7	GR9	Near Kyanjin	28.21268' N, 85.56897' E	3800	Rasuwa	Ranjeeta Odari and group
Larix himalaica	GymNepal 8	GR7	Lower than Kyanjin	28.2063° N, 85.6229° E	3500	Rasuwa	Ranjeeta Odari and group
Ephedra gerardiana	GymNepal 9	GR1	On the way to Kyanjin	28.2063° N, 85.6229° E	3500	Rasuwa	Rabin Dahal
Cycas pectinata	GymNepal 10	GL10	Godawari	85°22' 59.4” E; 27°35'38.2” N	1515	Lalitpur	Ranjeeta Odari

Cycas revoluta	GymNepal 11	GL11	Kirtipur	27.6807 N, 85.2835 E	1524	Kathmandu	Ranjeeta Odari
Abies pindrow	GymNepal 12	GW11	Near Jumla	29.2522° N, 82.1659° E	2514	Jumla	R.C. Poudel
Cupressus torulosa	GymNepal 13	MG3	Near Muktinath	28.8167° N, 83.8713° E	3800	Mustang	Mitesh Shrestha
Cedrus deodara	GymNepal 14	GM11	Nagarkot		2175	Bhaktapur	Ranjeet Odari, Sujata Ghimire, Chanu Thapa
Ginkgo biloba	GymNepal 15	GK14	Kirtipur	27.6807 N, 85.2835 E	1524	Kathmandu	Ranjeeta Odari
Podocarpus nerifolius	GymNepal 16	GK16	Kirtipur	27.6807 N, 85.2835 E	1524	Kathmandu	Ranjeeta Odari
Picea smithiana	GymNepal 17	GW15	Near Rara lake	29.5239' N, 82.0788' E	3000	Mugu	R.C. Poudel
Gnetum montanum	GymNepal 18	GT6	Near Dhor Barahi temple	28°3'49"N 84°2'41"E.	1800	Tanahu	R.C. Poudel
Cryptomeria japonica	GymNepal 19	GM15	Nagarkot		2175	Bhaktapur	Ranjeeta Odari
Araucaria columnaris	GymNepal 20	GM16	Godawari	85°22' 59.4” E; 27°35'38.2” N	1515	Lalitpur	Ranjeeta Odari, Prithvi Gurung
Araucaria bidwilli	GymNepal 21	GM17	Kirtipur	27.6807' N, 85.2835' E	1524	Kathmandu	Prithvi Gurung
Juniperus horizontalis	GymNepal 22	GM19	Godawari	85°22' 59.4” E; 27°35'38.2” N	1515	Lalitpur	Ranjeeta Odari, Prithvi Gurung
Ephedra intermedia	GymNepal 23	MG2	Mustang	28.8167° N, 83.8713° E	2500	Mustang	Mitesh Shrestha
Thuja orientalis	GymNepal 24	CG1	Satdobato	27.6561° N, 85.3277° E	1400	Lalitpur	Ranjeeta Odari

Cupressus macrocarpa	GymNepal 25	CG2	Satdobato	27.6561° N, 85.3277° E	1400	Lalitpur	Ranjeeta Odari
Cupressus sampevirens	GymNepal 26	CG3	Satdobato	27.6561° N, 85.3277° E	1400	Lalitpur	Ranjeeta Odari
Cupressus lusitanica	GymNepal 27	CG4	Kirtipur	27.6807' N, 85.2835' E	1524	Kathmandu	Ranjeeta Odari
Taxus wallichiana	GymNepal 29	GT1			3500	Kavrepalanchowk	R.C. Poudel
Taxus contorta	GymNepal 30	GT2			3500	Manang	R.C. Poudel
Taxus mairei	GymNepal 31	GT3			3500	Kavrepalanchowk	R.C. Poudel
Taxus media	GymNepal 32	GT4			3500	Makwanpur	R.C. Poudel
Cephalotaxus	GymNepal 33	GT5			1400	Kathmandu	R.C. Poudel
Araucaria angustifolia	GymNepal 34	GM18	Godawari		1515	Lalitpur	Ranjeeta Odari

Nanodrop reading of DNA concentration

Table 0.2 Nanodrop reading of DNA concentration of the samples studied

Sample ID	**Sample Type**	**260/280**	**Conc. (ng/ul)**		**Sample ID**	**Sample Type**	**260/280**	**Conc. (ng/ul)**		**Sample ID**	**Sample Type**	**260/280**	**Conc. (ng/ul)**
GR1	dsDNA	2.05	693.2		GM7	dsDNA	1.684	1271.2		GM1	dsDNA	2.146	81.35
GR2	dsDNA	2.099	1780.4		GM8	dsDNA	1.934	10.8		GM2	dsDNA	2.077	56.5
GR3	dsDNA	2.048	625.55		GM9	dsDNA	1.519	6.85		GM3	dsDNA	2.085	40.55
GR4	dsDNA	1.63	24.5		GM10	dsDNA	1.959	19.8		GM4	dsDNA	2.183	62.9
GR5	dsDNA	2.051	87.1		GM11	dsDNA	1.935	7.3		GM5	dsDNA	2.08	88.5
GR6	dsDNA	1.946	81.95		GM13	dsDNA	2.138	49.35		GM6	dsDNA	2.033	939.25
GR7	dsDNA	2.143	1417.4		GM14	dsDNA	1.942	468.95		GL11	dsDNA	-2.004	1.85
GR8	dsDNA	1.868	12.4		GM15	dsDNA	1.88	362.9		CG4	dsDNA	1.926	36.85
GR9	dsDNA	1.827	29.75		GM16	dsDNA	3.187	2.85		GT1	dsDNA	1.867	32.1
GL10	dsDNA	1.933	679.75		GM17	dsDNA	-2.416	0.55		GT2	dsDNA	1.79	33.5
GW11	dsDNA	2.02	77.85		GM18	dsDNA	1.906	524.55		GT3	dsDNA	1.938	25.85
GW12	dsDNA	2.076	728.1		GM19	dsDNA	2.133	735.65		GT4	dsDNA	1.827	40.8

