ASSESSMENT OF GENETIC DIVERSITY, DNA BARCODING AND CHEMICAL ANALYSIS OF *NEOPICRORHIZA SCROPHULARIIFLORA* (PENNELL) HONG OF NEPAL

M.Sc. Thesis
2014

Submitted to
CENTRAL DEPARTMENT OF BIOTECHNOLOGY
Tribhuvan University
Kirtipur, Kathmandu, Nepal

For the partial fulfillment of the requirement for degree in
Master of Science in Biotechnology

Gaurav Chandra Gyanwali
Exam Roll No.: BT 031-067
T.U. Regd. No.: 5-1-283-24-2002

ASSESSMENT OF GENETIC DIVERSITY, DNA BARCODING AND CHEMICAL ANALYSIS OF *NEOPICRORHIZA SCROPHULARIIFLORA* (PENNELL) HONG OF NEPAL

M.Sc. Thesis
2014

Submitted to
CENTRAL DEPARTMENT OF BIOTECHNOLOGY
Tribhuvan University
Kirtipur, Kathmandu, Nepal

Gaurav Chandra Gyanwali
Exam Roll No.: BT 031-067
T.U. Regd. No.: 5-1-283-24-2002

Supervisors
Dr. Jyoti maharjan
Dr. Sangita Shrestha
Dr. Rajani Malla
Prof. Dr. Tribikram Bhattarai

ACKNOWLEDGEMENT

I owe special gratitude to higher authorities of NAST; Prof. Dr. Surendra Raj Kafle, Vice Chancellor, NAST, Prof. Dr. Prakash Chandra Adhikari, Secretary, NAST and Mr. Ishwor Prasad Khanal, Science Faculty Chief, NAST for giving me the opportunity to conduct this research work at Molecular Biotechnology Laboratory, NAST.

I would like to extend my sincere and deep gratitude to my respected supervisors Dr. Jyoti Maharjan, Senior Scientific Officer, Molecular Biotechnology Unit, NAST; Dr. Sangita Shrestha, Chief Scientific Officer, Molecular Biotechnology Unit, NAST; Dr. Rajani Malla, Associate Professor and Head of Department, Central Department of Biotechnology and Dr. Tribikram Bhattarai, Professor, Central Department of Biotechnology for their valuable guidance, immense help, constructive and timely suggestions during my research. I'm highly indebted to them.

My special thanks are due to Dr. Rosa Ranjit, Senior Scientific Officer, Natural Product Research Laboratory (NPRL), NAST for guiding me in chemical as well as HPLC analysis during my research work. Her valuable guidance and suggestions were really helpful. I'd also like to thank Dr. Sajan Lal Syaula for his help during HPLC analysis. My sincere thanks also go to Rita Mandal, PhD scholar, NPRL, NAST for suggesting and helping me out during chemical analysis.

I'd like to thank my teachers of CDBt, TU; Dr. Mohan Kharel, Dr. Ganga Kharel, Mr. Balhari Poudel, Dr. Krishna Das Manandhar, Dr. Giri Raj Tripathi, Dr. Sampoornananda Jha, Mrs. Jarina Joshi, Mrs. Smita Shrestha and Dr. Deepak Panta for their guidance and staffs of Central Department of Biotechnology for their kind support. My special thanks are also due to Mr. Mohan Shrestha and Mrs. Ellen Pradhan for their helpful attitude which made lab work less daunting.

I'm highly indebted to Ms. Smita Shrestha, Assistant Research Fellow, Molecular Biotechnology Unit, NAST for her immense help and words of encouragement during my research work. I'd also like to thank Mr. Jagat K.C. Shrestha, Research Assistant, Molecular Biotechnology Unit, NAST for his help and suggestions during data analysis. I'd also like to thank Mrs. Poonam Yadav, PhD scholar, Molecular Biotechnology unit, NAST for her help and support. I express my sincere gratitude and thanks to my friends Surendra Neupane, Nabin Munankarmi, Sunita Khanal, Gorakh Raj Giri, and Rajesh Lamichhane for their friendly support throughout the entire period spent at NAST. I'd also like to express my sincere gratitude to my friend Surendra Raj Sharma for all his words of encouragement and immense help for everything. Special thanks are also due to my friends Arjun Paudel and Parshu Rai for their help and support. Besides, I'd also like to express my sincere thanks to my friend Pradeep Shrestha for helping me with statistical analysis. Finally, I'd like to thank all my friends of 2nd batch as well as my seniors and juniors of Central Department of Biotechnology. I'll treasure all the moments spent with you at the department.

Last but not the least; I'd like to thank my mother Mrs. Renu Gyanwali, my father Mr. Ishwor Chandra Gyanwali and my brother Er. Saurav Chandra Gyanwali for their immense support throughout my research work. I'd also like to thank my cousin Er. Vaskar Gnyawali for his constructive and inspiring words during my research. Their relentless trust in me has always pushed me during my tough times and they are the unending source of my inspiration all these years. I'd like to thank my relatives and all those who have directly or indirectly helped me in my endeavor.

LIST OF ABBREVIATIONS

µg	:	micro gram
µl	:	micro liter
ADP	:	Adenosine Diphosphate
AFLP	:	Amplified Fragment Length Polymorphism
AICc	:	Akaike's Information Criteria corrected
AMOVA	:	Analysis of Molecular Variance
ANOVA	:	Analysis of Variance
AP-PCR	:	Arbitrarily Primed – Polymerase Chain Reaction
asl	:	Above sea level
ATCC	:	American Type Culture Collection
BF	:	Bayesian Factors
BIC	:	Bayesian Information Criteria
BLAST	:	Basic Local Alignment Search Tool
BOLD	:	Barcode of Life Data System
CAMP	:	Conservation Assessment and Management Planning
CAPS	:	Cleaved Amplified Polymorphic Sequences
CBOL	:	Consortium for Barcode of Life
CI	:	Consistency Index
CI_c	:	Consensus Fork Index
CITES	:	Convention on International Trade in Endangered Species of Wild Fauna and Flora
COI	:	Cytochrome *c* oxidase I gene
cpDNA	:	Chloroplast DNA
CS	:	Charakh Samhita
CTAB	:	Cetyltrimethyl Ammonium Bromide
D	:	Dice coefficient of similarity
d.f.	:	Degree of freedom

DAF	:	DNA Amplification Fingerprinting
ddNTP	:	dideoxy Nucleoside Triphosphate
DNA	:	Deoxyribonucleic Acid
dNTP	:	deoxy Nucleoside Triphsophate
DPPH	:	1, 1-diphenyl-2-picryl-hydrazyl
DPR	:	Department of Plant Resources
DT	:	Decision Theoretic Risk
EST	:	Expressed Sequence Tags
ETS	:	External Transcribed Spacer
GSH	:	Glutathione
G_{ST}	:	Genetic differentiation among populations
GTR+I	:	General Time Reversal with invariable sites
H	:	Nei's Genetic Diversity
HBsAg	:	Hepatitis B surface Antigen
HBV	:	Hepatitis B virus
HPLC	:	High Performance Liquid Chromatography
I	:	Shannon's Diversity Index
I_B	:	Band Informativeness
IC_{50}	:	Half maximal Inhibitory Concentration
IGS	:	Intergenic Spacer
ISSR	:	Inter Simple Sequence Repeat
ITS	:	Internal Transcribed Spacer
IUCN	:	International Union for Conservation of Nature
J	:	Jaccard Coefficient of similarity
K2P	:	Kimura 2-parameter
KCA	:	Kanchenjunga Conservation Area
LGT	:	Lateral Gene Transfer
LNP	:	Langtang National Park
LPS	:	Lipopolysaccharide

LRT	:	Likelihood Ratio Test
MAPs	:	Medicinal and Aromatic Plants
matK	:	Maturase K
MCA	:	Manaslu Conservation Area
MEGA	:	Molecular Evolutionary Genetics Analysis
mg	:	milli gram
MHA	:	Mueller Hilton Agar
mL	:	milli liter
mm	:	milli meter
mM	:	milli Molar
MP	:	Maximum Parsimony
MS	:	Mean Square
N_2	:	Nitrogen molecule
NADH	:	Nicotinamide Adenine Dinucleotide - Hydrogen
NADPH	:	Nicotinamide Adenine Dinucleotide Phosphate – Hydrogen
NAST	:	Nepal Academy of Science and Technology
NCBI	:	National Centre for Biotechnology Information
NJ	:	Neighbor Joining
N_m	:	Gene flow
nm	:	Nanometer
NO	:	Nitric Oxide
NOR	:	Nucleolar Organizing Region
NPB	:	Number of Polymorphic Bands
NTFPs	:	Non Timber Forest Products
NTS	:	Non-transcribed Spacer
O_2	:	Oxygen molecule
OH	:	Hydroxide molecule
OTU	:	Operational Taxonomic Unit

PCoA	:	Principal Co-ordinate Analysis
PCR	:	Polymerase Chain Reaction
PI	:	Percent Inhibition
PIC	:	Polymorphic Information Content
PP	:	Posterior Probabilities
PPB	:	Percent Polymorphic Bands
QTL	:	Quantitative Trait Loci
RAPD	:	Random Amplified Polymorphic DNA
rbcL	:	Ribulose-1, 5-bisphosphate carboxylase – Large Subunit
rDNA	:	ribosomal Deoxyribonucleic Acid
RFLP	:	Restriction Fragment Length Polymorphism
RNS	:	Reactive Nitrogen Species
ROS	:	Reactive Oxygen Species
R_P	:	Resolving Power
SCAR	:	Sequence Characterized Amplified Region
SEM	:	Standard Error of Mean
SM	:	Simple Matching coefficient of similarity
SPR	:	Subtree-Pruning-Regrafting
SS	:	Sum of Squares
SSR	:	Simple Sequence Repeat
STS	:	Sequence Tagged Sites
TE	:	Tris EDTA
TM	:	Trade Mark
TNB	:	Total number of Bands
TPM1uf	:	Kimura 3-parameter with unequal frequency
TPM2uf+G	:	Kimura 3-parameter with unequal frequency + Gamma Distribution
TVM+I	:	Transversion Model with invariable sites

UBC	:	University of British Columbia
UPGMA	:	Unweighted Pair Group Method with Arithmetic Mean

LIST OF FIGURES

Numbers at the branches indicate the frequency of occurrence of that branch in all trees (bootstrap support).

Figure 4.10 A strict consensus phylogenetic tree for *rbcL* sequences using TPM2uf+G model under corrected Akaike's Information Criteria (AICc). Branch lengths are the expected number of substitutions per site.

Figure 4.11 A maximum parsimony phylogenetic tree for *rbcL* sequence to depict evolutionary relationship between *N. scrophulariiflora* accessions and reference sequences retrieved from NCBI nucleotide database with GeneBank accession numbers within parentheses. Numbers at the branches indicate the frequency of occurrence of that branch in all trees (bootstrap support).

Figure 4.12 A strict consensus phylogenetic tree for *trnH-psbA* sequences using TPM1uf model under corrected Akaike's Information Criteria (AICc). Branch lengths are the expected number of substitutions per site.

Figure 4.13 A maximum parsimony phylogenetic tree for *trnH-psbA* sequence to depict evolutionary relationship between *N. scrophulariiflora* accessions and reference sequences retrieved from NCBI nucleotide database with GeneBank accession numbers within parentheses. Numbers at the branches indicate the frequency of occurrence of that branch in all trees (bootstrap support).

Figure 4.14 A strict consensus phylogenetic tree using GTR+I model under corrected Akaike's Information Criteria (AICc). Branch lengths are the expected number of substitutions per site.

Figure 4.15 A maximum parsimony phylogenetic tree for **ITS sequence** to depict evolutionary relationship between *N. scrophulariiflora* accessions and reference sequences retrieved from NCBI nucleotide database with GeneBank accession within parentheses. Numbers at the branches indicate the frequency of occurrence of that branch in all trees (bootstrap support).

Figure 4.16 Distribution of K2P distances between *N. scrophulariiflora* accessions and outlier *matK* sequences. A 100% intraspecific K2P distribution between accession sequences is observed.

Figure 4.17 Distribution of K2P distances between *N. scrophulariiflora* accessions and outlier *rbcL* sequences. A 100% intraspecific K2P distribution between accession sequences is observed while an overlapping 33% and non

LIST OF TABLES

LIST OF PLATES

Table of Contents

ABSTRACT

Forty-three samples of *Neopicrorhiza scrophulariiflora* (Pennell) Hong collected from Eastern, Central and Western Nepal were studied for genetic diversity and DNA barcoding studies. Inter Simple Sequence Repeat (ISSR) marker system was employed for the assessment of genetic diversity within and among *N. scrophulariiflora* populations. Out of total 197 amplified bands generated using 10 ISSR primers, 180 bands (91.37%) were polymorphic while 17 (8.63%) were monomorphic. Five primers (UBC 807, 835, 848, 850 and 868) generated ten population-specific ISSR markers for three populations. Using NTSYS-pc version 2.21, Cluster analysis was performed and phenogram based on Jaccard's similarity coefficient and UPGMA clustering method was found to be the best for deducing genetic relationship between *N. scrophulariiflora* accessions. The average value of Jaccard's coefficient of similarity for all accessions was found to be 0.289. The mean genetic similarity between MCA and LNP was 20%, LNP and KCA was 15% and MCA and KCA was 20%. Accessions from Manaslu Conservation Area (MCA) and Kanchenjunga Conservation Area (KCA) clustered together while accessions from Langtang National Park (LNP) clustered separately in phenogram. The result was supported by 3D-plot and Principal Coordinate Analysis (PCoA). The first two axes of PCoA analysis showed 37.63% and 27.24% of total variation respectively with cumulative variance of 64.87%. Population genetic diversity parameters such as Nei's gene diversity (H) and Shannon's diversity index (I) were computed using POPGENE v. 1.32 and was found to be 0.2320 and 0.3636 respectively for MCA, 0.2284 and 0.3413 respectively for LNP and 0.2439 and 0.3836 respectively for KCA. At the species level, H and I values for *N. scrophulariiflora* was found to be 0.2070 and 0.3282 respectively. The mean genetic differentiation between populations (G_{ST}) over all loci was found to be 0.4676 indicating 46.76% variation among populations. The average gene flow (N_m) value was found to be 0.6501. AMOVA revealed the among population variation coefficient, Φ_{PT}, to be 0.501 indicating 50.1% variation among populations. DNA barcoding was performed using three chloroplast (cp) DNA loci, *viz*, matK*, rbcL* and *trnH-psbA* and one nuclear DNA loci the Internal transcribed spacer (ITS). Of the three cpDNA loci, *trnH-psbA* was found to have four variable sites within the sequence while rest of the loci lacked variable sites. Inter-specific and intra-specific divergences and Best Match/Best Close Match values were calculated using TaxonDNA software in order to identify potential barcode for *N. scrophulariiflora*. Analyses of *matK, rbcL, trnH-psbA* and ITS sequences showed high rate of species identification potential for *matK* (75%) followed by ITS (72%). Phylogenetic analyses were performed using model averaged phylogeny (jModelTest 2.1.4) and parsimony based method (MEGA 5.2). The best fitted nucleotide substitution model was computed based on corrected Akaike's Information Criteria (AICc) using jModelTest and was found to be TVM+I for *matK,* TPM2uf+G for *rbcL,* TPM1uf for *trnH-psbA* and GTR+I

for ITS. Maximum Parsimony (MP) trees were constructed for all loci using MEGA 5.2 with Consistency index (CI) value 0.933 for *matK* while MP trees for remaining loci lacked CI values. *matK, trnH-psbA* and ITS resolved all the accessions into a monophyletic clade but *trnH-psbA* sequence showed high intra-specific distance compared to *matK* and ITS thus *matK* and ITS are proposed as potential barcodes for *N. scrophulariiflora*. Antibacterial activity of methanolic rhizome extract of *N. scrophulariiflora* was assessed and was found that *Staphylococcus aureus, Klebsiella pneumoneae*, and *Pseudomonas aeruginosa* were sensitive. IC_{50} values were computed for extracts from three different locations using DPPH method and LNP sample was found to be more potent antioxidant (1.742) followed by KCA (2.398) and MCA (4.141). Quantification of Picroside I and II using HPLC method revealed that Picrosides I and II content were highest in MCA extract (678.705 µg/mL and 2015.947 µg/mL respectively) followed by KCA (426.501 µg/mL and 1029.281 µg/mL respectively) and LNP (203.623 µg/mL and 751.637 µg/mL respectively). Present study has produced valuable insights for the conservation and sustainable utilization of *N. scrophulariiflora* of Nepal Himalaya.

Keywords: *Neopicrorhiza scrophulariiflora*, Kutki, ISSR, DNA barcode, *matK, rbcL, trnH-psbA*, ITS.

CHAPTER 1 – INTRODUCTION

1.1 Introduction

Nepal is a small landlocked country in Southeast Asia extending along the Himalayas within 26^{o} 22' to 30^{o} 27' N latitude and 80^{o} 4' to 88^{o} 12' E longitude. With a varied altitudinal range starting from 60 meters to 8848 meters, varied topography, climatic differences and abundant ecological habitats, a rich floral and faunal diversity is found in small territory with an area of 147,181 sq. km of this country. It has 35 forests types, 75 vegetation units and 118 ecosystems (Gyawali, 2008; DPR, 2012). Nepal is ranked as 9^{th} among the Asian countries for its floral wealth with an estimated 9,000 species of flowering plants (Bhattarai *et al.*, 2013). So far only 6,653 species have been reported (Kunwar *et al.*, 2010). It is estimated that there are about 1950 species of medicinal and aromatic plants in Nepal (Ghimire *et al.*, 2008).

Medicinal and Aromatic Plants (MAPs) possesses medicinal value and thus provides health benefits to the consumer. MAPs are the main constituents of many traditional medicines like Ayurveda, Chinese and Tibetan Medicine, Homeopathy, Unani etc. A wide number of medicinal drugs have been developed from different medicinal plants. For eg: Digitoxin from *Digitalis purpurea*, Podophyllotoxin from *Podophyllum hexandrum* and *Podophyllum peltatum*, Quinine from *Cinchona ledgeriana*, Vinblastine and Vincristine from *Catharanthus roseus* etc (Gyawali, 2008). Various MAPs are used for treating various ailments for example in fever, root juice of *Rauwolfia serpentina*, young leaf of *Ricinus communis,* stem infusion of *Swertia spp*, rhizome and root decoction of *Neopicrorhiza scrophulariiflora* etc are used; in case of diarrhea and dysentery, fruit pulp of *Aegle marmalos*, fruit juice of *Phyllanthus emblica,* milky juice of *Ficus bengalensis* with sugar etc are used; in case of cough and cold, wood decoction of *Acacia catechu*, leaf juice of *Euphorbia royleana*, chewing of fruit of *Piper pepuloides*, plant decoction of *Trachyspermum ammi* etc. are used (Gyawali, 2008). Such traditional knowledge can be the basis for potential drug discovery in combating several human ailments (Katiyar *et al.*, 2012).

Nepal, having varied geo-climatic conditions, is a storehouse of various high value medicinal plants species. The approximate number of medicinal and aromatic plants (MAPs) reported in Nepal is 1950 (Ghimire, 2008). Many species of medicinal plants found in Nepal, such as *Swertia chirayita* (Roxb. Ex. Fleming) H. Karst, *Neopicrorhiza scrophulariiflora* (Pennell) Hong, *Podophyllum hexandrum* Royle, *Taxus wallichiana* Zucc., *Nardostachys grandiflora* DC etc., are highly demanded and traded (DPR, 2012). The world market for herbal medicines and their derivatives is estimated to have an annual growth rate between 5 and 15%. The total global herbal drug market is estimated to be US $62 billion and is expected to grow up to US $5 trillion by 2050 (Joshi *et al.*, 2004). Traditional Chinese medicine has its market share value of US $4.3 billion and is

projected to be 12.1 billion by 2020 (Nagpal and Karki, 2004). This has also resulted into higher demand of medicinal plant trade for Tibetan Medicines (Pandey, 2006).

Among all these medicinal plants found in Nepal, *Neopicrorhiza scrophulariiflora* (Pennell) Hong is one of the high value medicinal plants. It has been prioritized for research and development, agro-technology development and has been banned for export without identification and certification by Government of Nepal (DPR, 2012). It belongs to tribe Veroniceae of family Scrophulariaceae, which is one of the largest families found in Nepal (Shakya *et al.*, 1997; DPR, 2012). *Neopicrorhiza scrophulariiflora* is commonly known as "Kutki" and is found at a height of 3000 – 5000m above sea level (Luper, 1998). *Neopicrorhiza scrophulariiflora* possesses many medicinal properties and has been mentioned in Ayurveda (Katuka) and Chinese and Tibetan medicine. In Ayurveda, it is used in respiratory disorders, in anorexia, as appetizer, in abdominal swellings, in skin disorders, in wound healing, in obesity, in poisoning, vaginal and uterine disorders, purifying breast milk as mentioned in Charakh and Susrut Samhita (Smit, 1968). It has been officially listed in Pharmacopoeia of the People's Republic of China for the treatment of fever, jaundice, hemorrhoids and dysentery by Traditional Chinese and Tibetan Medicine (Liu *et al.*, 2011). Similarly, it has also been implicated in many health benefits like hepatoprotection, viral hepatitis, regulated bile secretion, antimalarial action etc due to many properties like antioxidant, stimulation of liver regeneration, anti-inflammatory, choleretic (Luper, 1998), immunomodulatory action (Smit *et al.*, 2000) etc.

Due to enormous health benefits and high value, it is one of the highly traded plants in Nepal. A recent report showed that in fiscal year 2066/2067 B.S. alone, a total of 2,171,522.3 Kg of NTFPs were collected from national forests (DPR, 2012). Nepal is main supplier of 'kutki' (66%) followed by India (19%) and Bhutan (14%) (Olsen, 2005). In fiscal year 2063/2064, a total of 13,364 kg of *N. scrophulariiflora* rhizomes were collected (Department Of Forestry, undated). Large collection of *N. scrophulariiflora* for trade has rendered it 'Vulnerable' (Shrestha and Joshi, 1996; Bhattarai *et al.*, 2002). While being more commonly traded than a related plant *Picrorhiza kurrooa*, TRAFFIC the wildlife trade monitoring network, has identified the need to monitor *N. scrophulariiflora* with a view to inclusion in the CITES list (Mulliken, 2000). It is classified as endangered species in the Chinese Plant Red Book (Fu, 1992). Apart from domestic market, the unprocessed rhizomes of this plant are traded to many countries including India, China, Japan, and Netherlands (Olsen, 2005). The health benefits of 'kutki' is attributed to its active constituents like Picrosides I, II, III and kutkosides which is collectively known as 'kutkin' along with cucurbitacin glycosides, apocynin, drosin etc (Luper, 1998).

Because of the impact of global warming and anthropogenic activity, natural populations of this species have suffered from rapid decline (Liu *et al.*, 2011). In Nepal, over collection of this plant by the traders due to high market demand from Indian Pharmaceutical Industry and Chinese market demand for their traditional system of medicine and the impact of global warming and changing climatic conditions can be considered as additional threats for its vulnerability. The survival index of a species is greatly determined by the percentage of polymorphism and the gene flow between populations. Thus, understanding genetic diversity of rare, vulnerable and endangered plants is vital for devising effective conservation practices (Sebastian *et al.*, 2010; Liu *et al.*, 2011). Study of biodiversity is greatly aimed at preservation of existing genetic diversity. Survival of a species depends on the maintenance of genetic variability within and among populations that accommodate new selection pressures brought about by environmental changes. Information on how genetic variation is distributed among the existing population of an endangered species can be used in designing recovery programs (Sebastian *et al.*, 2010).

Genetic diversity can be assessed in various ways. A wide range of methods have been employed to study genetic diversity of plants by many researchers in the past. One method is to use the genetic markers which are identifiable DNA sequences found at specific locations in the genome and are transmitted by the standard laws of inheritance from one generation to another (Semagn *et al.*, 2006). Among many techniques, Inter Simple Sequence Repeat (ISSR), a Polymerase Chain Reaction (PCR) - based marker technique, involves amplification of DNA segment present at an amplifiable distance in between two identical microsatellite repeat regions oriented in opposite direction (Zietkiewicz *et al.*, 1994). ISSR has numerous benefits which combines the benefits of Amplified Fragment Length Polymorphism (AFLP) and microsatellite analysis with the universality of Random Amplified Polymorphic DNA (RAPD). It has high reproducibility usually attributed to factors like use of longer primers (16-25 bps) which results in high melting temperature (45 – 60^{o}C) thus leading to stringency. The process is simple, quick and efficient and there is no need of radioactivity as well (Reddy *et al.*, 2002). ISSRs have been used in estimation of genetic diversity at inter- and intra-specific level including crop species like *Oryza* spp. (Joshi *et al.*, 2000), wheat (Nagaoka and Ogihara, 1997) as well as in the study of medicinal plants like *Dendrobium* sp. (Yang *et al.*, 2010), *Swertia* spp. (Joshi and Dhawan, 2007; Zhang *et al.*, 2007; Tamhankar *et al.*, 2009), *Podophyllum hexandrum* (Alam *et al.*, 2009; Naik *et al.*, 2010), *Rheum* spp. (Wang *et al.*, 2012a; Wang *et al.*, 2012b) etc.

Apart from studying the genetic diversity of a plant, its identification is also crucial. Identification of plant species is important in plant ecology, agriculture, conservation of endangered plants, forensic enforcements etc. Identification of plant species is generally

done based on morphology which requires special training in plant morphology and taxonomy or on-hand identification tools. But recently, a DNA based marker called "DNA Barcode" has been developed which is a molecular and bioinformatics tool aimed at identifying species (Hebert *et al.*, 2003). The sequence is compared to existing standard sequences in a database system called Barcode of Life Data system (BOLD) to identify species. This system allows people without expert taxonomic knowledge to identify plant from materials that may not be flowering, or may be in seed or seedling stage, plant fragments in foods, medicines etc. The variability in a single or few standard markers is used to discriminate biological entity such that the inter-specific variation exceeds intra-specific variation by a distance called 'barcoding gap' (Meyer and Paulay, 2005; Wiemers and Fiedler, 2007). *Neopicrorhiza scrophulariiflora* (syn. *Picrorhiza scrophulariiflora*) and a closely related species *P. kurrooa*, are similar to each other in that rhizomes of both species are comparable with regard to their morphology, histology, accumulated metabolites, bitterness activity and traditional use (Zhang *et al.*, 1965; Wang *et al.*, 2004). Morphologically also these two plants appear similar but the differences lie in the corolla and stamen lengths. *Neopicrorhiza scrophulariiflora* has longer, 4-lobed, bilabiate corolla while in *P. kurrooa* it is shorter, 5-lobed and nearly actinomorphic. The stamens in *N. scrophulariiflora* are slightly didynamous and equal in length to corolla but in *P. kurrooa* stamens are many times longer than corolla. The necessity to generate *N. scrophulariiflora* specific barcode originates from the fact that processed rhizome powder, traded as "Kutki", is a mix of rhizomes of *P. kurrooa* and *N. scrophulariiflora*. *Picrorhiza kurrooa* is enlisted in CITES Appendix II while *N. scrophulariiflora* is a vulnerable plant (Olsen, 2001; DPR, 2012). Identification of *N. scrophulariiflora* by barcode marker not only documents Nepalese *N. scrophulariiflora* species but also helps prevent illegal trade and collection of otherwise CITES listed *P. kurrooa* by differentiating it from *N. scrophulariiflora* species.

The present research is aimed at understanding the genetic diversity and structure within and among populations of *N. scrophulariiflora* of Nepal. Studies at molecular level of this species haven't been performed so far for Nepalese populations. In order to devise an effective conservation strategy of this high value medicinal plant, knowledge of its present state of polymorphism is imperative. Therefore, we employed ISSR-PCR marker system to assess genetic diversity as well as to search species-specific marker for authentication. Besides, with the advent of molecular taxonomy, the proper genetic marker necessary to identify *N. scrophulariiflora* species is equally necessary. In the present study, as suggested by Consortium for Barcode of Life (CBOL) Plant working group (2009), *matK* and *rbcL*, the two core barcode markers along with supplementary barcode markers *trnH-psbA* and Internal Transcribed Spacer (ITS) were assessed to determine the potential barcode markers to identify *N. scrophulariiflora*. Furthermore,

the study was also targeted to quantify couple of active constituents that impart health benefits to consumer thereby assessing which population possessed these compounds in higher quantity than other.

1.2 Research hypothesis

1) H_0: The vulnerability of *N. scrophulariiflora* populations of Nepal is not due to lack of genetic diversity; H_1:. The vulnerability of *N. scrophulariiflora* populations of Nepal is due to lack of genetic diversity.

2) H_0: There is no nucleotide sequence variation in all tested loci of different *N. scrophulariiflora* populations of Nepal; H_1: There is nucleotide sequence variation in all tested populations of Nepal.

H_0: Null Hypothesis; H_1: Alternate Hypothesis

1.3 Objectives

The overall objective of this study is to characterize Nepalese *N. scrophulariiflora* populations of Eastern, Central and Western Nepal at molecular level (using PCR based ISSR and DNA sequencing based Barcoding technique) and chemical level (using HPLC technique) and to study their antimicrobial and antioxidant properties. Specific objectives of this project are:

- To employ ISSR marker technique for the assessment of population genetic diversity and structure of populations under study and to optimize ISSR-PCR reaction and cycling parameters for *N. scrophulariiflora* and use the optimized parameters for ISSR profiling of *N. scrophulariiflora* collected from three development regions of Nepal and also to use the ISSR generated molecular data for assessing genetic diversity using NTSYS-pc software and population genetic diversity using POPGENE and GenAlex software.
- To generate *matK, rbcL*, *trnH-psbA* and ITS sequences for barcoding and phylogenetic studies.
- To assess the antimicrobial and antioxidant property of rhizome extract of *N. scrophulariiflora* collected from different locations.
- To quantify Picrosides I and II in rhizome extracts in different populations using standard samples of Picrosides I and II.

1.4 Justification of the Study

Medicinal and Aromatic plants (MAPs) of Nepal suffer a great threat of depletion and endangerment due to various anthropogenic activities pertaining to the illegal trade as

well as prevailing climatic change impacts. *Neopicrorhiza scrophulariiflora*, one of many such high value medicinal plants, is not only used locally for traditional medicines but also for international trade in bulk amounts (Olsen, 2001; Ghimire *et al.*, 2005; DPR, 2012). Harvesting for traditional use by amchis is limited and conservation friendly but harvesting for trade by collectors is unscientific and indiscriminate beyond its regenerative ability (Ghimire *et al.*, 2005) rendering it to be a "Vulnerable" species (Shrestha and Joshi, 1996; Bhattarai *et al.*, 2002). Besides, the unprocessed rhizome is mixed with rhizome of CITES listed similar plant *P. kurrooa* of Indian origin (Olsen, 2001).

In this scenario, understanding genetic diversity and structure at the population and specific level is imperative to know the potentiality of this species to withstand anthropogenic and climatic impact and allows formulation of effective conservation strategy. In addition, sequence based study generates direct evidence of genetic variation existing among different populations. This allows positive identification of species that is needed to many end users of different disciplines including scientists and enforcement agencies. Thus, the present study is aimed at studying genetic diversity within and among populations of *N. scrophulariiflora* so that populations with lesser diversity could be tailored with priority based conservation practices. The trade of two related species of *N. scrophulariiflora* and *P. kurrooa* by mixing their powdered rhizome poses threat to both, *N. scrophulariiflora* being vulnerable and *P. kurrooa* being CITES listed. *Neopicrorhiza scrophulariiflora* specific barcode markers can be used to identify *N. scrophulariiflora* which in turn can also help identify *P. kurrooa*.

Similarly, antimicrobial studies in medicinal plants can potentially lead to identification of lead compounds which can be modified and further developed into a new therapeutic agents. Antioxidant studies help in identifying medicinal plants with high antioxidant property that can quench harmful free radicals which are implicated in many diseases like liver cirrhosis, atherosclerosis, cancer, diabetes etc. (Govindarajan *et al.*, 2003). Determination and quantification of active constituents is necessary part of research related to medicinal plants. Comparative analysis of desirable chemical compounds in different populations enables selection of elite genotypes and best environment for cultivation of *N. scrophulariiflora* for business purpose.

1.5 Scope of the Study

Genetic diversity study of medicinal and aromatic plants is highly imperative in the present scenario of the world. On one side, medicinal plants are associated with economy of individuals and as a whole with that of country while on the other side, overexploitatioin for trade has rendered many such plants vulnerable to endangerment. Thus, studies related to diversity and genetic structure of plants has been widely acclaimed for devising effective conservation strategies around the world. Thus, this

study has opened a new avenue for research on many other endangered and critical medicinal plants. Barcoding studies have been recently visualized as an effective tool for identification of plants. Besides the general use of identification, it has been very helpful with identifying adulterants of closely related medicinal plants. Barcoding studies of most if not all medicinal plants of Nepal can someday provide valuable data and methods to rapidly identify medicinal plants banned for trade and collection thus aiding enforcement of conservation law. Study of antimicrobial and antioxidant property along with identification of new active compounds can aid in development of potential new and effective drugs and ayurvedic formulations particularly aimed at treating specific conditions.

CHAPTER 2: LITERATURE REVIEW

2.1 The Family Scrophulariaceae

The Scrophulariaceae family, commonly known as figwort or snapdragon family, is a family of flowering plants comprising of annual or perennial herbs. The family is named after one of its included genera, *Scrophularia* (Internet visit 1, 2013). Members of the family are cosmopolitan in distribution, ranging from temperate to high mountains. The family is large with around 4500 species and 220 genera (Hong *et al.*, 2014). A list of all the genera can be found on the website of Royal Botanical Garden, Kew (Internet visit 2, 2013). Members of this family are non-medicinal but mostly ornamental except some genera such as *Picrorhiza, Neopicrorhiza, Digitalis* etc. It also includes another famous genus *Paulownia*. Plants of the family are mostly herbs while some are shrubs and trees (eg: *Paulownia*). The leaves are alternate, opposite or sometimes whorled and are simple to pinnately divided. The family is characterized with irregular, bilaterally symmetrical flowers with 4-5 sepals and petals. The flowers are bisexual and have two pairs of anther-bearing stamens, and a sterile fifth stamen – a taxonomically important feature. All these parts are attached at the base of the ovary. The fruit type is usually a 2-chambered capsule.

2.2 Taxonomic notes on genus *Picrorhiza* and *Neopicrorhiza*

P. kurrooa and *N. scrophulariiflora* belong to family Scrophulariaceae. This family is arranged in the order Scrophulariales, subclass Asteridae and Class Dicotyledonae (Smit, 1968). This genus *Picrorhiza* was considered monotypic with *P. kurrooa* as its only species. But in 1943 Pennell distinguished another species called *P. scrophulariiflora* (Pennell, 1943). Later, Hong described separate genus called *Neopicrorhiza* in 1984 (Hong, 1984). *N. scrophulariiflora* and *P. scrophulariiflora* are generally used interchangeably. Hence, these names are referred to as basionyms. Since the rhizome of both of the species is comparable to their morphology, histology, accumulated metabolites, bitterness, and activity, the use of basionym for this species is justifiable (Smit, 1968).

Plate 2.1 *Neopicrorhiza scrophulariiflora* as found in natural habitat. The plants seen here were collected from Langtang National Park, Rasuwa. Picture on the right shows the capsule (fruit) of the plant along with the rhizome.

Later, Bentham described the genus and species in "Scrophularineae Indicae". In 1876, Bentham and Hooker described *Picrorhiza* in "Genera Plantarum" as a monotypic species until later Pennell distinguished another species, based on Hooker's "Flora of British India" (Pennell, 1943). Hooker observed that the flowers of *P. kurrooa* were dimorphic, which was a new observation. He observed two forms where one form had long stamens and a short corolla with five sub-equal lobes and another one had short stamens and a bilabiate corolla, the upper lip of which is longer while the lower lip consisted of three shorter lobes (Hooker, 1885). Pennell later observed that sample collections from the Western Himalaya had long stamens while those collected from Eastern Himalaya and Yunnan had short stamens. Based on these observations, he characterized two species: dry western type (*P. kurrooa*) and moist eastern Himalaya type (*P. scrophulariiflora*) (Pennell, 1943).

Since Bentham was first to give written description of species, the accepted species name is "*Picrorhiza kurrooa* Bentham" or alternatively, "*Picrorhiza kurrooa* Royle ex Bentham". But, according to the International Code of Botanical Nomenclature article 42, the correct assigned name of the species is "*Picrorhiza kurrooa* Royle" (Smit, 1968). The name *Picrorhiza* is derived from Greek where "picros" means bitter and "rhiza" means root (Smit, 1968).

Later, Hong (1984) observed additional differences in pollen grains of both species and used as evidence to place both taxa into two distinct genera: *Picrorhiza* and *Neopicrorhiza*. Both of these genera are monotypic. Hong observed that *P. kurrooa* had corolla almost in actinomorphic form with a short corolla tube and five equal lobes along with stamens which were three times longer than corolla (Hong, 1984). Currently, *Neopicrorhiza scrophulariiflora* (Pennell) Hong is assigned as official name (Brummit, 1992).

2.3 Description of *N. scrophulariiflora*

The plant is characterized by following features: Perennial herb with stout and elongated rhizome. Leaves are all basal and rosulate, spikes terminal. Flowers bracteates, bracteoles absent, calyx lobed to near base, corolla 2-lipped; upper lip much longer than lower, emarginated. Stamens 4; anterior 2 exceeding lower lip; posterior 2 slightly shorter than upper lip; anther locules basally divaricate, apically confluent. Ovary 2-loculed; ovules are numerous. Capsule is apically septifragal. Seed surface is hyaline reticulate (Hong, 1984).

N. scrophulariiflora is similar to another plant *P. kurrooa* in morphology however; the differences that distinguish these two genera are due to the corolla and stamens (Table 2.1).

Table 2.1 Differences in features between *P. kurrooa* and *N. scrophulariiflora*

Features	***P. kurrooa***	***N. scrophulariiflora***
Corolla	4 – 5 mm long; 5 – lobed; Nearly actinomorphic	9 – 10 mm long; 4 – lobed; Bilabiate
Stamens	Many times longer than corolla	Slightly didynamous; Equaling corolla

2.4 Distribution of *N. scrophulariiflora*

2.4.1 World Distribution

Both the genera of *Picrorhiza* and *Neopicrorhiza* are high altitude plants. Distribution of both plants is confined to Himalayan region. *P. kurrooa* is mostly distributed along Western Himalaya at an altitudinal range of 3000-4300 m asl (Pennell, 1943). It is distributed along Pakistan (Jammu and Kashmir) to India (Himachal Pradesh to Uttar Pradesh) (Smit, 1968).

N. scrophulariiflora is distributed mainly through Eastern Himalaya to the mountains of Yunnan at an altitude of 4300 – 5200 m asl (Pennell, 1943) and from Uttar Pradesh to Southwest China, 3600 – 4800 m (Polunin and Stainton, 1990) and through Nepal, Sikkim, and Bhutan (Smit, 1968).

2.4.2 Distribution in Nepal

In Nepal, *N. scrophulariiflora* is distributed throughout western, central and eastern Nepal (Hara *et al.*, 1982; Press *et al.*, 2000) with climatic condition ranging from subalpine to alpine. The distribution altitude varies from 3500 – 4800 m (Hara *et al.*, 1982; Press *et al.*, 2000; Ghimire *et al.*, 2008). Different authors have reported distribution of *N. scrophulariiflora* in different locations of Nepal. Distribution of this genus in Eastern Nepal has been described by Ohba and Akiyama (1992) to be around Jaljale Himal (Jaljale - Tin Pokhari at 4030 – 4130 m asl, around Banduke at 4150 m asl). In Central Nepal, it is found to be distributed along Rasuwa district and Langtang area (Bhattarai, 1993; Suwal, 1993) around Gosainkunda area and Laurivinayak area (Malla *et al.*, 1976). In western Nepal, Kihara and Olsen reported its distribution in Gorkha district along with Thaple Himal and Murkhaghari (Kihara, 1955; Olsen, 1998). Similarly, it has also been reported in Jumla district by Manandhar (1986) (Figure2.2). *Neopicrorhiza* is a monotypic genus with *N. scrophulariiflora* as its only species (Hong, 1984).

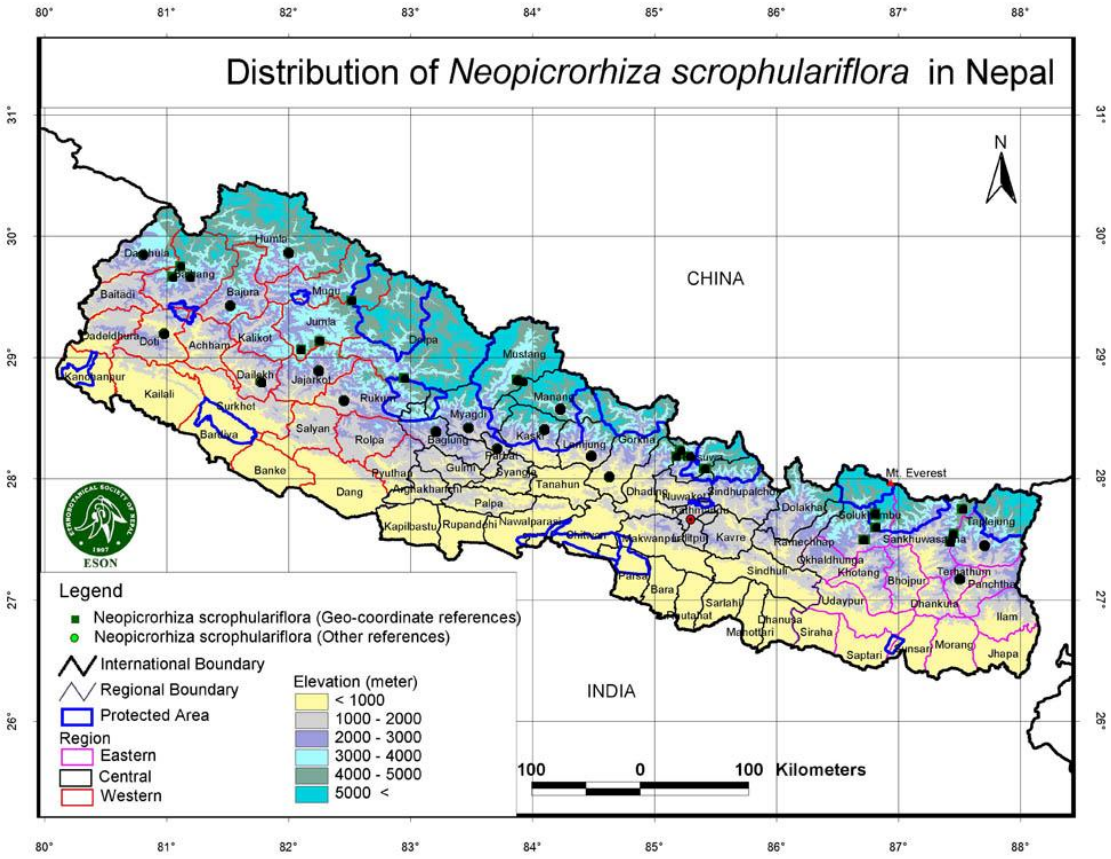


Figure 2.1 Distribution of *Neopicrorhiza scrophulariiflora* in different locations of Nepal.