GW13	dsDNA	1.978	50.3		MG1	dsDNA	1.776	18.5		GT5	dsDNA	1.502	5.55
GK14	dsDNA	1.796	41.65		MG2	dsDNA	1.362	61.3		GT6	dsDNA	1.897	501.9
GW15	dsDNA	1.964	34.7		MG3	dsDNA	2.099	81.7					
GK16	dsDNA	1.972	46.2		MG4	dsDNA	2.049	1024.05					
GW17	dsDNA	2.07	70.1		MG5	dsDNA	2.108	70.25					
GW18	dsDNA	2.074	1286.25		MG6	dsDNA	1.946	50.05					
GW19	dsDNA	2.009	700.65		MG7	dsDNA	1.128	4.35					
GW20	dsDNA	2.096	724.4		CG1	dsDNA	1.868	629.65					
GW21	dsDNA	1.876	6.6		CG2	dsDNA	1.973	58.2					

PCR and Sequencing details of DNA samples

PCR reaction mixtures for amplification of markers

Table: PCR reaction mixture for *ITS* region amplification from nuclear DNA

Components	**Volume in µl**	**Final concentration**
Master Mix (2X)	5	1X
Taq polymerase (5U/µl)	0.05	0.25U/10 µl*
Magnesium chloride (25mM)	0.4	1 mM *
Forward primer (ITSP5/ u1) (10pmoles)	0.4	0.4 pmoles
Reverse primer (ITSU4) (10pmoles)	0.4	0.4 pmoles
Bovine serum albumin (1mg/ml)	1	0.1 µg/µl
Template	1	
Nuclease Free water	1.75	
Total	10	

Note: * added for those samples not amplified by basic composition

Table PCR reaction mixture for amplification of *MatK* region from chlroplast DNA

Components	**Volume (µl)**	**Final concentration**
Master Mix (2X)	5	1X
Taq polymerase (5U/µl)	0.05	0.25U/10 µl*
Magnesium chloride (25mM)	0.4	1 mM *
Forward primer(MatK _Gym/MatK _eph_gym) (10pmoles)	0.4	0.4 pmoles
Reverse primer(MatK _Gym/MatK _eph_gym) (10pmoles)	0.4	0.4 pmoles
Bovine serum albumin (1mg/ml)	1	0.1 µg/µl
Template	1	
Nuclease Free water	1.75	
Total	10	

Table : PCR reaction mixture for *ycf1* region amplification from chloroplast DNA

Components	**Volume in µl**	**Final concentration**
Master Mix (2X)	5	1X

Forward primer(ycf1gF) (10pmoles)	0.4	0.4 pmoles
Reverse primer(ycf1gR) (10pmoles)	0.4	0.4 pmoles
Template	1	
Nuclease Free water	3.2	
Total	10	

Table PCR reaction mixture for *petN-psbM* region amplification from chloroplast DNA

Components	**Volume in µl**	**Final concentration**
Master Mix (2X)	5	1X
Forward primer(petNF) (10pmoles)	0.4	0.4 pmoles
Reverse primer(psbMR) (10pmoles)	0.4	0.4 pmoles
Template	1	
Nuclease Free water	3.2	
Total	10	

PCR conditions

Table :PCR condition for ITS amplification from nuclearDNA

Steps	Initial Denaturation	Denaturation	Annealing	Extension	Final Extension	Hold
Cycles	1	35			1	1
Temperature	94 ˚C	94 ˚C	55 ˚C	72 ˚C	72 ˚C	4 ˚C
Time	4mins	40sec	40sec	70sec	7mins	Infinite

Table :PCR condition for amplification of *MatK* from chloroplast DNA

Steps	Initial Denaturation	Denaturation	Annealing	Extension	Final Extension	Hold
Cycles	1	35			1	1
Temperature	94 ˚C	94 ˚C	52 ˚C	72 ˚C	72 ˚C	4 ˚C
Time	4mins	30sec	30sec	45sec	10mins	Infinite

Table :PCR condition for *ycf1* amplification from chloroplast DNA

Steps	Initial Denaturation	Denaturation	Annealing	Extension	Final Extension	Hold
Cycles	1	35			1	1
Temperature	94 ˚C	94 ˚C	52 ˚C	72 ˚C	72 ˚C	4 ˚C
Time	4mins	30sec	40sec	1min	10mins	Infinite

Table :PCR condition for *petN-psbM* amplification from chloroplast DNA

Steps	Initial Denaturation	Denaturation	Annealing	Extension	Final Extension	Hold
Cycles	1	30			1	1
Temperature	94 ˚C	94 ˚C	50 ˚C	72 ˚C	72 ˚C	4 ˚C
Time	4mins	30sec	30sec	40sec	7mins	Infinite

Table 0.3 PCR and Sequencing efficiency of the samples studied

Species	**Lab code**	**PCR**				**Sequencing**			
		ITS	*MatK*	*ycf1*	petpsbM	*ITS*	*MatK*	*ycf1*	petpsbM
Abies spectabilis	GR5	√	√	√	NA	√	√	√	NA
Juniperus recurva	GR8	√	√	NA	√	√	√	NA	√
Pinus wallichiana	GR2	×	√	√	NA	×	√	√	NA
Pinus roxburghii	GR4	×	√	√	NA	×	√	√	NA
Tsuga dumosa	GR3	√	√	√	NA	√	√	√	NA
Juniperus indica	GR6	√	√	NA	√	√	√	NA	√
Juniperus squamata	GR9	√	√	NA	√	√	√	NA	√
Larix himalaica	GR7	√	√	√	NA	√	√	√	NA
Ephedra gerardiana	GR1	√	√	NA	NA	√	√	NA	NA
Cycas pectinata	GL10	√	√	NA	NA	×	√	NA	NA
Cycas revolute	GL11	√	√	NA	NA	√	√	NA	NA
Abies pindrow	GW11	√	√	√	NA	√	√	√	NA