2.5 Traditional uses of *N. scrophulariiflora*

N. scrophulariiflora (commonly known as kutki) is used traditionally in many different systems of medicine. WHO defines traditional medicine as: "the sum total of the

knowledge, skills and practices based on the theories, beliefs and experiences indigenous to different cultures, whether explicable or not, used in the maintenance of health as well as in the prevention, diagnosis, improvement or treatment of physical and mental illness" (WHO, 2013). Traditional medicinal system encompasses various medical systems like Ayurveda, Traditional Tibetan and Chinese medicine, the Japanese Kampo, the Unani etc. Similarly, folk medicine, ethnomedicine, community medicine etc are also considered as traditional medicine (Gyawali, 2008).

N. scrophulariiflora/P. kurrooa is known as "*Katuka*" in *Ayurved*. In Sanskrit, various other names have been given to it such as *Katuka-rohini, Katuki, Matsyapitta, Tikta* etc. In Nepal, it is known as *Kutki, Kaduki, Katuki, Katuko* etc. Various vernacular names are used by various tribes such as *Hong len (Amchi), Hung gung (Bhotia), Kutaki (Gurung), Hunling (Sherpa), Kuraki (Tamang)* etc.

Medicinal use of kutki (as *"Katuka"*) is mentioned in both *Charakh Samhita* (CS) as well as *Susrut Samhita* (SS). In *Susrut Samhita*, it is mentioned as *katurohini* and *rohini* (as indicated in *kapha* and *vata* disorders, in respiratory disorders, in anorexia, as appetizer, as digestive and in abdominal swellings) or as *patoladigana* (in *kapha* and *vata* disorders, in anorexia, in fever, in skin disorders, in vomiting, in pruritus and in poisoning), or as *nyagrodhadigana* (in wound healing, as astringent, in fractures, in blood disorders, in burning sensation, in obesity and in vaginal and uterine disorders) and as *mustadigana* (in *kapha* disorders, vaginal and uterine disorders, purifying breast milk and as digestive) group of drugs. Besides that, both *Charakh* and *Susruta* have grouped *katuka* in bitter groups: *tikta skandha* and *tikta vargah* respectively (Smit, 1968).

In Traditional Chinese and Tibetan Medicine, it has been listed for the treatment of fever, jaundice, hemorrhoids and dysentery (Chinese Pharmacopoeia Commission, 2005).

2.6 Chemical Constituents of *N. scrophulariiflora*

Most of the publications on the biological activities and constituents of *Picrorhiza* deal with the species *P. kurrooa*, while *N. scrophulariiflora* (basionym *P. scrophulariiflora*) is scarcely described. However, the identification of the raw material has usually not been elaborated in these publications. Quite often, plant material is obtained through a chain of buyers and sellers, and ultimately there is no information on the origin of the material. Generally, no distinction is made between samples from Nepal (*N. scrophulariiflora*) and India (*P. kurrooa*). A proper identification can only be made when selective criteria are found to distinguish between rhizomes from both species, e.g. microscopically, chemically, or by DNA analysis. However, these criteria have not been

met so far. The difficulty in distinguishing both species from each other based on rhizomes only makes interpretation of experimental results on *Picrorhiza* species ambiguous. This uncertainty has to be taken into consideration when referring to publications on *P. kurrooa*. In cases where the origin of the material in not clearly described, there is a possibility that the material investigated is erroneously substituted for *N. scrophulariiflora*/*P. scrophulariiflora* (Smit, 1968).

Genus *Picrorhiza* has been widely studied for its phytochemistry and a total of 132 constituents from different parts of plants have been characterized. Chemical study on *P. kurrooa* rhizomes showed presence of iridoids (Gupta, 2001), acetophenones (Basu *et al.*, 1971), cucurbitacins (Stuppner and Wagner, 1989a; Stuppner and Wagner, 1989b; Stuppner *et al.*, 1990). Similarly, *P. kurrooa* also contains pikuroside, veronicoside, phenol glycosides, cucurbitacin glycosides and 4-hydroxyl-3-methoxy-acetophenone (Bantawa *et al.*, 2010; Gaikwad *et al.*, 2010) whereas *P. scrophulariiflora* contains cyclopentanoid monoterpenes, caffeoyl glycosides, phenylethanoid glycoside and plantamajoside (Li *et al.*, 1998; Wang *et al.*, 2004; Zhu *et al.*, 2008). The different classes of compounds found in *N. scrophulariiflora* are briefly discussed below.

2.6.1 Cucurbitacins

Cucurbitacins are steroid and are included in triterpene group of compounds. They are especially known for their bitterness and toxicity. They are characterized by the cucurbitane skeleton and are mainly tetracyclic and oxygenated at different sites. The basic structure is 9β-methyl-19-norlanosta-5-ene, also known as 10*α*-cucurbit-5-ene (Figure 2.2) (Lavie and Glotter, 1971).

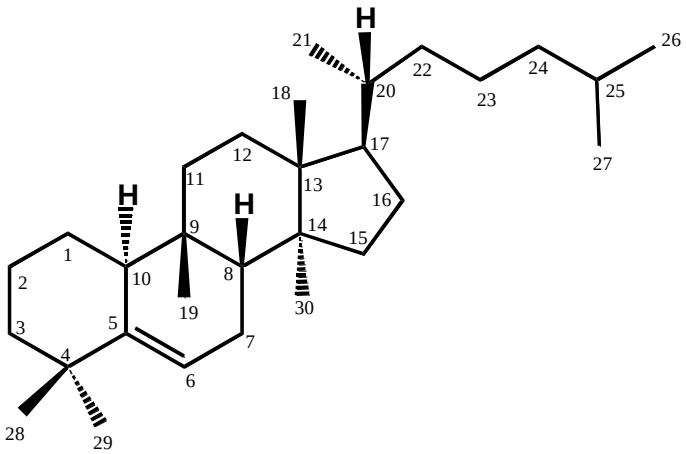


Figure 2.2 Basic skeleton of the 10*α*-cucurbit-5-ene skeleton (9β-methyl-19-norlanosta-5-ene).

From both *P. kurrooa* and *P. scrophulariiflora*, many different cucurbitacin glucosides have been isolated. Studies on *P. kurroa* roots using different solvents like methanol, ethanol, butanol, ethyl acetate, and chloroform have yielded different types of

cucurbitacin glycosides including arvenin III (Sah and Varshney, 2013). Similarly, from *P. scrophulariiflora* also many different cucurbitacin glucosides have been obtained from roots using different solvents. For e.g., from roots using butanol, 2β-glucopyranosyloxy-3,16,20,22-tetrahydroxy-9-methyl-19-norlanosta-5,24-diene has been isolated (Wang *et al.*, 1993) while from rhizome using butanol 2-β-glucosyloxy-3,16,20,25-tetrahydroxy-9-methyl-19-norlanosta-5,23-diene-22-one (Hu and Yang, 2005) along with other cucurbitacins have been isolated. Similarly, other cucurbitacins have also been isolated from stem using ethyl acetate or ethanol (Sah and Varshney, 2013). In *N. scrophulariiflora*, cucurbitacin glycosides like scrophoside A along with other cucurbitacins have been reported (Kim *et al.*, 2006).

2.6.2 Iridoid Glucosides

Iridoids represent a large group of monoterpenes and are found accumulated in number of plants mainly belonging to subclass Asterideae. They consist of basic cyclopentano-[c]-pyran skeleton, which is formed during ring closure of 8-*epi*-iridodial and usually occur as glucosides (Smit, 1968).

From *N. scrophulariiflora,* many different iridoid glycosides have been isolated using different solvents and from different parts of plants like roots, rhizomes, stems etc. Iridoid glycosides like Picroside I and II, kutkoside, picrorosides A, B and C has been reported (Kim *et al.*, 2006) from the plant along with four non-glycosidic iridoids like piscrocins D, E, F and G. Other iridoid glycosides like piscrosides A and B were also isolated from roots of *N. scrophulariiflora* (Wang *et al.*, 2006). Study performed on *P. scrophulariiflora* identified new group of compounds called caffeoyl glycosides (scrocaffeside A, B and C) together with two caffeic acid derivatives, 4-O-beta-D-glucopyranosyl caffeic acid and 4-methoxy caffeic acid (Zhu *et al.*, 2008). A different class of compound called secoiridoid glycosides has been isolated from *P. scrophulariiflora* and named as picrogentioside A, B and C with a preliminary immunomodulatory effect *in vitro* (Zou *et al.*, 2008). From the roots and rhizomes, a crystalline product “kutkin” or “Picroliv” is obtained which is a mixture of two major C9-iridoid glycosides, i.e. Picroside I (6-*O*-trans-cinnamoylcatalpol) and “Kutkoside” (10-*O*-vanilloylcatalpol) in the ratio of 1: 1.5 (Dhawan, 1995). Kutkin is a mixture of pircoside I and picroside II. The chemical structures of picroside I, picroside II and kutkoside are shown in Figure 2.3 below.

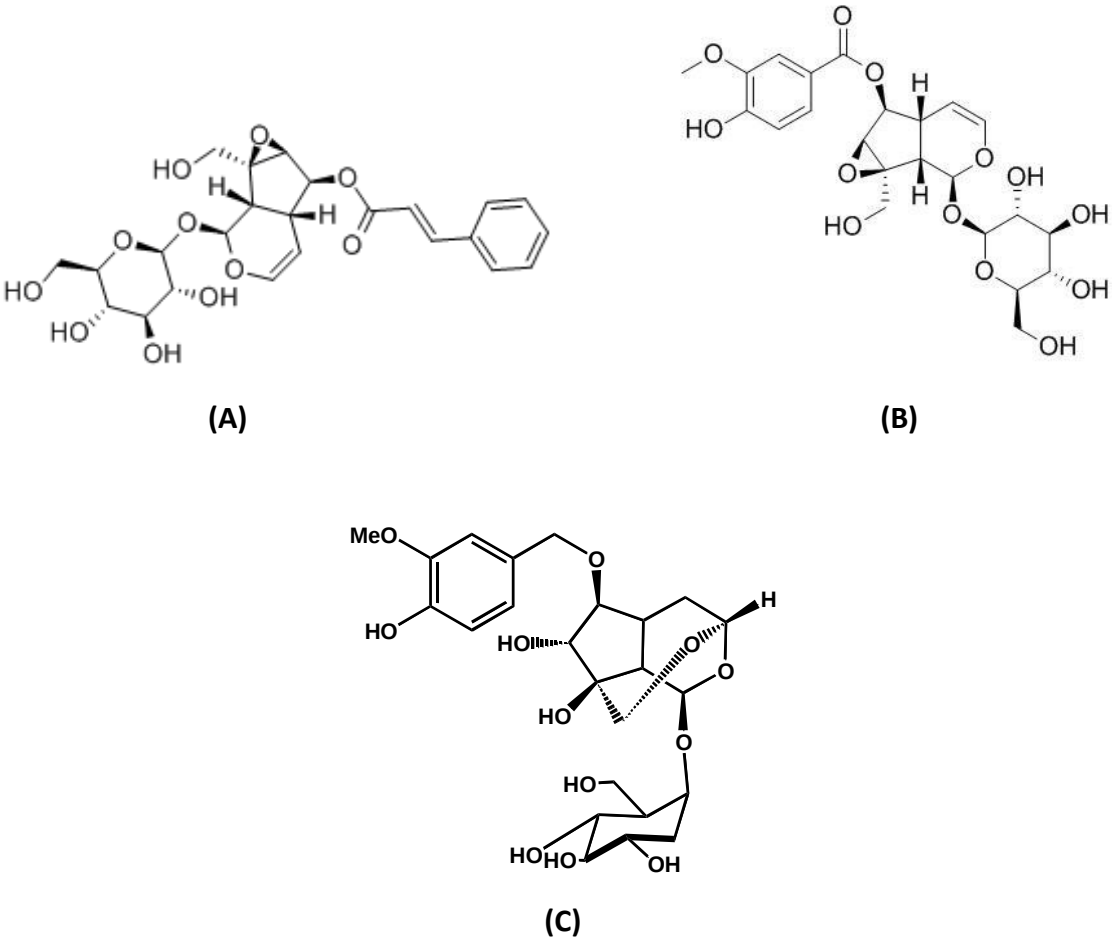
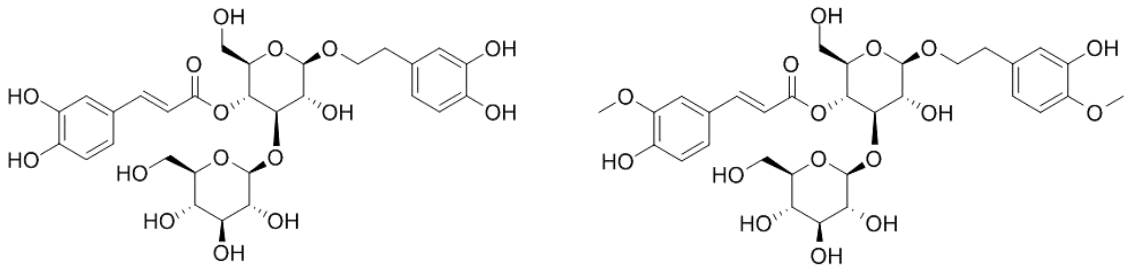


Figure 2.3 Chemical structures of (A) Picroside I, (B) Picroside II and (C) Kutkoside.

2.6.3 Phenylethanoids

Another classes of compounds found in *N. scrophulariiflora* and *P. kurrooa* are phenylethanoids and phenylpropanoids. In studies involving *P. scrophulariiflora,* phenylethanoid glycosides isolated from root involved scrosides A, B, and C and plantamajoside (Li *et al.*, 1998), scroside G (Huang *et al.*, 2009), and plantainoside D (Wang *et al.*, 2004). Similarly, scrosides D and E were found in stem along with phenylpropanoid glycoside such as hemiphroside (Huang *et al.*, 2004). Other phenyethanoid glycosides from stem included plantainoside D4 (Zou *et al.*, 2007).

From *N. scrophulariiflora,* phenylpropanoid glycosides such as scrophulosides A and B have been isolated along with androsin (Kim *et al.*, 2006). Chemical structures of plantamajoside and hemiphroside A are shown in Figure 2.4.



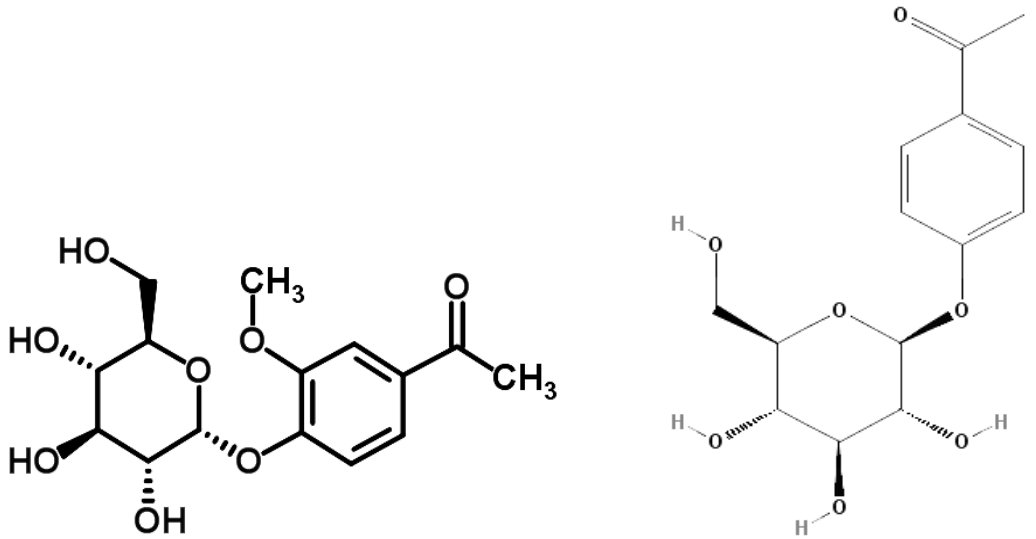
(A) **(B)**

Figure 2.4 Chemical structures of (A) Plantamajoside and (B) Hemiphroside A.

2.6.4 Phenolic Compounds

Monocyclic phenolic compounds are precursors and degradation products of lignin, which provides sturdiness to the plant and a physical defense barrier against parasites (Steinegger and Hänsel, 1988) and they also act as antifungal agents providing direct protection to the plant (Kokubun and Harborne, 1995; Aoyama *et al.*, 1997).

From *Picrorhiza*, many different phenolic compounds have been isolated. Studies involving *P. scrophulariiflora* showed the presence of phenolics like androsin (Wang *et al.*, 1993), picein (Wang *et al.*, 2004), catechin and luteolin (Huang *et al.*, 2009) in roots while in rhizome phenolic glycoside like umbelliferon was found (Xie *et al.*, 2007). Chemical structures of androsin and picein are shown in Figure 2.5 below.



(A) **(B)**

Figure 2.5 Chemical structures of (A) androsin and (B) picein.

2.7 Medicinal Properties of *N. scrophulariiflora*

Kutki is used for treatment of various diseases. *P. kurrooa* is used for treatment of liver disorders, fever, asthma and jaundice (Smit *et al.*, 2000; Sturm and Stuppner, 2001;

Bhandari *et al.*, 2008), gastrointestinal and urinary disorders, leukoderma, snake bite, scorpion sting and inflammatory (Weinges *et al.*, 1972; Visen *et al.*, 1998; Verma *et al.*, 2009), hepatoprotective (Dhawan, 1995; Vaidya *et al.*, 1996; Singh *et al.*, 2005), anti-inflammatory (Singh *et al.*, 1993), immunomodulatory (Puri *et al.*, 1992), free radical scavenging (Chander *et al.*, 1992), gastric ulcer (Ray *et al.*, 2002; Banerjee *et al.*, 2008), anti-allergic and anti-anaphylactic activities (Baruah *et al.*, 1998), and anti-hepatitis-B surface antigen activity (Mehrotra *et al.*, 1990) while *P. scrophulariiflora* is used for antidiabetic, antiasthmatic, cardioprotective, antioxidant and antiradical properties (Indian Herbal Pharmacopoeia, 2002), antiulcer (Anadan *et al.*, 1999), anticancer property (Joy *et al.*, 2000), and a selective enhancer of neuron growth (Li *et al.*, 2000). The various medicinal and/or health benefits conferred by Kutki is due to various biological activities discussed below.

2.7.1 Anti-inflammatory and Immunomodulatory Property

Kutki possesses anti-inflammatory effects among other health benefits. Its extract is found to have inhibitory effects on pro-inflammatory cells such as neutrophils, macrophages and mast cells (Pandey and Das, 1989). It was suggested that the extract inhibits membrane-mediated activation (β-adrenergic receptors) of these cells. However, no effect was found on prostaglandin production. Apocynin, a catechol, is found as one of the minor constituents in extract of kutki. This apocynin was found to exhibit powerful anti-inflammatory effects on a variety of inflammatory models. Apocynin prevented the formation of ulcerative lesions in rats injected intracutaneously with Freund's complete adjuvant (Hart *et al.*, 1992) and reduced swelling in collagen-immunized rats. Apocynin was found to inhibit neutrophil oxidative burst *in vitro* without affecting beneficial activities such as chemotaxis, phagocytosis, and intracellular killing of bacteria (Simons *et al.*, 1990; Stolk *et al.*, 1994). In *in vivo* animal models, apocynin inhibited lipopolysaccharide (LPS)-induced emphysema in hamsters (Stolk *et al.*, 1994). In another study, significant inhibition of acute inflammation *in vivo*, demonstrated by acute carragenan-induced paw-inflammation model, was observed which was attributed to possible inhibition of the classical pathway of complement activation (Smit *et al.*, 2000).

Studies regarding immunomodulatory property of kutki extract showed that a biopolymeric fraction (named RLJ-NE-205) may enhance non-specific, humoral and cellular immunity in a mouse model suggesting therapeutic usefulness in immunocompromised patients and as an adjuvant during vaccination programs in order to reduce the number of non-responders to vaccines (Gupta *et al.*, 2006). Picroliv, a standardized iridoid fraction obtained from kutki, significantly enhanced antileshmanial efficacy of combination of miltefosine plus paromomycin. This combinatorial treatment

also demonstrated significant phagocytic and lymphocyte proliferation response and prevented acceleration of production of Reactive Oxygen Species (ROS), Reactive Nitrogen Species (RNS) and Hydrogen peroxide (H_2O_2) (Sane *et al.*, 2011).

2.7.2 Cholerectic Property

Several hepatotoxins, including paracetamol and ethynylestradiol, have a cholestatic effect on the production of bile. Kutki has been shown to reverse acetaminophen and ethynylestradiol-induced cholestasis, maintaining both bile volume and flow (Luper, 1998).

2.7.3 Hepatoprotective Property

Kutki has been shown to protect liver cells from a wide variety of injuries including *Amanita* poisoning, carbon tetrachloride, galactosamine, ethanol, aflatoxin B1, acetaminophen, thioacetamide, oxytetracycline, monocrotaline etc induced liver damage both *in vitro* and *in vivo* experiments. When compared to such protective effects of other potent compound such as silymarin, the effect was found to be similar or in many cases superior to it (Luper, 1998). Kutki also has a clinical implication in viral hepatitis. In a study performed in India on 32 patients diagnosed with acute viral hepatitis (HBsAg negative) and 15 patients in the treatment group were given kutki root powder. Liver enzymes (AST and ALT) and bilirubin were significantly lower in the treatment group. Similarly, number of days required for serum bilirubin to drop to average value of 2.5 mg % was 27.4 for treatment groups while 75.9 days for placebo groups. Another study suggested that it possessed anti-HBsAg activity and it also inhibited purified HBV antigens prepared from HBsAg carriers (Luper, 1998).

2.7.4 Antimicrobial Property

Research on antimicrobial property of *P. kurrooa* has suggested that this plant possesses certain degree of this property. In a study, it was found that ethanol extract of *P. kurrooa* showed high antibacterial activity against *Staphylococcus aureus, Bacillus cereus, Escherichia coli, Klebsiella pneumoneae, Salmonella typhi* and *Streptococcus pyogenes* (Kumar *et al.*, 2010). Methanolic rhizome extract however showed high antibacterial activity against *S. aureus* and *Pseudomonas aeruginosa*. Aqueous extract did not show antibacterial activity. In another study, it was found that methanolic extract of *P. kurrooa* showed significant antibacterial activity against *S. aureus, Bacillus subtilis* and *E. coli* strains. Similarly, the aqueous extract showed significant activity against some fungi such as *Aspergillus niger* and *Candida albicans* (Sharma and Kumar, 2012). Similar result was found in another study where ethanolic extract of *P. kurrooa* showed activity against *B. subtilis, S. aureus, P. aeruginosa, E. coli* and fungi like *A. niger,*

C. albicans and *Malasseiza furfur* (Usman *et al.*, 2012). However, antimicrobial study involving *N. scrophulariiflora* of Nepal has not been reported so far.

2.7.5 Antioxidant Property

Most of the studies on antioxidant property as well as the free radical scavenging potential have been done in either *P. kurrooa* or *P. scrophulariiflora*. Most literatures are dedicated to the studies of *P. kurrooa* as these studies have mostly been carried out by Indian authors. The free radical scavenging potential is not limited to morphological differences between *P. kurrooa* and *N. scrophulariiflora* but is actually inherent to the chemical constituents present in them. Both these species might share same chemical constituents however this needs to be verified by research.

Free radicals have been identified as prime factor in causation of various ailments like cancer, cardiovascular diseases, liver diseases, brain dysfunctions (Chander *et al.*, 1992; Govindarajan *et al.*, 2003; Jagetia and Baliga, 2004; Bhandari *et al.*, 2010; Meng *et al.*, 2012). Reactive oxygen species (ROS) have been known to cause tissue injury through covalent binding and lipid peroxidation which augment collagen synthesis and fibrosis (Govindarajan *et al.*, 2003).

Kutki possesses significant antioxidant activity *in vitro* which may contribute to the hepatoprotective effect by reducing lipid peroxidation and free radical damage (Luper, 1998). Chander *et al.* (1992) found that the active constituents, picroside I and kutkoside, inhibited the non-enzymatic generation of O_2-methosulphate NADH system, inhibited oxidative malonaldehyde generation by both ascorbate-Fe^{2+} and NADPH-ADP-Fe^{2+} systems, and scavenged superoxide (O_2) anions generated in a xanthine-xanthine oxidase system. In other words, it demonstrated antioxidant activity similar to that of superoxide dismutase, metal-ion chelators, and xanthine oxidase inhibitors. Similarly, kutki was found to be effective against *Plasmodium berghei* (Singh and Banyal, 2011).

Study done on free radical scavenging potential of *P. kurrooa* by assaying lipid peroxidation, reduced glutathione (GSH), 1, 1-diphenyl-2-picryl-hydrazyl (DPPH) radical scavenging, nitric oxide scavenging and superoxide scavenging by Govindarajan *et al.* (2003) demonstrated the *in vitro* efficacy of alchoholic extract of *P. kurrooa* as hepatoprotective and immunomodulatory drug considering its antioxidant and free radical scavenging ability. Lipid peroxidation is caused by ferrous sulphate either through ferryl-perferryl complex or through OH radical by Fenton reaction. The study showed that alchoholic extract of *P. kurrooa* inhibited $FeSO_4$ - induced lipid peroxidation in a dose dependent manner. Similarly, the extract inhibited oxidation of reduced glutathione, reduced DPPH activity, inhibited superoxide dismutase activity and inhibited production of nitrite (Govindarajan *et al.*, 2003). Similarly, the

hepatoprotective action of picroliv glycosides may be due to prevention of lipid peroxidation and free radical generation during liver damage (Chander *et al.*, 1992).

Similarly, Nitric Oxide (NO) is an essential bioregulatory molecule involved in processes like signal transmission, immune response, vasodilatation, control of blood pressure etc. However, high level of NO is implicated in various pathological conditions including cancer (Jagetia and Baliga, 2004). NO may not interact directly with DNA or proteins but in aerobic condition, NO becomes unstable and reacts with oxygen to produce intermediates such as NO_2, N_2O_4, N_3O_4, stable nitrate and nitrite and peroxynitrite when reacting with superoxide. These intermediates are genotoxic and ultimately lead to carcinogenesis (Jagetia and Baliga, 2004). In a comparative study of many medicinal plants *P. kurrooa* was also found to scavenge NO projecting it as an important medicinal plant possessing antioxidant activity (Jagetia and Baliga, 2004).

2.8 Antioxidant assay using DPPH (1,1-diphenyl-2-picryl-hydrazyl) radical

Assessments of antioxidant properties of natural compounds are very important because of their uses in medicine, food and cosmetics (Mishra *et al.*, 2012). In living systems various metabolic processes and environmental stresses generate various reactive species, reactive oxygen species (ROS) being the main. Increased level of ROS can damage structure of biomolecules and modify their functions and lead to cellular dysfunction and even cell death. This cumulative effect has been implicated in various health problems such as cancer, age related diseases and cardiovascular diseases (Mishra *et al.*, 2012).

DPPH assay is routinely practiced for assessment of free radical scavenging potential of an antioxidant molecule and is considered as one of the standard and easy colorimetric methods for the evaluation of antioxidant properties of pure compounds (Mishra *et al.*, 2012). DPPH (Figure 2.6) is a stable radical in solution and appears purple in color absorbing strongly at 517 nm (Mishra *et al.*, 2012). This assay is based on the principle that DPPH on accepting a hydrogen (H) atom from scavenger molecule, i.e, antioxidant, is reduced to $DPPH_2$ from DPPH and hence the purple color changes to yellow with relative decrease of absorbance at same wavelength. The color change is measured spectrophotometrically and is used to determine the antioxidant properties (Mishra *et al.*, 2012).

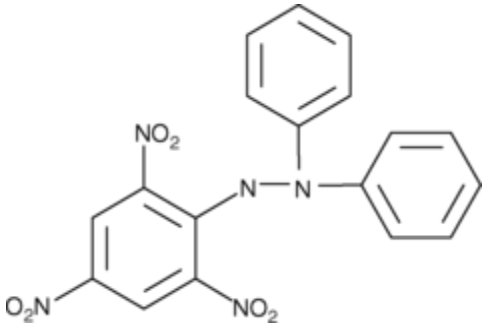


Figure 2.6 Chemical structure of 1, 1–diphenyl–2–picryl-hydrazyl (DPPH).

DPPH is a well-known radical and scavenger for other radicals. Therefore, rate reduction of a chemical reaction upon addition of DPPH is used as an indicator of the radical nature of that reaction. This property allows visual monitoring of the reaction, and the number of initial radicals can be counted from the change in the optical absorption at 517 nm (Figure 2.7).

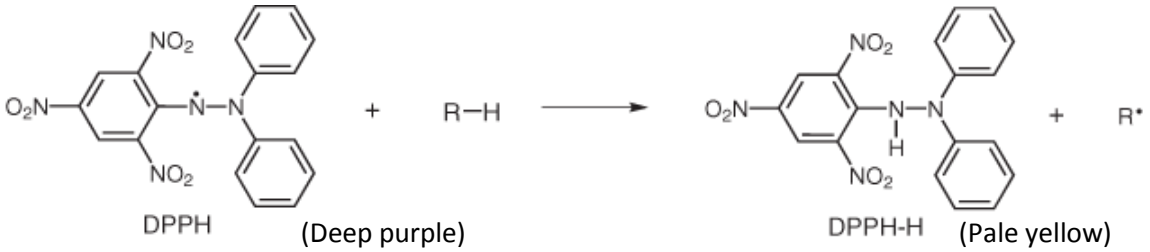


Figure 2.7 chemical reactions between DPPH molecule and a hydrogen donor.

This assay is in common use to measure the antioxidant content of various food items including wheat grains and bran, vegetables, conjugated linoleic acids, herbs, edible seed oils, and flours in different solvent systems including ethanol, aqueous acetone, methanol, aqueous alchohol and benzene (Cheng *et al.*, 2006).

Many studies regarding the use of HPLC method for detection of active constituents in *Picrorhiza* extracts have been performed by different authors. In a study involving *P. kurrooa* extract and its free radical scavenging effect, online HPLC-DPPH and colorimetric DPPH method was used (Bhandari *et al.*, 2010). The comparative study showed that colorimetric method showed very less free radical scavenging effect while HPLC-DPPH method showed high activity. In the study, methanol: water (80:20) solvent system was used as mobile phase. In another study, HPLC method was used to determine sugars like xylose, xylitol, mannitol, glucose sucrose and picrosides I and II in *P. kurrooa* and *P. scrophulariiflora* using acetonitrile: water (78:22) as mobile phase (Bhandari *et al.*, 2008). Picroside accumulation in *P. kurrooa* at different altitudes was assessed using HPLC system in a study performed by Katoch *et al.* (2011) and it was found that the picroside content decreased as the altitude decreased. Using methanol: water (2:3) as mobile phase, they also found out that picrosides I and II were accumulated more in rhizomes>root>inflorescence>flowers. In another study, picroside I was determined

from plasma using HPLC method (Dwivedi *et al.*, 1997). The drug was extracted from rabbit plasma samples after precipitation with methanol and separated using acetonitrile: 0.1M acetic acid (25:75) as mobile phase in HPLC system and picroside I was quantified.

2.9 Genetic Diversity

Genetic diversity refers to any variation in the nucleotides, genes, chromosomes, or whole genomes of organisms. Genetic diversity when described elementarily is the differences in the sequences of nucleotides that form the DNA within the cells of the organism (Harrison *et al.*, 2013). Most organisms are diploid; having two sets of chromosomes, and therefore has two copies of each gene. However, some organisms can be haploid, triploid, or tetraploid. Within any single organism, there may be variation in allele number (two or more alleles) for each gene. This variation is introduced either through mutation of one of the alleles, or as a result of sexual reproduction. During sexual reproduction, offspring inherit alleles from both parents and these alleles might be slightly different, especially if there has been migration or hybridization of organisms, so that the parents may come from different populations and gene pools. Also, when the offspring's chromosomes are copied after fertilization, genes can be exchanged in a process called sexual recombination. Harmless mutations and sexual recombination may allow the evolution of new characteristics (Harrison *et al.*, 2013). Besides having distinct combinations of genes, species may also have variation in the shape and composition of the chromosomes and in the total number of chromosomes present. Examination of these features of the chromosomes (karyotyping) can also provide insight in describing genetic diversity.

Analysis of genetic diversity can be applied to many different studies including population genetics, evolutionary genetics, ecological genetics, biodiversity conservation, genebank management (Pullin, 2002; Groom *et al.*, 2006) etc. Genetic studies can identify alleles that might affect the ability of the organism to survive in its existing habitat, or might enable it to survive in more diverse habitats. This is the basis of natural selection. Some alleles can confer a selective advantage on the host organism, such that it is more likely to survive than if it did not have the particular allele (Harrison *et al.*, 2013). The presence of unique genetic characteristics distinguishes members of a given population from those of any other population. Large populations usually have a greater diversity of alleles compared to small populations. This diversity of alleles indicates a greater potential for the evolution of new combinations of genes and, subsequently, a greater capacity for evolutionary adaptation to different environmental conditions. In small populations, the individuals are likely to be genetically, anatomically, and physiologically more homogeneous than in larger populations and less able to adapt

to different environmental conditions (Harrison *et al.*, 2013). Genetic diversity is therefore a key component for conservation efforts associated with population management (Andayani *et al.*, 2001). Populations have to survive under a lot of selection pressures. Apart from natural selection pressure, they also have to cope up with impacts of global warming and anthropogenic activities. Natural population of plants and animals suffer a rapid decline because of these various effects. Therefore, understanding genetic variation present in plant population is very important and necessary. Understanding genetic variation within and among populations is more essential for the establishment of effective conservation practices for rare and endangered plants (Liu *et al.*, 2011).

2.10 Tools for Assessment of Genetic Diversity

Genetic diversity assessment at different taxonomic levels can be performed by using various marker techniques. Genetic markers are identifiable DNA sequences found at specific locations in the genome and are transmitted by the standard laws of inheritance from one generation to another. Genetic markers fall into one of the three broad classes, namely, those which are based on visually assessable traits (morphological and agronomic traits), those based on gene product (biochemical markers) and those relying on DNA assay (molecular markers) (Chawla, 2010).

2.10.1 Morphological Markers

Morphological markers, also called phenotypic markers, are qualitative traits found in nature which can be scored visually. Such markers can be either dominant or recessive. Morphological markers are still used dominantly in breeding programs but they contain many limitations such as: they have large effects on phenotype hence are undesirable in breeding programs, they mask the effects of linked minor genes, they are highly influenced by environment, they cover only limited fraction of the genome of an organism etc (Chawla, 2010).

2.10.2 Biochemical Marker – Allozymes (Isozyme)

Isozymes are structurally different molecular forms of an enzyme with same catalytic function. A mutation in DNA results in an amino acid being replaced by another thus altering net electric charge of protein as well as the conformation. This affects the migration rate of proteins in electric field which can be detected by gel electrophoresis. Two or sometimes more than two loci (isoloci) can be distinguished for an enzyme. Isozymes have been used to study phylogenetic relationships to estimate genetic variability and taxonomy, to study population genetics and developmental biology, characterization of plant genetic resources management and plant breeding etc (Staub *et al.*, 1996). Isozymes have been used as reliable genetic markers in breeding and

genetic studies of plant species (Heinz, 1987; Erskine and Muehlbauer, 1991; Freville *et al.*, 2001).

2.10.3 Molecular Genetic Markers

Molecular markers are not normal genes with biological effects but are similar to constant landmarks in the genome. Markers and genes lie close to each other on same chromosome and stay together in each generation of plants. The relative position of markers in a chromosome and their closeness to any specific genes can be used to create a linkage map (Semagn *et al.*, 2006). Such maps can be used to study associations between economically important traits or Quantitative trait loci (QTL). Besides, such markers can also be linked to qualitative traits and can also be employed for the assessment of genetic diversity, phylogeny and diagnostic development (Mondini *et al.*, 2009; Arif *et al.*, 2010; Lübberstedt and Varshney, 2013).

An ideal molecular marker should have following desirable traits (Agarwal *et al.*, 2008; Jonah *et al.*, 2011):

1. It should be polymorphic and evenly distributed throughout the genome
2. It should provide adequate resolution of genetic differences
3. It should generate multiple, independent and reliable markers
4. It should be simple, quick, and inexpensive
5. It should need small amounts of tissues
6. It should have linkage to distinct phenotypes
7. It should require no prior information about the genome of an organism
8. It should be codominantly inherited in order to distinguish homozygosity and heterozygosity
9. It should have easy exchange of data between laboratories

It is difficult to find a single marker possessing all the properties mentioned above; hence, the selection should be determined by type of study undertaken.

Genetic markers can be divided into three categories: 1) non-PCR-based techniques or hybridization based techniques, 2) PCR-based techniques and 3) DNA sequencing based techniques.

2.10.3.1 Non-PCR based Techniques

a. Restriction Fragment Length Polymorphism (RFLP)

RFLP markers were first used in 1975 to identify DNA sequence polymorphisms for genetic mapping of a temperature-sensitive mutation of adenovirus serotypes

(Sambrook *et al.*, 1975). Later it was used for human genome mapping (Botstein *et al.*, 1980) and adapted to plant genome as well (Helentjaris *et al.*, 1986; Weber and Helentjaris, 1989). RFLPs are inherited in Mendelian fashion. The general principle is if two organisms differ in a distance between sites of cleavage of particular restriction endonuclease then the length of fragments produced will differ when the DNA is digested with a restriction enzyme. The similarity of patterns generated can be used to differentiate species and even strains from one another. The differential profile is generated due to nucleotide substitutions or DNA rearrangements like insertion or deletion or single nucleotide polymorphisms. Some of the differences in DNA sequences at the restriction sites can result in gain, loss, or relocation of a restriction site (Semagn *et al.*, 2006; Agarwal *et al.*, 2008; Jonah *et al.*, 2011). An overview of the steps performed in RFLP is shown in Figure 2.8.

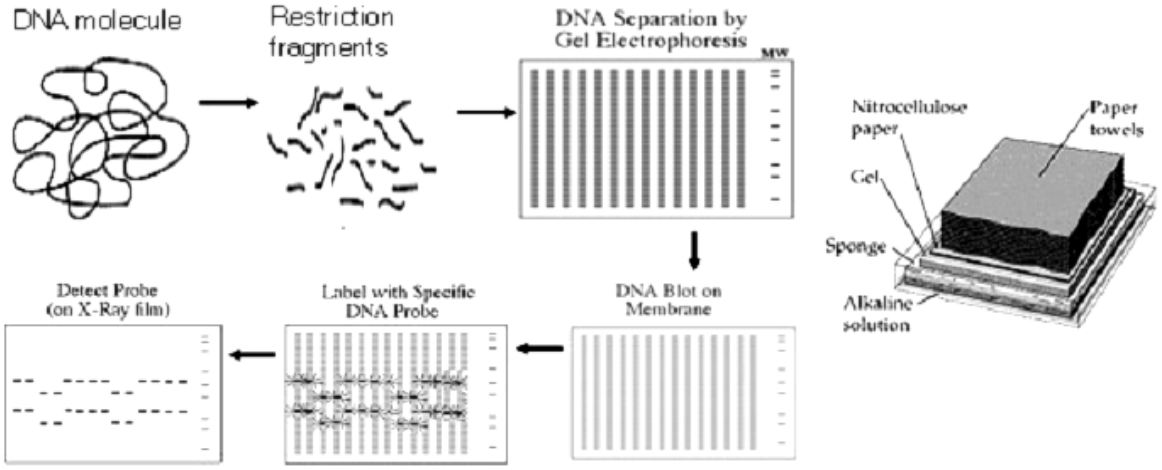


Figure 2.8 Outline of different steps performed in Restriction fragment length polymorphism (RFLP) marker.

Some limitations of RFLP are it requires high quantity and quality of DNA (Potter and Jones, 1991; Roy *et al.*, 1992; Young *et al.*, 1992) and it depends on the development of specific probe libraries for any species. Automation is not easy and level of polymorphism is low hence only few loci are detected per assay. Apart from being time consuming, laborious and expensive (Yu *et al.*, 1993) it usually requires radioactively labeled probes (Semagn *et al.*, 2006).

2.10.3.2 PCR based Techniques

The PCR-based techniques are of two types depending on the primers used for amplification:

1. Arbitrary or semi-arbitrary primed PCR (eg: AP-PCR, DAF, RAPD, AFLP, ISSR)
2. Site-targeted PCR (eg: EST, CAPS, SSR, SCAR, STS)

The commonly used PCR-based DNA marker systems are Random Amplified Polymorphic DNA (RAPD), Amplified Fragment Length Polymorphism (AFLP) and Inter Simple Sequence Repeat (ISSR). More recent techniques involve Simple Sequence Repeats (SSRs) or microsatellites (Staub *et al.*, 1996; Gupta and Varshney, 2000). The major limitations of these methods are low reproducibility of RAPD, high cost of AFLP and the need to know the flanking sequences to develop species specific primers for SSR polymorphisms (Reddy *et al.*, 2002).

a. Inter Simple Sequence Repeat – PCR (ISSR-PCR)

Inter simple sequence repeat (ISSR) technique is a PCR based technique which involves amplification of DNA segment present at an amplifiable distance in between two identical microsatellite repeat regions oriented in opposite direction (Zietkiewicz *et al.*, 1994). The technique uses microsatellites, usually 16-25 bp long, as primers in a single primer PCR reaction targeting multiple genomic loci to amplify mainly the inter-SSR sequences of different sizes. The microsatellite repeats can be di-, tri-, tetra-, or penta-nucleotides (Reddy *et al.*, 2002). The primers can be either unanchored (Meyer *et al.*, 1993; Gupta *et al.*, 1994; Wu *et al.*, 1994) or anchored at 3' or 5' end with 1 to 4 degenerate bases extended into flanking sequences (Zietkiewicz *et al.*, 1994) as shown in Figure 2.9.