Cupressus torulosa	MG3	√	√	NA	√	√	√	NA	√
Cedrus deodara	GM11	√	√	√	NA	×	√	√	NA
Ginkgo biloba	GK14	√	√	NA	NA	√	√	NA	NA
Podocarpus nerifolius	GK16	√	√	NA	NA	√	√	NA	NA
Picea smithiana	GW15	×	√	√	NA	×	√	√	NA
Gnetum montanum	GT6	√	√	NA	NA	√	√	NA	NA
Cryptomeria japonica	GM15	√	√	NA	√	√	√	NA	√
Araucaria columnaris	GM16	√	√	NA	NA	√	√	NA	NA
Araucaria bidwilli	GM17	√	√	NA	NA	×	√	NA	NA
Juniperus horizontalis	GM19	√	√	NA	√	√	√	NA	√
Ephedra intermedia	MG2	√	√	NA	NA	√	√	NA	NA
Thuja orientalis	CG1	√	√	NA	√	√	×	NA	√
Cupressus macrocarpa	CG2	√	√	NA	√	√	√	NA	√
Cupressus sampevirens	CG3	√	√	NA	√	√	√	NA	√
Cupressus lusitanica	CG4	√	√	NA	√	√	√	NA	√
Taxus wallichiana	GT1	√	√	NA	NA	√	√	NA	NA

Taxus contorta	GT2	√	√	NA	NA	√	√	NA	NA
Taxus mairei	GT3	√	√	NA	NA	√	√	NA	NA
Taxus media	GT4	√	√	NA	NA	√	√	NA	NA
Cephalotaxus	GT5	√	√	NA	NA	√	√	NA	NA
Araucaria angustifolia	GM18	√	√	NA	NA	√	√	NA	NA
Cupressus torulosa	GW12	√	√	NA	√	×	√	NA	√
Cedrus deodara	GW13	×	√	√	NA	×	√	√	NA
Tsuga dumosa	GW17	√	√	√	NA	×	√	√	NA
Juniperus indica	GW19	√	√	NA	√	√	√	NA	√
Abies pindrow	GW20	√	√	√	NA	×	×	√	NA
Juniperus indica	GW18	√	√	NA	√	√	√	NA	√
Pinus wallichiana	GW21	×	√	√	NA	×	√	√	NA
Juniperus indica	GW22	√	√	NA	√	√	√	NA	√
Picea smithiana	GM1	×	√	√	NA	×	√	√	NA
Abies pindrow	GM2	√	√	√	NA	×	√	√	NA
Tsuga dumosa	GM3	√	√	√	NA	×	√	√	NA

Pinus wallichiana	GM4	×	√	√	NA	×	√	√	NA
Pinus roxburghii	GM5	×	√	√	NA	×	√	√	NA
Juniperus indica	GM6	×	√	NA	√	√	√	NA	√
Juniperus indica	GM7	×	√	NA	√	√	√	NA	√
Pinus roxburghii	GM8	×	√	√	NA	×	√	√	NA
Pinus wallichiana	GM9	×	√	√	NA	×	√	NA	NA
Pinus roxburghii	GM10	×	√	√	NA	×	√	√	NA
Cedrus deodara	GM13	√	√	√	NA	√	√	√	NA
Cupressus torulosa	GM14	√	√	NA	√	√	√	NA	√
Ephedra gerardiana	MG7	×	√	NA	NA	×	√	NA	NA
Ephedra intermedia	MG1	√	√	NA	NA	√	√	NA	NA
Juniperus indica	MG4	√	√	NA	√	×	√	NA	√
Pinus wallichiana	MG5	×	√	√	NA	×	√	√	NA
Cryptomeria japonica	MG6	×	√	NA	NA	×	√	NA	NA

*NA: not applicable

Herbarium Study of Gymnosperms collected from different parts of Nepal: National Herbarium and Plant Laboratory(KATH), Godawari, Lalitpur

Table 0.4 Herbarium Study of Gymnosperms collected from different parts of Nepal: National Herbarium and Plant Laboratory(KATH), Godawari, Lalitpur

SN	Family	Species	Habit	Local name	Altitude	Distribution	Locality	Remarks
1	Cycadaceae	***Cycas pectinata*** Griff.,	Shrub	Jaggar, Thakal, Kalwal	300-450 m	Himalaya (Nepal, Sikkim), Assam (Khasia), Manipur, Burma, Yunnan.	Central **Makwanpur** : Phurke Chaur,Paos . East **Sankhuwasabha**-Pilwa; Jhapa-Ganjabari Garwa : North east facing slope of Siwalik hill(220m) **Illam**: Chulachuli (1200m)	
		Cycas revoluta Thunb.	Shrub	Kalwal	1300m		Central, **Kathmandu**-Nayabajar	Cultivated
2	Ginkgoaceae	***Ginkgo biloba*** L.	Tree	Baal kumari, Ginko	4200ft	China	Central, **Kathmandu**-New road Tundikhel CDB,TU,Kirtipur	Cultivated
3	Podocarpaceae	***Podocarpus neriifolius*** D. Don, ; in Lamb., Descr. Gen. Pinus 2: 21 (1824); Prodr. Fl. Nepal.: 55 (1825).	Tree	Gunsi	1000-1100 m;	Himalaya (Nepal, Sikkim), Assam, Burma, China, Taiwan, Malaysia, New Guinea, Fiji.	Central **Kathmandu** Sankhu West **Kaski** : Phewa lake ,Pokhara(3000ft) Siwanta 850- pass1260 ,Tsinging 1200m CDB,TU ,Kirtipur	
4.	Taxaceae	***Taxus wallichiana***Zucc., Abh. Akad.	Tree	Louth salla	2300-4300m	Afghanistan, Himalaya (Kashmir to	Central **Kaski** Chinkhola village; Kaski-Banthanti East	

		Muench. 3: 803, t. 5. 1837-43; Abh. Bayer. Akad. Wiss. 3: 803, t. 5 (1843).				Bhutan), Assam, N. Burma, Indo-China, W. China, Malaysia	**Solukhumbu** :Lamujo-Chumoa	
		Taxus contorta Griff.	Tree	Louth salla	2600-2800m		West **Darchula** – Tipulchyakti **Humla**- Kallas Central **Baglung** - Nisyaldhor	
5.	Araucariaceae	***Araucaria bidwillii*** Hook., ; London J. Bot. 2: 503, t. 18, 19 (1843).	Tree	Monkey Puzzle Tree काँडेसल्लो Kande sallo, Thingre salla	2050m	1200 m; Australia.	Central **Bhaktapur** Nagarkot- Bhangeri Durbar	Cultivated
		Araucaria columnaris Hook.	Tree		4200ft		Central **Kathmandu** –Bahadur bhawan campus	Cultivated
6.	Pinaceae	***Abies densa*** Griff. ex Parker, ; Ind. Forester 53 (12): 686 (1927).	Tree		3000-4000m	E. Nepal to Assam, and S.E. Tibet.	East **Taplejung**: Wolangchung Gola – Thasa kharka, Wolangchung gola(3110m)-Tamur Yangma Khola junction (2580m) - Cheni Kharka (2820m)- Thasa kharka (3260m),, Deuma kharka (3855m) – flow junction)3575m)- Wolangchung gola (3110m) **Solukhumbu:** Sarkari pati (3350m)-Ringmo(2700m)-Phera (2500m)-Beni(2350m) **Sankhuwasabha** : Around Cha Ding Kharka (3970m) to whole Shiptoon pass (4127m), Milke danda, Chhippon (2980m)- Dobati	

							kharka (3320m)- Gidde (3500m),, Around Jaljale (3500m),, Numbuk/Mumbuk (3500m)- Barun khola (3170m)- Phemathang Kharka /Tematan Kharka (3380m) -Yangri Kharka (3540m),,Khongma/Kauma (3500m)- Shipton pass (Keke La) (4120m)- Tutu La (4070m) -Cha Ding Kharka (3970m)- Numbuk/Mumbuk (3500m) Central, **Janakpur**: east of Neju (3651m)	
		Abies pindrow Royle, ; Illustr. Bot. Himal.: 350, t. 86, f. a-c (1836).	Tree	Thingre salla	2100-2500m	Eastern Afghanistan to W. Nepal.	West **Humla** : Chankheli- Rimi,, Darma –Loha chaur(2400m) **Mugu** –Khatyarkhola(2650m) **Darchula** : Chheti- Mechchra(3020m),, Khaptad – Pujarigaon (1800m),,Near Kantel god (2250m) **Bajhang** : Rosia range (2730m), Dhalaun-Rasa (2600m),, Dhalaun(2400m)- a pass (2880m) –Ghat Khola (1980m)- Jima (2230m)-Rasa (2260m)	
		Abies spectabilis (D. Don) Mirb., ; Mem. Mus. Hist. Nat. Paris 13: 70 (1825).	Tree	Himalayan Silver Fir तालिसपत्र Talish patra (Nep.);Talisa, Talish Patra (San); Buk (Dolpa), Ba-lu (Tib); Kye (Manang) ,	2400-4400m	Himalaya (Chitral to C. Nepal).	West **Humla** ; Marghor- Durpa **Dolpa** :Phoksundo- Gandala(4000m) **Darchula** :Khaptad(3000m) **Bajhang** : Naya Odar(3240m) **Jumla** :GothiChauar (2700m),,Maharigaon(2800m),, Bibeya(3540m)- (3510m)-Chakare pani (3670m)- a saddle(3600m)- Deula deuli (3850m) **Myagdi** : above ward and Chukhola	

				Gobre sallo			(3100m),, Chaubi/Charbi/Chanbi/Chambi(3000m) **Jajarkot** : Tuntupauwa (3000m) **Kalikot**; near Mabu pass(2900m) **Bajura and Bajhang:** Pategaon(2190m)-Ghodi lekh Deurali(3350m)- Badi gaon (2230m) **Mugu** ; 3450m **Doti** : patan(Khaptad) (3500m) **Manang** : Gyabru (3400m),,Chungar (3950m), Tilicho (3900m) ,Chame(2600m)-Kaleku (2680m)-Bhartang (2770m)-Pisang(3090m) Central **Kaski**; Ghodepani-Poon Hill, Banthanti (2600m),,Ghodepani -Poon Hill(3100m) **Ramechhap**:Lapchane dada (3000m),Thodung (3050m) **Rasuwa** : Chandanbari (3380m),,Chadanbari -Gosaikunda (3550m-12000ft),,Langtang(3400m)-Jhade (11,000ft),,Rupchet kharka-Balchegaon (3400m) Dhunche Camp(1970m)-Shin Gompa (3260m),,Foprang (3180m),, Dhunche Deurali (2800m),, Syabru (2240m, 6820ft) -Shin Gompa (3200m) **Dolakha** :Chhumigaliyo (3550m),Duri Kharka (2700m),Kalinchown (3600m) **Sindhupalchowk** : Nelenbu (3800m),Tharu pati pass (3300m),Thade pati (3500m) **Kavre**: 2300m	

							Nuwakot : Gangabhanjyang (3300m) **Parbat** :Banthanti (2650m) -Gorepani Deurali (3170m) **Okhaldunga**: Maile -Okhaldhunga(2900m) **Gorkha** –Lungdang Gompa(3150m)-Torogompa glacier (3820m),,Thangmanang Kharkha (3730m) -Syallagaon (3330m)-Samagaon(3320m)-Sama Gompa(3380m) **Baglung** : Dhorpatan(2800m) East **Solukhumbu** ; Namche bazar –Phurte(3550m), Gohim-Junbesi, Sanasa-Tengboche (3450m) Phurte -Thame(3200-3500m) ,Sete (2500m)-Kata Bisana(3000m)-Goyem (3400m)-LAmjurea Bhanjyang (3200m)-Taktor (3000m) **Illam**- Chhintapu(10,700ft)	
		Cedrus deodara (Roxb. ex D. Don) G. Don, ; in Loudon, Hort. Brit.: 388 (1830).	Tree	देवदा? Devdaru, Diyar, Devdaar	2000-2500m	Afghanistan, Himalaya (Kashmir to W. Nepal).	West **Kalikot**;Dillikot- Odan Gaon(1100m),Chilkhaya (2000m),Khallagadh -Ranchula(1800m) **Dolpa**:Ranga Gaon-Thulibheri (2400m), Dunai (2090m)-Juphal airport (2480m),Thalagaon (2820m)-Jangra khola(2170m) ,Ankhe (2680m)-Rahagaon (2870m)-a saddle (2480m)-Dunai (2090m) **Jumla**: around 2350m,Sija(2400m),Tatopani (2110m) **Humla** :Darma Gaon (2200m)	