ISSR combines the benefit of AFLP and microsatellite analysis with the universality of RAPD. ISSR has high reproducibility possibly due to the use of longer primers (16-25 mers) as compared to RAPD primers (10 mers) thus allowing use of high annealing temperature (45-60^{o}C) leading to stringency and the primers aren't proprietary (as in SSR-PCR). It also has a distinct advantage of simplicity, quickness and efficiency and there is no use of radioactivity at all. The amplified products are 200 – 2000 bp long and can be detected by agarose and polyacrylamide gel electrophoresis (Reddy *et al.*, 2002). ISSRs are inherited as dominant markers following simple Mendelian inheritance (Gupta *et al.*, 1994; Tsumura *et al.*, 1996; Ratnaparkhe *et al.*, 1998; Wang *et al.*, 1998). In some cases, however, they are also inherited as co-dominant markers which can be used as marker to distinguish between homozygotes and heterozygotes (Wu *et al.*, 1994; Akagi *et al.*, 1996; Wang *et al.*, 1998; Sankar and Moore, 2001).

It has been found that the microsatellites region change considerably than other types of DNA and thus the polymorphism is also higher in them. During replication, the slippage of DNA polymerase coupled with failure to repair mismatches creates hypervariability within SSRs (Levinson and Gutman, 1987). Events like mutation at the priming site and/or insertion/deletion resulting in prevention of amplification of a fragment gives presence or absence of polymorphisms. Similarly, when unanchored primers are used (i.e only SSRs are used as primers) then the primers tend to slip within

the repeat units during amplification leading to smears depending upon nature of primers used. Anchoring 1 to 4 degenerate nucleotides at either 3' or 5' end of primer ensures annealing to the ends of a microsatellite in template. Further, 5' anchored primers give more number of bands and higher degree of polymorphism (Reddy *et al.*, 2002). It has been found that di-nucleotide repeats, anchored either at 3' or 5' end give high polymorphism (Nagaoka and Ogihara, 1997; Blair *et al.*, 1999; Joshi *et al.*, 2000). In general, it is found that primers with (AG), (GA), (CT), (TC), (AC), (CA) repeats show higher polymorphism than primers wih tri-, or tetra-nucleotide repeats (Reddy *et al.*, 2002).

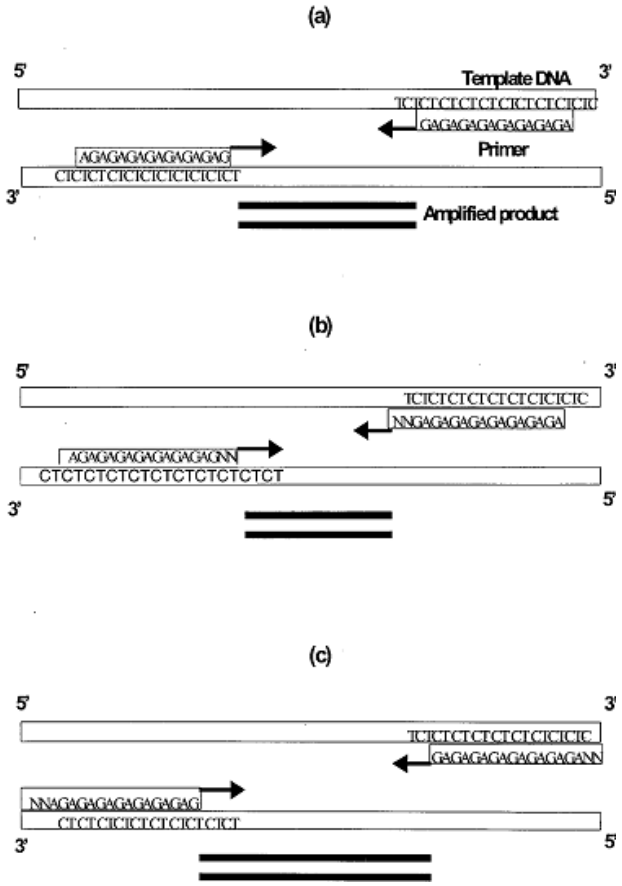


Figure 2.9 A schematic representation of a single primer $(AG)_8$, unanchored (a), 3'-anchored (b) and 5'-anchored (c) targeted towards $(TC)_n$ repeat used to amplify inter simple sequence repeat region flanked by two inversely oriented $(TC)_n$ sequences (Reddy *et al.*, 2002).

ISSRs are used for varied purpose one of them being estimation of genetic diversity at inter- and intra-specific level. It has been used in crop species like rice (Joshi *et al.*, 2000), wheat (Nagaoka and Ogihara, 1997) etc. Similarly it has been used in the study of genetic structure of different medicinal plants like *Dendrobium* spp. (Yang *et al.*, 2010), *Swertia* spp. (Joshi and Dhawan, 2007; Zhang *et al.*, 2007; Tamhankar *et al.*, 2009), *Podophyllum hexandrum* (Alam *et al.*, 2009; Naik *et al.*, 2010), *Rheum* spp. (Wang *et al.*,

2012a; Wang *et al.*, 2012b). ISSR has also been used as a DNA marker to study the genetic diversity and polymorphism in *P. kurrooa* (Rawat *et al.*, 2013) and *N. scrophulariiflora* (Liu *et al.*, 2011).

A comparison between different markers frequently used in various studies is enlisted in Table 2.2 below.

Table 2.2 Comparison of five widely used DNA markers in plants (Semagn *et al.*, 2006)

Markers / Features	RFLP	Microsatellite	RAPD	AFLP	ISSR
Genomic Abundance	High	Medium	Very high	Very high	Medium
Part of genome surveyed	Low copy coding regions	Whole genome	Whole genome	Whole genome	Whole genome
Amount of DNA required	High	Low	Low	Medium	Low
Type of polymorphism	Single base changes, insertion, deletion	Changes of length of repeats	Single base changes, insertion, deletion	Single base changes, insertion, deletion	Single base changes, insertion deletion
Level of polymorphism	Medium	High	High	Very high	High
Effective multiplex ratio	Low	Medium	Medium	High	Medium
Marker index	Low	Medium	Medium	High	Medium
Inheritance	Codominant	Codominant	Dominant	Dominant	Domina nt
Detection of alleles	Yes	Yes	No	No	No
Ease of use	Labour intensive	Easy	Easy	Difficult initially	Easy
Automation	Low	High	Medium	Medium	Medium
Reproducibility	High	High	Intermediate	High	Medium to high
Types of	Low copy	Specific	Usually 10 bp	Specific	Specific

probes/primers	genomic DNA or cDNA clones	repeat DNA sequences	random nucleotides	sequences	repeat: DNA sequence
Cloning and/or sequencing	Yes	Yes	No	No	No
Radioactive detection	Usually yes	No	No	Yes/no	No
Development/start-up costs	High	High	Low	Medium	Medium
Utility for genetic mapping	Species specific	Species specific	Cross specific	Cross specific	Cross specific
Proprietary rights status	No	No (some are licensed)	Licensed	Licensed	No

2.10.3.3 DNA sequencing-based Markers

A number of DNA sequencing methods are currently available which includes Sanger dideoxynucleotide chain termination method and its variants (enzymatic methods), Maxam & Gilbert method, Pyrosequencing™ method, and single molecule sequencing with exonuclease (Franca *et al.*, 2002). Among these techniques, Sanger method (Sanger *et al.*, 1977) is still the most widely used technique. Briefly, in the presence of DNA polymerase, catalytic polymerization of deoxynucleoside triphosphates (dNTP) onto DNA occurs. The polymerization extends until the enzyme incorporates a modified nucleoside called terminator or dideoxynucleoside triphosphate (ddNTP) into the growing chain thus terminating the polymerization process. Four reactions for four bases are carried out and the resulting mixture of different-sized DNA fragments are resolved by electrophoresis on polyacrylamide gel in four parallel lanes and the patterns are read to produce a sequence (Franca *et al.*, 2002).

Amplification of a certain region within DNA using Polymerase Chain Reaction (PCR) is subjected to one of these techniques to generate DNA sequence which can be studied within organisms to determine polymorphism. A number of genes and DNA regions of plants and animals from nuclear as well as organellar genomes are available for phylogenetic estimation. Various sequencing based markers such as nuclear ribosomal DNA (nrDNA) like Internal Transcribes spacer (ITS), chloroplast DNA markers (*matK, rbcL, trnH-psbA*) and mitochondrial (*COXI*) are widely applied in the phylogenetic reconstruction, DNA Barcoding and diagnostic development for various applications (Shrestha, 2013).

2.10.4 Application of ISSR Technique

ISSR-PCR technique has been widely applied to various scientific researches. ISSR technique has also been widely used in studies involving genetic diversity and phylogenetic analysis. Estimation of genetic diversity at inter- and intra- specific level of medicinal plants and crops has been performed using ISSR technique, such as, rice cultivars (Suh *et al.*, 2010; Steele *et al.*, 2013), citrus cultivars (Fang *et al.*, 1997), medicinal plants like *Swertia* spp. (Joshi and Dhawan, 2007), *N. scrophulariiflora* (Liu *et al.*, 2011) etc.

Similarly, ISSR technique has also been used as an important tool for characterization of germplasm and identification of varieties, hybrids, parental sources in plant breeding, germplasm management (Reddy *et al.*, 2002). In studies involving mapping of Japanese and European larch genomes, ISSR along with AFLP and RAPD markers were used (Arcade *et al.*, 2000) suggesting possible use of ISSR in genome mapping. ISSR marker has also been used in gene tagging, for e.g. in rice, ISSR marker was converted to a sequence tagged site (STS) marker to identify fertility restoration gene, *Rf-1* (Akagi *et al.*, 1996) and also ISSR has been used to tag male sterility gene in rice (Hussain *et al.*, 2000).

2.11 DNA Barcoding

Identification of plant species is critical in varied biological disciplines such as plant ecology, environment, agriculture, conservation biology etc. Plants are generally identified using morphological characteristics requires special training in plant taxonomy besides on-hand identification tools (like books, flora information, scientific papers etc.) and most importantly suitable plant materials preferably with the reproductive structures are required. In the absence of reproductive structures it is almost impossible for many species to be correctly identified.

The term “DNA barcode” was first coined by Paul Hebert in 2003. DNA barcoding is a method of species identification and recognition using specific region of DNA using sequence data (Hebert *et al.*, 2003). The main aim of barcoding is to use DNA sequences of standard regions of DNA, commonly called barcode regions, instead of relying only on morphology as a tool for plant identification. The barcode sequences provide unique identifier for a species, compared with barcodes of items in a supermarket. Such sequences are stored in a database such that samples can be identified by comparing sample sequences against the known sequences in the database. This system of identification allows people without expert taxonomic knowledge to identify plant materials and especially from materials that may not be flowering, or may be seeds or seedlings, plant fragments in foods or medicines or even from gut of animals, timber

samples etc. All the barcode data generated during barcode studies are deposited in a database called Barcode of Life Data System or BOLD.

The barcode of life data system (BOLD) (www.barcodinglife.org) provides an integrated bioinformatics platform that supports all phases of the analytical pathway from specimen collection to tightly validated barcode library. It is a repository for the specimen and sequence records that forms the basic data unit of all barcode studies. Likewise, it is also a workbench that aids the management, quality assurance and analysis of barcode data. It also provides a vehicle for collaboration across geographically dispersed research communities by coupling flexible security and data entry features with web-based delivery (Ratnasingham and Hebert, 2007).

The variability in a single or few standard markers is used to discriminate biological entity. Hence, this method relies on the assumption that the genetic variation between species exceeds that of within species. DNA barcoding analysis reflects the distributions of intra- and inter-specific variabilities separated by a distance called 'DNA barcoding gap' (Meyer and Paulay, 2005; Wiemers and Fiedler, 2007).

In DNA Barcoding, it is often claimed that interspecific genetic variation exceeds intraspecific variation such that a clear gap exists which enables assigning of individuals to their species with a negligible error rate. This error is supposed to be due to incipient species with incomplete lineage sorting (Hebert *et al.*, 2003; Hebert *et al.*, 2004; Barrett and Hebert, 2005). Hence, it is devised that degree of divergence between sequences above a given threshold by at least 10 times would indicate specific distinctiveness, whereas divergence below such threshold would indicate taxonomic identity among samples (Wiemers and Fiedler, 2007). This approach results in errors of false positives and false negatives. False positives often result when populations within one species are genetically quite distinct, for example, in distant populations with limited gene flow or in allopatric populations with interrupted gene flow while in contrast false negatives occur when little or no sequence variation in the barcoding fragment is found between different, reproductively isolated population groups (Mayr, 1969).

DNA barcoding is carried out, in animals, using mitochondrial gene encoding cytochrome c oxidase subunit 1 (*CO1*). It is a favorable region for most animal species including some fungal species belonging to groups Ascomycota, Basidiomycota and Chytridiomycota (Chen *et al.*, 2010). However, *CO1* gene in plants aren't useful for identifying them due to the low amounts of variation in the genes as well as due to the variable mitochondrial genome (Chase *et al.*, 2005; Kress *et al.*, 2005; Chase *et al.*, 2007; Pennisi, 2007; Fazekas *et al.*, 2008). Thus it is important to screen either single or multiple regions in nuclear and plastid genomes in plants to identify appropriate DNA barcodes as shown by the studies depicted in Figure 2.10.

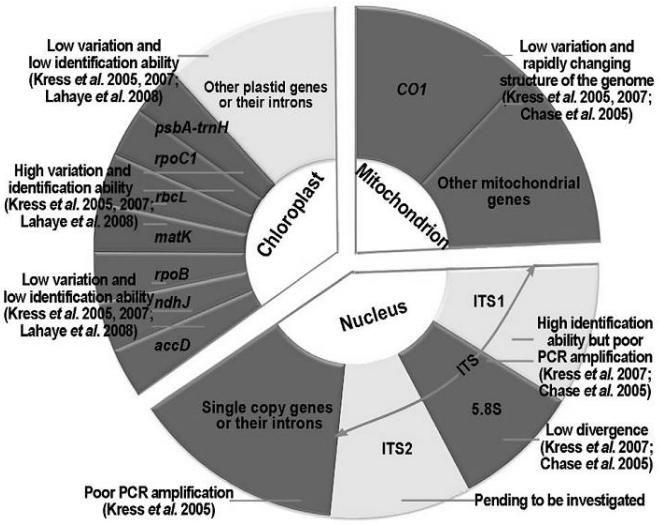


Figure 2.10 Genes from three genomes in plants that can be potential barcodes (Chen *et al.*, 2010).

In plants, many studies have been assessed demonstrating the effectiveness of DNA barcoding for species identification (Steven and Subramanyam, 2009; Moniz and Kaczmarska, 2010) but a barcode similar to *COI* still remains elusive (Kress *et al.*, 2005).

An appropriate DNA region or combination of DNA regions used in plant barcoding should have the properties of routine amplification with universal primers, easy sequencing via single-pass sequencing, appropriate length (i.e. 300-800 bp) (Kress *et al.*, 2005), sufficient variability to separate closely related species and exhibition of variability within species (Chase *et al.*, 2005; Cowan *et al.*, 2006; Newmaster *et al.*, 2008). A number of candidate regions have thus been studied extensively for this purpose. The nrITS and plastid *trnH-psbA* intergenic spacer have been proposed as potential DNA barcoding regions for flowering plants (Kress *et al.*, 2005). The plastid gene *rbcL* has been proposed as a core DNA barcoding gene to discriminate plant species at genus level in a tiered method in which a highly variable locus can be implemented when necessary (Newmaster *et al.*, 2006). According to Chase et al. (2007), due to low level of variation in plastid DNA, three regions are necessary for identification and thus proposed two combinations, *viz.*, *rpoC1, rpoB*, and *matK* and *rpoC1*, *matK* and *psbA-trnH* as markers for land plants. Meanwhile, Kress and Erickson (2007) proposed combination of *trnH-psbA* and *rbcL* as two-locus global barcode for land plants. Studies performed by Newmaster et al. (2008) tested seven chloroplast loci for barcoding *Compsoneura* of family Myristicaceae and proposed *matK* and *trnH-psbA* as promising barcodes for the nutmeg family. Similarly, in order to discriminate cycads (Cycadaceae), ITS was found to be most suitable as a barcode (Sass *et al.*, 2007). Fazekas *et al.* (2008) performed a study using eight chloroplast DNA regions (*rbcL*, *matK*, *rpoC1, rpoB*, 23S rDNA, *trnH-psbA*, *atpF-atpH, psbK-psbI*) and a mitochondrial *CO1* gene in order to identify species and

found 23S rDNA to be able to resolve species to a mere 7% and *trnH-psbA* to a 59%. Using other chloroplast markers like *rbcL*, *trnSGG* and *trnH-psbA* in Hymenophyllaceae, Nitta (2008) found *trnH-psbA* as a candidate barcode with greatest possibility. The Consortium of Barcode of Life (CBOL) Plant working Group compared the performance of seven candidate plastid regions and recommended the combination of *rbcL* and *matK* as core barcode for land plants (CBOL Plant Working Group, 2009). The choice of these two markers as core barcode was based on the consideration that *rbcL* region was easy to recover while *matK* had discriminatory power. A research performed by Chen et al. (2010) proposed ITS2 as a novel universal barcode for plant identification. A research performed in cycadales suggested multigene combinations for its identification (Nicolalde-Morejón *et al.*, 2011). A study performed in *Alnus* (Betulaceae) showed *trnH-psbA* and ITS to be the best combination (Ren *et al.*, 2010). The chronological timeline of use of different barcode markers is depicted in Figure 2.11.

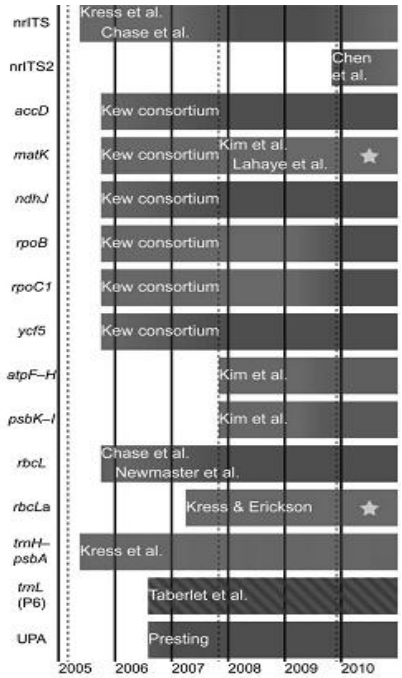


Figure 2.11 A schematic timeline of different markers used as plant barcodes. (Source: (Hollingsworth *et al.*, 2011).

2.12 Barcode Markers

2.12.1 Maturase K (*matK*)

matK is one of the most rapidly evolving coding sections of the plastid genome (Hilu and Liang, 1997) and can be regarded as the closest plant analogue to the CO1 animal barcode (Hollingsworth *et al.*, 2011). Using the currently available 'universal' primer pair (3F/1R) (Kim, unpublished), PCR and sequencing success of approximately 70% were observed in angiosperms. Similarly, using another primer pair (390F/1326R) (Cuenoud *et*

al., 2002), the amplification and sequencing success was increased by another ca. 10%. However, *matK* is not recoverable from some bryophyte and fern groups with current primer sets (de Groot *et al.*, 2011; Kuo *et al.*, 2011). This low success of *matK* has been attributed to low amplification and sequencing rate due to low universality of currently available primers and mononucleotide repeats (Yu *et al.*, 2011). The *matK* gene is approximately 1570 bp in length and codes for maturase protein. The coding region of *matK* lies within an intron of the chloroplast trnK gene. Being a coding region and the very high evolutionary rate of *matK* has made it usable in phylogenetic reconstruction at taxonomic levels such as Order or Family and sometimes to Genus or species (Wolfe, 1991; Hilu *et al.*, 2003; Muller *et al.*, 2006; Chase *et al.*, 2007; Lahaye *et al.*, 2008). However, in a study performed by Yu *et al.* (2011) a newly designed primers 472F and 1248R showed strong amplification (93.1%) and sequencing (92.6%) success. *matK* gene along with the position and orientation of different primers are shown below in Figure 2.12.

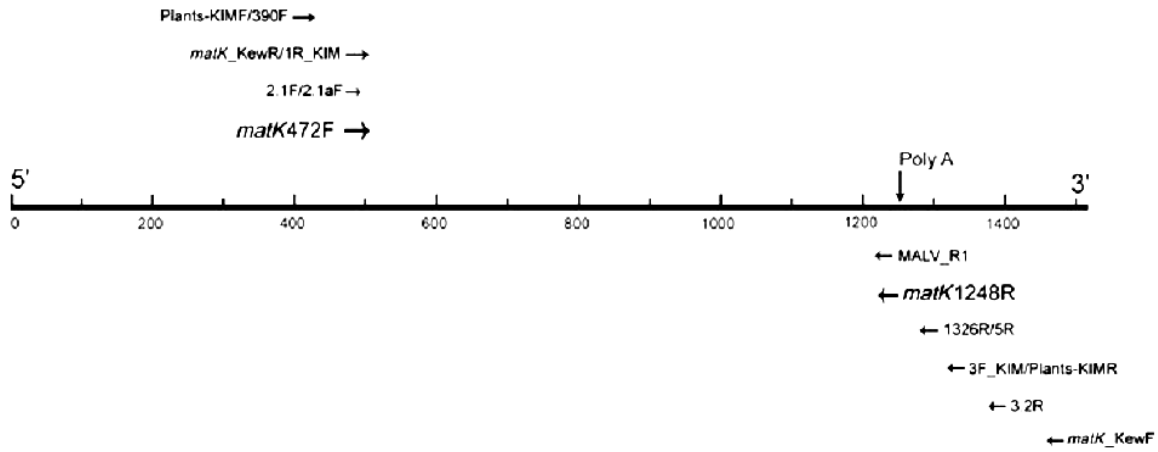


Figure 2.12 The position and orientation of *matK* primer pairs used in different studies along the entire length of *matK* gene of *Arabidopsis thaliana* (1515 bp). (source: (Yu *et al.*, 2011)

2.12.2 Ribulose-1, 5-bisphosphate carboxylase - Large Subunit (*rbcL*)

The most common gene used to provide sequence data for plant phylogenetic studies is plastid-encoded *rbcL* gene (Chase *et al.*, 1993; Donoghue *et al.*, 1993). It is a single copy gene and is approximately 1430 bps in length, is free from length mutations except at the far 3' end and has a fairly conservative rate of evolution. However, the ability of *rbcL* to resolve phylogenetic relationships below family level is often poor (Doebley *et al.*, 1990). Because of its slow synonymous nucleotide substitution rate in comparison with the nuclear genes and its functional constraint that reduces the evolutionary rate of nonsynonymous substitutions (Wolfe *et al.*, 1987), *rbcL* is considered to be more useful than isozymes (Gastony, 1988) and RFLP at higher taxonomic levels (Hasebe *et al.*, 1994). The function of the *rbcL* gene is to code for the large subunit of ribulose 1, 5 –

bisphosphate carboxylase/oxygenase or RUBISCO. Phylogeny, based on *rbcL* sequences, has been obtained at family level (Zurawski and Clegg, 1984; Morgan and Soltis, 1993). In Zingiberales, however, *rbcL* is used for interordinal or intrafamilial level (Smith *et al.*, 1993). The advantages of this gene are that it is easily amplified and sequenced in most land plants and it is regarded as a benchmark locus in phylogenetic investigations by providing a reliable placement of a taxon into a plant family and/or genus (Levin *et al.*, 2003; Kress and Erickson, 2007). Different primers and their binding sites result in varied amplicon size as depicted in Figure 2.13.

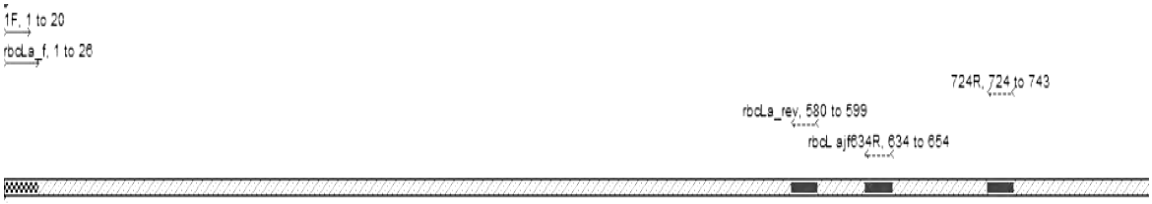


Figure 2.13 Location of binding of different primers in *rbcL* region of *Arabidopsis thaliana*.

2.12.3 *trnH-psbA* Intergenic spacer

Besides the *matK+rbcL* combination being used as core barcodes, *trnH-psbA* is a widely used supplementary barcode marker. It is straightforward to amplify across land plants and is one of variable Intergenic spacers in plants (Shaw *et al.*, 2007) (Figure 2.14).

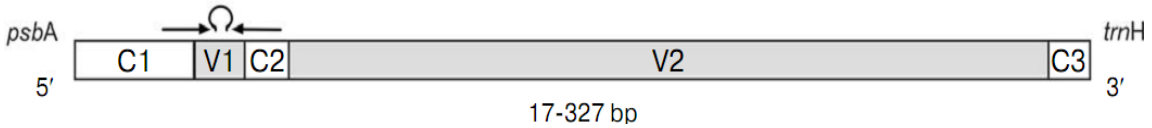


Figure 2.14 General scheme of organization of the *trnH-psbA* intergenic spacer. Also shown are conserved (C1, C2, C3) and variable (V1, V2) regions. (Source: (Degtjareva *et al.*, 2012).

2.12.4 Internal Transcribed Spacer (ITS)

Internal transcribed spacer (ITS) is a non-functional RNA situated between structural ribosomal RNAs (rRNA) on a common precursor transcript. This common transcript contains 18S rRNA, ITS1, 5.8S rRNA, ITS2, 26S rRNA as shown in Figure 2.15. The tandem repeat structure and extremely high copy number of nrDNA (Rogers and Bendich, 1987) make it especially easy to detect or clone in laboratory.

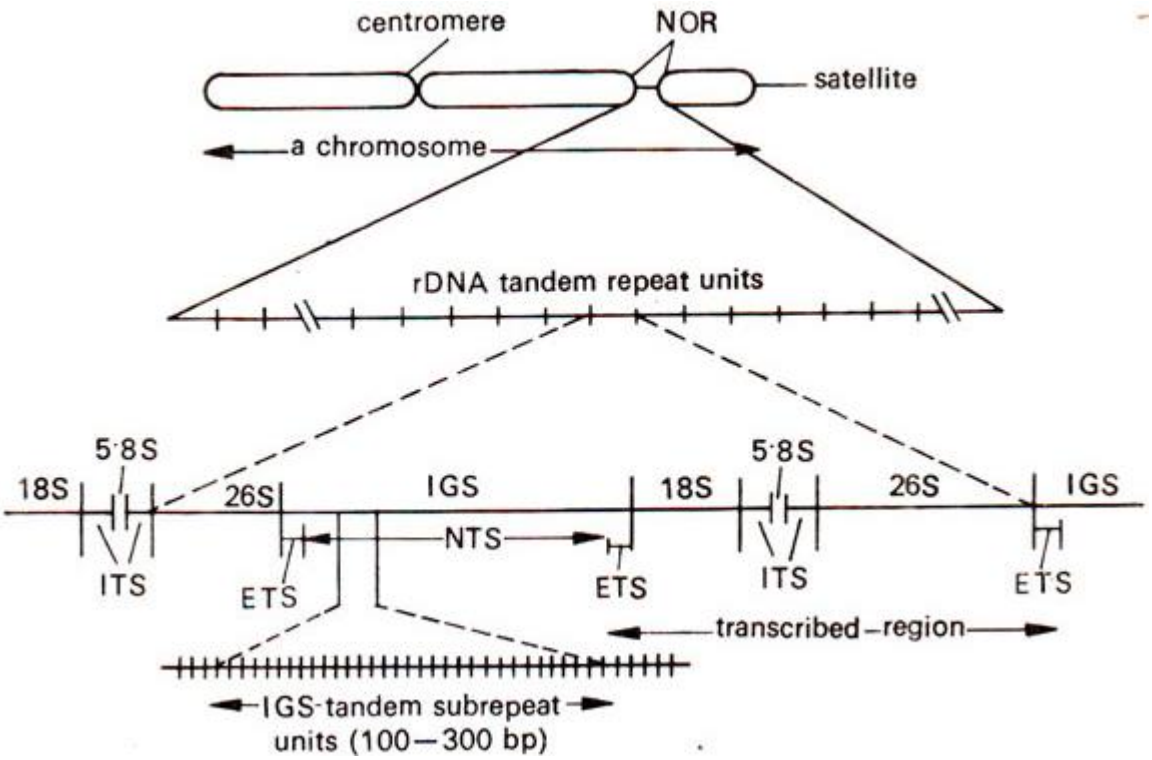


Figure 2.15 Organization of ribosomal DNA (rDNA) in nucleolar organizing region (NOR) of an eukaryotic chromosome (IGS = intergenic spacer; NTS = non-transcribed spacer; ETS = external transcribed spacer; ITS = internal transcribed spacer; 18S, 5.8S, 26S = different genes responsible for 18S rRNA, 5.8S rRNA and 26S rRNA) (Source: (Internet visit 3, 2014).

The different characteristics and properties of barcode markers are summarized in Table 2.3.

Table 2.3 Characteristics of barcode markers (Hollingsworth *et al.*, 2011)

Marker	**Genomic source**	**Type**	**Amplicon length range (bases)**	**Frequency of amplicons with mononucleotide repeats ≥ 10 bases**
nrITS	Nuclear	Transcribed spacers and 5.8S gene	407 – 1630	0.013
matK	Plastid	Protein coding	862 – 910	0.235
rbcL	Plastid	Protein coding	654 – 654	0.000
trnH-psbA	Plastid	Intergenic spacer	226 – 934	0.296

2.13 Applications of Plant DNA Barcoding

Plant DNA barcoding studies has a multiple spectrum of applications. Besides providing information related to species-level taxonomy and the taxonomic process of defining and delimiting species, it can also assist in a process of identifying unknown specimens to known species (Hollingsworth *et al.*, 2011). Apart from these routine uses, barcoding has been used to identify cryptic species in bryophytes (Long *et al.*, 2006; Miwa *et al.*, 2009). Barcoding has also been used in ecological forensics where it is used to identify plant roots, seedlings, or cryptic life stages, e.g. fern gametophytes (Schneider and Schuettpelz, 2006; Gonzalez *et al.*, 2009; Kesanakurti *et al.*, 2011). DNA barcoding can also provide identifications where material has been processed in one way or another, such as analyzing diet of herbivores (Jurado-Rivera *et al.*, 2009; Staudacher *et al.*, 2009; Valentini *et al.*, 2009; Stech *et al.*, 2011), food products (Jaakola *et al.*, 2010), or components of herbal medicine (Srirama *et al.*, 2010). Another emerging application of DNA barcoding in plants is the identification of protected plant species in trade which is enlisted in Convention on International Trade in Endangered Species of wild fauna and flora (CITES). For e.g. Ogden et al. (2008) developed SNP genotyping approach based on *matK* DNA barcodes to distinguish between traded timber products of Ramin (*Gonostylus* spp.) from other con-familial species which are not CITES protected.

2.14 Phylogenetics Study using DNA Sequences

A phylogenetic tree is a graphical representation of the evolutionary relationships among species, genes, genomes or any operational taxonomic unit (OTU) that shares a common ancestor. Phylogenetic relationship is drawn after analysis of similarities and differences among biological entities. These relationships are represented as a branching diagram, or called phylogenetic tree or cladogram, with branches joined by nodes and leading to terminals at the tip of the tree (Harrison and Langdale, 2006). Phylogenetic tree can distinguish three main type of relationship: monophyly, paraphyly and polyphyly (Hennig, 1966). Monophyletic and paraphyletic groups have a single evolutionary origin. Hence, monophyletic groups include all the descendants from a single ancestor as well as the ancestor. But if one lineage emerging from a monophyletic group is removed then what remains is called paraphyletic group. In contrast to this notion, polyphyletic groups result from convergent evolution and thus the characters found in this group are absent in recent common ancestor (Kitching *et al.*, 1998).

Inferring phylogenetic relationship is not straightforward. Similarity of sequences alone cannot be an accurate indicator of degree of relatedness because rates of divergences for a sequence can vary due to realistic evolutionary scenarios such as random chance or difference in mutation rates or selective pressures. Similarly, phylogenetic

reconstruction using branching pattern is oversimplification of evolutionary process and can lead to misinterpretations. For eg: In sexually reproducing species, events like hybridization can lead to mixing and matching between species through Lateral Gene Transfer (LGT) such that DNA can move across large evolutionary distances. Furthermore, duplication, deletion, domain shuffling, gene conversion events etc. can complicate a single genome. Phylogenetic inference can also be complicated by occurrence of convergent and parallel evolution which can lead to lineages becoming similar in some features, often called homoplasy. Hence, only traits that are homologous and in which the observed similarities are not due to convergence or parallel evolution should be compared and contrasted. Sequence data, compared to morphological state data, have several advantages (Harrison and Langdale, 2006). First, the cost of sequencing has dropped thus making sequence data less expensive to gather. Second, sequence data allows study of molecular basis of evolution. Third, discrete and well-defined nature of character traits (i.e., 4 nucleotides, 20 amino acids) makes quantifying trait evolution straightforward and lastly, the principles that apply for molecular sequence data apply to other types of data as well.

Sequence-based phylogenetic analysis comprises of four steps:

1. Selection of sequence of Interest: which maybe whole gene, a region of gene (coding or noncoding), regulatory region for a gene, transposable element, or even a whole genome or amino acid sequence data
2. Identification of homologs
3. Alignment of sequences
4. Construction of phylogeny

Methods of analysis of alignment is determined by the length of computational time and degree of rigor required (Harrison and Langdale, 2006). The main techniques are distance, parsimony and likelihood (including Bayesian analysis).

Distance Methods, such as neighbor joining (NJ), distance and minimum evolution; calculate pair wise distances between sequences and finally group sequences that are most similar. This method is computationally simple and fast. The only disadvantage of this method is it doesn't allow analysis of contribution of a character to particular grouping. The output is often a single tree limiting the examination of conflicting tree topologies. Hence, it can be said that distance based trees are useful in making an initial tree (Harrison and Langdale, 2006).

Parsimony methods assume that shared characters in different entities result from common descent. Thus, groups are built on the basis of such shared characters, and the simplest explanation for the evolution of characters is taken to be the correct or most

parsimonious one. Considering multiple characters, different groupings may be possible, or equally parsimonious, and accordingly multiple trees are generated. In such case, either a strict consensus tree should be built or a majority rule consensus tree should be built. A strict consensus tree includes topologies that are not contradicted in any of the initial trees. If the strict consensus tree is unresolved then it is likely that data used to build the tree are phylogenetically uninformative. While in case of majority rule consensus tree nodes that are consistent in half to all of the most parsimonious trees are displayed and the percentage of trees in which a given topology exists is shown on the branches. Majority rule trees are not informative about phylogeny because node of any one tree that contradicts the other is collapsed (Harrison and Langdale, 2006).

Likelihood method computes the probability of fitting of data set to a tree obtained from the data using a specified model of sequence evolution. The first step is to compare the data against a set of models of sequence evolution and choose the best that describes the observed pattern of sequence variation [for e.g., using software like jModelTest (Posada, 2009)]. The model is then used in likelihood analysis. The analysis is initiated with a specified tree derived from the input dataset and the branches are swapped on the starting tree until a tree with highest likelihood score i.e. the best probability of fitting the data is gained (Harrison and Langdale, 2006).

A phylogenetic tree should provide evidence of ancestor-descendant relationship. When such relationship is absent then the direction of change must be inferred by rooting the trees. The rooting can be done by two ways: Outgroup rooting (Maddison *et al.*, 1984) compares the characters in the group of interest (called in-group) with those in a group that is closely related to but not in the in-group (called out-group). These differences are used to infer the direction of character change in the resultant tree. In second way, duplicated genes are used where sequences from one gene clade are used to root another (Simmons *et al.*, 2000) such that it can reveal unexpected relationships amongst the species or genes in main clades with ambiguous rooting (Gogarten *et al.*, 1989; Brown and Doolittle, 1995). An example of rooted versus unrooted tree can be seen in Figure 2.16 below.

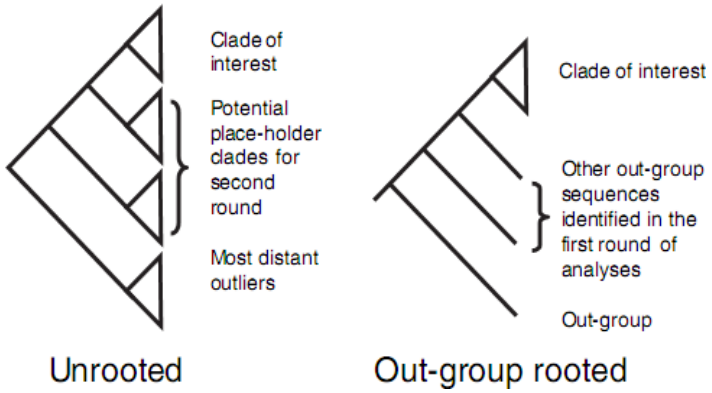


Figure 2.16 Outgroup rooting of a gene family tree. (Source: (Harrison and Langdale, 2006)

2.14.1 Model Selection

Models of DNA substitution process are fundamental to statistical phylogenetic inference. A model of sequence evolution translates the information in a set of sequences into phylogenetic information that can be directly interpreted by a biologist. However, different models make different sets of assumptions about how DNA evolves, resulting in differences of interpretation about the phylogenetic information contained in the sequence. The answer to which model is most appropriate for analyzing a set of DNA sequences is "Model Selection" while any information related to the trustworthiness of the selected model is called "Model Plausibility (Model Adequacy)" (Brown, 2013). Phylogenetic inferences are based upon the use of statistical approach to select an appropriate model of sequence evolution (Sullivan and Joyce, 2005). The more complex the model, the better it fits into the data. More complex models require more parameters. This can give rise to errors. If very simple models of evolution are used that basically ignore fundamentally important processes occurring in nature then it results in "systematic errors (bias)". On the other hand, using very complex models result in "stochastic error (inflated variance)"(Brown, 2013). Hence, model selection methods try to find an approximate model (Burnham and Anderson, 1998) such that these two sources of errors are balanced (Figure 2.17) despite the fact that they are based upon different theoretical foundations. Model Selection Criteria are based upon: Frequentist hypothesis Testing (Likelihood Ratio Test – LRT), Information theory (Akaike's Information Criterion – AIC), Decision Theory (Decision Theoretic Risk – DT or R), Mix of Bayesian motivation and information theory (Bayesian Information Criterion – BIC), Bayesian model selection (Bayesian Factors – BF), Bayesian Hypothesis Testing or Model Averaging (Posterior Probabilities – PP).

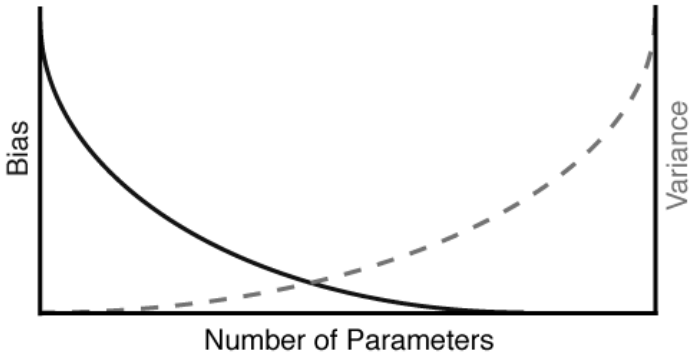


Figure 2.17 The balance between model selection bias and variance according to the number of parameters in a model (Brown, 2013).

The close evolutionary relationship between two different species of *P. kurrooa* and *N. scrophulariiflora* and distribution of both species to a confined region in Himalayas makes it imperative to study the genetic differences not only between these two species but also among different locations of their distribution. Considering both species being monotypic and their high value as medicinal plant, the impact of ensuing global warming and unwarranted collection has rendered both plants endangered and vulnerable. In Nepal, molecular work related to study of medicinal plants has only a recent history. Current study of diversity assessment and intraspecific phylogeny was perceived and carried out using both the PCR-based molecular marker ISSR and sequence based DNA marker. The active constituents that confer medicinal value to the plant was assayed and quantified and also the antioxidant and antibacterial assay of the extract obtained from the rhizome was carried out in order to elucidate its antioxidant potential and antibacterial property.

CHAPTER 3: MATERIALS AND METHODS

3.1 Plant Materials

A total of 43 plant samples of *N. scrophulariiflora* were collected from three different geographical locations of Nepal representing Western, Central and Eastern development regions during the months of June to August, 2013 (Table 3.1, Figure 3.1).