							Mugu: Rara – Gumgadhi (2700m),Tum (2900m) **Baitadi** :Tadi Gaon (1660m) **Kaski**: Pokhara (1500m) Central **Bhatkapur** Nagarkot –Bhangeri Durbar **Kathmandu** : Singhadurbar , CDB(TU) Kirtipur,Balaju (5000ft),Nagarjun (4800ft) Lalitpur :Patan **Dolakha** – Jiri(2100m)	
		Larix griffithiana Carriere, ; Traite Gen. Conif: 278 (Jun. 1855). Orr in Not. Roy. Bot. Gard. Edinburgh 18: 130, t. 241, f. 1 (1933). Stearn in 150th. Ann. Vol. Roy. Bot. Gard. Calcutta: 115 (1938). Kitam. in H. Kihara, Fauna & Fl. Nepal Himal.: 81 (1955). H. Hara, Photo-Alb.	Tree	Langtang sallo, Bogre salla ,Lekh salla	1400-3900m	Himalaya (Nepal, Sikkim, Bhutan), Assam	East **Taplejun**g ; Kypra -Pheri ,Ghunsa khola (3000m) ,Ghunsa(3300m)-Charmalung kharka (3620m)Rampuk kharka(3660m) ,Gybala(2630m)-Nag pokhari khola (2620m)-Phale (2900m)-Ghunsa (3300m) Central **Rasuwa**: Sherpa gaon -Langtang area(3000m)	
		Larix himalaica Cheng & L. K. Fu, ; Acta Phytotax. Sinica 13: 84, f. 26, 1-6 (1975).	Tree	Langtang sallo	2400-3600m	C. Nepal, S. Tibet.	Central **Rasuwa** –Langtang , Mu gomba, Sherpa Gaon -Lantang area/south of /Langtang khola (3000m-3400m) ,Lama hotel(2480m)-Below Ghode Tabela	

							(2850m),Kyangjin gompa (3800m) **Gorkha** : Thangmanang kharka (3730m)-Syallagaon (3330m) Samagaon (3320m)-Sama gompa(3380m) ,Lungdang Gompa(3150m)-Torogompa glacier (3820m)	
		Picea smithiana (Wall.) Boiss., ; Fl. Orient. 5: 700 (1884).	Tree	Jhule sallo, Jure salla	2300-3600m	Afghanistan, Himalaya (Kashmir to C. Nepal).	West **Kalikot** – near Dillikot **Jumla**: Dori Lekh(3000m),Nyorgadh,Gothichaur (2760m) ,Tila Nadi (2310m) -Bibeya (3540m) ,Chive patan (3200m) **Humla** : Durpa (2900m),Kallas (2750m) **Mugu** : Rara area (2900m) **Darchula** : Khaptad-Pujari gaon (2600m) **Doti** : Chhoyagaon -Khaptad (2550m) **Dolpa** : Daja (3620 m)- Pugmi/Pamgma(3250m)-Sumduwa(3100m) ,Lukhor(2960m)- a saddle (3550m)-Bange(3840m) ,Ankhe,(Suligadh) (2550m **Bajhang** : Jimkot(1480m) -Sonar gaon (1520m) -Daruhgaon (1650m) -Meltadi (1830m) -Khaptad NP(2980m) **Dadeldhura** : Gayardoti (1600m) **Bajura** : Birseni (1560m)-Shalkatiya (1700m)-Naulakot (1810m)-Pirekot (1900m)-Pandusen (2020m) Porakhya deurali (2650)-Porakya (2540m) Central **Baglung** – Nisyaldhor(2550m)	

							Manang; Bagarchhap (2030m) -Danake(2120m)-Timang Besi (222om) -Latamanang (2320m)-Koto(2510m)-Chame (2600-2860m) , Marsyangdi khola : Humde(3320m) -Pisang (3130m)-Deurali Danda (3050m)-Bardang (3320m-2850m)-Chame (2630m)	
		***Pinus* roxburghii** Sarg., ; Silva N. Amer. 11: 9 (1897).	Tree	Khote sallo, Aule sallo	1100-2100m	Afghanistan, Himalaya (Kashmir to Bhutan).	West **Kanchanpur**- Katunje **Dailekh:** Ranimatta (2170m)-Rato khola Ghodabas(1660m)-Tal Pokhari (1570m) – Pani tanki (1290m)-Dungesar(720m) **Surkhet**; Chigad Gaon- Surkhet(2700m) Central **Dolakha** – Singati East **Solukhumbu** ; Wanku- Chochomigaon	
		Pinus wallichiana A. B. Jacks., ; Kew Bull. 1938: 85 (1938).	Tree	गेब|] सल्ला Gobre sallo, Rani sallo	1800-3300m	Afghanistan, Himalaya (Kashmir to Nepal).	West **Darchula** - Thin **Humla** – Kallas Central **Baglung** – Nisyaldhor **Rasuwa**; Dhunche – Chandanbari East **Solukhumbu** – Jorsale	
		Tsuga dumosa (D. Don) Eichler, ; in Engler, Pfl.-fam. II-l: 80 (1887).	Tree	Thinge salla, Chune sallo	2100-3600m	Himalaya (Kumaun to Bhutan), N. Burma.	West **Darchula** – Palaa(2250m)Khaptad (9000ft)-Pujarigaon(2400m) **Bajura** – Ghodilekh (3000m) ,Dugwa- Ghodilekh (2400m) , Birseni (1560m)- Shalkatiya (1700m) -Naulakot (1810m)-Pirekot (1900m)-	

							Pandusen (2020m) -Porakya deurali (2650m) **Dailekh** : Panipokhari – Haude lekh (2400m) ,Chhanna – Bhabase (2500m)- Mabu pass (2400m) Kaski **Dolpa** : 3500m – Chaarikot (3080m) – saddle (3120m)- Rimi (2890m) **Darchula** : Kasoti – Chheti (2500m) **Bajhang** : Meltadi (1500m),Ghodilekh (3030m) **Manang** : Bhratang- Chame(2900m) - Bhatang – Pisang (2730m-2970m) **Kalikot** : Chankebada – Badarigaon , ,Pankipokhari (2020m)-Bajangeneta (2220m) -Beuli (2440m)- A pass (2880m) – Chaukebada (2340m) **Humla;** Chankheli – Rimi(3000m) **Jajarkot** : Harisallaa lekh (2630m), Chantheri (2680m) ,Samaila (2100m) ,Tutnupauwa – Ghattechari (2800m) Central **Mustang** : Marpa – Thumagaon (2450m), ,Kalopani (2530m) – Dhampus (2480m)- Kokethanti (2500m)- Larjung (2520m) – Khanti (2540m) – Tukuche (2590m) **Rasuwa** ; Dhunche- Chandanbari- Gosainkunda (2800m),,Syabru (2300m) ,Haku 2080m), Sherpagaon- Lantang (9000ft)	

							,South of Lantang khola (Ghoda tabela) (2950m-3048m)-Lama Hotel (2480m) ,Parbati Kund (near Gatlang 2590m)- Yure Khaka (3300m),Sherpa gaon -Langtang (9000ft) ,Domjechauri (2840m) **Gorkha** : Dunje (2270 m), Ripche (2365m) ,Ripche (2390m) – Dumje Thumje (2310 m) -Lungdang Gompa (3150m) **Dhading** : below Singla (2720m) **Ramechhap** : Pati (2600m) **Nuwakot** : Tharepati (3300m) **Dolakha**; Lamabagar- Hum(2100m) ,Deurali (2350m),Chisapani -Sikri gaon (1400-1880m) **Parbat** : Banthanti(2650m) -Gorepani (3170m) East **Solukhumbu** ; Junbesi- Sallung, Ghat -Mujo (2500-2600m), Chumoa (2680m), Namche bajar -Toktok (2940m) **Sankhuwasabha** : Jiijibuk (2930m) **Taplejung** : Bhuje lekh (6000ft) **Okhaldhunga** : Maili (2900m) **Illam**- Maipokhari	
7.	Taxodiaceae	***Cryptomeria japonica*** (L. f.) D. Don, ; Trans. Linn. Soc. 18: 167, t. 13, f. 1 (1841).	Tree	कत्लेसल्ला Katle salla, Dhupi salla, Dhupi	1300-2600m	Japan, cultivated elsewhere (Nepal, Sikkim, Burma, Malaysia).	Central **Bhaktapur**: Nagarkot – Bhangeri durbar **Kathmandu** : Naxal Gokarna (4500ft), Kirtipur (4500ft)	

8.	Cupressaceae	***Cupressus torulosa*** D. Don in Lamb, Descr. Gen. Pinus. 2: 18. (1824);. Prodr. Fl. Nepal.: 55 (1825).	Tree	राजसल्लोf Raj sallo	1800-3300m	Himalaya (Kashmir to C. Nepal).	West **Darchula**- Sunchera (1800m) ,Khandeshwori -Kantel Gadh (2170m) **Dolpa** : 3500m – Chaarikot (3080m)–saddle (3120m)-Rimi (2890m), Palang (Thuli bheri) (1700m) ,Daja (3620m)- Pugmi (3250m) -Sumduwa (3100m) ,Near Dunai (2090m) **Jajarkot**- Maina Gaon (1800m) Central **Kathmandu**: Ratnapark (4200ft), Kesharmahal (4400ft) **Bhaktapur**:Nagarkot , Bhangeri Durbar (2050m) **Palpa** : Durbar Square (1300m) **Mustang** ; Kalopani (2530m) – Dhampus (2480m)- Kokethanti (2500m)- Larjung (2520m) – Khanti (2540m) –Tukuche (2500-2590m) ,Lete -Tukuche (9500ft) ,Chhatang -Tukuche (2500m), Chhatag -Larjung (2550m) ,Kabhre (1720m) – Tallo Ghansa (1980m) – Kaiku (2180m) -Lete (2470m) -Kalopani (2530m) ,Marpha – Pak Kharka (3050m),Marpha (2640m) -Alubari kharka (3580m-3630m) – Yak kharka(3800m)- Tukuche (2500m),,Jomsom (roadside 2750m) ,Marpha -Tukuche (2600m), Marpha – Thuman (2500m)	

		Juniperus indica Bertol., ; Misc. Bot. 23: 228, t. 1 (1862); in Mem. Acad. Sci. Bologna ser. 2, 1: 228, t. 1 (1862).	Tree Shrub	धुपि Dhupi (Nep.), Pamu	3700-4100		West **Darchula** ; Dipukang- Api **Dolpa**- Daja (3620m)-Pugmi (3250m)-Samduwa (3100m), Kagmara base camp(4650m)- the junction of Kanjirowa base camp (3820m) , Suligadh (3100m) ,Bheding – Na gaon (14000ft)- Chabuk (4300m) , ,,Phoksundo (3600m)– Gandala (4190m) ,Polam (3170m)- Ringmigaon (3630m)-Phoksundo (3612-3620m), Gompa (3260-3630m) Humla : Poom – Kallas (3000m),Simikot (2900m), ,Near to Burma (Karnali river 9000ft) **Manang** : 3666m- 3800m, Pisang (3120m), Pisang(3090m-3130m)-Humre (3210m-3360m)- Deorali danda (3050m) –Bardang (2850m), Chame (2680m) ,Thanchowk -Chame (2730m), Around Bhimtang (3800m) ,, Thanti (4080m) –Tenki (3650m), Phe (4500m) ,Khangsar Khola (3650m – 3700m)- Tilicho base camp(4000m) **Mustang**: Marpha – Thuman (2500m), Chele (3050m)-Samar(3660m) –Bhena (3820m) –Yamda La(3990m)-Syangboche (3840m), ,Chelengaon – Ghilling gaon (3400-4000m) ,Bhena (3670m)- Samar (3450m)-Dozong La	