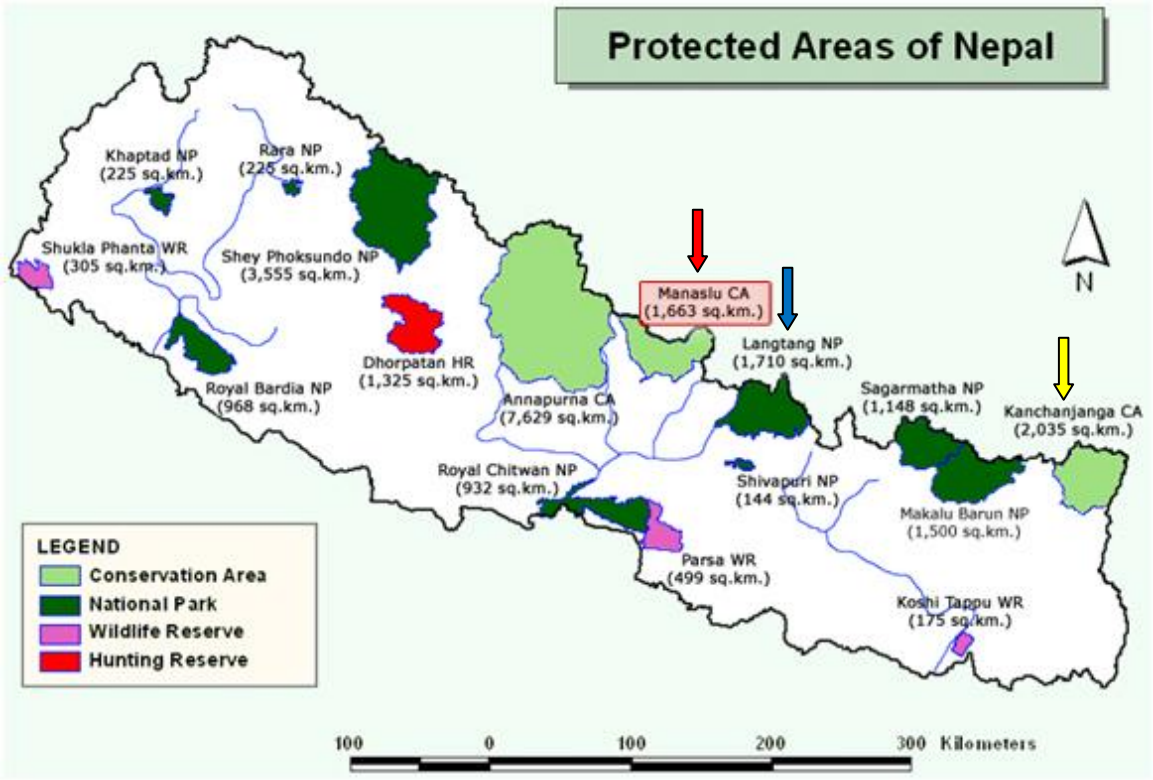


Figure 3.1 Collection sites of different accessions of *N. scrophulariiflora.* Collections were made from three different conserved sites (Kanchenjunga Conservation Area, Langtang National Park and Manaslu Conservation Area) representing three development regions: Eastern, Central and Western respectively.

The collected plants were designated a generic accession code according to the criteria of Molecular Biotechnology Unit, Nepal Academy of Science and Technology (NAST). The leaf samples were placed in paper pouches which were stored in dried silica gel beads in air tight plastic container. The root samples were collected in a plastic bag. All these samples were brought to Molecular Biotechnology Unit Laboratory of NAST for subsequent molecular and chemical analysis.

Herbarium samples were mounted appropriately in Herbarium sheets and stored properly at Molecular Biotechnology Unit laboratory of NAST. The voucher specimens were identified by cross tallying with previously identified Herbarium samples of KATH Central Herbarium, Godavari. One copy of Herbarium sample from each locality of Western, Central and Eastern Nepal have been submitted to KATH, Godavari.

Table 3.1 Details of samples collected from different locations of Nepal

S. N.	Sampling Location	Develop-ment Region	No. of samples collected	Altitude	Latitude	Longitude	Accession code
1.	Bhajiyo, MCAP	Western	9	4200 m	N28^{o}37.678'	E85^{o}8.013'	ME18 - ME27*
2.	Chosyung Lake, MCAP	Western	4	4236 m	N28^{o}35.161'	E85^{o}05.755'	ME28 - ME31
3.	Buddha Mandir, LNP	Central	10	4181 m	N28^{o}05.219'	E85^{o}23.464'	L1 - L10
4.	Dudh Pokhari, KCAP	Eastern	20	4200 m	N27^{o}37.411'	E87^{o}59.680'	K1- K20

*Note: Due to insufficient yield of DNA, sample ME 24 was removed from study.

3.2 DNA Extraction

The CTAB DNA extraction protocol of Graham *et al.* (1994) with slight modifications was used to extract DNA from silica gel preserved dried leaves. Approximately 200 mg dried leaves were ground in sterile mortar and pestles in the presence of liquid N_2. The powdered samples were treated with 1.5 mL of CTAB buffer and the contents were poured into 1.5 mL sterile Eppendorf tubes. The tubes were then incubated at 65^{o}C for 15 minutes. All the tubes were centrifuged at 8,000 g for 15 minutes at room temperature. The supernatants were transferred to a clean and sterile Eppendorf tubes and equal volume of Chloroform: Isoamyl Alcohol (24:1) was added. The tubes were inversed several times for five minutes and subjected to centrifugation at 8,000 g for 2 minutes. Upper aqueous layer of each sample was then transferred to new sterile Eppendorf tube and re-extracted with chloroform: Isoamyl alcohol (24:1). The tubes were centrifuged at 8,000 g for 2 minutes and the upper layers were transferred to new tubes. The samples were then treated with 1/10th volume of Ammonium acetate (7.5 M) followed by addition of equal volume of ice cold absolute ethanol. The tubes were stored at - 20^{o}C overnight to allow precipitation of DNA. Next day, tubes were centrifuged at 8,000 g at 4^{o}C for 20 minutes. The supernatants were discarded and the DNA pellets were quickly washed twice with ice cold 70% Ethanol. Tubes were centrifuged at 8,000 g for 1 minute to get rid of salts. Remaining Ethanol was pipetted out and the pellets were allowed to dry briefly in Laminar Air Flow. Finally, pellets were

re-suspended in TE buffer with RNase and stored at – 20^{o}C for future use.The stored DNA was used for both ISSR and DNA barcoding experiments.

3.3 DNA Quantification

Quantification of DNA was carried out by spectrophotometric method using Biophotometer (Eppendorf-AG22331, Germany). Assessment for both the concentration and purity was carried out.

3.4 Gel Electrophoresis

Extracted DNA as well as the amplified products of ISSR-PCR and DNA barcodes was analyzed using Agarose Gel (1.5% in 1X TAE) in EMBI TEC (Santiago, CA) gel tank. For staining and visualization purpose, 5 µL Ethidium Bromide (10 mg/mL, Promega) was added to the molten gel. A Total volume of 12 µL was loaded into wells which consisted of 2 µL 6X Loading Dye and 10 µL DNA samples or PCR amplified products. A constant current of 50V (8.47 V/cm) was applied for two and half hours. Then, the gels were visualized using Syngene (Syngene Bio-imaging, UK) gel documentation system.

3.5 PCR Optimization

3.5.1 ISSR-PCR Optimization

3.5.1.1 Components Optimization

ISSR-PCR reaction conditions were optimized by varying several parameters such as template DNA concentration, $MgCl_2$ concentration, primer concentration, dNTPs concentration, and *Taq* polymerase concentration (Table 3.2).

Table 3.2 Concentrations of various components used for optimization of ISSR-PCR

S.N.	Components	Concentrations
1.	DNA	10, 15, 20, 25 ng
2.	$MgCl_2$	0.5, 1.0, 1.5, 2.0, 2.5, 3.0 mM
3.	Primers	0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6 µM
4.	dNTPs	0.1, 0.2, 0.3, 0.4, 0.5 mM
5.	*Taq* Polymerase	0.5, 1.0, 1.5, 2.0 U/µL

3.5.1.2 Cycle Optimization

In order to choose best ISSR cycling condition for *N. scrophulariiflora*, three different ISSR-PCR cycling conditions was assessed (Table 3.3).

Table 3.3 Different cycling conditions used for the selection of best ISSR-PCR condition for *N. scrophulariiflora*

S.N.	Program	Cycling conditions		Reference
1.	Program 1	94°C	3 mins	(Tamhankar *et al.*, 2009)
		94°C	30 secs	
		51 °C	45 secs	
		72 °C	2 mins	
		72 °C	5 mins	
		4 °C	5 mins	
2.	Program 2	94 °C	5 mins	(Xiao *et al.*, 2006)
		94 °C	30 secs	
		50 °C	45 secs	
		72 °C	90 secs	
		72 °C	5 mins	
		4 °C	10 mins	
3.	Program 3	94 °C	5 mins	(Liu *et al.*, 2011)
		94 °C	30 secs	
		54 °C	45 secs	
		72 °C	2 mins	
		72 °C	5 mins	
		4 °C	5 mins	

3.6 Primer Screening

Using the optimized ISSR-PCR cycling and reaction conditions, 40 ISSR primers (UBC set 9, University of British Columbia, Vancouver, Canada) were screened using genomic DNA of one of the *N. scrophulariiflora* samples (ME20) from MCAP region. Out of 40 primers, 10 primers that produced multiple scorable bands (Table 4.4) were selected for ISSR-profiling of 43 *N. scrophulariiflora* accessions. Primers producing four or more distinct bands in the gel electrograph were selected for profiling reactions.

3.7 ISSR-PCR Profiling

The extracted DNA of all 43 accessions was used in ISSR profiling experiments for the study of genetic diversity and population genetic study among various populations under study.

3.7.1 Genetic Diversity Assessment based on ISSR Profiling

The polymorphic band was identified as a band that is either present in DNA samples of at least one accession or absent in at least one accession. Individual accessions were screened for presence or absence of observed total number of polymorphic bands after amplification with 10 primers. Bands were scored visually. Molecular sizes of amplified bands were determined by comparing the respective position of bands with 100 bp plus DNA ladder (Gene ruler ^{TM}, Fermentas #SM0323). Total number of bands (TNB), Number of Polymorphic bands (NPB) were recorded.

Population specific markers were also investigated to find out genetic distinctiveness of *N. scrophulariiflora* populations from particular region. The data were analyzed statistically and documented in order to study the genetic diversity of *N. scrophulariiflora* populations based on ISSR markers.

3.7.1.1 Estimation of Percent polymorphism (PP), Polymorphic Information Content (PIC), Band Informativeness (I_B) and Resolving Power (R_p)

Banding pattern created by each primer was used to compute Percent polymorphism (PP), Polymorphic Information Content (PIC), Band Informativeness (I_B), and Resolving power (R_P). Each of them was calculated using Microsoft Excel (2007).

Polymorphic Information Content (PIC) values were calculated using the formula: $PIC = 1 - \sum(P_{ij})^2$, where P_{ij} is the frequency of the i^{th} pattern due to j^{th} primer summed across all patterns revealed by the primers (Botstein *et al.*, 1980). For each primer, Percent Polymorphism (PP) was calculated as NPBs/TNBs (where, NPBs = number of polymorphic bands, TNBs = total number of bands).

The ability of primer combinations to differentiate between accessions, by calculating Resolving Power (R_P) according to Prevost and Wilkins (1999), was computed using the equation, $R_P = \sum I_B$ (I_B is band Informativeness calculated as $I_B = 1 - [2*(0.5-P)]$, where P is the proportion of accessions containing band).

3.7.1.2 Genetic Diversity Analysis using Similarity Matrices and Phenograms

The binary data matrix generated by 10 primers from 43 samples was analyzed using Numerical Taxonomy and multivariate system (NTSYS-PC, v. 2.21i, Exeter Software,

Setauket, New York, USA). Bands were scored as '1' for presence, '0' for absence and '9' for amplification failure denoted as missing data (Transue *et al.*, 1994). Similarity indices were calculated by the software using SIMQUAL algorithm. Based on the similarity matrices, clustering was done using Sequential, Agglomerative, Hierarchial and Nested (SAHN) method using Unweighted Pair Group Method of Artihmetic Average (UPGMA) algorithm (Sneath and Sokal, 1973). Phenograms were used to show relationship between different populations based on these matrices. Similarity matrices were created based on three different coefficients as presented in Table 3.4 below.

Table 3.4 Coefficients used to create similarity matrices

Coefficient	Formula	Reference
Simple Matching (SM)	$Sij = \frac{a+d}{a+b+c+d}$	(Sokal and Michener, 1958)
Dice's Coefficient of Similarity (D)	$Sij = \frac{2a}{2a+b+c}$	(Dice, 1945; Nei and Li, 1979)
Jaccard's Coefficient (J)	$Sij = \frac{a}{a+b+c}$	(Jaccard, 1908)

Where,

Sij = the similarity between 2 individuals, i and j;

a = the number of bands present in both i and j;

b = the number of bands present in i and absent in j;

c = the number of bands present in j and absent in i

d = the number of bands absent in both i and j

Similarly, using NTSYS-PC cophenetic correlations were also calculated. Matrices were then compared using Mantel test (Mantel, 1967) based on obtained SM, Dice and Jaccard similarity coefficients. In order to construct consensus tree, Consensus Fork Index or CI_c values were calculated. Using these parameters the best fit matrix and tree was computed.

A 3D-plot was also computed that was used to study the distribution of *N. scrophulariiflora* accessions. Using NTSYS – PC v. 2.21i, DCenter values were first calculated in order to calculate the Eigen vector values. The Eigen vector values were then used to calculate a 3D plot which represents the respective distribution of various populations with each other.

3.7.1.3 Principal Coordinate Analysis (PCoA)

Relationship between populations of *N. scrophulariiflora* accessions was also studied using Principle Coordinate Analysis (PCoA) module of NTSYS-PC. The similar analysis was performed using GenALEx ver 6.5 software (Peakall and Smouse, 2006; Peakall and Smouse, 2012) by standardized covariance method.

3.7.1.4 Population Genetic Analysis

Using POPGENE version 1.32 (Yeh *et al.*, 1997), within and among diversity between three populations of *N. scrophulariiflora* was obtained using binary matrix obtained from ISSR profiling data. POPGENE is windows based standalone software used to analyze genetic variation within and among natural populations using co-dominant or dominant markers and quantitative traits based on haploid or diploid data. Genetic relatedness between populations of *N. scrophulariiflora* was calculated using Nei's genetic identity (I_{xy}) and genetic distance (D_{xy}) based on unbiased measures (Nei, 1973). This relatedness was illustrated with a phenogram based on UPGMA by using POPGENE software and was compared to phenogram obtained from NT-SYS PC.

a. Genetic Structure

Nei's distance was calculated based on the following formula.

$$d_{ij} = -ln\left(\frac{\sum_k \left|X_{ki}X_{kj}\right|}{\sqrt{\sum_k X_{ki}^2 X_{kj}^2}}\right)$$

Where, summation indicates the summation of all loci (k) between 2 individuals, i and j.

Genetic variation within and among populations of *N. scrophulariiflora* was also assessed with another parameter called Shannon's information index using the formula, $I = -\sum P_i * \log_2 P_i$ (Lewontin, 1995) where I refers to diversity and P_i is the frequency of a particular ISSR band. Similarly, Nei's genetic diversity (Nei, 1973) was used to calculate within and among genetic diversity of three populations using the formula, $H = 1 - \sum X_i^2$ where (if the marker is dominant), X_i is the population frequency of each allele (1 and 0) at any locus i.

b. Genetic Differentiation

Likewise, heterozygosity was measured as total and mean heterozygosity. Total heterozygosity, $H_T = 2q_i(1-q_i) + var(q_i)$ (where, q_i is the frequency of the null allele at i^{th} locus in a population) was calculated using Lynch and Milligan's correlation (1994).

Mean heteozygosity within population (H_S) was calculated using

$$H_S = \sum_{i=1}^{i=k} p^2$$

Where, p is the mean frequency of the i^{th} allele (1 or 0) at the locus k in each population and the value is averaged over all populations.

Diversity among populations was calculated using equation $D_{ST} = H_T - H_S$ and the coefficient of population differentiation was calculated using $G_{ST} = D_{ST}/H_T$. The relative magnitude of genetic differentiation among subpopulations, G_{ST}, was determined according to Nei (1987). Number of Immigrants (N_m, gene flow per generation) for each locus was calculated using the formula, $N_m = 0.5\ (1 - G_{ST})/ G_{ST}$. G_{ST} values were also used to estimate the mean value across loci (McDermott and McDonald, 1993).

c. Analysis of Molecular Variance (AMOVA)

Genetic variation was also evaluated from the binary data matrix, for all ISSR loci, by analysis of molecular variance (AMOVA) using the statistical program GENEALEx 6.5 (Peakall and Smouse, 2006; Peakall and Smouse, 2012). AMOVA is a statistical model for computing molecular variation in a single species by estimating population differentiation directly from molecular data (Excoffier *et al.*, 1992). AMOVA is based on Euclidean distances between pairs of vectors which are any kind of raw molecular data expressed as Boolean vector p_i and is given by, $E = (\epsilon^2_{XY}) = n\ [1\text{-}2n_{XY}/2n]$, where, $2n_{XY}$ is the number of markers shared by two individuals (X and Y) and n is the total number of polymorphic markers (Tsuda *et al.*, 2004). Molecular variance within the population was calculated as an indication of intra-population genetic variation. The significance of variance components was tested by calculating their probabilities, based on 999 random permutations using the program GENALEx v 6.5 (Peakall and Smouse, 2006; Peakall and Smouse, 2012).

3.8 DNA Barcoding

3.8.1 Samples

Selected eight samples of DNA used in ISSR experiment were used for DNA barcoding experiments. The samples were given various codes before being sent for sequencing (Table 3.5).

Table 3.5 Sample names and codes for various barcoding loci used for DNA sequencing

S.N	Barcoding loci	Sample code	Sample
1.	ITS	INS 1, INS 2	ME19, ME20
		INS 6, INS 7, INS 8	L4, L5, L6

		INS 11, INS 12, INS 13	K16, K17, K18
2.	*matK*	MNS 19, MNS 20	ME 19, ME 20
		MNS 4, MNS 5, MNS 6	L4, L5, L6
		MNS 16, MNS 17, MNS 18	K16, K17, K18
3.	*rbcL*	RNS 19, RNS 20	ME 19, ME 20
		RNS 4, RNS 5, RNS 6	L4, L5, L6
		RNS 16, RNS 17, RNS 18	K 16, K 17, K 18
4.	*trnH-psbA*	TPNS 19, TPNS 20	ME 19, ME 20
		TPNS 4, TPNS 5, TPNS 6	L4, L5, L6
		TPNS 16, TPNS 17, TPNS 18	K16, K17, K18

3.8.2 PCR Optimization and Amplification

3.8.2.1 ITS-PCR Optimization

The amplification of DNA templates were done in BioER thermal cycler (BioER, China). To assess the optimum parameters for best amplification, experimentation with varying PCR reaction parameters were conducted.

ITS sequences were amplified using 2X Master mix (Promega). The components of Master Mix (Promega) were as tabulated in Table 3.6.

Table 3.6 Components of 2X Master Mix obtained from Promega

S.N.	Components	Concentration
1.	*Taq* DNA Polymerase (pH 8.5)	50 U/mL
2.	dNTPs (dATP, dGTP, dCTP, dTTP)	400 µM (each)
3.	$MgCl_2$	3 mM

Components in master mix were predefined hence only template DNA concentration and primer concentration could be subjected to experimentation for optimization study. Different DNA concentrations (10, 15, 20, 25, 30, 40, 50 ng) were tested while different primer concentrations that were experimented were 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, and 0.8 µM as shown in Table 3.7. Other components were used after diluting according to the concentration required.

Table 3.7 Optimization of different components for ITS Amplification

S.N.	Components	Concentrations used
1.	Template DNA	10, 15, 20, 25, 30, 40, 50 (ng)
2.	Primer	0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 µM

3.8.2.2 Amplification of ITS, *matK, rbcL* and *trnH-psbA* regions

Four standard DNA barcode regions were chosen for the study. Two of the regions were coding sequences (*matK* and *rbcL*) and are considered core barcodes for flowering plants as suggested by CBOL Plant working group (2009) while two were non-coding intergenic spacers (ITS and *trnH-psbA)*. The primers used for amplification of these loci are summarized in the Table 3.8.

Table 3.8 Primers used for amplification of different barcode loci

Locus	Primers	Sequence (5'-3')	Reference
ITS	ITS 1 (F) ITS 4 (R)	TCCGTAGGTGAACCTGCGG TCCTCCGCTTATTGATATGC	(White *et al.*, 1990)
matK	472F (F) 1248R (R)	CCCRTYCATCTGGAAATCTTGGTTC GCTRTRATAATGAGAAAGATTTCTGC	(Yu *et al.*)
rbcL	rbcLa-F (F) rbcLa-R (R)	ATGTCACCACAAACAGAGACTAAAGC GTAAAATCAAGTCCACCRCG	(Levin *et al.*, 2003) (Kress and Erickson, 2007)
trnH-psbA	psbA3_f (F) trnHf_05 (R)	GTTATGCATGAACGTAATGCTC CGCGCATGGTGGATTCACAATCC	(Sang *et al.*, 1997) (Tate and Simpson, 2003)

Amplification was done using 2X PCR Master Mix (Promega, USA). PCR program and the PCR components used were similar for *matk, rbcL* and *trnH-psbA* while the program for nrITS was different. The PCR amplification was carried out in 25 µL reaction mixture. In case of MgCl2 concentration for *matk, rbcL* and *trnH-psbA* amplification, a further 1 mm was added to original 1.5mM from master mix making the final concentration 2.5 mm.

The amplification of DNA templates were done in BioER thermal cycler (BioER, China). The different components used for different barcode loci are tabulated in Table 3.9.

Table 3.9 Different reaction conditions used for amplification of different barcode loci

S.N	Barcode loci	Components	Concentration used
1.	*matK,* *rbcL,* *trnH-psbA*	2X Master Mix	1X (12.5 µL)
		Primers (Table 3.7)	0.4 µM
		tDNA	25 ng
		$MgCl_2$ (added)	1 mM (Total 2.5 mM)
2.	nrITS	2X Master Mix	1X (12.5 µL)
		Primers (Table 3.7)	0.4 µM
		tDNA	25 ng
		$MgCl_2$	1.5 mM

The cycling conditions used for amplification of different barcode loci are presented in Table 3.10.

Table 3.10 Different cycling conditions used for amplification of different barcode loci

S.N.	Barcode loci	Cycling conditions	No. of cycles
1.	*matk* *rbcL* *trnH-psbA*	94°C – 2 mins	35 Cycles
		94 °C – 35 secs	
		52 °C – 40 secs	
		72 °C – 55 secs	
		72 °C – 7 mins	
		4 °C – hold	
2.	ITS	94°C – 3 mins	Step 1
		50 °C – 1 min	
		72 °C – 2 mins	

		94 °C – 30 secs	39 cycles
		50 °C – 30 secs	
		72 °C – 75 secs	
		72 °C – 5 mins	
		25 °C – 2 mins	
		4 °C – 20 mins	

The amplified products were visualized after electrophoresis using 1.5% Agarose gel in TAE buffer (1X) at 80V for half an hour with Ethidium Bromide (EtBr) staining (0.5 µg/mL) and photographed under ultraviolet light using Syngene Gel doc system (Syngene Bio-imaging, UK).

3.8.3 DNA Sequencing

Among the collected accessions, 8 out of 43 samples, representing three populations under study (MCA, LNP and KCA) were selected for DNA sequencing. Accessions ME19 and ME20 from MCAP; L4, L5 and L6 from LNP; and K16, K17 and K18 from KCAP were sequenced (Table 3.5).

A total of 20 µL each of amplified DNA samples were sent for bidirectional sequencing to Macrogen (Macrogen Inc.), South Korea.

3.8.4 Sequence Analysis

The sequence data generated by Macrogen, Korea after sequencing of all the samples were retrieved and assembled using CodonCode Aligner version 4.0.4 software (CodonCode Corporation Centerville, MA 02632, USA). Consensus sequences were generated for each locus of individual sample.

Using the consensus sequences, the most similar sequences were searched in the NCBI nucleotide database using BLAST (Altschul *et al.*, 1990). The most similar sequences were aligned with the assembled sequences using CodonCode Aligner version 4.0.4 and adjusted manually. Multiple sequence alignment was performed using ClustalW (Thompson *et al.*, 1994) option of Molecular Evolution Genetics Analysis (MEGA) Version 5.2 (Tamura *et al.*, 2011). The intraspecific and interspecific variation of each barcoding region was characterized by calculating Kimura 2-parameter (K2P) distances in MEGA Version 5.2.

3.8.5 Model based Phylogeny

For model based phylogeny, the statistical selection of best-fit model of nucleotide substitution was done using jModelTest v. 2.1.4 (Darriba *et al.*, 2012) using Akaike's Information Criteria corrected (AIC_c). Maximum likelihood scores were calculated using 88 nucleotide substitution models in jModelTest v. 2.1.4 and base trees were searched using Nearest Neighbor Interchange (NNI) algorithm.

3.8.6 Phylogenetic Analysis based on Maximum Parsimony (MP)

Maximum Parsimony (MP) trees based on *matK, rbcL, trnH-psbA* and ITS sequences were constructed in MEGA v. 5.2 with 1000 bootstrap replications to determine clade support. Branch swapping algorithm used was Subtree-Pruning-Regrafting (SPR) algorithm. Condensed tree with 50% cutoff value was used to depict the phylogenetic relationship.

3.8.7 Barcoding Gap and Candidate Barcode Selection

The barcoding gaps were calculated by comparing intra- and interspecific divergences using K2P distances of each candidate locus using TaxonDNA version 1.7.8 (Meier *et al.*, 2006). "Best match/Best close match" method was used to identify the best barcode for interspecific and intraspecific discrimination ability.

3.9 Antimicrobial Susceptibility Testing

3.9.1 Microbial Strains

The following microorganisms from American Type Culture Collection (ATCC) were used as test organisms in antimicrobial assay (Table 3.11). Among them, two were Gram positives (*Bacillus subtilis, Staphylococcus aureus*) and six were Gram negatives (*Pseudomonas aeruginosa, Enterobacter sp*, *Salmonella typhimurium*, *Escherechia coli, Klebsiella pneumonia, Serratia sp*). The cultures were diluted to achieve optical densities corresponding to 1.5×10^6 colony forming units (CFU/mL). The strains were kindly provided by Professor Jae Kyung Sohng, Sun Moon University, South Korea.

Table 3.11 ATCC strains of bacteria used in the Antimicrobial Assay

Test Microbial Strains	ATCC accession	Remarks
Bacillus subtilis	6051	Gram positive
Staphylococcus aureus	25925	Gram positive
Pseudomonas aeruginosa	2785	Gram negative
Enterobacter sp.	23373	Gram negative

Serratia sp.	13880	Gram negative
Salmonella typhimurium	14028	Gram negative
Escherechia coli	25922	Gram negative
Klebsiella pneumoniae	700603	Gram negative

3.9.2 Methodology

Antimicrobial susceptibility assay was performed as previously described by (Irshad *et al.*, 2012) with modifications as required. Glycerol stocks of the ATCC strains were incubated at 37^{o}C for half an hour to facilitate revival. Under the laminar air flow, a loopful of stock suspensions was transferred to the respective labeled culture broth using sterile loop. The standard strains were then incubated at 37^{o}C for 24 hours. Culture broths after 24 hours were maintained at 0.5 McFarland turbidity standards before culturing them on Nutrient Agar plates for antimicrobial assay. The culture tubes were aligned with the 0.5 McFarland tube against black and white stripes drawn on white paper manually. The visibility of the stripes was compared. Further dilutions of the culture suspension, if needed, were done using sterile Nutrient broth.

Using sterile cotton swabs, 0.5 McFarland bacterial solutions were spread on surface of Mueller Hilton Agar (MHA) prepared for each organism. Separate swabs were used for different test plates. The plates were allowed to dry for 5 minutes under laminar air flow. Using sterile and separate cork borers (6 mm diameter), five holes were bored in each plate. On any three of the holes, 20 µL of respective extract solutions from different locations were loaded, one of the holes was loaded with methanol as a negative control while Chloramphenicol disc (50 µg) was kept as positive control. The plates were left at room temperature for one hour to facilitate diffusion of extract solution as well as the antibiotic. The plates were then placed in an incubator at 37^{o}C overnight. The diameter of zones of inhibition was measured with standard scale. All studies were performed in triplicate and finally the mean value was recorded.

3.10 Chemical Analysis

3.10.1 Solvents and Chemicals

All the chemicals including solvents were of HPLC grade. Analytical grade methanol and water were purchased from Merck (Merck Company Pvt. Ltd., India).

3.10.2 Preparation of Extract from Dried Rhizomes

Mature plants of *N. scrophulariiflora* were collected from different locations (Figure 3.1, Table 3.1). A total of 43 plant rhizomes were collected from three different locations (MCA, LNP and KCA). Each rhizome collected was at least 5 meters apart to ensure that the rhizomes were not derived from same rhizome. The plant materials were divided into separate groups according to the locations they were collected. Rhizomes from same location were grouped as one and analyzed for their properties. Each group was air-dried at room temperature and ground to a fine powder using a grinder. The weight of the powder was recorded.

The extraction was performed with cold extraction method. Each of the powdered rhizomes were mixed with 400 mL of Methanol and left for a week at room temperature. The methanolic extracts were filtered using Whatman filter no. 4 and the filtrate was dried by evaporation at reduced pressure and temperature (40^{o}C). The dried filtrate was left at room temperature in order to facilitate evaporation of remnant methanol traces. A second extraction was performed using same method. The extracts were mixed and left for final evaporation at room temperature and weighed.

3.10.3 Preparation of Stock Solutions and Calibration of Standard Solutions

Pure reference compounds of Picrosides I and II were purchased from Sigma-Aldrich (St. Louis, MO, USA). Stock solutions (1 mg/mL) of pure reference compounds were prepared in HPLC grade water (Merck Co. Pvt. Ltd, Mumbai, India) and stored in refrigerator at 4^{o}C. From stock solutions, working solutions of 1 µg/mL, 10 µg/mL, 100 µg/mL, 500 µg/mL and 1000 µg/mL were prepared by dilution with HPLC grade water (Merck Co. Pvt. Ltd, Mumbai, India). These working solutions of both reference compounds were analyzed separately. A 10 µL aliquot of the working solutions of different concentrations of each reference compounds were used.

The reference compounds were analyzed by Shimadzu HPLC system (Kyoto, Japan) consisting of LC – 20AD pump, DGU – 20A3R degasser unit, CTO – 20A column oven, and SPD – M20A detector and CBM – 20Alite as system controller. LabSolutions software v. 5.51 (Shimadzu, Kyoto, Japan) was used for data analysis and data processing. The samples were analyzed at 32^{o}C on VP-ODS C-18 column (150 mm × 4.6 mm, 5 µm particle size). Detection was done by Photo Diode Array (PDA) at 270 nm. Analysis was carried out using a mobile phase of methanol: water (2:3) at a flow rate of 0.7 mL/min. A standard calibration curve for both reference compounds were plotted using the software.

3.10.4 HPLC Analysis of Root Extracts and Quantification of Picroside I and Picroside II

25 mg of different root extracts was weighed and dissolved in 1 mL of HPLC grade methanol (Merck Co. Pvt. Ltd., Mumbai, India). The extracts were filtered using 0.45 µm filter paper. Extract solutions (10 µL) were injected manually into HPLC system. Data were analyzed using LabSolutions software v. 5.51. Using the software, the concentrations of Picrosides I and II were determined for samples from different locations using the standard calibration curves.

3.11 Antioxidant Assay

3.11.1 Chemicals

1,1-diphenyl, 2-picryl hydrazyl (DPPH) (Sigma Aldrich, USA) was kindly provided by Dr. Laksmaiah Sreerama, Minnesota State University, Minnesota, USA. Methanol and Gallic acid was purchased from Merck Co. Pvt. Ltd. (Mumbai, India). Extract solutions were prepared by weighing 25 µg methanolic extract in 1 mL of methanol.

3.11.2 DPPH Radical Scavenging Assay

Stock solutions of 50 mg/mL sample solutions from different locations were diluted serially to make final concentration of 0.1, 0.5, 1, 5, 10, 20, 30, 40 and 50 mg/mL working solutions. Gallic acid was used as standard. A stock solution of 1 mg/mL Gallic acid was serially diluted to make standard solutions of 10, 20, 30, 40, 50, 60, 70, 80, 90 and 100µg/mL. DPPH solution was prepared by dissolving 4 mg DPPH in 100 mL methanol (final concentration equivalent to 1 mM). The DPPH solution was wrapped immediately in Aluminum foil and kept in dark to prevent spontaneous oxidation.

The assay was performed as described by (Formagio *et al.*, 2013) with slight modifications. Control solution was prepared using 10 µL Methanol mixed in 990 µL of DPPH solution and was stored in dark at room temperature for 30 minutes. 10 µL of various sample concentrations were added to 990 µL of fresh DPPH solution. The mixture was shaken and left to stand at room temperature in the dark for 30 minutes. The process was repeated for Gallic acid. After 30 minutes, using methanol as a blank, the absorbance was measured at 517 nm using spectrophotometer. All experiments were performed in triplicates.

The IC_{50} (concentration required for 50% inhibition of DPPH) values were calculated using the following equation in Microsoft Office Excel 2007 (Microsoft Corporation, USA):

$$\mathrm{IC}\ 50 = \mathrm{Exp}\ (\ln(\mathrm{conc} > 50) - \{\frac{(\mathrm{PI} > 50 - 50)}{(\mathrm{PI} > 50 - PI < 50)} \times \ln\frac{\mathrm{conc} > 50}{\mathrm{conc} < 50}\}$$

Where, IC_{50} = Inhibitory concentration; PI = Percentage Inhibition and conc = concentration of sample solution.

A graph of percentage inhibition (% I) versus the extract concentration in mg/mL was depicted. The PI of DPPH (% I) was calculated using the following equation:

$$\%\, I = \left(\frac{A_0 - A}{A_0}\right) \times 100$$

Where, A_0 = absorbance of DPPH (control) and A = absorbance of sample with DPPH.

3.11.3 Statistical Analysis

The results were expressed as Mean±SD and student's t-test and one-way ANOVA were used to assess statistical significance of the data obtained from the experiments using GraphPad Prism version 5 for Windows (GraphPad Software, San Diego California USA, www.graphpad.com). P-values <0.05 were considered to be significant.

CHAPTER 4: RESULTS

4.1 Ecological Information of Samples

A total of 43 samples were collected from three different populations of *N. scrophulariiflora* representing three different geographical regions (Eastern, Central and Western Development regions) of Nepal (Table 3.1).

4.2 DNA Extraction and Quantification

DNA from all the accessions were successfully extracted using CTAB method (Graham *et al.*, 1994). DNA quantification and estimation was done using Biophotometer (Eppendorf, Germany). DNA concentrations of all forty three samples were estimated which ranged from 46.9 µg/mL to 626.0 µg/mL (Appendix III).

4.3 ISSR Profiling

4.3.1 Cycling Conditions

Of the three ISSR-PCR cycling conditions described by Liu *et al.* (2011), Tamhankar *et al.* (2009) and Xiao *et al.* (2006) and assessed for the selection of best condition for *N. scrophulariiflora* ISSR profiling, PCR program with following cycling conditions according to (Xiao *et al.*, 2006) was found to be the best for *N. scrophulariiflora* (Table 4.1).

Table 4.1 Reaction conditions selected for ISSR profiling of *N. scrophulariiflora* (Xiao *et al.*, 2006).

Initial Denaturation	94^{o}C for 2 mins	
Denaturation	94^{o}C for 30 secs	40 cycles
Annealing	50^{o}C for 45 secs	
Elongation	72^{o}C for 90 secs	
Final Elongation	72^{o}C for 5 mins	
Hold	4^{o}C for 10 mins	

4.3.2 Reaction Conditions

All of the ISSR-PCR reaction conditions were optimized using sample ME20 and Primer 809. A negative control was also maintained for each experiment. The different parameters tested are presented in Table 4.2.

Table 4.2 Various parameters used for the PCR-optimization and selection of optimized parameters.

S.N	PCR parameters	Range tested	Selected optimized condition	Remarks
1	DNA concentration (ng)	10, 15, 20, 25	20 ng	Most of the concentration produced the reproducible bands yet 20 ng of template DNA was selected as optimum condition because it is the minimum concentration that produces discernible bands (Plate 4.1 A).
2	$MgCl_2$ concentration (mM)	0.5, 1.0, 1.5, 2.0, 2.5, 3.0	2 mM	Intense bands were observed at 1.5 – 3.0 mM concentrations and 2.0 mM $MgCl_2$ concentration was taken as optimum concentration for subsequent experiments as it revealed more number of bands (Plate 4.1 B).
3	Primer concentration (µM)	0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6	0.7µM	Multiple banding patterns were observed from primer concentration of 0.2 µM – 1.6µM. For subsequent ISSR-PCR profiling experiments 0.7µM concentration was selected as it produced darker and more number of bands (Plate 4.1 C).
4	dNTPs concentration (mM)	0.1, 0.2, 0.3, 0.4, 0.5	0.2mM	ISSR banding patterns showed that 0.2 mM concentration of dNTPs produced dark bands. So 0.2 mM concentration was selected as optimum concentration (Plate 4.1 D)

5	*Taq* polymerase concentration (U)	0.5, 1.0, 1.5, 2.0	1 U	As 1U of *Taq* polymerase showed better bands it was chosen for subsequent profiling reactions (Plate 4.1 E).

From the optimization experiments, the final optimized reaction conditions for ISSR-profiling of *N. scrophulariiflora* was determined to be 20 ng of genomic DNA, 2.0 mM of $MgCl_2$, 0.7 µM of primer concentration, 0.2 mM of dNTP concentration, and 1 U Taq polymerase buffered with 10X PCR Reaction Buffer (Fermentas, 100 mM Tris-HCl pH 8.8 at 25^{o}C, 500 mM KCl, 0.8% Nonidet P_{40}).

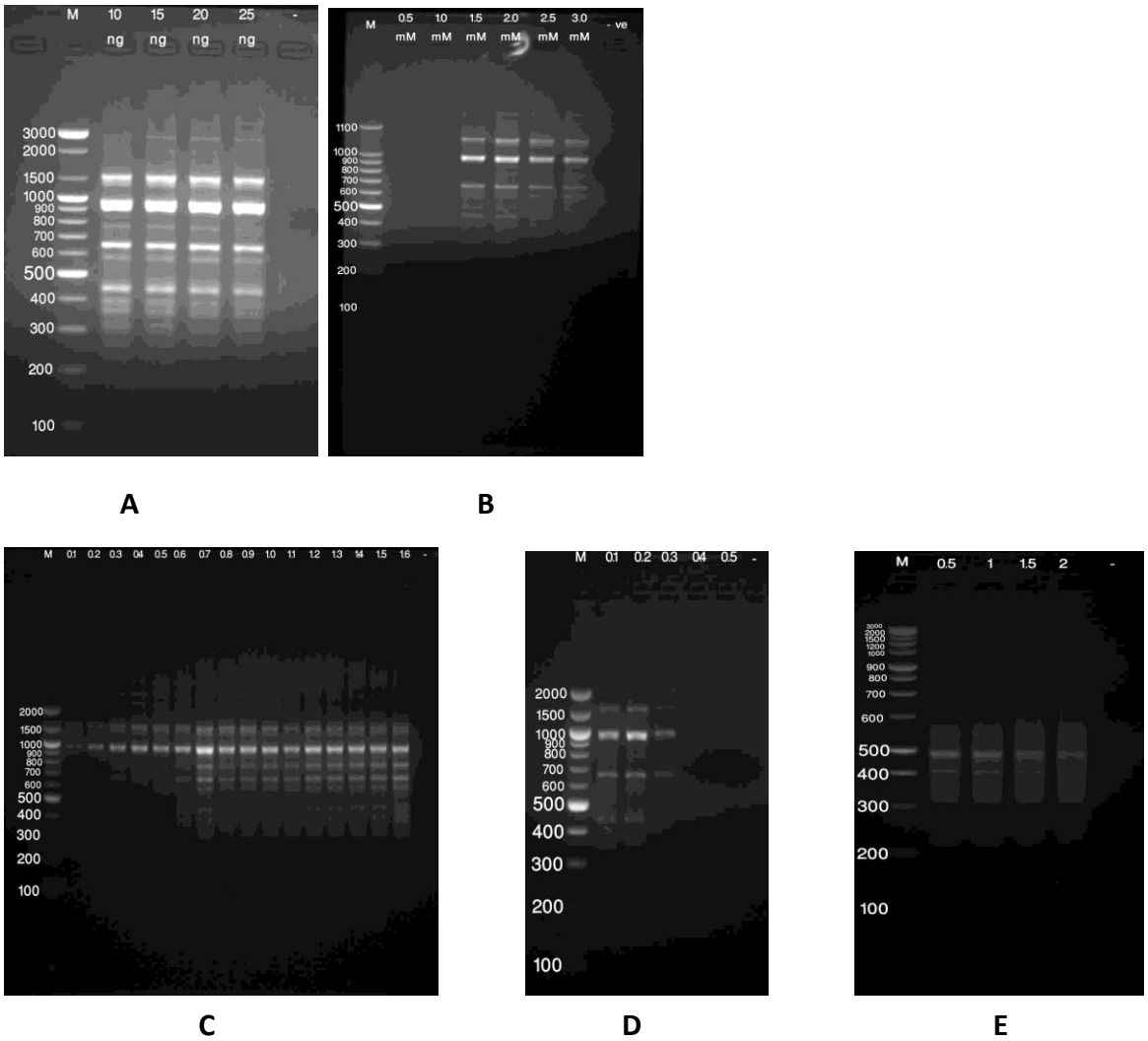


Plate 4.1 Gel picture of ISSR-PCR optimization with varying PCR reaction parameters using primer UBC 809 and DNA sample of ME20. Lanes marked 'M' are 100 bp+ molecular weight markers; (A) template DNA (10 – 25 ng) used; (B) $MgCl_2$ (0.5 – 3.0

mM); (C) primer concentrations of UBC 809 (0.1 µM – 1.6 µM) used; (D) dNTPs concentration used for the experiment (0.1 – 0.5 mM); (E) *Taq* polymerase experimented with (0.5 – 2 U).

4.4 Primer Screening

Using the optimized ISSR-PCR reaction and cycling conditions, total of 40 UBC primers were screened.

Out of 40 primers, only 10 primers were selected for ISSR profiling and diversity assessment based on presence or absence of bands and number of amplified bands for each primer (Plate 4.2; Table 4.3). Primers that produced high number of clear and crispy bands were chosen for further study.