							(3410m) –Takam La (3340m)- Tsaile /Chaile/Chele (2990m) Phalyak (3140m-3720m) – Sangda pass(a ridge of Jeula Danda 4440m)-Pongio kharka (4410m)-Sangda village(4350m), ,Marpha(2640m)-Alubari kharka(3580m)-Pak Kharka (3600m), Jomsom(2700m), Tukuche (2590m) –Yak kharka (3850m), Chele –Samar (3600m) ,Between Tukuche and Marpha (8200ft) ,Ghemi (3520m) – Ghemi La (3770m) – Nyi La (4010m)- Tamagaon (3710m) – Syangboche (3800m)-Yamda La (3900m) – Bhena (3820m) –Samar (3660m) **Jumla** : Narayani gadh/ Nyari gadh (2670m), Nyor gadh – Sipty (2700m), Rankha – Tatopani (2150m), Maharigaon (2900m), ,,Bibeya (3540m)- Chakare pani (3670m)- a saddle (3600m)- Deula Deuli (3850m) **Mugu** : Rara NP (2900m), Jhyari gaon – Rara (2600m) **Bajhang** : Saipal (3878m), Dubai pass - Sunigad (3100m-4050m) **Baglung** : Dhorpatan (2800m) Central **Rasuwa** – Above Kyangjin Gompa(3500, 3800m-4060m), Chandanbari (13, 600 ft) Lirung Glacier Moraine(3910m)- Kyangjin gompa(3800m), Jaisul Kunda (4350m) ,Lamtang (13,200ft / 3800m)	

							Ramechhap : Thare (4000m-4150m) **Gorkha** :Lungdang Gompa (3150m) – Torogompa Glacier (3820m), Sama gompa (3380m) – Samdo (3680m)- Dharam Sala (4270m) **Dolakha**; Nagaon- Chobuk (4000m), Kalinchowk (3800m) East **Solukhumbu** ; Phorse Tenga-Dhole,,Khumbu (4500m) ,Namche bajar-Thyangboche (11,825 ft – 12,900 ft) ,Phortse Tenga –Dole (Dudh Koshi Valley 3930m), ,Bhote Koshi valley (4530m),Pangboche (3950m) ,Thame khola valley , Khure kharka (4273m), ,Beni Kharka (Dambuk 3970m) –Yurigolcha (4700m), ,Dingboche (4400m) **Sankhuwasabha**: 3900m, Around Ta Dasa – Mera Ding near Makalu Barun CA (3950m-4030m), Dasa (3900m)-Jak kharka (4100m)-Lyang Mo Le Kharka (4260m) –Merek (4450m) Upper Barun khola above Mera **Taplejung** : Ghunsa(3300m)- Charmalung kharka (3620m)- Rampuk kharka (3660m), ,Yangma (4030m) –lowermost moraine (3840m) – Mangyangbuk (3520m)-Thasa kharka (3260m),	

							,Bomrang –Deurali (4280m)	
		Juniperus recurva Buch.-Ham. ex D. Don, ; Prodr. Fl. Nepal.: 55 (1825).	Tree	धुपि Dhupi	3300-4600m	Chitral, Himalaya (Kashmir to Bhutan), Assam, Burma, W. China.	West **Manang** : Pisang – Bhartang (3100m) , Bimtang (3530m)- Suggi khola (2570m)- Ghwa (2360m)- Tiliche (2160m) **Mustang** : Alu bari Kharka (3580m) – Above Yak Kharka (4500m) , Ghami (3410m)- a hill (4090m)- a ridge (4350m)- Zhaite/Chaite (3610m), ,Phalyak (3140m)-Sangda pass (4440m)-Pongio kharka (4410m) Ghemi (3520m) –Ghemi La (3770m)-Nyi La (4010m)- Tama gaon (3710m)-Syangboche La (3850m) –Syangboche (3800m)-Yambda La (3990m)-Bhena (3820m)- Samar (3660m) Central **Rasuwa**: Jaisul kunda (4350m), Gosaikunda (10,500ft-13,650ft-15000ft) , Langtang valley (13,000ft)	

							Chandanbari (3100m, 3300m, 3500m), Rupchet kharka –Bhalchegaon (3600m), Changbu glacier front above Kyangjin (4000m), Around Kyangjin Gompa (3800m)-Langtang (3380m), Solang Pati (3480m) **Gorkha** : Lhogaon (3100m)- Khang sang Fu Gompa (3330m)-Chuwa kharka (3560m) – Thangmanang Kharka (3730m), , Lungdang Gompa (3150m) – Torogompa Glacier (3820m) **Nuwakot** : Rupchet (3600m) **Dhading** : Chhakang (10,000ft) **Sindhupalchowk**: Helambu **Dolakha** : Chhusa-Chhulema, Bheding (3300m) ,Bhitte Kharka – Patlo Pokhari (3600m), ,Kalinchowk –Rupthang (11,600ft), Chhusa – Chhulema (3300m) **Ramechhap** : Botase kharka (4500m)-Koshing kharka (4000m), Neju (3600m), ,Deorali (2700m)- Thodung (3000m)-Serdingma (3400m), ,Chhu Ningma (4040m)-Bike Bhanjyang (4100m)- Jata pokhari (4220m), Man Danda (12,500ft) – Jata Pokhari (13,500ft), Serdingma (3400m)-Dubi Kharka (3720m) ,Thare Og (4150m) – Gyajo La (4877m) – Neju (3651m) ,Man Danda / Nan Danda (12,500ft)- Jata Pokhari (13,500ft), Jhyare Gajo La- Neju (4100m), ,Baula pokhari (3960m) – a peak (4100m)-	

							Chhu Ningma (4040m) East **Solukhumbhu**: Phaphlu-Dudhkunda(3960m) ,Taktor(3000m)- Junbesi (2650m) – Mopung (2800m), ,Sheli (8700)–Junbesi (11,750ft),Pangbuche (13,000ft), ,Beni kharka(3970m) –Sarkari pati (3350m), ,Dudh koshi valley(Mong La to Phortse Tenga , Khumjung to Gokyo 3940m), ,Langmuche - Thame (3874m), Thesebu khola (3651m) ,Thorpe kharka (3545m), Above Namche bajar (3500m), ,Sanasa – Tengboche (3450m), **Sankhuwasabha** : Lamine odar – Ekle Pokhari (3900m), Khamling (3450m), Above Pa khola , ,Khamluy Kharka (3450m) ,Yangle (Barun valley 11,200ft) ,Cha Ding Kharka (3970m)- Tutu La (4070m) -Thulo Pokhari (3980m) -Shipton pass (Keke La) (4127m)- Sano pokhari (3930m) – Khongma (3510m), ,Barun Khola between Repu kharka and lower Barun Glacier (3950m), Yangri kharka (3500m)- Ne kharka (3650m) –Refuk / Ripuck kharka (3840m)-Ta Dasa (3900m), ,Numbuk(3500m)-Barun Khola (3170m)-Phemathang kharka (3380m)- Yangri kharka (3540m)	

							,Jaljale Himal, SHuwan kharka (near Panch pokhari 4180m)- 4360m- Topke Gola (3570m) **Taplejung**: Thasa Kharka(3260m) – Seon Kharka (4030m)-Nango La(4590m)	
		Juniperus squamata Buch.-Ham. ex D. Don, ; in Lamb., Descr. Gen. Pinus 2: 17 (1824); Prodr. Fl. Nepal.: 55 (1825).	Shrub	धुपि Dhupi	3300-4400m	Afghanistan, Himalaya, N. Burma, W. China.	West **Dolpa**- :Phoksundo (3600m/12,500ft) – Gandala (4170m) Phoksundo –she gompa (4000m), **Jumla** : Bibeya(3540m)- (3510m)-Chakare pani (3670m)- a saddle(3600m)- Deula deuli (3850m)-Gompa (3930m) –Kharka (4070m)- a pass (4490m)-Hanka (4380m) **Mustang**:Samar Gaon (3800m)- Syangboche (3840m),Above Muktinath (3600-4000m),Thorungse Fedi (4000m), **Manang** : Chame(2730m) –Bhartang(2970) -Pisang (3090m-3230m)- Humre (3210m)- Manang (3360m) – Gunsang (3720m) – Yak kharka (3840 -3870m) – Alubari (3750m)- Marpha (2650m)- Tukuche (2580m), , Phedi(4400m) - Thanti (4080-4100m) – Tenki (3650m) –Manang (3550m), Central **Rasuwa** : Langtang valley (13,000ft), Kyangjing, Langtang khola (3600-3900m), Changbu glacier (Above Kyangjin3700m) , Dhunche –ShinGompa (3260m)- Gosainkunda (4300m) **Ramechhap** : Thodung (3050m) **Sindhupalchowk** : Tharu pati pass (3300m) East	

							Sankhuwasabha : Yangri kharka (3540m)- (3500m) Chainpur north west 11,500ft (Lekh Lame – Sain gaon 12,000ft)	
		Thuja orientalis	Tree				West: **Banke** : Nepalgunj	Cultivated
9.	Ephedraceae	***Ephedra gerardiana*** Wall. ex Stapf, ; Akad. Wiss. Wien, Math. Naturwiss. Kl. Denskchr. 56(2): 75, t. 3, f. 18 (1889); H. Hara et al., Enum. Flow. Pl. Nep. 1: 24 (1978).	Shrub	Somlata, (Nep); Saankhaapaa, Chhe (She), Kaagh Charo	3700-5200m	Afghanistan, Himalaya (Karakoram to Bhutan).	West **Manang** : Yak Kharka – Thorung Phedi **Jumla** (2300m-2600/ 7,600ft) ,Lihi – Tatopani (2130m) ,Dawa/Daua (2600m) **Dolpa** : Phoksundo–Palan(3500)-Gandala (4170m -4210m), Thuli Bheri (Dunai 2000-2100m), ,Polam (3170m)- RingmiGaon (3630m)-Phoksundo Tal (3612-3620m), ,Doizum (3000m)-Kharka (3200m)-Garpung khola (3750m)- Chune (4020m) **Bajhang** : Naya odar –Topu (3220m)- Dhuli (3333m) Central **Ramechhap** : Around Neju **Rasuwa** : Kyangjin (12,500ft) East **Solukhumbu:** Tangnag – Sabai Tscho	
		Ephedra intermedia Schrenk & Meyer, ; Versuch & Monogr. Gatt. Ephedra: 88 (1846).	Shrub		2700-4000m	Afghanistan, Kashmir, ?Nepal.	West **Mustang**: Muktinath	
10.	Gnetaceae	**Gnetum montanum** Markgr., ; Bull.	Woody Climber	Bhotey Lahara	300-1800m	Himalaya (Nepal to Bhutan), Assam,	Central **Makawanpur** – Hetauda Farm Area (1700ft, 500m),	

		Jard. Bot. Buitenz. ser. 3, 10: 466, t. 8 (1930).				Burma, Thailand, Indo-China, W. & S. China	,Phurke Chaur (390m) **Chitwan-** Along Mul Khola, Near Paidi Bas Choki, Churia Hills, Eastern Boundary (390m) **Tanahu** : Dhorphirdi -7 (near Dhorbarahi mandir, 660m) East **Sankhuwasabha** : Waleng, near Seduwas Seduwa khola (1800m), Between Noom and Sedhuwa (Arun valley 780m), Below Noom (Arun river 2700 ft) **Illam**:Chulachuli region-between Kosi and Kankai Mai **Jhapa**: Ratuwa Khola (1000ft)	