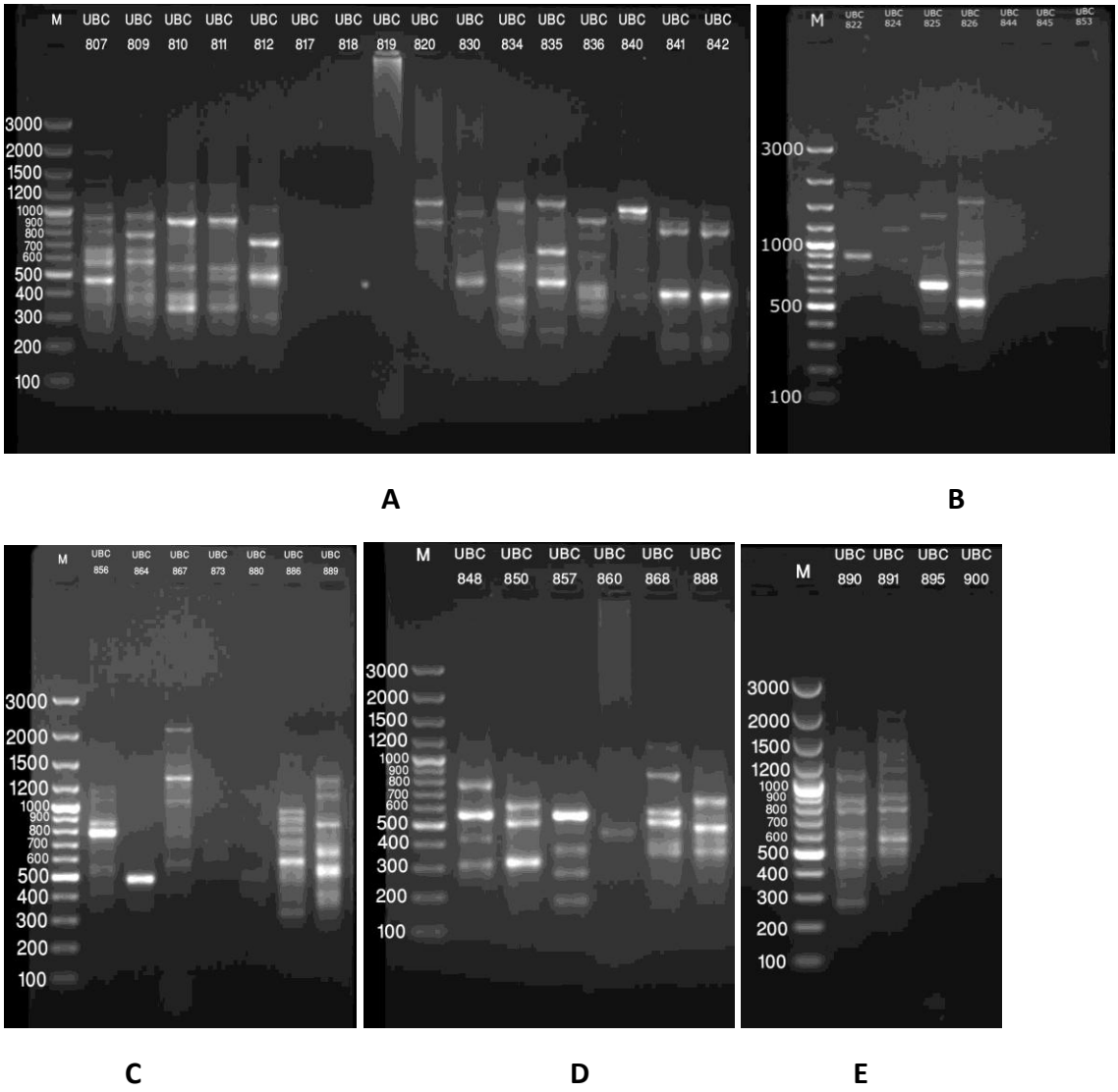


Plate 4.2 Gel photograph of primer screening using ME 20 as a template DNA. Lanes marked 'M' represent 100 bp+ molecular weight markers. Pictures (A) (B) (C) (D) (E) represents different UBC primers used in the experiment.

Table 4.3 ISSR primers selected from primer screening experiment for further study.

S.N.	Primer Code	5' – 3' sequence	GC content (%)	Length
1.	UBC 807	AGA AGA GAG AGA GAG AGA GT	52.9	20mer
2.	UBC 809	AGA GAG AGA GAG AGA GG	52.9	17mer
3.	UBC 811	GAG AGA GAG AGA GAG AC	52.9	17mer
4.	UBC 834	AGA GAG AGA GAG AGA GYT	47.2	18mer
5.	UBC 835	AGA GAGA GAG AGA GAG YC	52.8	18mer
6.	UBC 848	CAC ACA CAC ACA CAC ARG	52.8	18mer
7.	UBC 850	GTG TGT GTG TGT GTG TYC	52.8	18mer
8.	UBC 856	ACA CAC ACA CAC ACA CYA	44.4	18mer
9.	UBC 868	GAA GAA GAA GAA GAA GAA	33.3	18mer
10.	UBC 886	VDV CTC TCT CTC TCT CT	50	17mer

Where, Y = T/C (pyrimidine), R = G/A (Purine), V = G/C/A (all but T)

4.5 ISSR profiling and population-specific ISSR Markers for *N. scrophulariiflora*

ISSR profiling of three different populations of *N. scrophulariiflora* with selected 10 primers showed that certain sequences were found in one population while it was absent in other populations. Primers UBC 807, 835, 848, 850 and 868 were found to produce unique bands. These bands are found to be population specific and can be useful to differentiate one population from other. These bands can be used for population specific discrimination of *N. scrophulariiflora* and also as a fingerprint for this plant (Table 4.4).

Table 4.4 Population-specific ISSR marker of *N. scrophulariiflora* (+ = present; - = absent).

Primers	Marker (bp)	Populations obtained from		
		MCA	LNP	KCA
UBC 807	450	+	-	-
UBC 835	1000	-	+	-
UBC 848	490	-	+	-
	320	-	+	-
UBC 850	1600	+	-	-
	490	-	+	-
	390	-	+	-
UBC 868	750	-	+	-
	650	-	-	+
	400	-	-	+

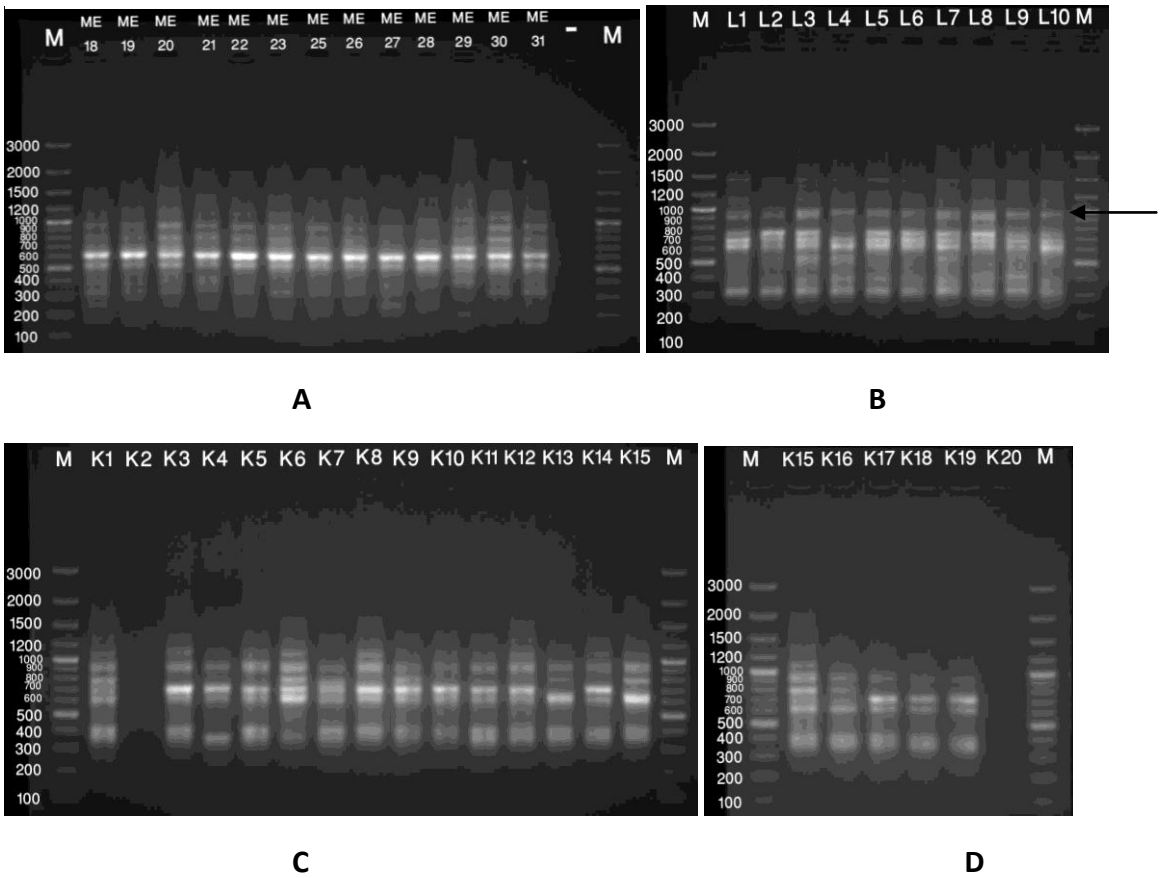


Plate 4.3 ISSR profile of all accessions using UBC 835. Lanes marked 'M' are 100 bp+ DNA marker. (A) Accessions of MCA ; (B) Accessions of LNP; (C) and (D) Accessions of KCA.

Dark arrow indicates population specific marker for *N. scrophulariiflora* from LNP (1000 bp).

4.6 ISSR Profiling and Assessment of Genetic Diversity

Out of forty ISSR primers screened, only ten UBC primers were used to study the genetic diversity within and between plant populations of *N. scrophulariiflora*. A total of 197 bands were amplified with ten primers in all forty three accessions. The band size ranged from 100-2000 bp. The average number of bands per primer was 19.7. Out of total amplified bands, 180 (91.37%) were found to be polymorphic while 17 (8.63%) were found to be monomorphic. The polymorphism percentage ranged from 78% to 95% for individual primers. Table 4.5 summarizes the data obtained from the analysis.

Table 4.5 ISSR profiles of the amplified products of ten ISSR primers. Total number of amplified band (TNB), number of polymorphic band (NPB), Percentage of polymorphic band (PPB) and range of amplicon size of ten ISSR primers used to generate ISSR profiles of forty three *N. scrophulariiflora* accessions.

Primers	TNB	NPB	PPB	Amplicon size (bp)
UBC 807	24	22	91.67	100-1300
UBC 809	26	24	92.00	290-1700
UBC 811	18	17	94.44	250-1900
UBC 834	19	15	**78.95**	200-2000
UBC 835	16	15	93.75	250-1500
UBC 868	19	17	89.47	300-2000
UBC 848	16	15	93.75	250-1600
UBC 850	21	20	**95.24**	300-1750
UBC 856	13	12	92.31	300-1200
UBC 886	25	23	92.00	100-1500
Total	197	180		

4.6.1 Polymorphic Information Content (PIC), Band Informativeness (I_B) and Resolving Power (R_p)

Table 4.6 summarizes the data obtained from all analyzed bands. Polymorphic Information Content (PIC) score was calculated for each primer and was found to be in the range of 0.839 for UBC 856 to 0.928 for UBC 850 with an average of 0.901. The average band informativeness (I_B) for 10 primers was found to be 0.456. I_B value ranged from 0.238 for UBC 886 to 0.685 for UBC 834. The maximum resolving power (R_P) of 10 primers ranged from 13.023 for UBC 834 to 5.953 for UBC 886 with an average of 8.512 (Table 4.6).

Table 4.6 Polymorphic Information Content (PIC), Band Informativeness (I_B) and Resolving power (R_P) of ten ISSR primers used in the study

Primers	PIC	Range of I_B	Average I_B	R_P
UBC 807	0.899	0.000 – 1.349	0.304	7.302
UBC 809	0.915	0.000 – 1.070	0.263	6.837
UBC 811	0.901	0.000 – 1.349	0.401	7.209
UBC 834	0.908	0.000 – 1.907	**0.685**	**13.023**
UBC 835	0.912	0.000 – 1.256	0.567	9.070
UBC 868	0.907	0.000 – 1.488	0.487	9.256
UBC 848	0.896	0.000 – 1.860	0.683	10.930
UBC 850	**0.928**	0.000 – 1.488	0.434	9.116
UBC 856	**0.839**	0.000 – 1.581	0.494	6.419
UBC 886	0.907	0.000 – 1.116	**0.238**	**5.953**
Average	**0.901**		**0.456**	**8.512**

4.6.2 Genetic Diversity Estimate based on Cluster Analysis

Based on the binary data matrix of ISSR loci, different similarity indices were calculated using the SIMQUAL algorithm of NTSYS-pc ver 2.21i software. The similarity indices were calculated using Simple Matching (SM), Jaccard's (J) and Dice (D) coefficients. The coefficients of similarity for SM ranged from (0.561 – 0.990), J ranged from (0.043 – 0.951), and D ranged from (0.083 – 0.975) with an average coefficient of similarity 0.747, 0.289, and 0.421 respectively (Appendix 2, 3 and 4). Using the three similarity

coefficients, phenograms were generated that depicted the genetic relationship between 43 accessions collected from three different geographical locations. The phenogram topologies obtained using SM, J, and D coefficients were compared to each other.

4.6.2.1 Phenogram generated from Simple Matching Coefficient (SM)

The genetic similarities calculated from Simple Matching coefficient among 43 accessions ranged from 0.561 (between K6 and L7) to 0.990 (between ME30 and ME31). The topology of phenogram from Similarity Matching coefficient showed three major clusters (i.e. Cluster I, II and III). Cluster I comprised of MCA population while Cluster II comprised of KCA population. Cluster III comprised of population from LNP.

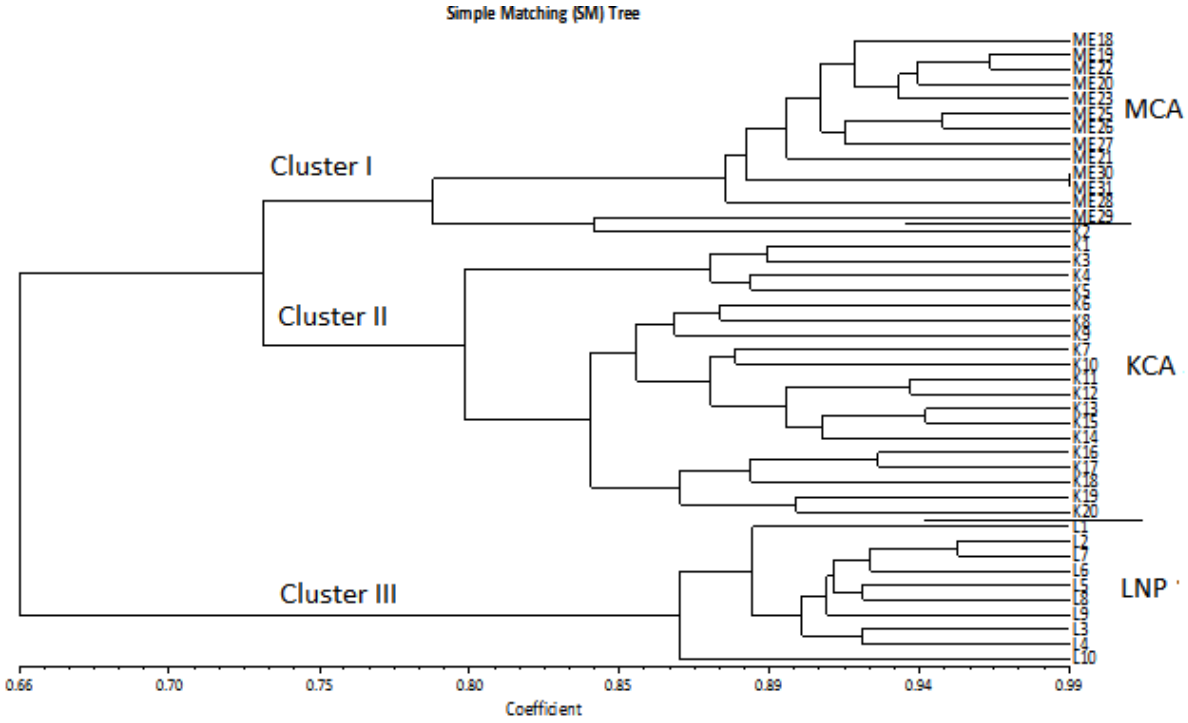


Figure 4.1 Phenogram depicting relationship between forty-three *N. scrophulariiflora* accessions based on Simple Matching (SM) coefficient using UPGMA clustering method.

4.6.2.2 Phenogram generated from Jaccard Coefficient (J)

The genetic similarities calculated from Jaccard coefficient among 43 accessions ranged from 0.043 (between L5 and K2) to 0.951 (ME30 and ME31). The topology of the phenogram generated from Jaccard coefficient showed four clusters (i.e. Cluster I, II, III and IV). Cluster I comprised of accessions from MCA, Cluster II comprised accessions

from KCA, Cluster III comprised accessions from LNP and Cluster IV comprised of accession K2 from KCA.

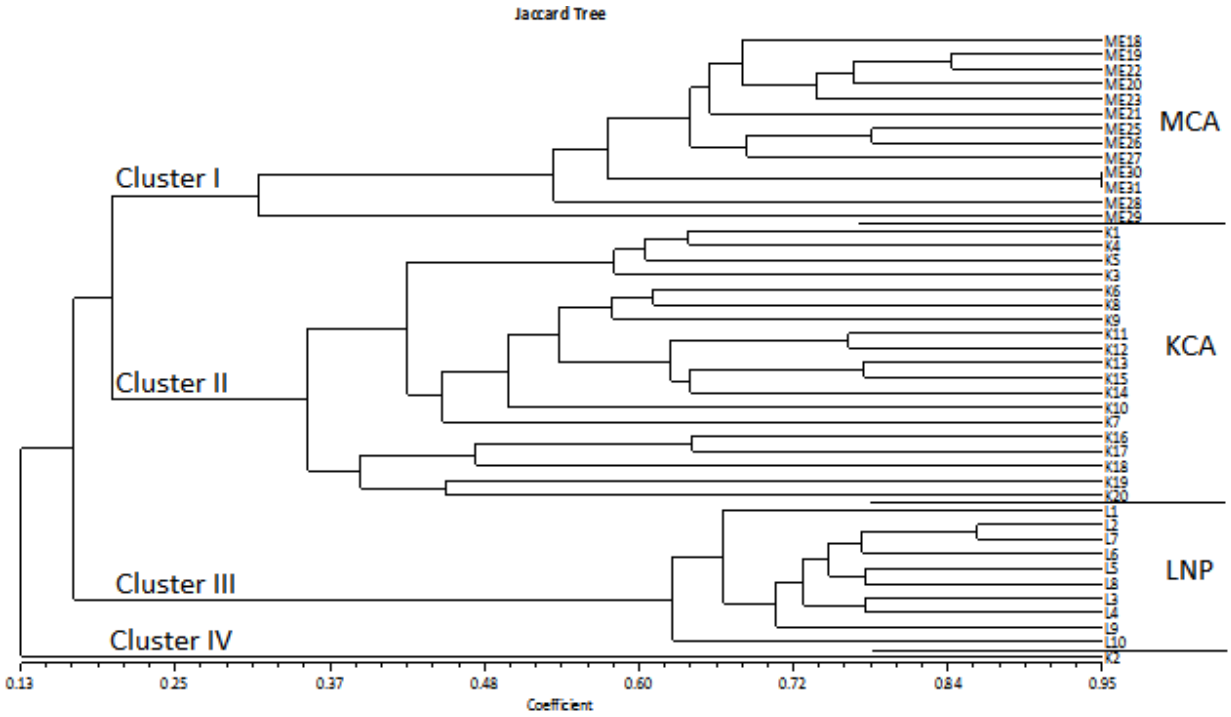


Figure 4.2 Phenogram depicting relationship between forty-three *N. scrophulariiflora* accessions based on Jaccard's (J) coefficient using UPGMA clustering method.

4.6.2.3 Phenogram generated from Dice Coefficient (D)

The genetic similarities calculated from Dice coefficient among 43 accessions ranged from 0.083 (between L5 and K2) to 0.975 (ME30 and ME31). The topology of the phenogram generated from Dice coefficient showed four clusters (i.e. Cluster I, II, III and IV). Cluster I comprised of accessions from MCA, Cluster II comprised accessions from KCA, Cluster III comprised accessions from LNP and Cluster IV comprised of accession K2 from KCA.

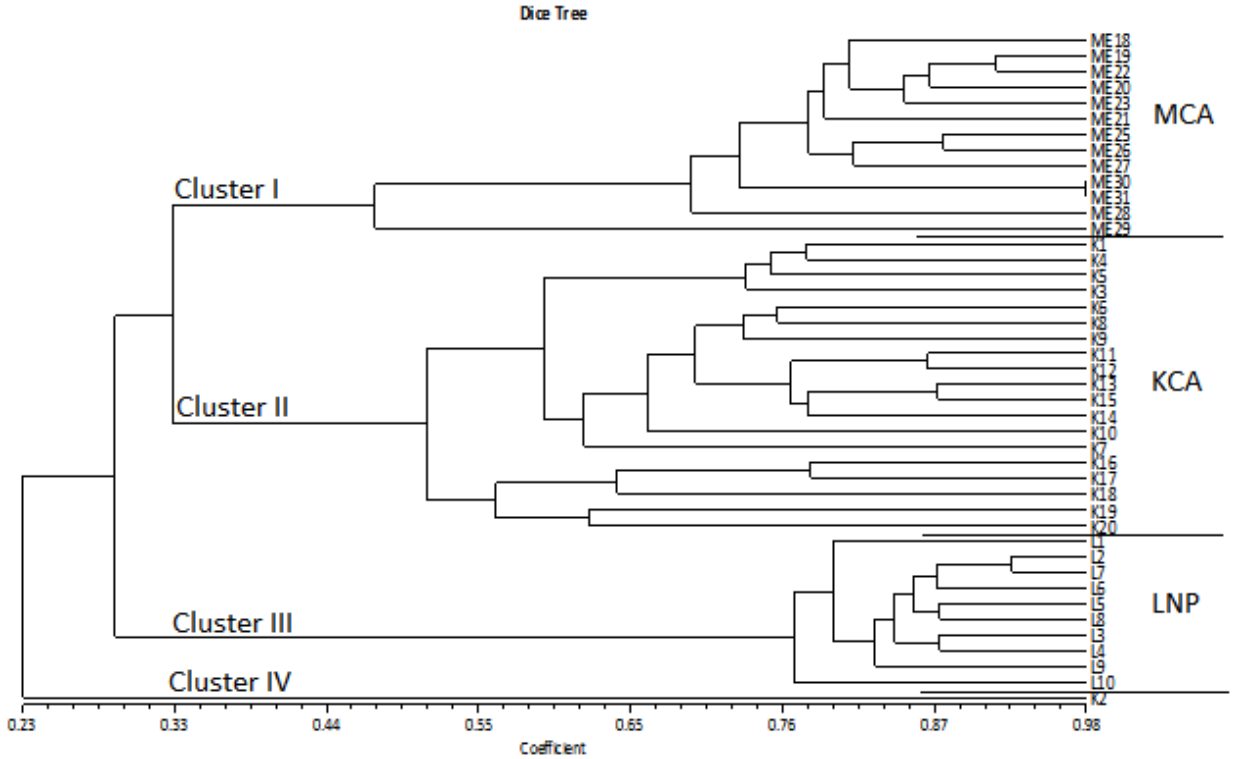


Figure 4.3 Phenogram depicting relationship between forty three *N. scrophulariiflora* accessions based on Dice (D) coefficient using UPGMA clustering method.

Correlation values, obtained by comparing original matrices, were calculated using 2-way Mantel test (Mantel, 1967) as shown in Table 4.7. The result of Mantel test showed highest correlation between Jaccard and Dice similarity matrix (0.99278) and thus was most significant.

Table 4.7 Correlation coefficients from 2-way Mantel test between similarity matrices

Coefficient	Simple Matching	Jaccard	Dice
Simple Matching	*****		
Jaccard	0.88608	*****	
Dice	0.88949	0.99278	*****

Similarly, clustering based on Unweighted Pair Group Method of Arithmetic Averages (UPGMA) for Jaccard coefficient was observed to give a high cophenetic correlation value of 0.96609 and comparatively, the lowest cophenetic correlation value was obtained for Simple Matching coefficient (Table 4.8). Because of highest correlation value and comparison of standard chart for goodness of fit (Table 4.9), Jaccard coefficient is best suited for studying the genetic relatedness among various *N. scrophulariiflora* accessions under study.

Table 4.8 Cophenetic correlation coefficient values (r) of similarity matrices (Simple Matching, Dice and Jaccard) and cluster computed by UPGMA module using MXCOMP (matrix comparison) option of NTSYS-PC.

Clustering Method	Simple Matching	Dice	Jaccard
UPGMA	0.95469	0.95071	0.96609

Table 4.9 Test of goodness of fit criteria based on correlation values (Rohlf, 2009)

Range	Interpretation
0.9 < r	Very good fit
0.8 < r < 0.9	Good fit
0.7 < r < 0.8	Poor fit
r < 0.7	Very poor fit

Where, r = cophenetic correlation coefficient value.

The study showed that Jaccard similarity with UPGMA clustering yielded highest correlation coefficient value but the difference between Jaccard, Dice and Simple Matching coefficients were not wide (Table 4.8). Hence, the three coefficients show decreasing order (J>SM>D) of use in interpreting the genetic relationship among *N. scrophulariiflora* accessions.

For evaluating the trees constructed using UPGMA clustering method and three similarity coefficients, consensus indices (CI_C) values were also calculated (Table 4.10). Highest consensus fork index, CI_C, was found for Dice and Jaccard coefficients (1.0000) while CI_C values for both SM and J (0.7561) and SM and D were lower (0.7610).

Table 4.10 Consensus fork Index (CI_C) values among UPGMA based phenograms produced by different similarity coefficients of *N. scrophulariiflora* accessions.

	Simple Matching	Jaccard	Dice
Simple Matching	*****	0.7561	0.7610
Jaccard		*****	1.0000
Dice			*****

On the basis of Jaccard's similarity coefficient (Appendix 3), genetic similarity was estimated within and among *N. scrophulariiflora* populations and pair wise comparison between all the collected accessions was performed (Table 4.11). It was observed that

percent similarity values for samples collected from MCA ranged from 26 – 95%. Similarly, the within population similarity for LNP ranged from 59 – 86%. In KCA populations, the similarity ranged from 13 – 77%. The result obtained indicated the diverse nature of the samples within the same population. Pair wise comparison of all the accessions revealed that least similarity was observed between L5 and K2 (4%) hence they were most distant, followed by ME29 and K20 (6%) and ME29 and L8 (9%). The most similar accessions were found to be ME 30 and ME 31 (95%) from MCA region. Percent similarity indices based on similarity matrix generated by Jaccard coefficient for different populations are shown in Table 4.11.

Table 4.11 Percent Similarity observed between different *N. scrophulariiflora* populations based on Jaccard similarity matrix within and among populations.

Pop/Pop	MCA (%)	LNP (%)	KCA (%)
MCA	26 - 95	9 – 26	6 – 32
LNP		59 – 86	4 – 28
KCA			13 - 77

Where, pop = populations

4.6.2.4 Cluster Analysis

Phenogram obtained using UPGMA clustering based on Jaccard's similarity matrix using NTSYS-PC v. 2.21i showed the genetic relationship of *N. scrophulariiflora* accessions collected from different locations. Four major clusters were seen at a similarity level of 0.34 which is marked by dotted line in Figure 4.4. Clusters I and II are separated at a similarity level of 0.199 and clusters III is separated from clusters I and II at a similarity level of 0.179.

Cluster I contains accessions from MCA population while Cluster II contains all the accessions from KCA population except K2 which is separated as a distinct branch, Cluster IV. Cluster III shows accessions from LNP population. Cluster IV contains a single accession of K2. It is segregated from rest of populations. The phenogram depicting this information are presented in Figure 4.2.

Based on Jaccard similarity matrix (Appendix 3), the range of similarity level for MCA population was found to be 0.262 (between accessions ME27 and ME29) to 0.952 (between accessions ME30 and ME31) with a mean similarity level of 0.580. For LNP population, the range was found to be 0.586 (between accessions L1 and L10) to 0.857 (between accessions L2 and L7) with a mean similarity level of 0.701. Similarly, for KCA population, the range was found to be 0.133 (between accessions K2 and K20) to 0.771 (between accessions K13 and K15) with a mean similarity level of 0.408. Similarly, the

similarity level for MCA and KCA ranged from 0.086 (between accessions ME29 and L8) to 0.259 (between accessions ME25 and L4) with a mean similarity level of 0.196. For populations LNP and KCA the similarity level was found to be in the range of 0.043 (between accessions L5 and K2) to 0.282 (between accessions L5 and K3) with a mean similarity level of 0.150. For populations MCA and LNP, the range of similarity level was found to be 0.061 (between accessions ME29 and K20) to 0.324 (between accessions ME19 and K9) with a mean similarity level of 0.196.

4.6.2.5 3D Plot

3D plot with Eigen vectors was constructed for all *N. scrophulariiflora* accessions using Jaccard similarity matrix. The accessions were grouped based on comparative dispersion and scatterings of population clusters.

The plot diagram shows that KCA and MCA population clusters are close to each other than LNP population cluster. Accessions of KCA population are more scattered while in case of MCA and LNP most of the accessions are less scattered (Figure 4.4).

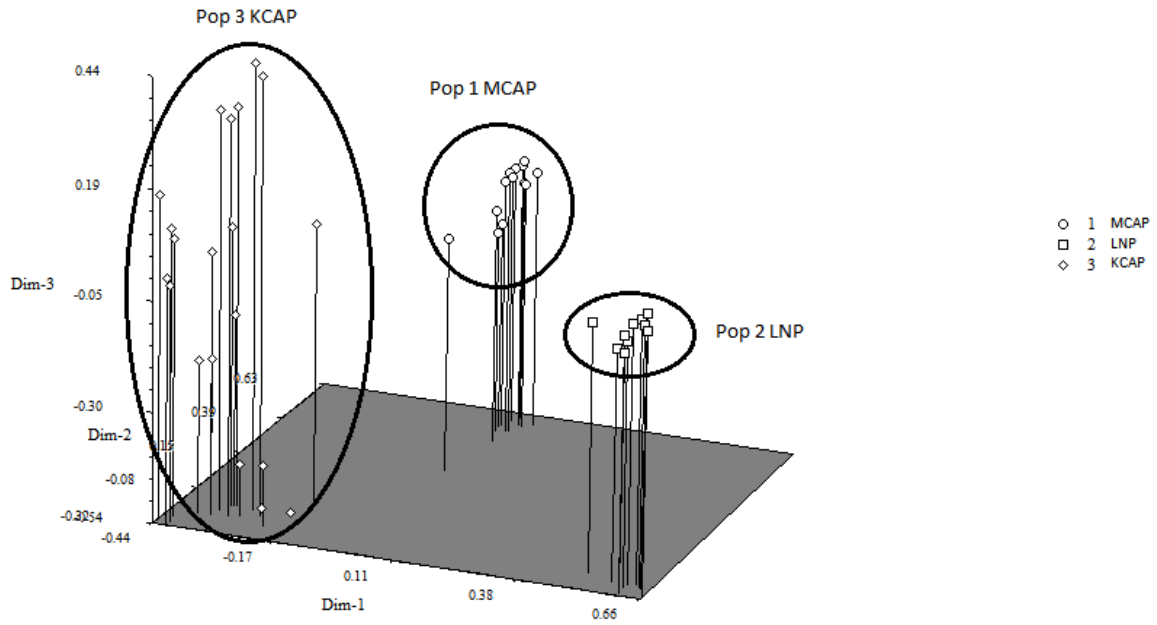


Figure 4.4 A 3D Plot constructed using NTSYS-PC version 2.21i to study dispersion of *N. scrophulariiflora* accessions.

4.6.3 Principal Coordinate Analysis (PCoA)

Principal Coordinate Analysis (PCoA) was calculated based on Nei's genetic distance (Nei, 1973) using GenALEx version 6.5 software (Peakall and Smouse, 2006; Peakall and Smouse, 2012). The PCoA analysis revealed three distinct clusters of three populations of *N. scrophulariiflora* (Figure 4.5).

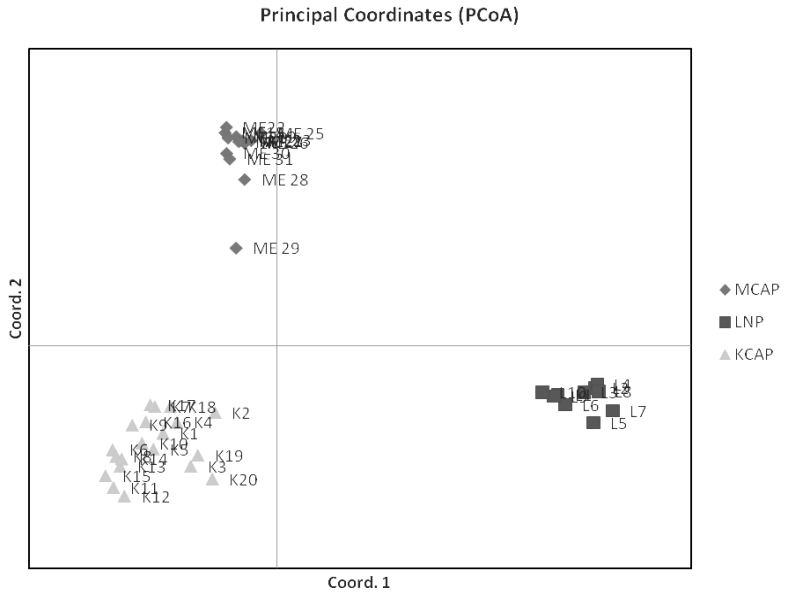


Figure 4.5 Principal Coordinate Analysis (PCoA) plot based on standardized covariance method.

The percentages of variation as seen in the three axes were calculated which is a criteria to select significant axes that can provide information on total variance of a dataset (Table 4.12). The first two PCoA axes accounted for 64.87% of the variance.

Table 4.12 Percentage of variation as seen by first three axes in PCoA analysis

Axis	**1**	**2**	**3**
%	37.63	27.24	14.58
Cum %	37.63	64.87	79.45

4.6.4 Study of Population Genetic Structure of *N. scrophulariiflora*

4.6.4.1 Genetic Identity and Distance among Populations of *N. scrophulariiflora*

Nei's (1972) original measure of genetic Identity (I_{xy}) and genetic distance (D_{xy}) along with Nei's (1978)) unbiased measure of genetic identity and distance were estimated to illustrate the genetic relationship among the populations under study by using POPGENE Version 1.32 (Yeh *et al.*, 1997).

Highest genetic identity was observed between populations of MCA and KCA populations (0.9006 as Nei's original measure and 0.9047 as Nei's unbiased measure) and thus the least genetic distance (0.1047 as Nei's original measure and 0.1002 as Nei's unbiased measure) was observed between these populations. Similarly, lowest genetic identity was observed between MCA and LNP populations (0.7920 as Nei's original

measure and 0.7961 as Nei's unbiased measure) hence, highest genetic distance (0.2332 as Nei's original measure and 0.2280 as Nei's unbiased measure) was observed between them. Similar result was shown by the phenogram where MCA and KCA populations were clustered together whereas LNP population was clustered separately. The results are shown in Figure 4.6 and 4.7.

```
=========================================
pop ID    1         2         3
=========================================
1           ****    0.7920    0.9006
2         0.2332      ****    0.7954
3         0.1047    0.2290      ****
=========================================

Nei's genetic identity (above diagonal) and genetic distance (below diagonal).


   ********************************************************************************
   **                                                                            **
   **         Dendrogram Based Nei's (1972) Genetic distance: Method = UPGMA     **
   **            --Modified from NEIGHBOR procedure of PHYLIP Version 3.5        **
   **                                                                            **
   ********************************************************************************



                               +------------------------pop1
  +----------------------------1
--2                            +------------------------pop3
  !
  +-------------------------------------------------------pop2
```

Figure 4.6 Dendogram based on Nei's original measure of genetic identity and genetic distance. Pop 1 refers to MCA population, Pop 2 refers to LNP population and Pop 3 refers to KCA population as obtained using POPGENE version 1.32 (Yeh *et al.*, 1997)

```
========================================
pop ID    1         2         3
========================================
1           ****    0.7961    0.9047
2         0.2280      ****    0.7992
3         0.1002    0.2241      ****
========================================

Nei's genetic identity (above diagonal) and genetic distance (below diagonal).


   ****************************************************************************************
   **                                                                                    **
   **         Dendrogram Based Nei's (1978) Genetic distance: Method = UPGMA             **
   **            --Modified from NEIGHBOR procedure of PHYLIP Version 3.5                **
   **                                                                                    **
   ****************************************************************************************



                                +-----------------------pop1
  +-----------------------------1
--2                              +-----------------------pop3
  !
  +-------------------------------------------------------pop2
```

Figure 4.7 Dendogram based on Nei's unbiased measure of genetic identity and genetic distance. Pop 1 refers to MCA population, Pop 2 refers to LNP population and Pop 3 refers to KCA population as obtained using POPGENE version 1.32 (Yeh *et al.*, 1997).

4.6.4.2 Genetic Diversity within and among Populations of *N. scrophulariiflora* and at Species Level

Assuming Hardy-Weinberg equilibrium, POPGENE v. 1.32 was used to calculate the variety of genetic diversity parameters, such as, the percentage of polymorphic bands (PPB), Nei's genetic diversity (H), Shannon's Information Index (I) and Gene differentiation (G_{ST}).

Of the total 197 DNA bands, 180 (91.37%) DNA bands were observed to be polymorphic at the species level. The percentage of polymorphic DNA bands for individual populations varied from 66.27% (for LNP population) to 96.26% (for KCA population) with an average of 83.81%. Nei's gene diversity (H) ranged from 0.2284 (for LNP population) to 0.2439 (for KCA population) with an average of 0.2348 at population level and 0.2070 at species level. Shannon's information index (I) ranged from 0.3413 (for LNP population) to 0.3836 (for KCA population) with average estimate of 0.3628 at population level and 0.3282 at species level (Table 4.13).

From this it can be said that, the highest level of genetic diversity occurred in population of KCA (PPB, 96.26%; H, 0.2439; I, 0.3836) while the lowest level occurred in LNP population (PPB, 66.27%; H, 0.2284; I, 0.3413).

Table 4.13 Genetic variability within populations of *N. scrophulariiflora* as observed with POPGENE v. 1.32 using ISSR data.

Population	Sample number (N)	Number of Polymorphic bands	PPB (%)	H	I
MCA	13	80	88.89	0.2320	0.3636
LNP	10	55	**66.27**	**0.2284**	**0.3413**
KCA	20	103	**96.26**	**0.2439**	**0.3836**
Mean	14.33	79.33	83.81	0.2348	0.3628
Species Level	43	180	**91.37**	**0.2070**	**0.3282**

Where, PPB = Percentage of polymorphic bands; H = Nei's genetic diversity; I = Shannon's information Index.

4.6.4.3 Total Genetic Differentiation

Genetic differentiation and diversity estimates revealed by POPGENE are summarized in Table 4.15. The data were generated using the ISSR binary matrix in POPGENE. The heterozygosity within populations (H_S) was observed to be in the range of 0.0593 to 0.823 (at loci UBC 807 and 809 respectively) with an average of 0.190. The total diversity (H_T) ranged from 0.1322 to 0.3180 (at loci UBC 809 and 835 respectively) with an average of 0.222. The mean genetic differentiation between populations (G_{ST}) was found to be 0.468 within range of 0.2652 (at loci UBC 834) to 0.660 (at loci UBC 807). The average gene flow, N_m, was found to be 0.650 ranging from 0.2573 (at loci UBC 807) to 1.385 (at loci UBC 834).

Table 4.14 Total **g**enetic differentiation and diversity indices of *N. scrophulariiflora* shown by different primers.

Primers	H_T	H_S	G_{ST}	N_m
UBC 807	0.1747	**0.0593**	**0.6602**	**0.2573**
UBC 809	**0.1322**	**0.823**	0.3775	0.8247
UBC 811	0.1780	0.1228	0.3105	1.1105

UBC 834	0.2525	0.1855	**0.2652**	**1.3853**
UBC 835	**0.3180**	0.1326	0.5829	0.3578
UBC 848	0.2947	0.1415	0.5199	0.4617
UBC 850	0.2641	0.1320	0.5001	0.4997
UBC 856	0.2170	0.1152	0.4690	0.5661
UBC 868	0.2484	0.1111	0.5528	0.4046
UBC 886	0.1374	0.0773	0.4377	0.6422
Mean	0.2217	0.19003	0.46758	0.6501
SD	0.06439	0.22505	0.12297	0.35718

Where, N_m = Gene flow; H_T = Total Heterozygosity/diversity; H_S = Mean heterozygosity/gene diversity within population; G_{ST} = Mean Genetic differentiation.

4.6.5 Analysis of Molecular Variance (AMOVA)

The AMOVA showed 50% variation among and within populations ($p < 0.001$) (Table 4.15). The among population value of the indicator of genetic differentiation, Φ_{PT}, for three populations was 0.501. This suggested that genetic differentiation of *N. scrophulariiflora* within and among populations was same irrespective of their distribution and geographical location.

Table 4.15 Analysis of molecular variance (AMOVA) for 43 individuals in three populations from three regions with degree of freedom (d.f.), sum of squares (SS), mean square (MS), estimated variance, percent variance (%) and significance (n = 999 permutations)

Source of variation	**d.f.**	**SS**	**MS**	**Estimated variance**	**Total variance (%)**	**P – value**
Among Populations	2	510.424	255.212	17.341	50	<0.001
Within Populations	40	691.250	17.281	17.281	50	<0.001
Total	42	1201.674		34.622	100	

4.7 DNA Barcoding

4.7.1 PCR Amplification and Sequencing

The universal primers specific for four different barcode loci (*matK, rbcL, trnH-psbA*, ITS) (Table 3.7) with specific reaction and cycling conditions (Tables 3.8 and 3.9) were used for PCR amplification. All primers successfully amplified the barcode loci. However, only eight samples (ME19, ME20, L4, L5, L6, K16, K17 and K18) representing three different populations were selected for sequencing. On visualization with gel electrophoresis, the approximate band sizes were found to be 800, 600, 400, and 700 bps for *matK, rbcL*, *trnH-psbA* and ITS respectively (Plate 4.4). Sequencing was done by Macrogen Inc., South Korea.

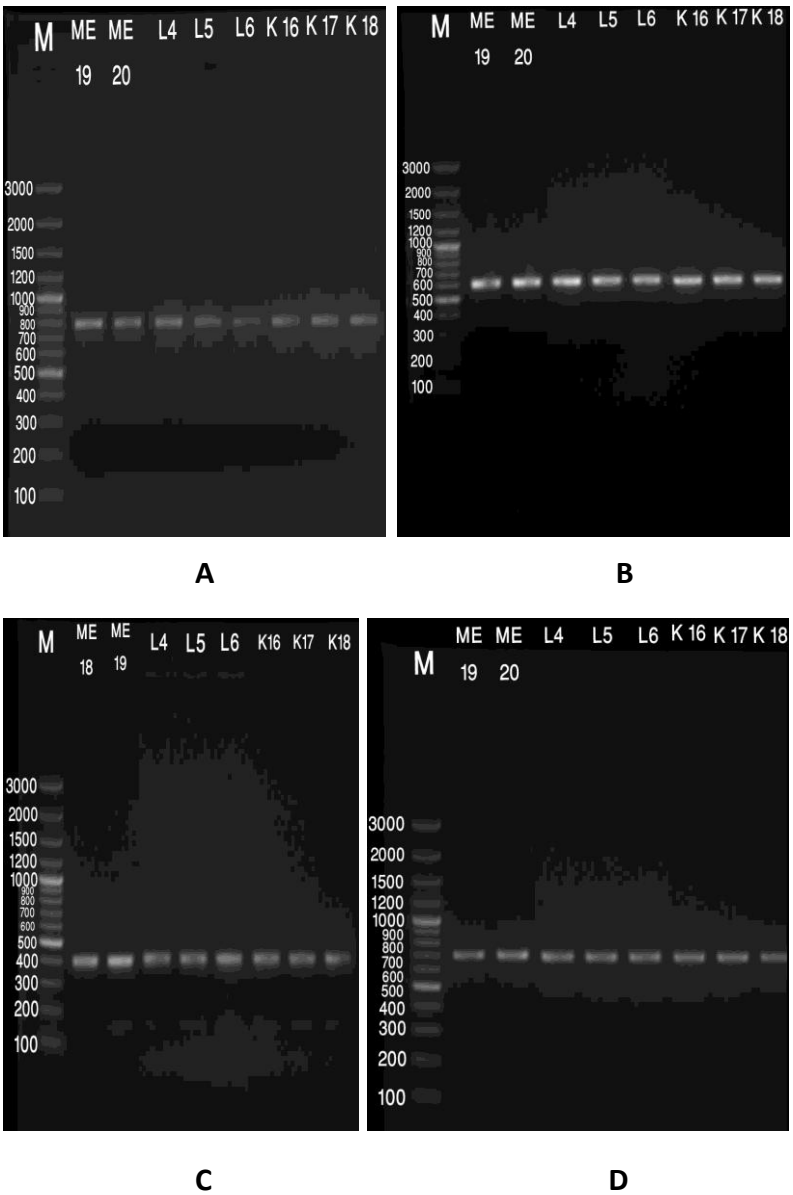


Plate 4.4 DNA bands of amplified barcode regions. (A) represents amplified *matK* sequences; (B) represents amplified *rbcL* sequences; (C) represents amplified *trnH-psbA*

sequences and (D) represents amplified ITS sequences. M represents 100 bp+ molecular weight markers.

4.7.2 Sequence Data Analysis

Sequence data from Macrogen Inc., South Korea was retrieved from Macrogen website and analyzed using CodonCode Aligner v. 4.0.4. Consensus sequences were generated by using CodonCode Aligner software after end trimming and assembling of the reverse and forward sequence of respective accessions of *N. scrophulariiflora.* Reference sequences retrieved from National Center for Biotechnology Information (NCBI) nucleotide database were used for editing. ITS sequence of *N. scrophulariiflora* [(GI: 498542122, (Li and Sun, 2012)] was used as reference sequence to edit ITS sequences while for *trnH-psbA*, sequences of *P. kurrooa* [GI: 242199403, (Albach and Meudt, 2010)] and *N. scrophulariiflora* [GI: 498542275, (Li and Sun, 2012)] were used. In case of *matK*, *matK* sequence of *Veronica spicata* [GI: 379038358, (de Vere *et al.*, 2011)] was used while for *rbcL*, *rbcL* sequence of *Veronica alpina* [GI: 557836325, (Ndiribe *et al.*, 2013)] was used. ITS sequence of accession K18 and *trnH-psbA* sequence of L5 lacked contig assembly and thus were excluded from phylogenetic study. The total lengths of sequences of various barcode regions after editing are presented in Table 4.16.

Table 4.16 DNA sequence lengths (bp) of four DNA regions of *N. scrophulariiflora* accessions from different populations after editing with CodonCode Aligner software.

Regions / Loci	MCA	LNP	KCA
matK	759	759	759
rbcL	555	555	555
trnH-psbA	255	255	255
ITS	597	597	597

4.7.3 BLAST Results

Basic Local Alignment Search Tool (BLAST) (Altschul *et al.*, 1990) was performed with sequences of assembled contigs of all the barcode loci for all accessions to obtain sequences producing significant alignment.

The sequences of ITS of all the accessions showed significant homology with database sequences of *N. scrophulariiflora* (accession KC413448.1) from China with total score value of 1057 and 99% identity. Likewise, *trnH-psbA* sequences of all accessions showed

98% sequence identity with database sequence of *N. scrophulariiflora* (accession KC413527.1) from China.

Alignment result for *rbcL* showed sequence identity with *Veronicastrum brunonianum* (NCBI accession JQ933251.1) with 99% sequence identity and total score of 1003. The alignment was significant as depicted by E-value of 0.0. In case of *matK* sequence, the consensus sequences of query accessions showed sequence similarity with *Veronica spicata* (NCBI accession JN894225.1). The total score was 1251 with 96% sequence identity thus producing significant alignment (E-value 0.0). Both *Veronica spicata* and *Veronicastrum brunonianum* belong to tribe Veroniceae to which *N. scrophulariiflora* also belongs (Table 4.17).

Table 4.17 BLAST result of query sequences of barcode loci of all accessions with NCBI nucleotide sequence database.

Barcoding loci	Query Accessions	Identity with	Total score	Query cover	E - value	Identity	NCBI Accession number
ITS	ME19, ME20, L4, L5, L6, K16, K17	*Neopicrorhiza scrophulariiflora*	1057	97%	0.000	99%	KC413448.1
matK	ME 19, ME20, L4, L5, L6, K16, K17, K18	*Veronica spicata*	1251	99%	0.000	96%	JN894225.1
rbcL	ME19, 20, L4, L5, L6, K16, K17, K18	*Veronicastrum brunonianum*	1003	100%	0.000	99%	JQ933251.1
trnH-psbA	ME19, 20, L4, L6, K16, K17, K18	*Neopicrorhiza scrophulariiflora*	494	100%	2*e*-136	98%	KC413527.1

4.7.4 Characteristics of the Barcode Loci

Of the total 8 samples selected for study, the PCR success rate for all the barcode loci were 100% while the sequencing efficiency was 100% for *matK* and *rbcL* while 87.5% for ITS and *trnH-psbA*. The aligned sequence lengths for four barcode regions were 759, 555, 255 and 597 bps for *matK, rbcL, trnH-psbA* and ITS respectively. No any variable sites were observed in *matk, rbcL* and ITS sequences however four variable sites were observed in *trnH-psbA* sequences (Table 4.18).

Table 4.18 Evaluation of four DNA markers

Parameters	***matK***	***rbcL***	***trnH-psbA***	**ITS**
Percent success of PCR	100%	100%	100%	100%
Percent success of sequencing	100%	100%	87.5%	87.5%
Aligned length (bp)	759	555	255	597
No. of variable sites	0	0	4	0
No. of parsimoniously informative sites	0	0	4	0
No. of samples	8	8	7	7

Table 4.19 Variable sites in *trnH-psbA* region between accessions of *N. scrophulariiflora*. Numbers above the sites are their positions in multiple sequence alignments.

Position / **Samples**	**94***	**107***	**130***	**227***
ME19	G	A	G	T
ME20	G	A	G	T
L4	T	C	T	G
L6	T	C	T	G
K16	T	C	T	G
K17	T	C	T	G
K18	T	C	T	G

Note: * indicates variable sites

Further analysis was performed regarding sequence differences between Nepalese *N. scrophulariiflora* and database sequences of *N. scrophulariiflora* and *P. kurrooa* using MEGA v. 5.2 as presented in Table 4.21.

Table 4.20 Sequence differences among accessions of *trnH-psbA* and ITS sequences of *N. scrophulariiflora* and NCBI database sequences of *N. scrophulariiflora* and *P. kurroa*.

Barcode loci	Database sequence (NCBI accession)	Conserved sites	Variable sites
trnH-psbA	*N. scrophulariiflora* (KC413527.1)	134	151
	P. kurrooa (FJ848090.1)	136	149
ITS	*N. scrophulariiflora* (KC413448.1)	595	2
	P. kurrooa (AF509813.1)	585	12

4.8 Phylogenetic Analysis

4.8.1 Selection of Best-fit Model of Nucleotide Substitution

jModelTest v. 2.1.4 (Darriba *et al.*, 2012) was used to determine the best fit model for nucleotide substitutions for the different accessions used in the study using Corrected Akaike's Information Criteria (AIC_c) (Table 4.22).

Table 4.21 Best fit model of nucleotide substitution as determined using jModelTest.

Locus	***matK***	***rbcL***	***trnH-psbA***	**ITS**
Best Fit Model	TVM+I	TPM2uf+G	TPM1uf	GTR+I
AIC_c value	3322.2697	3952.9944	1909.7908	2892.2900

Where, TVM = Transversion Model; GTR = General Time Reversal; TPM1uf and TPM2uf = Kimura 3-parameter (K81) with unequal base frequencies (uf), +I = invariable sites; +G = Gamma Distribution.

Lower AIC values signify the best model among all others. jModelTest based calculation showed the lowest AICc score to be 3322.27 for TVM+I Model for *matK* sequence. Similarly, TPM2uf+G model was found to be the best nucleotide substitution model for

rbcL sequence with lowest score of 3952.99. For *trnH-psbA* sequence, TPM1uf model showed lowest score of 1909.79 and hence was found to be the best model. For ITS sequence, GTR+I model with score value of 2892.29 was found to be the best model of nucleotide substitution.

4.8.2 Phylogenetic Tree Construction

Phylogenetic trees were constructed to depict the evolutionary relationship between different accessions of *N. scrophulariiflora*. Using jModelTest (v. 2.1.4), a model based phylogenetic tree was constructed and using MEGA (v. 5.2) a maximum parsimony (MP) tree was constructed for different barcode loci. Using best fit nucleotide substitution model from jModelTest (v. 2.1.4), strict consensus trees were generated with confidence level of 95%.

Phylogenetic analyses of *matK, rbcL, trnH-psbA* and ITS were carried out using reference sequences retrieved from NCBI nucleotide database. *matK* sequences of different accessions were compared with *matK* sequences of *Veronica scutella* (NCBI gb JN894218.1), *V. spicata* (gb|JN894225.1|), *V. beccabunga* (gb|JN894414.1|) and *V. anagallis-aquatica* (gb|JN894700.1|). Likewise, *rbcL* sequences were compared with that of sequences of *Veronica officinalis* (gb|HQ590322.1|), *V. peregrina* (gb|HQ590323.1|), *V. serpyllifolia* (gb|HQ590.324.1|), and *Veronicastrum brunonianum* (gb|JQ933251.1|). *trnH-psbA* sequences were compared with those of *Picrorhiza kurrooa* (gb|FJ848090.1|), *Wulfenopsis amherstiana* (gb|FJ848089.1|), *Neopicrorhiza scrophulariiflora* (gb|KC413527.1|), and *Paederota lutea* (gb|FJ848091.1|). Similarly, ITS sequences of accessions were compared with that of *Picrorhiza kurrooa* (gb|AF509813.1|), *Neopicrorhiza scrophulariiflora* (gb|KC413448.1|), *Veronicastrum stenostachyum* (gb|FJ980431.1|) and *Wulfenopsis amherstiana* (gb|FJ848064.1|).

4.8.2.1 Phylogenetic Tree for *matK* Sequences

The strict consensus tree for *matK* sequences based on TVM+I model showed *N. scrophulariiflora* accessions in a single clade separated from outgroup sequences at branch length of 13 demonstrating that 13 nucleotide changes were necessary to separate the accessions of *N .scrophulariiflora* from outgroup sequences (Figure 4.8).

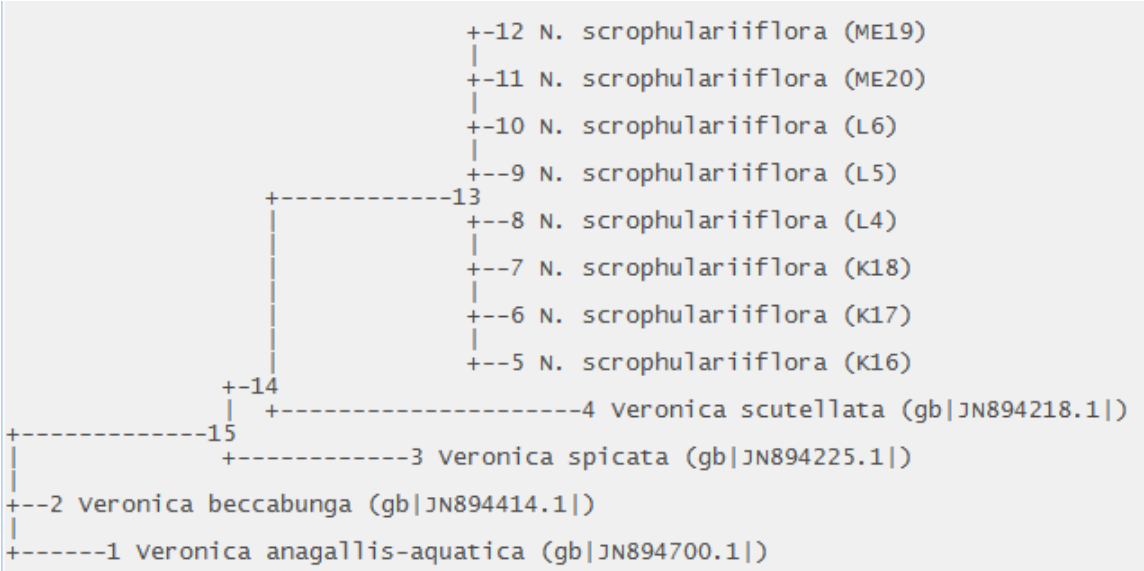


Figure 4.8 A strict consensus phylogenetic tree for ***matK* sequences** using TVM+I model under corrected Akaike's Information Criteria (AICc). Branch lengths are the expected number of substitutions per site.

Similarly, using MEGA v.5.2 software, a phylogenetic tree based on Maximum Parsimony with 1000 bootstrap replicates was also constructed. *matK* data set generated maximum parsimonious trees with CI value 0.933. The tree was less resolved because branches corresponding to partitions reproduced in less than 50% trees were collapsed (Figure 4.9).

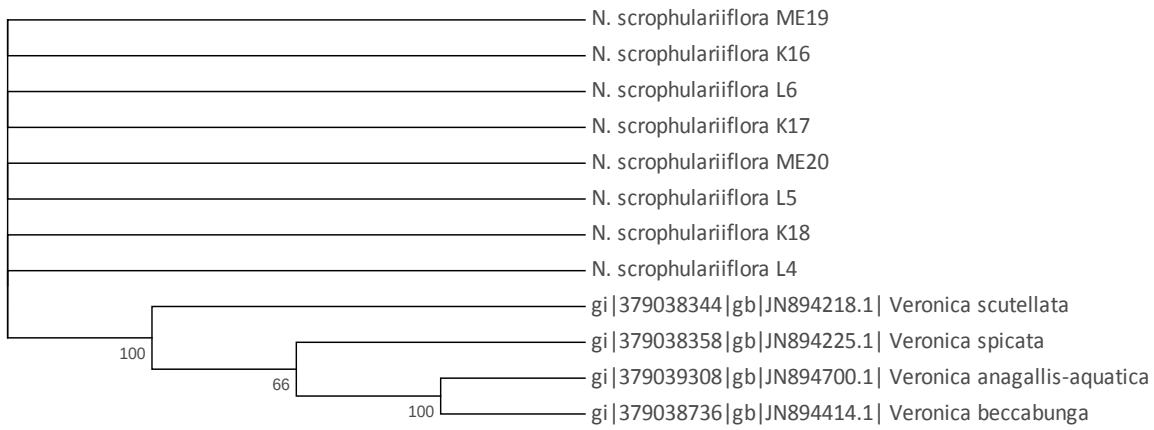


Figure 4.9 A maximum parsimony phylogenetic tree generated by MEGA for ***matK* sequence** to depict evolutionary relationship between *N. scrophulariiflora* accessions and outgroup sequences retrieved from NCBI nucleotide database with GeneBank accession numbers within vertical bars. Numbers at the branches indicate the frequency of occurrence of that branch in all trees (bootstrap support).

4.8.2.2 Phylogenetic Tree for *rbcL* Sequences

Out of four outgroup sequences used in the analysis, two (*viz*. *Veronica officinalis* and *Veronicastrum brunonianum*) formed clade together with the ingroup sequences of Nepalese samples whereas, two other outgroup sequences (*Veronica serpyllifolia* and *Veronica peregrine*) separated from the main clade. This showed that *rbcL* couldn't properly resolve accession sequences from that of outgroup sequences.

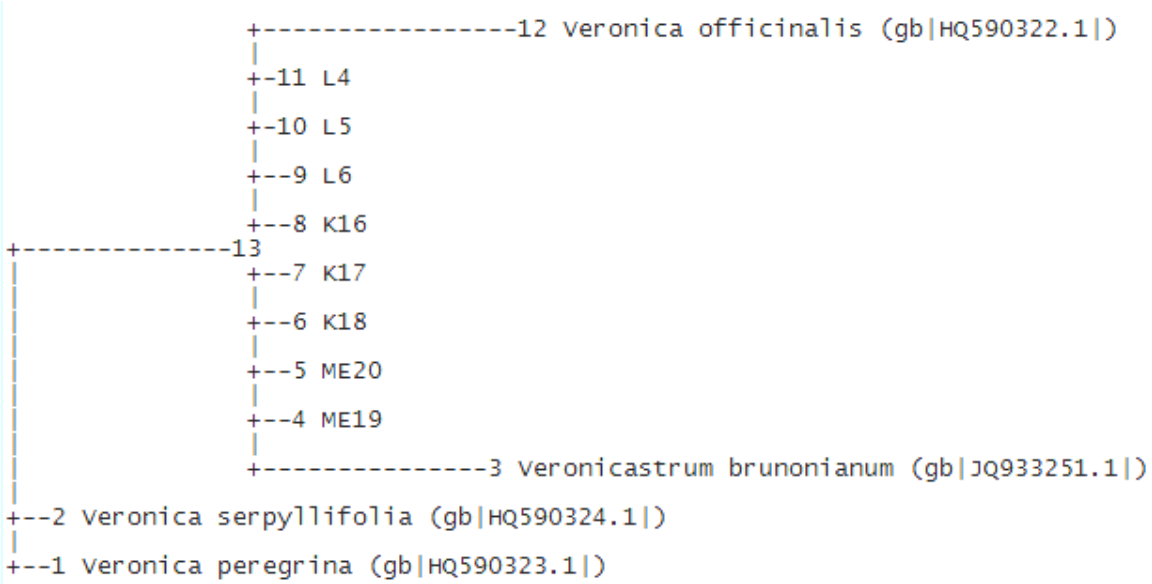


Figure 4.10 A strict consensus phylogenetic tree for ***rbcL* sequences** using TPM2uf+G model under corrected Akaike's Information Criteria (AICc). Branch lengths at the nodes represent the expected number of substitutions per site.

Similarly, maximum parsimony trees were generated using *rbcL* sequences with 1000 bootstrap replicates in MEGA (Figure 4.11). All accessions and two outgroup sequences (*Veronicastrum brunonianum* and *Veronica officinalis*) were placed in a single clade similar to that of strict consensus tree generated using jModelTest thus indicating the validity of result obtained. The consistency index value for *rbcL* based MP tree was 0. The tree was less resolved because branches corresponding to partitions reproduced in less than 50% trees were collapsed (Figure 4.11).

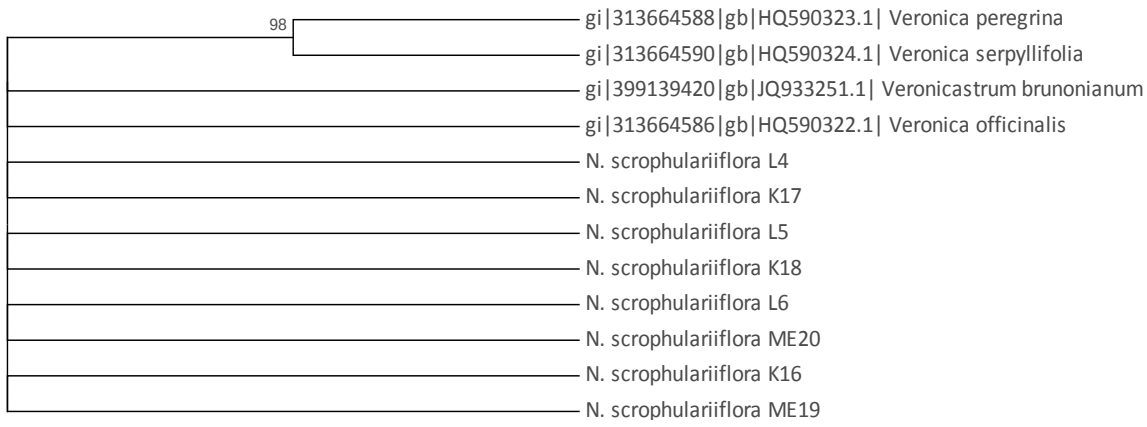


Figure 4.11 A maximum parsimony phylogenetic tree for ***rbcL* sequence** to depict evolutionary relationship between *N. scrophulariiflora* accessions and reference sequences retrieved from NCBI nucleotide database with GeneBank accession numbers within parentheses. Numbers at the branches indicate the frequency of occurrence of that branch in all trees (bootstrap support).

4.8.2.3 Phylogenetic Tree for *trnH-psbA* Sequences

In case of phylogenetic analysis using *trnH-psbA* sequences of *N. scrophulariiflora* accessions, a strict consensus tree using TPM1uf model showed accessions into a single clade at branch length of 13 (Figure 4.12). Accessions ME19 and ME20 formed a separate clade with branch length of 12. *P. kurrooa* and *N. scrophulariiflora* outgroup sequences were separated from accession sequences at branch length of 14 and 15 respectively.

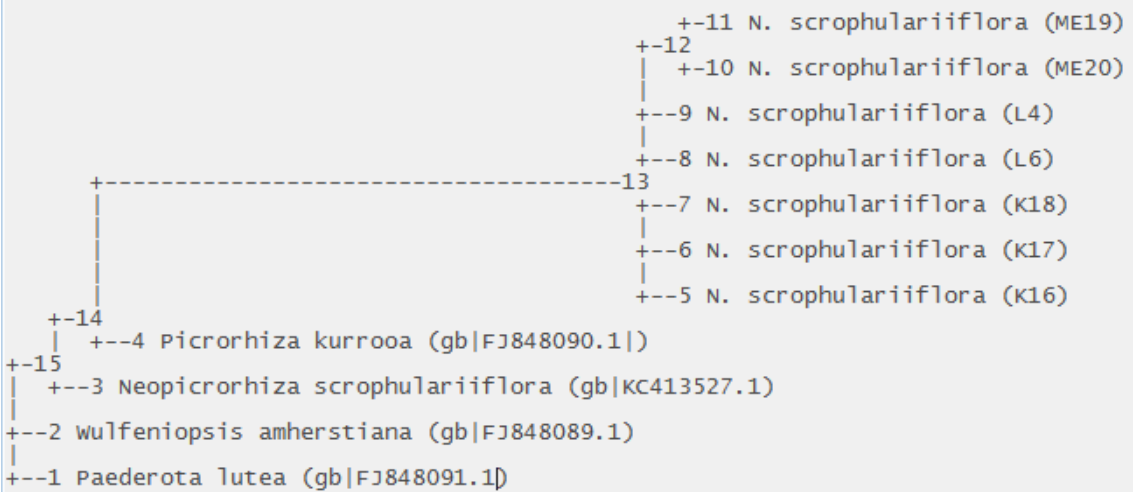


Figure 4.12 A strict consensus phylogenetic tree for ***trnH-psbA* sequences** using TPM1uf model under corrected Akaike's Information Criteria (AICc). Branch lengths at the nodes are the expected number of substitutions per site.

Using MEGA, maximum parsimony tree was constructed using *trnH-psbA* sequences. The MP tree had CI value of 0. Accessions ME19 and ME20 were resolved into a separate clade with 92% bootstrap value. While rest of accessions was resolved into a single clade due to branches corresponding to partitions reproduced in less than 50% trees and hence were collapsed (Figure 4.13).

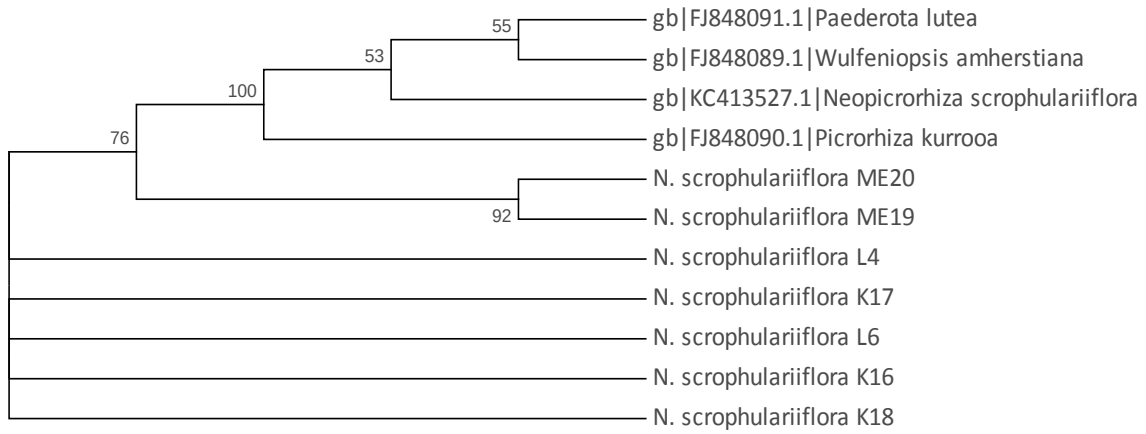


Figure 4.13 A maximum parsimony phylogenetic tree for ***trnH-psbA* sequence** to depict evolutionary relationship between *N. scrophulariiflora* accessions and reference sequences retrieved from NCBI nucleotide database with GeneBank accession numbers within parentheses. Numbers at the branches indicate the frequency of occurrence of that branch in all trees (bootstrap support).

4.8.2.4 Phylogenetic Tree for ITS Sequences

In case of ITS sequence, the strict consensus tree constructed using GTR+I model showed that the accession sequences along with outgroup sequence of *N. scrophulariiflora* were placed in a single clade with a branch length of 12 (Figure 4.14). ITS sequences were seen to resolve *N. scrophulariiflora* species into a single clade and even distinguish a closely related *P. kurrooa* ITS sequence. *Picrorhiza kurrooa* outgroup sequence separated from accession sequences at a branch length of 13.

Figure 4.14 A strict consensus phylogenetic tree using GTR+I model under corrected Akaike's Information Criteria (AICc). Branch lengths at the nodes are the expected number of substitutions per site.

The same result was observed with Maximum Parsimonious Tree using ITS sequences with 100 bootstrap replicates (Figure 4.15). The CI value for MP tree was 0. *Neopicrorhiza scrophulariiflora* accessions along with outgroup sequence of *N. scrophulariiflora* were placed on a single clade with 99% bootstrap value while *P. kurrooa* outgroup sequence was placed in a single clade with 100% bootstrap value (Figure 4.15).

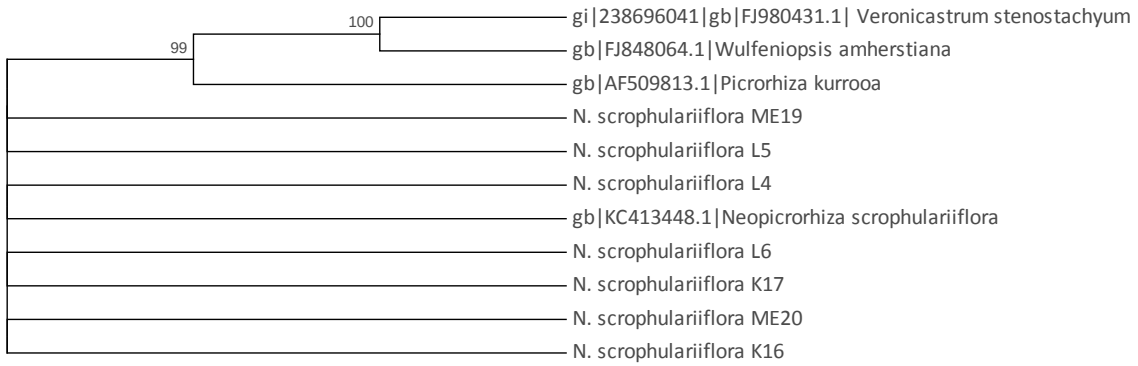


Figure 4.15 A maximum parsimony phylogenetic tree for **ITS sequence** to depict evolutionary relationship between *N. scrophulariiflora* accessions and reference sequences retrieved from NCBI nucleotide database with GeneBank accession within parentheses. Numbers at the branches indicate the frequency of occurrence of that branch in all trees (bootstrap support).

4.9 Intraspecific and Interspecific (Overall mean distance) between *N. scrophulariiflora* and Outgroup sequences

Intraspecific and Interspecific variation between different accessions of *N. scrophulariiflora* and outgroup sequences (mentioned in section 4.8.2) were calculated using MEGA v. 5.2 (Tamura *et al.*, 2011) as presented in Table 4.22.

Table 4.22 Intraspecific and Interspecific (overall mean distance) distances between accessions of *N. scrophulariiflora* compared with outgroups.

Loci	Intraspecific distances	Overall Mean distance (Interspecific)
matK	0.00	0.036
rbcL	0.00	**0.005**
trnH-psbA	0.007	**0.285**
ITS	0.00	0.036

4.10 TaxonDNA Results

TaxonDNA v. 1.7.8 (Meier *et al.*, 2006) was used to calculate the interspecific and intraspecific divergences and Best Match/Best Close Match values in order to identify a potential barcode for *N. scrophulariiflora* species.

Analysis of *matK, rbcL, trnH-psbA* and ITS sequences of accessions and four outgroup sequences (a total of five species) showed that species identification rate was high for *matK* (75%) and ITS (72.72%) while for *trnH-psbA* and *rbcL* it was 0%.

Barcoding gap was graphed based on distribution of variation of K2P distances of inter- and intraspecific differences for each barcode loci. A slight overlap between and inter- and intraspecific distances was observed in *rbcL* while in case of *matK* and ITS sequence only intraspecific K2P distribution was observed. *trnH-psbA* sequence lacked both inter- and intraspecific K2P variation (Figure 4.16, 4.17, 4.18, and 4.19).

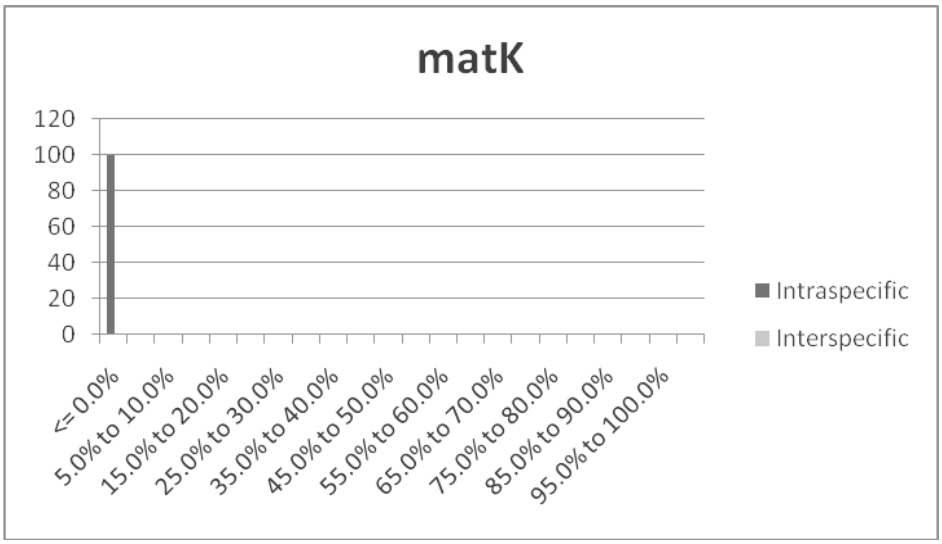


Figure 4.16 Distribution of K2P distances between *N. scrophulariiflora* accessions and outgroup *matK* sequences. A 100% intraspecific K2P distribution between accession sequences is observed.

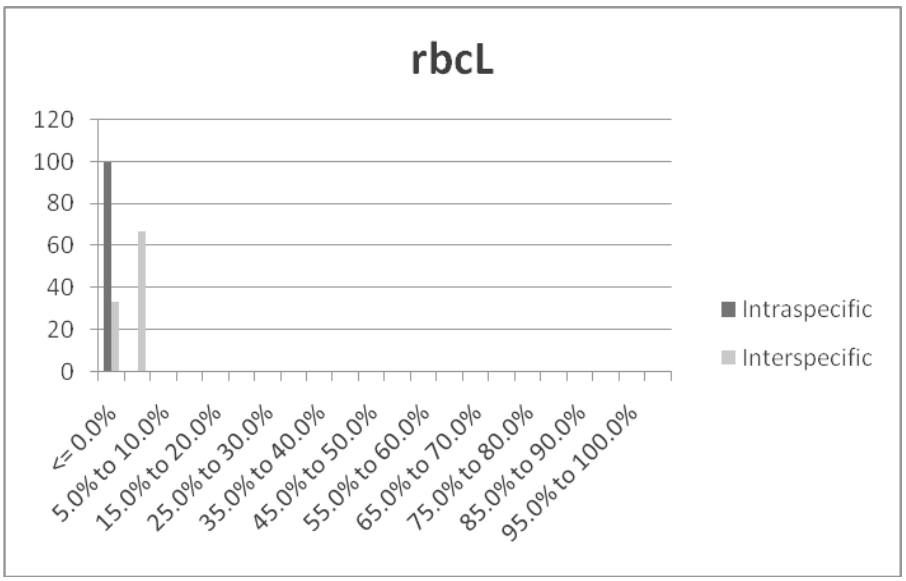


Figure 4.17 Distribution of K2P distances between *N. scrophulariiflora* accessions and outgroup *rbcL* sequences. A 100% intraspecific K2P distribution between accession sequences is observed while an overlapping 33% and non overlapping 66% K2P distribution between accessions sequences and outgroup sequences is observed.

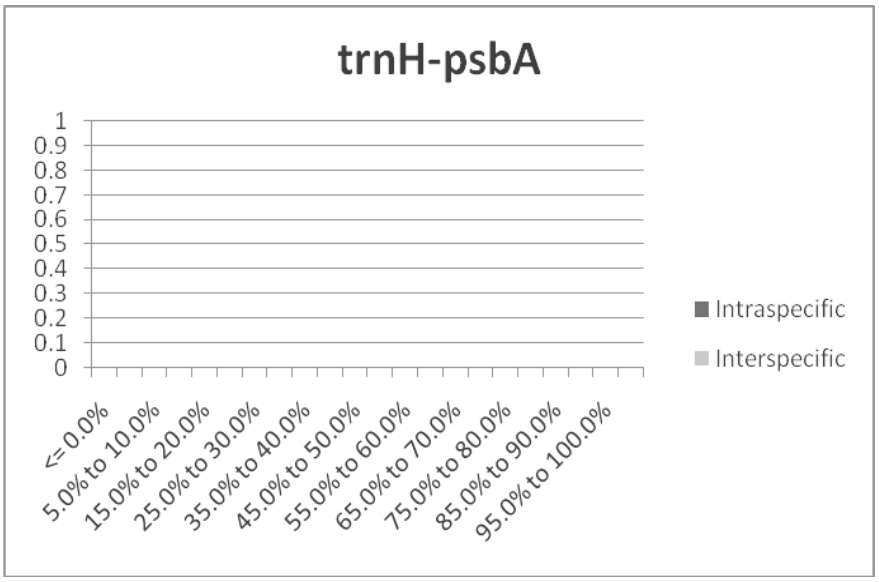


Figure 4.18 Distribution of K2P distances between *N. scrophulariiflora* accessions and outgroup *trnH-psbA* sequences. A 0% intraspecific and interspecific K2P distribution between accession sequences and outgroup sequences is observed.

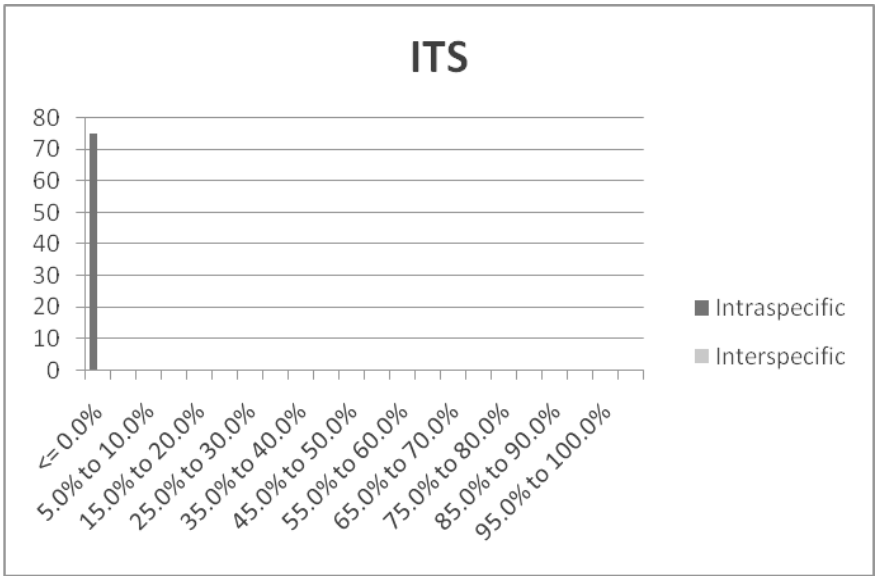


Figure 4.19 Distribution of K2P distances between *N. scrophulariiflora* accessions and outgroup ITS sequences. A 75% intraspecific K2P distribution between accession sequences is observed.

4.11 Biological and Chemical Characterization of *N. scrophulariiflora* Rhizome Extract

4.11.1 Yield of Methanolic Extract of Rhizome of *N. scrophulariiflora*

Methanolic extract of rhizome of *N. scrophulariiflora* was used for the antibacterial, antioxidant assay as well as for HPLC analysis. The yield of methanolic extracts of rhizomes from MCA, LNP and KCA are listed in Table 4.23. A maximum of 39.28% yield was obtained in case of MCA rhizome, whereas 33.99% was obtained for LNP and 31.47% for KCA.

Table 4.23 Percentage yield of *N. scrophulariiflora* rhizome extract in methanol from three different geographical regions.

S.N.	Location	Wt. of powder (gm)	Wt. of Extract (gm)	% yield (in methanol)
1.	MCA, Gorkha	61.91	24.317	39.28
2.	LNP, Rasuwa	61.90	21.039	33.99
3.	KCA, Taplejung	66.34	20.878	31.47

4.11.2 Antibacterial/Antioxidant/HPLC Analysis

4.11.2.1 Antibacterial Activity of Rhizome Extract

The antibacterial activity of the rhizome extracts from three different geographical locations was assayed by agar well diffusion method against seven different microorganisms. The antibacterial activity is expressed in terms of diameter of zone of inhibition in millimeter. The result of methanol extract in different test microorganisms is summarized in Table 4.24.

Table 4.24 Antibacterial activity of extracts of *N. scrophulariiflora* from different locations by agar well diffusion assay

	Zone of Inhibition (mm)					
Organism	**Chloramphenicol (+ve control)**	**Methanol (-ve control)**	**MCA**	**LNP**	**KCA**	**Gram Reaction**
Staphylococcus aureus	40	-	10	15	10	+ve

Klebsiella pneumoneae	42	-	11	13	8	-ve
Pseudomonas aeruginosa	35	-	12	15	9	-ve
Enterobacter	25	-	-	-	-	-ve
Serratia	22	-	-	-	-	-ve
Escherichia coli	23	-	-	-	-	-ve
Salmonella typhimurium	22	-	-	-	-	-ve
Bacillus subtilis	23	-	-	-	-	+ve

All the test microorganism (as enlisted in table above) were sensitive to Chloramphenicol, used as a positive control and resistance to methanol used as a negative control.

The methanolic extract of plant collected from different locations showed varying degree of antibacterial sensitivity. Of the three extracts from different locations, the extract from LNP was found to have highest antibacterial activity with inhibition zone of 15 mm for *S. aureus*, 13 mm for *K. pneumoneae* and 15 mm for *P. aeruginosa*. Similarly the extract for MCA showed inhibition zone of 10 mm for *S. aureus,* 11 mm for *K. pneumoneae* and 12 mm for *P. aeruginosa*. KCA extract showed inhibition zone of 10 mm for *S. aureus,* 8 mm for *K. pneumoneae* and 9 mm for *P. aeruginosa.* Out of eight tested organisms, five microorganisms were found to be resistant to the methanolic extract of *N. scrophulariiflora*. This result showed that the antibacterial activity of *N. scrophulariiflora* rhizome extracts from different locations were of moderate type.

4.11.2.2 Estimation of Antioxidant Activity

The antioxidant property of methanolic extracts of *N. scrophulariiflora* collected from different locations was determined by their ability to scavenge stable 1,1-diphenyl-2-picrylhydrazyl (DPPH) radicals. The effect of increasing concentration of different extracts on inhibition of DPPH solution gave information on the antioxidant property of these extracts (Table 4.25).

Table 4.25 Free radical scavenging activity of methanolic extract of *N. scrophulariiflora* from different locations using DPPH inhibition assay. Values are expressed as Mean±SEM of 3 replicates.

Extract Concentration (mg/mL)	Inhibition (%)		
	MCA	LNP	KCA
0.1	8.058±0.3652	18.65±0.9210	12.62±2.969
0.5	6.793±0.1112	18.37±0.4128	16.01±0.03176
1.0	9.873±0.07939	26.90±1.524	21.81±0.7145
5.0	55.78±2.477	93.78±0.1588	74.39±3.795
10.0	92.90±1.080	95.93±0.03176	95.46±0.3017
20.0	95.68±0.04764	95.82±0.06352	96.62±0.04764
30.0	95.63±0.01588	95.74±0.01588	96.42±0.03175
40.0	95.68±0.04764	95.52±0.1429	96.45±0.07939
50.0	95.54±0.03176	95.41±0.1429	96.48±0.03176

The result showed stoichometric reduction of DPPH radicals with increasing extract concentrations which quenches more DPPH radicals in exponential manner (figure 4.20).

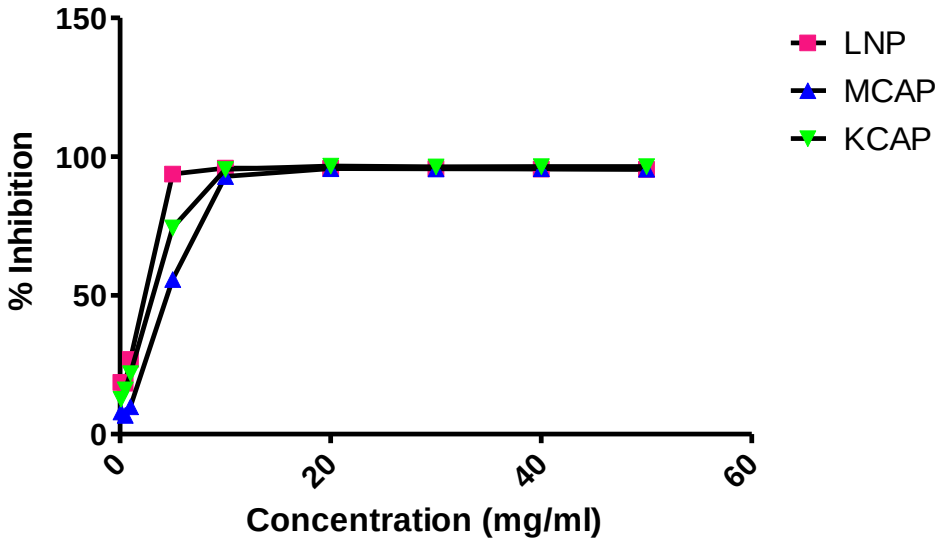


Figure 4.20 Graph representing the reduction of DPPH radicals (expressed as % Inhibition) by increasing concentrations of extracts (in mg/mL) obtained from different locations.

The antioxidant activity was expressed as IC_{50}, which is the concentration of antioxidant needed to reduce 50% DPPH radicals (Figure 4.21), i.e. it is a measure of inhibitory concentration.

The IC_{50} values of the extracts and the standard gallic acid was calculated in order to determine the antioxidant activity of rhizome extracts of different locations. Lower the IC_{50} value, higher will be the antioxidant activity of the sample. The IC_{50} values for methanolic extracts of MCA, LNP, KCA and gallic acid was found to be 4.141, 1.742, 2.398 and 0.094 respectively. Therefore, LNP extract was found to possess highest antioxidant activity followed by KCA and MCA extracts (Table 4.26, Figure 4.21).

Table 4.26 IC_{50} values of standard and methanolic *N. scrophulariiflora* extracts of different locations using DPPH assay. Values are presented as Mean±SEM.

S.N.	Sample	IC_{50} value (mg)
1.	Gallic Acid (Standard)	0.094±0.0051
2.	MCA extract	4.141±0.3201
3.	LNP extract	1.742±0.0396
4.	KCA extract	2.398±0.1290

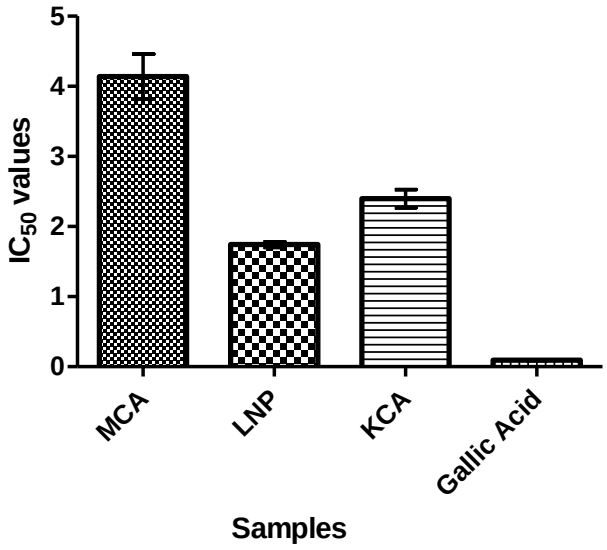


Figure 4.21 IC_{50} values for standard and different extracts. Bars represent the Mean±SEM values.

One way ANOVA was also performed to test the significance of test values between MCA, LNP and KCA. The data was found to be statistically significant (P=0.004) as presented in Table 4.27.

Table 4.27 One-way ANOVA table for IC_{50} values of KCA, LNP and MCA extracts

ANOVA Table	Sum of Square (SS)	Degree of freedom (df)	Mean sum of squares (MS)	R^2	P-value (<0.05)
Treatment (between values)	9.219	2	4.610	0.9272	0.004 (***)
Residual (within values)	0.7240	6	0.1207		
Total	9.943	8			

*** indicate highly significant result

4.11.2.3 Quantification of Picroside I and Picroside II

The identification and quantification of Picroside I and II in the rhizome extract of *N. scrophulariiflora* collected from three different geographical locations were performed with reference to standard Picroside I and II. The rhizome extracts were analyzed using Shimadzu HPLC system in order to determine their respective retention times and to plot a calibration curve. The calibration curve was plotted on the basis of different concentrations of both compounds (1, 10, 100, 500 and 1000 µg/mL) (Appendices X and XI). In order to carry out the experiments a solvent system of methanol: water (2:3) with elution rate of 0.7 mL/min was used. The standard chromatogram showed that Picroside II eluted first at 4.531 minutes and Picroside I eluted at 7.001 minutes (Figure 4.22 and 4.23).

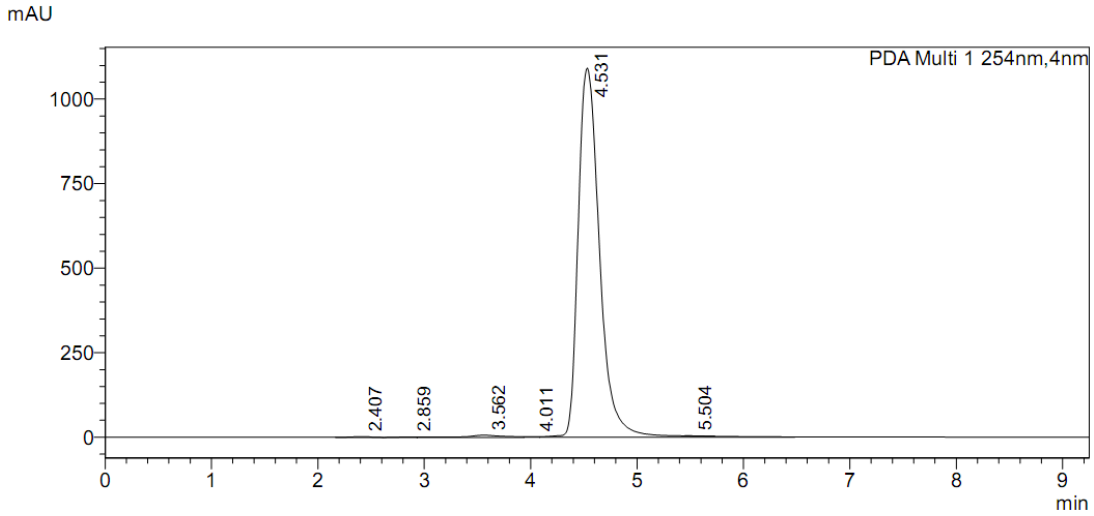


Figure 4.22 Standard chromatogram of Picroside II using HPLC with elution time marked at 4.531 minutes.

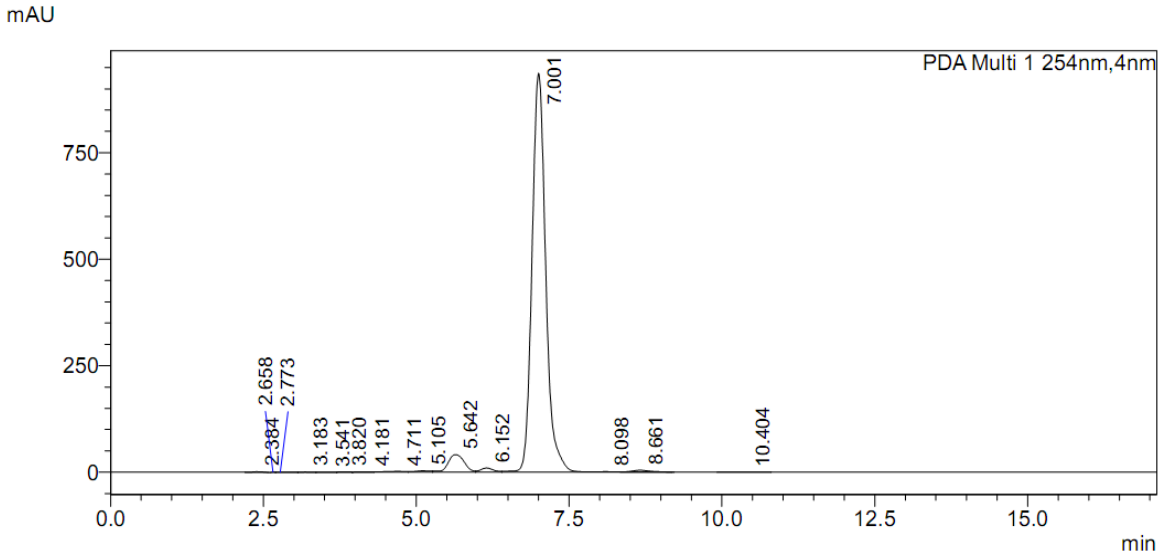


Figure 4.23 Standard chromatogram of Picroside I using HPLC with elution time marked at 7.001 minutes.

A compound regression equation with correlation coefficient values of 0.997 and 0.998 for Picrosides I and II determined the statistical significance of the calibration curve (Table 4.28).

Table 4.28 Summary of HPLC method for compounds Picrosides I and II.

S.N.	Compound	Compound Regression Equation	Correlation coefficient (R^2)
1.	Picroside I	f(x) = 55922.9x – 238516	0.99777
2.	Picroside II	f(x) = 37288.4x – 303615	0.99899

The HPLC chromatogram of extracts confirmed the presence of picrosides I and II in all the extract solutions by comparing them with the retention time of standard compounds (Figure 4.24).

The elution time of Picroside II was found to be 4.475, 4.495 and 4.470 minutes for MCA, LNP and KCA respectively while the elution time for Picroside I was found to be 6.983, 7.012 and 7.047 respectively (Figure 4.24).

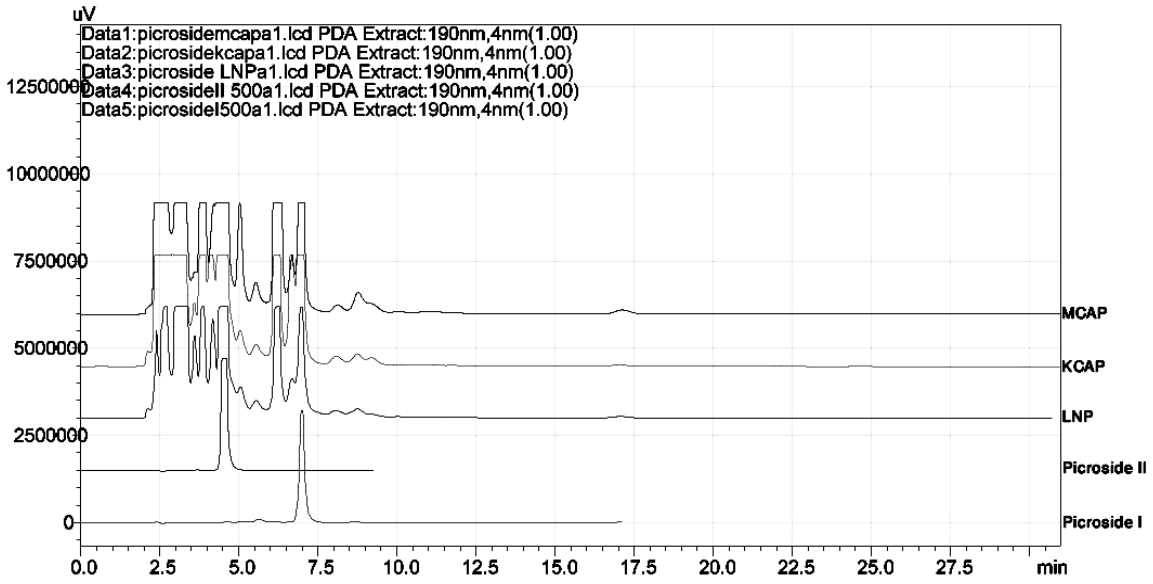


Figure 4.24 HPLC chromatogram of Picroside I and Picroside II content of extracts from different locations presented as overlapping of the peaks of standard compounds with that of peaks obtained from different extracts.

Based on the peak height, area and calibration curve of standard compounds, the concentration of Picrosides I and II were calculated. The highest amount of Picroside I and II was found in MCA extract followed by KCA and LNP extracts (Table 4.29).

Table 4.29 Concentrations of Picrosides I and II in different samples using HPLC system.

S.N.	Compound	Location	Concentration (µg/L)
1.	Picroside I	MCA, Gorkha	678.705
2.	Picroside I	LNP, Rasuwa	203.623
3.	Picroside I	KCA, Taplejung	426.501
4.	Picroside II	MCA, Gorkha	2015.947
5.	Picroside II	LNP, Rasuwa	751.637
6.	Picroside II	KCA, Taplejung	1029.281

CHAPTER 5 – DISCUSSIONS

Discussions

The growing demand of high value Nepalese MAPs has ultimately put pressure on the survival of many high value medicinal plants including *Neopicrohiza scrophulariiflora*. It is enlisted as a "Vulnerable (V)" species in IUCN category (IUCN, 2013). Conservation assessment and management planning (CAMP) workshop (2002) designated *N. scrophulariiflora* as a "Vulnerable (VU)" species as well. The vulnerability of this species is attributed to overexploitation by harvesting from wild populations and unscientific harvesting by collectors for trade (Ghimire *et al.*, 2005).

Apart from market demand, plants in natural population have to cope up with selection pressure. Plants must adapt themselves to changing climate, nutrients and environmental factors. Continuous evolution aids in acclimatization to different environmental conditions and thus the genetic structure of plants change over a period of time. Hence, assessing genetic structure of different species found on different locations not only provide information on plant's adaptive potential but also can aid in devising proper conservation strategy for genetically vulnerable plants. Apart from this, strategy for sustainable use of endangered and vulnerable plants can be devised which can greatly aid in socio-economic condition of a country.

The present study was conducted to study the genetic diversity and population structure of *Neopicrorhiza scrophulariiflora* samples from three different populations using the genetic structure information obtained using the PCR-based molecular marker technique called Inter-simple sequence repeat (ISSR). Also DNA sequence based barcode markers were generated for accessions coming from different populations and their phylogenetic analysis was performed. Furthermore, the antimicrobial activity and antioxidant property were assessed. In addition, out of several active constituents the plant possesses, two mostly noted active constituents, Picrosides I and II, were quantified using HPLC method. For the current study, the samples were collected from three protected areas of Nepal representing three development regions, namely, Manaslu Conservation Area (MCA), Langtang National Park (LNP) and Kanchenjunga Conservation Area (KCA).

5.1 ISSR profiling for the assessment of genetic diversity and population genetic structure of *N. scrophulariiflora*

ISSR profiling was performed by amplifying DNA template of various accessions collected from three different locations by 10 selected primers. Data matrix for genetic diversity study was created by scoring presence or absence of a band of certain size based on visualization of such bands on a gel electrophoresis. ISSR based PCR amplification has widely been used to study genetic diversity and structure of various medicinal plants and

furnishes valuable information for biological conservation. In a comparative study performed on two *Rheum* spp. to detect the genetic variation, ISSR markers demonstrated the presence of 98.86% polymorphism between them (Wang *et al.*, 2012a). Similar study on *Rheum officinale* demonstrated 95.24% polymorphism (Wang *et al.*, 2012b). Studies on *Curcuma longa* (Singh *et al.*, 2012), *Rauvolfia serpentina* (Pillai *et al.*, 2012), *Sinopodophyllum hexandrum* (Xiao *et al.*, 2006) etc. have also reported ISSR as a tool for studying genetic structure and measure of genetic diversity.

5.1.1 ISSR-PCR Optimization and Primer Screening

PCR is highly sensitive molecular technique in which a single DNA molecule can be amplified and/or single-copy genes can be extracted out of complex mixtures of genomic sequences. Since no single protocol will be appropriate to all situations, each new PCR assay requires optimization. Optimization also helps to rule out unwanted outcomes like co-detectable product or low yield, presence of nonspecific background bands due to mispriming, formation of primer-dimers etc (Innis and Gelfand, 1990).

In the present study two parameters were assessed for the generation of better amplification products, *viz.*, cycling and reaction conditions. Different reaction parameters like concentration of DNA template, concentration of magnesium ions, concentration of primers, concentration of deoxynucleotides (dNTPS), and concentration of DNA polymerase were assessed. The optimal DNA concentration for ISSR analysis of *N. scrophulariiflora* was chosen to be 20 ng. Amplification result as visualized in gel electrophoresis (Plate 4.1 A) showed multiple, dense and clearer bands in all range of DNA concentrations tested (Table 4.3) so 20 ng was ultimately chosen. Various studies have used 20 – 25 ng of template DNA (Shaw *et al.*, 2009; Zhang *et al.*, 2010; Liu *et al.*, 2011).

In order to determine the optimal concentration of magnesium chloride, a wide range of magnesium chloride concentrations were tested (Table 4.3) and 2.0 mM of Magnesium chloride concentration was found to be optimum. Determination of optimum concentration of Magnesium is necessary because it may affect primer annealing, strand dissociation temperatures of both template and PCR product, product specificity, formation of primer-dimer, and enzyme activity and fidelity (Innis and Gelfand, 1990). Excess magnesium results in non-specific products while insufficient magnesium results in reduced yield (Grunenwald, 2003).

Similarly, optimal primer concentration was also determined during the study. Of the tested range of primer concentration (0.1 µM to 1.6 µM) (Table 4.3), 0.7 µM was selected for our study. Concentrations of primers influence the amplification specificity and efficiency. Higher primer concentrations can result in nonspecific priming and formation of primer-dimers whereas lower primer concentrations may affect PCR

efficiency (Grunenwald, 2003). ISSR study on *N. scrophulariiflora* by Liu *et al.* (2011) used 0.5 µM primer concentration which is close to our present study.

dNTPs concentrations were tested in the range of 0.1 to 0.5 mM and the optimal concentration was found to be at 0.2mM (Table 4.3). It is suggested that dNTPs concentration should be used between 0.02 to 0.2 mM for optimal balance between yield, specificity and fidelity. Four dNTPs should be used at equivalent concentrations to minimize misincorporation errors. By lowering the dNTP concentrations both the specificity and fidelity of PCR can be increased (Innis and Gelfand, 1990). Low dNTP concentrations minimize mispriming at nontarget sites and reduce the likelihood of extending misincorporated nucleotides (Innis *et al.*, 1988). dNTPs are typically added in excess during PCR setup hence determination of lowest dNTPs concentration is necessary because higher dNTP concentration increases the error rate of DNA polymerase and actually inhibits *Taq* polymerase (Gelfand, 1989; Ohler and Rose, 1992).

Taq polymerase is the most common enzyme used for PCR amplification because of its thermostability and processivity. It has extension rate of 35 to 100 nucleotides per second at 72^{o}C which is the most common extension temperature (Innis *et al.*, 1988). *Taq* polymerase was optimized by experimenting with a range of concentrations starting from 0.5 units to 2 units (Table 4.3). The optimum concentration was found to be 1U (Plate 4.1 D). A recommended range of *Taq* concentration is between 1 and 2.5 units per 100 µL reaction volume (Lawyer *et al.*, 1989; Innis and Gelfand, 1990). If the enzyme concentration is too high, nonspecific background products may accumulate and if too low, an insufficient amount of desired product is yielded (Grunenwald, 2003).

Out of three ISSR cycling conditions tested as described by Liu *et al.* (2011), Tamhankar *et al.* (2009) and Xiao *et al.* (2006), cycling conditions described by Xiao *et al.* (2006) (Table 3.3) resulted in multiple, crispy and better quality bands and hence was chosen as optimum cycling condition for *N. scrophulariiflora* under study (Table 3.3).

Out of 40 UBC primers, 10 primers were selected on the basis of reproducibility, number of scorable bands and band quality and intensity. The primers thus selected were 17-20 mer in length (Table 4.4) and the GC content ranged from 33% to 53%. The optimal primer size is usually 18 to 28 bases long and a GC content of 40 to 60% assures higher melting temperature (T_m) and therefore increases specificity (Grunenwald, 2003). Longer primer length of ISSR compared to RAPD (10-mers) allows for the use of high annealing temperatures (45-60^{o}C) and thus results in higher stringency (Reddy *et al.*, 2002).

5.1.2 Identification of population-specific markers for *N. scrophulariiflora*

After amplification with optimized cycling and reaction condition, 10 primers out of 40 showed some peculiar bands specific to certain population while lacking in other.

Amplified ISSR fragments ranged in size from 100 – 2000 bps. Study on Chinese *N. scrophulariiflora* populations showed size range of 250 – 1800 bps (Liu *et al.*, 2011).

10 population specific markers were observed in the present study generated by 10 UBC primers. Band size of 450 bps (UBC 807) and 1600 bps (UBC 850) were found only in MCA populations. Band size of 1000 bps (UBC 835), 490 and 320 bps (UBC 848), 490 and 390 bps (UBC 850) and 750 bps (UBC 868) were unique to LNP populations. While in case of KCA population, band size of 650 and 400 bps (UBC 868) were peculiar.

Species-specific molecular markers are useful for various purposes. In the present study no such markers were revealed. However, population specific markers obtained from this study can be used to identify population specific samples and this can aid in monitoring illegal trade of this high value medicinal plant. Any collection of samples, for trade, from regions which are prone to illegal and over collection can be identified using these population specific markers. In order to generate species specific markers for use in molecular diagnostics, a more elaborate work, encompassing wide range of samples from varied geographical locations, should be undertaken. Another use of such markers can be as a fingerprint tool for authentication purposes. Population specific markers for *N. scrophulariiflora* can be used as DNA fingerprints for this plant species. ISSR markers have been found useful in authentication of real species being adulterated with multiple species (Harsulkar *et al.*, 2009; Dhanya and Sasikumar, 2010). Furthermore, SCAR markers (co-dominant markers) can be developed from ISSR markers for the rapid and accurate detection of individual plant or groups of plants of particular species (Li and Park, 2012; Riaz *et al.*, 2012).

5.1.3 Genetic Diversity Assessment using Similarity Coefficients and Cluster Analysis

ISSR profiling performed on Nepalese *N. scrophulariiflora* revealed 91.37% polymorphism across all 43 accessions. Out of 197 bands amplified by 10 primers, 180 bands were found to be polymorphic and 17 bands (8.63%) were monomorphic (Table 4.6). Study performed in Chinese populations showed 100% polymorphic bands at species level (Liu *et al.*, 2011). In the study, 82 bands generated during the study were all found to be polymorphic. The Nepalese population showed high genetic variation at the species level but compared to the Chinese study it was less.

Individual primers were also assessed for different parameters such as Polymorphic Information Content (PIC), Band Informativeness (I_B) and Resolving power (R_p). PIC is the probability of detection of polymorphism between randomly drawn genotypes or the probability of any solitary band being polymorphic and is calculated for each primer (Sarwat, 2012). Theoretically, PIC values can range from 0 to 1. A PIC value of 0 indicates that marker has only one allele while a PIC value of 1 indicates that the marker would

have infinite number of alleles. A PIC value of less than 0.25 indicates low polymorphism; a value between 0.25 and 0.5 indicates average polymorphism and a value higher than 0.5 indicates a highly polymorphic locus. Primers with higher PIC value are more useful than others for distinguishing accessions (Teklewood and Becker, 2006). All the primers used for the study showed high PIC value greater than 0.8, suggesting that these primers are useful especially for ISSR based analysis of genetic diversity in *N. scrophulariiflora*. ISSR markers have more polymorphic information content (PIC) than AFLP (Blair *et al.*, 1999).

Resolving power of a primer is its ability to distinguish between genotypes. In other words, it is an index developed to compare the values of different primers in terms of informative bands obtained in a given set of germplasm (Prevost and Wilkinson, 1999). Thus primers with higher R_p values have a greater capacity to discriminate plant genotypes (Prevost and Wilkinson, 1999). The highest R_p value was found to be 13.023 for UBC 834 indicating that among the 10 UBC primers, UBC 834 is best able to distinguish between accessions of *N. scrophulariiflora* while the lowest value was found for UBC 886 (5.953) (Table 4.7).

One of the commonly used methods in study of genetic diversity, within and among populations using a marker, is based on the comparison of individual genotypes' fingerprints within and between populations. In this case, a genetic similarity (or dissimilarity) matrix is constructed from pair-wise comparisons of individuals to characterize population structure. Such approach requires suitable methods for evaluating similarity between individuals (Dalirsefat *et al.*, 2009). Thus, the choice of an appropriate coefficient of similarity is very important and decisive point to evaluate clustering, true genetic similarity between individuals, analyzing diversity within populations and studying relationship between populations.

Coefficients of similarity between two individuals or accessions under study can be estimated based on presence or absence of bands using Simple Matching similarity coefficient (SM) (Sokal and Michener, 1958), Jaccard similarity coefficient (J) (Jaccard, 1908) and Dice similarity coefficient (D) equivalent to Nei-Li's coefficient (Dice, 1945). Comparison of these similarity coefficient matrices with their phenotypic illustrations is essential for selection of competent coefficient with proper goodness of fit thus leading to unbiased assessment. Similarity matrices were computed using SimQual (similarity for qualitative data) module of NTSYS-pc v. 2.21i software. Dendrograms were produced using unweighted pair-group mean arithmetic method (UPGMA). Similarity matrices based on J and D were highly correlated (r=0.993) according to mantel matrix correspondence test together with high correlation of their phenograms as given by Consensus index (r=1.00). The difference between cophenetic correlation (r) values of J, D and SM coefficients were not very different (0.966, 0.951, and 0.955 respectively).

Consensus index provides a relative estimate of similarities between dendrograms (Dalirsefat *et al.*, 2009). According to Rohlf (2009), if cophenetic correlation value (r) is such that $0.8 \leq r \leq 0.9$, the similarity coefficient is interpreted as good fit for determination of genetic variablility and all coefficients showed good fit however, J coefficient showed highest correlation value (Table 4.10) hence was evaluated as best for deducing genetic relationship among *N. scrophulariiflora* accessions using UPGMA module. Cophenetic correlation coefficient is a measure of how faithfully a dendrogram preserves the pairwise distances between the original unmodeled data points (Sokal and Rohlf, 1962; Rohlf and Fisher, 1968).

In the present study, Jaccard's similarity coefficient values ranged from 0.043 to 0.952 revealing high genetic diversity within *N. scrophulariiflora* accessions from three different geographical regions. In a study performed by Joshi and Dhawan (2007) in *Swertia chirayita*, the similarity values of Jaccard's coefficient ranged from 0.68 to 0.97 indicating lower diversity.

Clustering analysis was performed using SAHN (Sequential Agglomerative Hierarchical and Nested clustering) module of NT-SYS pc software in which successive merging was done to group individuals such that individuals with similar characteristics were placed in a single group. UPGMA method was used to graphically represent the relation between all the accessions. Clustering showed that forty-three accessions were separated into three major clusters with Cluster I representing MCA population, Cluster II representing KCA population and Cluster III representing LNP population. A separate cluster was formed by K2 accession (Figure 4.4). K2 is differentiated from rest of KCA samples which may be due to different genetic structure of K2 from rest of KCA accessions.

The mean genetic similarity among three populations of *N. scrophulariiflora* was observed to be 15% (between LNP and KCA), 20% (between MCA and LNP) and 20% (between MCA and KCA). The percent genetic similarity value explains why MCA and KCA accessions are clustered together in a phenogram while LNP is clustered separately considering the fact that LNP is located geographically between MCA and KCA. Genetic similarity between KCA and MCA is unexpected when their geographical distances are considered. However, this can be attributed to transfer of genetic material from MCA to KCA or vice versa by humans sometimes in the past. Other possible reason could be the sampling error.

One of the major applications of molecular genetic diversity analysis is in plant breeding. The genetic distances revealed among accessions under study can be the basis of selection of parents in conventional breeding program. Higher the genetic distance, more ideal will be the accessions to be used as parents in hybridization in order to

develop elite varieties (Prakash *et al.*, 2006). The high genetic diversity in *N. scrophulariiflora* populations may be due to cross pollination between individuals in the populations and natural evolutionary forces such as genetic drift, mutation, natural selection, climate change etc. (Slatkin, 1987). This can also indicate the good survival potential of the species in the existing environment which is brought about by the selection pressure of changing environment. Furthermore, greater genetic diversity can also confer traits like disease resistance, greater productivity, greater nutrient retention and ecosystem stability (Tilman, 2000).

3D Plot was constructed using NTSYS-pc v. 2.21i for all 43 accessions of *N. scrophulariiflora* (Figure 4.5) and analysis was performed with Principal Coordinate Analysis (PCoA). 3D plot represents the respective distribution of various populations with each other. It plots a perspective view of 3-dimensional space showing the objects as "points" on top of "pins" floating above a "baseboard". 3D plot analysis showed result similar to the UPGMA phenogram where KCA and MCA populations were close to each other and LNP population was separated from both these populations at a distance. 3D plot also showed the relative distribution of accessions within their populations. The first two axes of PCoA analysis showed 37.63% and 27.24% of total genetic variation respectively with cumulative variance of 64.87%. Principal Coordinate Analysis (PCoA) was calculated using GenAlex v. 6.5. PCoA analysis is a good indication of the genetic relationships among the species (Sebastian *et al.*, 2010).

5.1.4 Population Genetic Diversity and Structure of *Neopicrorhiza scrophulariiflora*

POPGENE v.1.32 was used to assess variation within and among populations of *N. scrophulariiflora* by computing Nei's (Nei, 1972) original measure of genetic identity (I_{XY}) and genetic distance (D_{XY}) and Nei's (Nei, 1978) unbiased measure of genetic identity (I_{XY}) and distance (D_{XY}). Both Nei's original and unbiased measure gave similar result.

Highest genetic Identity was observed between populations of MCA and KCA (0.9006) and thus least genetic distance (0.1047) (Figure 4.8). Likewise, the lowest genetic identity was observed between MCA and LNP populations (0.7920) and thus the highest genetic distance (0.2232). Similar data was observed based on Nei's unbiased measure. Highest genetic identity was observed in MCA and KCA populations (0.9047) thus there was least genetic distance (0.1002) and the lowest genetic identity was observed between MCA and LNP populations (0.7961) thus they had highest genetic distance (0.2280) (Figure 4.9). Similarity matrix based on Jaccard coefficient showed the similar result. The MCA and KCA populations were shown to be genetically closer and KCA and LNP populations to be distant although LNP lies between MCA and KCA geographically. According to Jaccard matrix (Appendix 3), the mean genetic similarity between MCA and

LNP is 0.196, between MCA and KCA is 0.196 and between KCA and LNP is 0.150 which is congruent with the result obtained from Nei's measure of genetic identity and distance.

5.1.4.1 Genetic Variation within Populations of *Neopicrorhiza scrophulariiflora* and at species level

According to Jaccard similarity matrix, it was observed that accessions L5 and K2 bore the least similarity (0.043) and ME30 and ME31 bore the most similarity (0.952) indicating diverse nature of individuals within the populations. Genetic diversity provides mechanism for individuals and populations to adapt to varying environment. The survival of species depends on the maintenance of genetic variability within and among populations that accommodate new selection pressures brought about by varying environment (Sebastian *et al.*, 2010). The extent and distribution of genetic diversity of a plant species depends upon its breeding system, ecological and geographical factors, environmental factors etc (Ramanatha and Hodgkin, 2002).

Variation analysis of populations at the species level was performed using indices like percentage of polymorphic bands (PPB), Nei's genetic index (H) and Shannon's diversity index (I) using POPGENE v. 1.32. Data indicate that accessions from KCA population are highly diverse with PPB, 96.26%; H, 0.2439 and I, 0.3836 while accessions from LNP population are least diverse with PPB, 66.27%; H, 0.2284 and I, 0.3413 (Table 4.15). The result obtained from POPGENE v. 1.32 is congruent with the one obtained from NTSYS-PC. Study at species level diversity showed PPB, 91.37%; H, 0.2070; and I, 0.3282. A similar study performed on Chinese *N. scrophulariiflora* using ISSR markers, Liu *et al.* (2011) found PPB, 100%; H, 0.4073; and I, 0.5917 at species level. Compared to Chinese populations, Nepalese *N. scrophulariiflora* populations are less diverse. Similarly, in another study involving *Rheum* spp. the value obtained were PPB, 98.29%; H, 0.3107 and I, 0.4677 for *Rheum palmatum* and PPB, 91.43%; H, 0.2848 and I, 0.4333 for *Rheum tanguticum* (Wang *et al.*, 2012a). In another study involving Chinese yam (*Dioscorea opposita*), the overall Shannon index (I) was 0.3191 (Zhou *et al.*, 2005) indicating high genetic diversity. It has been reported that narrowly distributed species tend to possess low level of gene diversity (He *et al.*, 2000).

5.1.4.2 Overall Genetic Differentiation

The overall genetic differentiation within and among populations was calculated using Nei's genetic diversity statistics. The coefficient of genetic differentiation among population (G_{ST}) values had a mean value of 0.468 indicating that 46.8% of total genetic diversity existed among populations and 53.2% genetic diversity existed within populations. The values indicate high level of population genetic diversity in *N. scrophulariiflora.* Such high level of population genetic diversity may be due to breeding system or geographical isolation of populations (Hogbin and Peakall, 1999). In previous

studies, the values obtained were *Dysosma versipellis* (H_T = 0.083; H_S = 0.046; G_{ST} = 0.445) (Qiu *et al.*, 2005), and *Rheum palmatum* (H = 0.2848, G_{ST} = 0.537) and *Rheum tanguticum* (H = 0.1415, G_{ST} = 0.497) (Wang *et al.*, 2012a). Factors like evolution, mating system, seed dispersal, geographic range as well as natural selection affect population genetic diversity and structure of species, however, breeding system is a major factor which affects genetic diversity among and within population (Hamrick and Godt, 1990). Similarly, mean gene flow (N_m) value was 0.650. The gene flow was found to be high in case of Nepalese populations (N_m = 0.6501) compared to that of Chinese population (N_m = 0.2189) (Liu *et al.*, 2011). In order to find out the true gene flow at population level, samples representing all the populations of Nepal should be studied. Breeding system of a species is described to be an important determinant of variability at both species and population level (Han *et al.*, 2007) and *N. scrophulariiflora* is an outcrossing species (Liu *et al.*, 2011). Outcrossing promotes gene flow among populations and as a result decreases genetic differentiation. In case of Nepalese *N. scrophulariiflora*, the gene flow was high and thus the genetic differentiation between populations was low compared to the Chinese populations. It may appear that *N. scrophulariiflora* populations in Nepal do not suffer such loss of diversity which may be due to two probable reasons: 1. the study undertaken hasn't encompassed all the populations of Nepal and 2. The collected samples were from the conserved areas (such as conservation areas and national park) where it is illegal to collect endangered medicinal plants for trade. This might have allowed sufficient time for outcrossing and rhizome propagation of *N. scrophulariiflora*.

5.1.4.3 Genetic Differentiation using Analysis of Molecular Variance (AMOVA)

Differences within and among accessions were also evaluated by means of AMOVA (Excoffier *et al.*, 1992). GenAlex v. 6.5 was used to compute AMOVA. AMOVA is used to describe how the marker variance is partitioned within and among populations (Stewart and Excoffier, 1996). The analysis showed that there is a 50 - 50% variation within and among populations of *N. scrophulariiflora*. G_{ST} value showed a 46.8% variation existed among populations and 53.2% variation existed within population but AMOVA result showed that 50 – 50% variation existed within and among populations. In a study performed in *Brassica oleracea*, AMOVA analysis showed among accession variation to be 49.47% of total variance and within accession variance to be 52.17% (Phippen *et al.*, 1997).

5.2 DNA sequence-based study of *N. scrophulariiflora*

5.2.1 Characteristics of Barcode Loci

PCR amplification was successful for all loci and all the forty three samples included in the study showed successful amplification but only those which gave a much brighter

and intense band were selected (Plate 4.3). A sum total of 8 samples (two from MCA, three each from LNP and KCA) were chosen for sequencing and further study. In a study performed on species identification of Potamogetonaceae (Du *et al.*, 2011), sample size ranging from three to eight were chosen for a single species.

The bidirectional sequencing success was found to be 100% for *matk* and *rbcL* loci while it was 87.5% for *trnH-psbA* and ITS. Evaluation of sequence quality and coverage in a study performed by CBOL Plant working group (2009) showed that high quality bidirectional sequence read is routinely obtained for *rbcL* sequence while *trnH-psbA* intergenic spacer posed greater problem in obtaining bidirectional read with few ambiguous bases due to high frequency of mononucleotide repeats which disrupts sequencing reads and *matK* required manual editing and produced fewer bidirectional reads.

The band size for *matK* after PCR amplification was found to be approximately 800 base pairs for all the accessions from different locations. Similarly, band size for *rbcL* was found to be approximately 600 bps. Band size for *trnH-psbA* was found to be approximately 400 bps and for ITS the band size was approximately 700 bps (Table 4.18). The amplicon length range in base pairs for *matK*, *rbcL*, *trnH-psbA* and ITS has been reported in the range of 862-910 bps, 654-654 bps, 226-934 bps and 407-1630 bps respectively and the median amplicon length in bases in completely sequenced plastid genomes is found to be 880, 654, 500 and 705 bps respectively (Hollingsworth *et al.*, 2011).

A sequence similarity search was performed using BLAST (Altschul *et al.*, 1990). The pairwise sequence matches were ranked by statistical significance. The significance scores help to distinguish evolutionary related sequences from unrelated ones (Xiong, 2006). In BLAST searches, the statistical indicator, *E*-value (expectation value) indicates the probability that the resulting alignment from a database search are caused by random chances. If $E < 1e\text{-}50$, it can be said with extremely high confidence that the database match is result of homologous relationship; if $0.01 < E < 1e\text{-}50$ then the match can be considered a result of homology and if $0.01 < E < 10$ then the match is not considered significant but may indicate distant homology (Xiong, 2006).

For ITS sequence of accessions, ITS sequence of *N. scrophulariiflora* (NCBI accession KC413448.1) was found to be highly significant. Similarly for *trnH-psbA* sequence, database sequence of *N. scrophulariiflora* (NCBI accession KC413527.1) was found to be significant. For *matK* and *rbcL* sequences, sequences corresponding to *N. scrophulariiflora* were not found in the database. Hence, *matK* sequence of *Veronica spicata* (NCBI accession JN894225.1) was found to be highly significant. For *rbcL* sequence, the sequence of *Veronicastrum brunonianum* was found to be highly

significant. This can possibly be due to that fact that *Veronica, Veronicastrum* and *N. scrophulariiflora* all belong to same tribe Veroniceae on a hierarchical scheme of classification and both *matK* and *rbcL* are highly conserved protein coding sequences. Besides *Neopicrorhiza* is a monotypic genus and thus *N. scrophulariiflora* is its only species (Hong, 1984). Our study might be the first study based on DNA barcoding of *N. scrophulariiflora* since no *matK* and *rbcL* sequences were found deposited in NCBI database. Thus, the *matK* and *rbcL* sequences of *V. spicata* and *V. brunonianum* were chosen for further study.

Pairwise and multiple sequence alignment were performed using the four barcode regions among the accessions and using the outgroup sequences using ClustalW (Thompson *et al.*, 1994) module of MEGA v. 5.2 (Tamura *et al.*, 2011). Among the four regions, only *trnH-psbA* was found to have four variable sites (Table 4.20 and 4.21). *matK*, *rbcL* and ITS was found to lack any variable sites within accessions. Computation of intraspecific distances between the barcode accessions sequences revealed highest intraspecific distance in *trnH-psbA* sequence (0.007). Rest of the sequences lacked any intraspecific distance. From this data, it can be said that *trnH-psbA* sequence maybe used to identify *N. scrophulariiflora* from different locations. Interspecific distance was highest for *trnH-psbA* sequence (0.285) and lowest for *rbcL* sequence (0.005). In order to assign a barcode for species identification, the interspecific variation must exceed the intraspecific variation with a threshold value of 10 fold (Wiemers and Fiedler, 2007). Hence, *rbcL* can't be used to identify *N. scrophulariiflora*. *trnH-psbA* sequence may be used as a potential barcode for *N. scrophulariiflora* but intraspecific variation between accessions of different locations are high compared to other barcode loci hence it is too divergent to allow identification and hence is not considered suitable for barcoding *N. scrophulariiflora* but maybe used for identification of *N. scrophulariiflora* from different locations within Nepal.. From the data, *matK* and ITS can be potential barcode for identifying *N. scrophulariiflora*.

5.2.2 Species Identification using DNA Barcoding

In order to further elucidate the potentiality of the barcodes among the four barcode loci chosen to be used for identification of *N. scrophulariiflora* species, TaxonDNA (Meier *et al.*, 2006) computation was performed to calculate the Best Match/Best Close Match values. These values indicate the species identification rate of all four barcode markers used in the present study. It was observed that the species identification rate was high for *matK* (75%) and ITS (72.72%) sequences while *rbcL* and *trnH-psbA* sequence didn't show species identification. The comparison was made with reference to respective barcode sequences of four other outgroup sequences obtained from NCBI BLAST results. But this can be further studied using large sample size and incorporating the closely related species *P. kurrooa* in future studies.

5.2.3 Molecular Phylogeny

Phylogenetics is the study of evolutionary history of living organisms using tree-like diagrams to represent pedigrees of the organisms (Xiong, 2006). Evolutionary history using fossil records and morphological information has been surpassed by the use of genes and other biological macromolecules to study evolutionary relationship. In phylogenetic studies, a group of taxa descending from single common ancestor is called clade or monophyletic group while group of taxa that share more than one closest common ancestor are referred to as paraphyletic group (Xiong, 2006).

5.2.3.1 Model based Phylogeny

All phylogenetic methods are based on assumptions about the process of DNA substitution. As a result, all phylogenetic inferences depend on their underlying substitution models. In order to draw inferences properly it is necessary to have confidence in substitution models (Goldman, 1993). Use of one or other nucleotide substitution model affects stages of phylogenetic inferences like estimates of phylogeny, substitution rates, bootstrap values, posterior probabilities or tests of the molecular clock (Posada and Buckley, 2004). When the model of evolution assumed is wrong, the phylogenetic methods become less accurate and inconsistent (Bruno and Halpern). Hence, in order to accurately predict the phylogenetic analysis, appropriate models should be determined and used.

The best fit model of nucleotide substitution for the barcode loci was determined using the software jModelTest v. 2.1.4 (Darriba *et al.*, 2012) using Akaike Information criteria corrected (AICc). Akaike information criteria, or AIC (Akaike, 1974; 1998), technically is an asymptotically unbiased estimator of the expected relative Kullback-Leibler information quantity or distance (K-L) (Kullback and Leibler, 1951) which represents the amount of information lost when using a model to approximate another model. In phylogenetics, AIC can be visualized as amount of information lost when we use a model, for example GTR, to approximate the real process of nucleotide substitution and thus a model with smallest AIC is preferred. When sample size is small compared to the number of parameters ($n/k < 40$), corrected Akaike Information criteria (AIC_c) is used.

For *matK* sequence, the best nucleotide substitution model was found to be TVM+I (Posada, 2003) (Tranversion Model with invariable sites) with the lowest AICc score of 3322.27. Common models of nucleotide substitution include parameters that describe base frequencies, the substitution rates among the four nucleotides or the distribution of the rate of evolution among sites (rate variation) and among lineages (the molecular clock) (Posada, 2003). Transversion model assumes that there are variable base frequencies with variable transversion rates but equal transition rates (AC, AT, CG, GT, AG=CT). For *rbcL* sequence, with the lowest AICc score of 3952.99, TPM2uf+G (Kimura 3-

parameter (K81) with unequal base frequencies (Kimura, 1981)) was found to be the best substitution model. This model assumes unequal base frequencies with substitution rates as AC=AT, CG=GT and AG=CT. For *trnH-psbA* sequence, TPM1uf (Kimura, 1981) was found to be the best substitution model with the lowest AICc score of 1909.79. TPM1uf is also a Kimura 3-parameter (K81) with unequal base frequencies. It assumes unequal base frequencies with substitution rates between nucleotides as AC=GT, AT=CG and AG=CT. For ITS, the best nucleotide substitution model was GTR+I (General Time reversal) with the lowest AICc score of 2892.29. GTR is the most general, neutral, independent, finite-sites, time-reversible model possible where base frequencies are unequal and substitution rates are described as AC≠ AG≠ AT≠ CG≠ CT≠ GT (Tavaré, 1986).

Based on the models, a phylogenetic strict consensus tree for each locus was created with a confidence interval of 95% using jModelTest v. 2.1.4. These trees were compared with Maximum Parsimony trees constructed for all the respective barcode loci. All model based trees were congruent with their respective parsimonious trees.

5.2.3.2 Phylogeny Analysis using Parsimony

Using MEGA v. 5.2 Maximum Parsimonious (MP) trees for each locus was constructed and Consistency Index (CI) of all parsimonious trees was computed. The value of CI for *matK* was found to be 0.933 while MP trees of other loci lacked CI. CI value indicates the realness of clades present in generated cladograms (Shrestha, 2001). CI is widely used method to measure homoplasy which corresponds to the minimum amount of possible evolutionary change divided by actual tree length (i.e. the number of actual genetic changes in the tree). Lower values of consistency indices indicate that many characters are contradictory to the evolutionary tree. When a maximum parsimony tree is generated, the given set of data may yield a set of very similar trees with same tree length and generating a consensus tree may be useful to represent common features of this set of trees (Rohlf, 1982).

In phylogeny, the relations between taxa are presented in the form of tree with recent taxa at terminal ends. Such arrangement can be of three types: monophyly, paraphyly and polyphyly. Monophyletic groups contain taxa descended from single common ancestor such that two taxa share a unique common ancestor not shared by any other taxa (Xiong, 2006). A paraphyletic group meanwhile contains taxa that share more than one closest common ancestors (Xiong, 2006). In a polyphyletic group, however, some lineages of a population join with some lineages of second population to form a clade. In interspecific phylogeny, polyphyletic groups are considered to be based on convergent characters while paraphyletic groups are considered to be based upon shared primitive characters (Lowe *et al.*, 2004).

Phylogenetic trees are evaluated by some estimate of confidence interval, more commonly using bootstrap method (Felsenstein, 1985). The bootstrap method is a non-parametric statistical analysis which is used to assess the confidence limits on phylogenics. It tests the monophyly of individual clades (Sanderson, 1989). In this method, a pseudoreplicate sample is generated with sample size similar to original data. Since, the sampling occurs with replacement it is likely that some characters appear more than once in the pseudoreplicate and some characters may not be represented at all. The newly generated data matrix is used to generate a tree and the process is iterated as many as 100-1000 times.

Comparing model averaged phylogeny and maximum parsimony trees, *matK, trnH-psbA* and ITS sequences of all *N. scrophulariiflora* accessions formed monophyletic clade. In case of *rbcL* sequences, all the accessions and outgroup sequences of *Veronica officinalis* and *Veronicastrum brunonianum* appeared as a single clade (Figure 4.13 and 4.14) indicating no sequence level variation even at generic level. For *rbcL* locus, no sequence were available in NCBI for *N. scrophulariiflora* as well as its sister taxon *P. kurrooa*. Incorporation of these two sequences in future studies can shed some light on whether *rbcL* can be used as a barcode for identification of *N. scrophulariiflora* or not but from the current data it's highly improbable that *rbcL* can be used for identification because it couldn't clearly resolve the accessions of *N. scrophulariiflora* from that of outgroup sequences. Based on *trnH-psbA* sequences, although all accessions under study formed a single monophyletic clade, MCA accessions formed a subclade inside the large clade indicating MCA population being diverged from two other populations. As this locus is variable, it should be further explored for its potential use as a diagnostic marker for *N. scrophulariiflora*. In case of ITS sequence, it was observed that all the accessions and the reference sequence of *N. scrophulariiflora* from nucleotide database a single monophyletic clade while *P. kurrooa* was shown as sister species of *N. scrophulariiflora*. Thus, ITS holds a great promise as a potential barcode for identification of *N. scrophulariiflora*. In case of *matK*, the accessions from different locations were clustered as single monophyletic clade. Thus, *matK* also holds a promise for identification of *N. scrophulariiflora*. But this result has to be further verified with samples from all over the country and thus this research has opened a new avenue for further research of *N. scrophulariiflora* in Nepal.

5.3 Antibacterial Activity of Rhizome Extract

Infectious disease caused by bacteria, viruses, fungi and parasites are still a major threat to public health (Cos *et al.*, 2006). In the past three decades, a lot of microbial resistance to antibiotics have been observed (Chopra *et al.*, 1997). This has led to research on development of new anti-microbial drugs in order to treat the infectious disease and

thus the attention has been drawn to medicinal plants in order to identify potential drugs since medicinal plants contain a wide range of substance that can be used to treat lots of infectious diseases with reduced side effects (Duraipandiyan *et al.*, 2006).

In the present study, the antibacterial assay of methanolic extract of *N. scrophulariiflora* from Nepal was carried out. A lot of research activities on antimicrobial property has been focused on closely related species, *P. kurrooa* (Rani and Khullar, 2004; Kumar *et al.*, 2010; Sharma and Kumar, 2012; Usman *et al.*, 2012). A comparison of active constituents of both these species showed similar composition (Smit, 1968).

The methanolic extract from LNP possessed highest activity among all the extracts. And all the extracts were active against *Staphylococcus aureus, Klebsiella pneumoneae* and *Pseudomonas aeruginosa* while no activity was observed with other strains (Table 4.24). Methanolic extracts from MCA showed high potency against *P. aeruginosa*, *K. pneumoneae, and S. aureus* in decreasing order as observed by inhibition zones of 12, 11 and 10 mm respectively in comparison to 35, 42 and 40 mm respectively of Chloramphenicol. Similarly, methanolic extract of LNP showed 15, 13 and 15 mm zone of inhibition respectively for *P. aeruginosa, K. pneumoneae* and *S. aureus* and methanolic extract of KCA showed 9, 8 and 10 mm respectively for *P. aeruginosa, K. pneumoneae* and *S. aureus* against 35, 42 and 40 mm respectively for Chloramphenicol. In studies concerning antimicrobial activities of *P. kurrooa* extracts, the activity was found from moderate to potent. In a study performed to test antimicrobial activity of *P. kurrooa* extract against standard strain of *Salmonella typhi* and its hospital isolates, the activity of methanolic extract of *P. kurrooa* was found to be moderate with ≥ 5-9 mm zone of inhibition (Rani and Khullar, 2004) while in another study the ethanolic extract showed high antibacterial activity against standard strain of *S. typhi* (13 mm) while the methanolic extract showed moderate activity (9 mm) (Kumar *et al.*, 2010). In different studies concerning antimicrobial activity of *P. kurrooa* root extract against *S. aureus*, *B. subtilis*, *E. coli*, *P. aeruginosa*, *K. pneumoneae*, it was found that *E. coli* was most susceptible to ethanolic extract (Usman *et al.*, 2012). The finding was in congruence with another study where ethanol extract showed high activity against *E. coli* and other strains such as *S. aureus, K. pneumoneae, S. typhi* and *S. pyogenes* (Kumar *et al.*, 2010). In another study, methanolic extract of *P. kurrooa* was found highly active against *S. aureus, B. subtilis*, and *E. coli* (Sharma and Kumar, 2012). Contrary to these findings, in the present study, methanolic extract of *N. scrophulariiflora* didn't show any activity against *E. coli*, *S. typhi, B. subtilis, Serratia* and *Enterobacter spp* (Table 4.24) which indicated that more polar extract possessed higher antimicrobial activity. And also the activity may vary for extracts of samples collected from different altitude.

The antimicrobial activity of *N. scrophulariiflora* and *P. kurrooa* may be due to presence of different glucosides such as picrosides and kutkosides and/or other active

constituents like cucurbitacins, phenylethanoids and phenolics (Section 2.6). Besides, other compounds like saponins and alkaloids may also be held responsible (Kumar *et al.*, 2010). Thus this study supports the traditional use of *N. scrophulariiflora* to treat infections due to presence of antimicrobial activity in the rhizome extract of the plant.

5.4 Antioxidant Activity of Rhizome Extract

N. scrophulariiflora is highly important medicinal plant with wide spectrum of medicinal property. Among them, antioxidant property is due to wide variety of active constituents present in the different parts of the plant which includes iridoid glucosides, cucurbitacins, phenylethanoids, phenolics, alkaloids etc (Section 2.6). Most research on antioxidant properties and/or the medicinal properties are directed to closely related species *P. kurrooa*. Although these two species can be differentiated on the morphological basis, they may still show many chemical constituents being sister species.

The effects of antioxidants, in the DPPH radical scavenging test, reflect the hydrogen donating capacity of the compounds. 1,1-Diphenyl-2-picryl-hydrazyl (DPPH) is a stable free radical which has an unpaired valence electron at one atom of nitrogen bridge (Eklund *et al.*, 2005). When the radical form of DPPH is scavenged by an antioxidant through the donation of hydrogen to form a stable DPPH molecule, a color change from purple to yellow and a decrease in absorbance at 517 nm occurs which allows measurement of the antioxidant property of the compound or extract tested (Kant *et al.*, 2013).

In the present study, methanolic extracts of rhizomes of *N. scrophulariiflora* from three different locations were evaluated for their antioxidant property and the IC_{50} values were found to be 4.14 mg/mL, 1.74 mg/mL and 2.39 mg/mL for MCA, LNP and KCA respectively. The inhibition of DPPH molecules at increasing concentrations (i.e., quenching of free DPPH molecules by the rhizome extracts) showed a stoichometric relation with increasing concentration (Table 4.26, Figure 4.23) indicating increasing antioxidant activity with increasing concentration.

Generally, the antioxidant property is expressed as IC_{50} values and is compared with IC_{50} value of a standard antioxidant in order to quantify the antioxidant potential of that substance. IC_{50} value is defined as the amount of antioxidant necessary to decrease the absorbance of DPPH by 50% of the initial absorbance (Mishra *et al.*, 2012). The IC_{50} values of Gallic acid (standard), LNP, KCA, and MCA extracts were found to be 0.094, 1.742, 2.398, and 4.141 mg/ml respectively (Table 4.27, Figure 4.24). From the observed values, the antioxidant property was highest in LNP extract followed by KCA and MCA. The lower the IC_{50} value of any antioxidant, the better it is at quenching free radicals. This result indicated that LNP extract possibly contains greater number of phenolic

functionalities. The experiments were carried out in triplicates and were assessed for statistical significance by performing one way ANOVA and the observed values were found to be statistically significant (Table 4.28). Study of antioxidant property of *P. kurrooa* methanolic extract showed IC_{50} value to be 5.7 mg/ml (Govindarajan *et al.*, 2003). Compared to the antioxidant property of *P. kurrooa*, the antioxidant properties of Nepalese *N. scrophulariiflora* were high.

It has been reported that the antioxidant property of Kutki is due to the presence of iridoid glucosides such as Picroside I, II and kutkoside. As it is evident in a study of *P. kurrooa* that picroside I and kutkoside present in the extract scavenged superoxide anion thus exhibiting potent antioxidant property (Chander *et al.*, 1992). In the present investigation, the antioxidant property of methanolic extract of *N. scrophulariiflora*, which is used as hepatoprotective, immunomodulatory, anti-asthmatic, antipyretic etc drug was determined *in vitro*. The result suggested that the use of this plant extract in treatment of many diseases including hepatic damage, immunomodulation etc may be due to its antioxidant and free radical scavenging ability. Free radicals are implemented in wide range of diseases including liver cirrhosis, atherosclerosis, cancer, diabetes etc (Govindarajan *et al.*, 2003). Antioxidants reduce the adverse effects of these diseases through free radical scavenging action.

5.5 Quantification of Picrosides I and II in Extracts

Picrosides I and II, are iridoid glucosides which are a large group of monoterpenes with basic cyclopentano-[c]-pyran skeleton (Smit, 1968). The health benefits of Picrosides I and II have been studied regarding properties such as cardiovascular protective property (Kim *et al.*, 2007; Meng *et al.*, 2012), antihepatotoxic property (Ansari *et al.*, 1988; Jeena *et al.*, 1999), hypoglycemic/anti-diabetic property (Joy and Kuttan, 1999), immunomodulatory property (Smit *et al.*, 2000), antimalarial property (Irshad *et al.*, 2011), leshmanicidal property (Mittal *et al.*, 1998), antioxidant property (Chander *et al.*, 1992; Govindarajan *et al.*, 2003), antimicrobial property (Rani and Khullar, 2004; Usman *et al.*, 2012) etc.

Using HPLC system, the quantity of picrosides I and II in methanolic extracts of *N. scrophulariiflora* from three geographical locations were determined. Calibration curve plotted using standard picrosides I and II showed high correlation coefficient (R^2) values (0.998 and 0.999 for Picrosides I and II respectively (Table 4.29)). The elution time for Picroside II was 4.5 minutes and for Picroside I it was 7 minutes (Figure 4.26). In previous studies, the elution time varied from 7.5 minutes for Picroside II, 9.5 minutes for Picroside I (Bhandari *et al.*, 2008); 1.56 minutes for Picroside II, 2.64 minutes for Picroside I (Bhandari *et al.*, 2010); 12.917 minutes for Picroside II, 22.293 minutes for Picroside I (Katoch *et al.*, 2011). This may be due to the different parameters such as

solvent system, pressure and flow of solvent used. Solvent systems used in various studies were ethanol: water (Bhandari *et al.*, 2008), methanol: water (Bhandari *et al.*, 2010; Katoch *et al.*, 2011) and acetonitrile: acetic acid (Dwivedi *et al.*, 1997).

Based on the calibration curve of standard Picrosides I and II, the highest concentration of picroside I and II was found in MCA followed by KCA and LNP. Picroside I (678.705 µg/L) and Picroside II (2015.947 µg/L) was observed in extract from MCA region while in KCA extract, picroside I (426.501 µg/L) and picroside II (1029.281 µg/L) was observed. In case of LNP extract, picroside I (203.623 µg/L) and picroside II (751.637 µg/L) was observed. In general, concentration of picroside II in methanolic extracts from all locations was higher than picroside I concentration. In a study concerning the effect of altitude on picroside content of *P. kurrooa*, it was found that accumulation of picrosides depended on altitude variation (Katoch *et al.*, 2011). Altitudinal changes have effect on ecological niche including various factors like temperature, UV radiation, abiotic and biotic factors. In different ecological niches, plants behave differently in term of biochemical aspects in order to better adapt to their environment (Katoch *et al.*, 2011).

CHAPTER 6 – CONCLUSION

6.1 Conclusion

Nepal is rich in floral and faunal diversity with many different species of plant and animals. Among the 1950 species of medicinal and aromatic plants found in Nepal, *N. scrophulariiflora* is one of them. Due to multitude of health benefits and being a high value medicinal plant, it is one of the highly traded medicinal plants of Nepal. As a result, it is rendered vulnerable in its natural habitat. Natural influences such as climate change are also believed to contribute to the vulnerability of many medicinal plants including *N. scrophulariiflora*.

It was found from the present study using ISSR marker that Nepalese *N. scrophulariiflora* exhibits high level of polymorphism despite being subjected to high degree of collection for trade. Among the samples collected from Eastern (KCA), Central (LNP) and Western (MCA) Nepal, accessions from KCA showed high degree of polymorphism (96.26%) followed by MCA (88.89%) followed by LNP (66.27%). The present study shows that high degree of diversity exists among populations (H, 0.2348; I, 0.3628) of *N. scrophulariiflora* than at species level (H, 0.2070; I 0.3282). At the species level, Nepalese *N. scrophulariiflora* exhibited 91.37% polymorphism with Nei's genetic diversity (H) as 0.2070 and Shannon's diversity index (I) as 0.3282 which is low as compared to Chinese populations (Liu *et al.*, 2011). Similarly, the genetic distance was observed to be highest for L5 and K2 accession (0.957) and lowest for ME30 and ME31 (0.048) based on Jaccard's coefficient of similarity. Populations with higher genetic distance can be used as parents to develop elite varieties (Prakash *et al.*, 2006). Mean genetic differentiation between populations (G_{ST}) values showed that Nepalese *N. scrophulariiflora* exhibited 46.8% diversity among populations and 53.2% diversity existed within population and because of high gene flow (N_m = 0.6501) there was decreased genetic differentiation compared to Chinese *N. scrophulariiflora*. Low level of gene flow introduces lesser number of new genes into a population; as a result, the species can't cope with environmental threats rendering it vulnerable and endangered.

Species identification is fundamental to many research endeavors in biological sciences. DNA barcoding has emerged as most advanced molecular tool for species identification. Using four DNA barcode loci (*matK, rbcL, trnH-psbA* and ITS) for identification of potential barcodes, it was found from the present study that species identification rate of *matK* was 75% while that of ITS was 72.72%, based on the best match option of TaxonDNA, that these two barcode loci can potentially identify *N. scrophulariiflora* species of Nepal. Phylogenetic analysis performed using model based and parsimony based methods, based on *matK* and ITS sequences revealed *N. scrophulariiflora* to be monophyletic and *P. kurrooa* to be its sister species. In *trnH-psbA* sequence based cladogram, all the accessions formed a monophyletic clade but due to intraspecific

variation within the species itself, smaller sub-clades were also visible inside the main clade where accessions from MCA formed a sub-clade within the main clade. As *rbcL* couldn't resolve *N. scrophulariiflora* accession sequences and outgroup sequences, it can't be used for identification of Nepalese *N. scrophulariiflora*. A total of 16 sequences were generated during the study which will be deposited in public databases like NCBI or BOLD in near future which will aid in future research for species identification, authentication, molecular phylogenetics etc. Furthermore, the *trnH-psbA* and ITS sequence showed significant difference with that of Chinese *N. scrophulariiflora* and Indian *P. kurrooa*.

Present DNA Barcoding as well as intra-specific phylogenetic study on *N. scrophulariiflora* is novel work performed in Nepal. The DNA sequences of various loci thus generated will be submitted to NCBI in near future. These sequences will be the reference sequence for various future investigations.

The antibacterial property of methanolic rhizome extract from three different locations showed that LNP extract had the highest activity followed by MCA and KCA against *Staphylococcus aureus, Klebsiella pneumoneae* and *Pseudomonas aeruginosa* suggesting presence of antibacterial compounds in extracts of *N. scrophulariiflora*. The antibacterial activity was moderate compared to the Chloramphenicol. Thus the traditional use of *N. scrophulariiflora* to treat infections is supported by the present study. The rhizome extract also possessed antioxidant property and it was found that LNP extract was most potent followed by KCA and MCA and as a whole the extract from three locations were better antioxidants than Indian *P. kurrooa* (Govindarajan *et al.*, 2003). It is reported that antioxidant property of *N. scrophulariiflora* is due to presence of iridoid glucosides like Picroside I, II, kutkoside along with other class of compounds like phenolic glycosides, and phenylethanoids found in rhizomes as well as in other parts such as roots and leaves. Hence, study of picrosides I and II in rhizome extracts of *N. scrophulariiflora* from different location and its quantification showed that highest amount of picrosides I and II were found in MCA followed by KCA and LNP. Therefore, cultivation of *N. scrophulariiflora* in MCA region is recommended for commercialization purpose. This might be due to suitable habitat and climatic conditions for this species. It can also be possible that MCA genotypes may be elite in terms of chemical compounds that can only be proven by cultivating MCA genotypes in other suitable habitat and their chemical properties assessed.

6.2 Recommendations

Based on the study, following recommendations are drawn for long term conservation and sustainable utilization of *N. scrophulariiflora* in Nepal.

1. This project was constrained with limited budget. Therefore, number of population under study was limited to three as well as number of individuals collected per population had to be limited to 8 – 10 samples. It is highly recommended that future population genetic studies be conducted with sufficient samples for representative populations. The present study shows that *N. scrophulariiflora* is highly diverse at the genetic level but the fact can't be overlooked that these samples were collected from conserved sites like National Parks and Conservation Area. Vast quantity of this medicinal plant is being illegally collected and traded from buffer zones surrounding National Parks and Conservation areas. This activity may sooner be extended to the conserved and protected areas thus threatening this species to extinction. Therefore, conservation planning and management for sustainable harvesting of this species should be performed.
2. It is recommended that the local people, the harvesters and people practicing traditional medicine like amchis be trained and equipped with the knowledge of importance of proper harvesting period to allow outcrossing and pollination within populations and to retain genetic diversity of species which is vital for long term survival of species. People should also be made aware of the fact that complete harvesting of a rhizome from a location wipes out the future prospect of regeneration of the plant from remaining rhizome. Thus, it's highly necessary that some part of the rhizome be left behind to allow regeneration of the plant in the future.
3. *In situ* conservation strategies of pocket area under high threat should be developed and the seeds and germplasm of this species should be conserved through *ex situ* methods in seed banks and gene banks. As present DNA barcoding study revealed, population level genetic differences seeds from diverse geographical location should be conserved to capture overall genotypic diversity of the species.
4. Further research on genetic diversity of *N. scrophulariiflora* incorporating more representative populations should be performed to understand existing diversity level in all these populations found in Nepal. This can aid in formulating conservation practices for not just few but all populations found in Nepal which can lead to effective regulation of illegal trade of this species.
5. Based on molecular genetic diversity assessment and genetic distances revealed from this study, breeding experiments should be performed for the development of elite cultivars of *N. scrophulariiflora*.

6. From the present study, it is recommended that *matK* and ITS can be potential barcode for discrimination of *N. scrophulariiflora* and *P. kurroa*. Further studies involving both species from diverse geographical location can add *trnH-psbA* as candidate barcode for their discrimination. Furthermore, detailed barcode studies can also confirm if *P. kurrooa* species is also found in various parts of Nepal.
7. Molecular techniques such as PCR-RFLP on potential barcode sequences can be utilized for molecular diagnostic development of *N. scrophulariiflora* and its practical utilization for various purposes.
8. *P. kurrooa* is a "CITES listed" species of India while *N. scrophulariiflora* has been listed as "Vulnerable" in Nepal. This is because Nepalese *N. scrophulariiflora* is being recklessly exploited to fulfill National and International demands. Therefore, Nepalese Government has to be very cautious and reexamine its conservation status. It is recommended that Nepalese species should also be considered for CITES listing.
9. It is highly recommended that a National level assessment of molecular genetic diversity status and barcoding studies of all the medicinal plants be performed by Nepalese Government to ensure the formulation of state of the art and innovative conservation practices.
10. It is also highly recommended to the Government of Nepal to ensure a complete ban on export and trade of raw medicinal plant material without processing.
11. Further studies on qualitative and quantitative aspects of chemoprospecting of *N. scrophulariiflora* from diverse geographical location need to be performed for the selection of elite genotypes for further utilization.
12. Good cultivation practices (seed, cuttings, tissue culture etc.) of *N. scrophulariiflora* should be developed and disseminated to end users for income generation.

References:

Agarwal M, Shrivastava N and Padh H (2008) Advances in molecular marker techniques and their applications in plant sciences. *Plant cell reports*. **27**(4): 617-31

Akagi H, Yokozeki Y, Inagaki A and Fujimura T (1996) *Microsatellite DNA markers for rice chromosomes. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*. **93**(7): 1071-7

Akagi H, Yokozeki Y, Inagaki A, Nakamura A and Fujimura T (1996) A codominant DNA marker closely linked to the rice nuclear restorer gene, RF-1, identified with inter-SSR fingerprinting. *Genome / National Research Council Canada = Genome / Conseil national de recherches Canada*. **39**(6): 1205-9

Akaike H (1974) *A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*. **19**(6): 716-723

Akaike H (1998) *Information Theory and an Extension of the Maximum Likelihood Principle. In: *Selected Papers of Hirotugu Akaike*. (Eds.) Parzen E, Tanabe K and Kitagawa G. Springer New York. 199-213.

Alam MA, Naik PK and Mishra GP (2009) *Congruence of RAPD and ISSR Markers for Evaluation of Genomic Relationship Among 28 Populations of *Podophyllum hexandrum* Royle from Himachal Pradesh, India. *Turk. J. Bot.* **33**: 1-12*

Albach DC and Meudt HM (2010) Phylogeny of Veronica in the Southern and Northern Hemispheres based on plastid, nuclear ribosomal and nuclear low-copy DNA. *Molecular phylogenetics and evolution*. **54**(2): 457-471

Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**: 403-10

Anadan N, Rekha R, Saravanan N and Devaki T (1999) *Protective effects of *Picrorhiza kurrooa* against HCl/ethanol induced ulceration in rats. *Fitoterapia*. **70**: 498-501

Andayani N, Morales JC, M.R.J. Forstner, Supriatna J and Melnick DJ (2001) Genetic variability in mtDNA of the silvery gibbon: implications for the conservation of a critically endangered species. *Conserv Biol*. **15**(3): 1545-48*

Ansari RA, Aswal BS, Chander R, Dhawan BN, Garg NK, Kapoor NK, Kulshreshta DK, Mehdi H, Mehrotra BN, Patnaik GK and Sharma SK (1988) Hepatoprotective activity of kutkin, the iridoid glycoside mixture of *Picrorhiza kurroa. Indian J. Med. Res*. **87**: 401-404

Aoyama M, Mori M, Okumura M, Doi S and Anetai M (1997) *Antifungal activity of 2',6'-dihydroxy-4'-methoxyacetophenone and related compounds. *Mokuzai Gakkaishi (J. Wood Res. Soc.)*. **43**: 108-111*

Arcade A, Anselin F, Rampant PF, Lesage MC, Pâques LE and Prat D (2000) Application of AFLP, RAPD and ISSR markers to genetic mapping of European and Japanese larch. *Theoret. Appl. Genetics*. **100**(2): 299-307

Arif IA, Bakir MA, Khan HA, Al Farhan AH, Al Homaidan AA, Bahkali AH, Sadoon MA and Shobrak M (2010) A brief review of molecular techniques to assess plant diversity. *International journal of molecular sciences*. **11**(5): 2079-96

Banerjee D, Maity B, Nag S, Bandyopadhyay S and Chattopadhyay S (2008) *Healing potential of *Picrorhiza kurrooa* (Scrofulariaceae) rhizomes against indomethacin-induced gastric ulceration: a mechanistic exploration. *BMC Complement. Alternative Medicine* **88**(3)

Bantawa P, Ghosh S, Bhandari P, Singh B, Ghosh P, Ahuja P and Mondal T (2010) *Micropropagation of an elite line of *Picrorhiza scrophulariiflora* (Pennell), an endangered high valued medicinal plant of the Indo-China Himalayan region. *Med Aro Plant Science Biotechnology*. **4**(1): 1-7

Barrett RDH and Hebert PDN (2005) Identifying spiders through DNA barcodes. *Canadian Journal of Zoology*. **83**(3): 481-491

Baruah CC, Gupta PP, Nath A, Patnaik LGK and Dhawan BN (1998) Anti-allergic and anti-anaphylactic activity of picroliv - a standardized iridoid glycoside fraction of *Picrorhiza kurrooa*. *Pharmacological Research*. **38**(6): 487-492

Basu K, Dasgupta B, Bhattacharya S and Debnath P (1971) *Chemistry and pharmacology of apocynin, isolated from *Picrorhiza kurrooa* Royle ex Benth. *Current Science*. **40**(22): 603-604

Bhandari P, Kumar N, Singh B and Ahuja PS (2010) Online HPLC-DPPH method for antioxidant activity of *Picrorhiza kurroa* Royle ex Benth. and characterization of kutkoside by ultra-performance LC-electrospray ionization quadrupole time-of-flight mass spectrometry. *Indian J Exp Biol*. **48**(3): 323-328

Bhandari P, Kumar N, Singh B and Kaul VK (2008) Simultaneous determination of sugars and picrosides in *Picrorhiza* species using ultrasonic extraction and high-performance liquid chromatography with evaporative light scattering detection. *Journal of Chromatography A*. **1194**: 257-61

Bhattarai KR, Måren IE and Chaudhary RP (2013) *Medicinal plant knowledge of the Panchase region in the Middle Hills of the Nepalese Himalayas*. 21.

Bhattarai N, Tandon V and Ved DK (2002) *Highlights and Outcomes of the Conservation Assessment and Management Planning (CAMP) Workshop. In: *Sharing Local and National Experience in Conservation of Medicinal and Aromatic Plants in South Asia, Proceedings of the Workshop.* (Eds.) Bhattarai N, Tandon V and Ved DK. Pokhara, Nepal.

Bhattarai NK (1993) Folk medicinal use of plants for respiratory complaints in Central Nepal. *Fitoterapia*. **64**: 163-170

Blair MW, Panaud O and McCouch SR (1999) *Inter-simple sequence repeat (ISSR) amplification for analysis of microsatellite motif frequency and fingerprinting in Rice (*Oryza sativa* L.). *Theoret. Appl. Genetics*. **98**(5): 780-792

Botstein D, White RL, Skolnick M and Davis RW (1980) *Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics*. **32**(3): 314-331

Brown J. (2013) Model Selection [Online] <http://treethinkers.org/tutorials/model-selection/>

Brown JR and Doolittle WF (1995) Root of the universal tree of life based on ancient aminoacyl-tRNA synthetase gene duplications. *Proceedings of the National Academy of Sciences of the United States of America*. **92**(7): 2441-5*

Brummit RK (1992) *Vascular plant families and genera. Kew, Royal Botanical Gardens**:** 667

Bruno WJ and Halpern AL (1999) *Topological bias and inconsistency of maximum likelihood using wrong models. *Mol Biol Evol.* **16**(4): 564-6

Burnham KP and Anderson DR (1998) **Model selection and inference: a practical information-theoretic approach*. Springer, New York.

CBOL Plant Working Group (2009) A DNA barcode for land plants. *Proc Natl Acad Sci USA*. **106**(31): 12794-7

Chander R, Kapoor NK and Dhawan BN (1992) Picroliv, picroside-I and kutkoside from *Picrorhiza kurrooa* are scavengers of superoxide anions. *Biochemical Pharmacology*. **44**(1): 180-183

Chase M, Cowan R, Hollingsworth P, van den Berg C, Madriñán S, Petersen G, Seberg O, Jørgsensen T, Cameron K, Carine M, Pedersen N, Hedderson T, Conrad F, Salazar G, Richardson J, Hollingsworth M, Barraclough T, Kelly L and Wilkinson M (2007) A Proposal for a Standardised Protocol to Barcode All Land Plants. *Taxon*. **56**(2)

Chase MW, Salamin N, Wilkinson M, Dunwell JM, Kesanakurthi RP, Haidar N and Savolainen V (2005) Land plants and DNA barcodes: short-term and long-term goals. *Philos Trans R Soc Lond B Biol Sci*. **360**(1462): 1889-95

Chase MW, Soltis DE, Olmstead RG, Morgan D, Les DH, Mishler BD, Duvall MR, Price RA, Hills HG, Qiu Y-L, Kron KA, Rettig JH, Conti E, Palmer JD, Manhart JR, Sytsma KJ, Michaels HJ, Kress WJ, Karol KG, Clark WD, Hedren M, Gart BS, Jansen RK, Kim K-J, Wimpee CP, Smith JF, Furnier GR, Strauss SH, Xiang Q-Y, Plunkett GM, Soltis PS, Swensen SM, Williams SE, Gradek PA, Quinn CJ, Eguiarte LE, Barrett SCH, Dayanandan S and Albert VA (1993) Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbcL*. *Annals of the Missouri Botanic Garden*. **80**: 528-80

Chawla HS (2010) *Introduction to Plant Biotechnology (3rd Ed.)*. Oxford & IBH Publishing Co. Pvt. Ltd., New Delhi.

Chen S, Yao H, Han J, Liu C, Song J, Shi L, Zhu Y, Ma X, Gao T, Pang X, Luo K, Li Y, Li X, Jia X, Lin Y and Leon C (2010) Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PloS one*. **5**(1): 0008613

Cheng Z, Moore J and Yu L (2006) *High-throughput relative DPPH radical scavenging capacity assay. *Journal of agricultural and food chemistry*. **54**(20): 7429-36

Chinese Pharmacopoeia Commission (2005) **Pharmacopoeia of the People's Republic of China*. People's Medical Publishing House, Beijing.

Chopra I, Hodgson J, Metcalf B and Poste G (1997) *The search for antimicrobial agents effective against bacteria resistant to multiple antibiotics. *Antimicrobial Agents and Chemotherapy*. **41**(3): 497-503

Cos P, Vlietinck AJ, Berghe DV and Maes L (2006) *Anti-infective potential of natural products: how to develop a stronger *in vitro* 'proof-of-concept'. *J Ethnopharmacol*. **106**(3): 290-302

Cowan R, Chase MW, Kress WJ and V S (2006) 300,000 Species to Identify: Problems, Progress, and Prospects in DNA Barcoding of Land Plants. *Taxon*. **55**: 611-616

Cuenoud P, Savolainen V, Chatrou LW, Powell M, Grayer RJ and Chase MW (2002) Molecular phylogenetics of Caryophyllales based on nuclear 18S rDNA and plastid *rbcL*, *atpB*, and *matK* DNA sequences. *American journal of botany*. **89**(1): 132-44

Dalirsefat SB, da Silva Meyer A and Mirhoseini SZ (2009) Comparison of similarity coefficients used for cluster analysis with amplified fragment length polymorphism markers in the silkworm, *Bombyx mori*. *J Insect Sci*. **9**: 1-8

Darriba D, Taboada GL, Doallo R and Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nat Meth*. **9**(8): 772-772

de Groot GA, During HJ, Maas JW, Schneider H, Vogel JC and Erkens RH (2011) Use of rbcL and trnL-F as a two-locus DNA barcode for identification of NW-European ferns: an ecological perspective. *PloS one*. **6**(1): 0016371

de Vere N, Rich TCG, Ford CR, Trinder SA, Long C, Moore CM, Satterthwaite D, Davies H, Allainguillaume J, Ronca S, Tatarinova T, Garbett H, Walker K and Wilkinson MJ (2011) Direct submission to NCBI.

Degtjareva GV, Logacheva MD, Samigullin TH, Terentieva EI and Valiejo-Roman CM (2012) Organization of chloroplast psbA-trnH intergenic spacer in dicotyledonous angiosperms of the family umbelliferae. *Biochemistry Moscow*. **77**(9): 1056-1064

Department Of Forestry (undated) *Hamro Ban, annual report for Fiscal Year 2063/2064 (In Nepali)*. Department of Forestry, Babarmahal, Kathmandu, Nepal.

Dhanya K and Sasikumar B (2010) *Molecular Marker Based Adulteration Detection in Traded Food and Agricultural Commodities of Plant Origin with Special Reference to Spices. *Current Trends in Biotechnology and Pharmacy*. **4**(1): 454-489

Dhawan BN (1995) Picroliv - a new hepatoprotective agent from an Indian medicinal plant, *Picrorhiza kurroa*. *Med. Chem. Res*. **5**: 595-605

Dice LR (1945) *Measures of the Amount of Ecologic Association Between Species. *Ecology*. **26**(3): 297-302

Doebley J, Durbin M, Golenberg EM, Clegg MT and Ma. DP (1990) Evolutionary analysis of the large subunit of carboxylase (*rbcL*) nucleotide sequence among the grasses (Gramineae). *Evolution; international journal of organic evolution*. **44**: 1097-1108

Donoghue MJ, Olmstead RG, Smith JFP and Almer JD (1993) Phylogenetic relationships of Dipsacales based on *rbcL* sequences. *Annals of the Missouri Botanic Garden*. **79**: 333-45

DPR (2012) *Plants of Nepal: Fact Sheet*. Government of Nepal, Ministry of Forests and Soil Conservation, Department of Plant resources, Kathmandu.

Du Z-Y, Qimike A, Yang C-F, Chen J-M and Wang Q-F (2011) Testing four barcoding markers for species identification of Potamogetonaceae. *Journal of Systematics and Evolution*. **49**(3): 246-251

Duraipandiyan V, Ayyanar M and Ignacimuthu S (2006) *Antimicrobial activity of some ethnomedicinal plants used by Paliyar tribe from Tamil Nadu, India. *BMC Complementary and Alternative Medicine*. **6**(1): 35

Dwivedi AK, Kulkarni D and Singh S (1997) Sensitive high-performance liquid chromatographic assay method for the determination of picroside I in plasma. *Journal of Chromatography B*. **698**: 317-320

Eklund PC, Langvik OK, Warna JP, Salmi TO, Willfor SM and Sjoholm RE (2005) *Chemical studies on antioxidant mechanisms and free radical scavenging properties of lignans. *Org Biomol Chem*. **3**(18): 3336-47

Erskine W and Muehlbauer FJ (1991) *Allozyme and morphological variability, Out crossing rate and core collection formation in lentil germplasm. *Theor. and Appl. Genet.* **83**: 119-125

Excoffier L, Smouse PE and Quattro JM (1992) *Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*. **131**(2): 479-91

Fang DQ, Federici CT and Roose ML (1997) *Development of molecular markers linked to a gene controlling fruit acidity in citrus. *Genome / National Research Council Canada = Genome / Conseil national de recherches Canada*. **40**(6): 841-9

Fazekas AJ, Burgess KS, Kesanakurti PR, Graham SW, Newmaster SG, Husband BC, Percy DM, Hajibabaei M and Barrett SC (2008) Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. *PloS one*. **3**(7): e2802

Felsenstein J (1985) *Confidence limits on phylogenies: An approach using the bootstrap. *Evolution; international journal of organic evolution*. **39**: 783-791

Formagio AS, Kassuya CA, Neto FF, Volobuff CR, Iriguchi EK, Vieira Mdo C and Foglio MA (2013) The flavonoid content and antiproliferative, hypoglycaemic, anti-inflammatory and free radical scavenging activities of *Annona dioica* St. Hill. *BMC Complement Altern Med*. **13**(14): 1472-6882*

Franca LT, Carrilho E and Kist TB (2002) A review of DNA sequencing techniques. *Quarterly reviews of biophysics*. **35**(2): 169-200

Freville H, Justy F and Olivieri I (2001) Comparative allozyme and microsatellite population structure in a narrow endemic plant species, *Centaurea corymbosa* Pourret (Asteraceae). *Molec Ecol* **10**: 879-889

Fu LG (1992) **Plant Red Book of China: Rare Threatened Plant*. Science Publishing House Press, Beijing.

Gaikwad PS, Bhope SG, Kuber VV and Patil MJ (2010) Validated TLC method for Simultaneous Quantitation of Kutkoside and Picroside-I from Kutki Extract. *Phytochemical Analysis*

Gastony GJ (1988) The Pellaea glabella complex: electrophoretic evidence for the derivations of the agamosporous taxa and a revised taxonomy. *Am Fern J*. **78**: 44-67

Gelfand DH (1989) *Taq* DNA polymerase, in PCR Technology: Principles and Applications for DNA Amplifcation. Ehrlich HA. New York, Stockton Press

Ghimire S, McKey D and Aumeeruddythomas Y (2005) Conservation of Himalayan medicinal plants: Harvesting patterns and ecology of two threatened species, *Nardostachys grandiflora* DC. and *Neopicrorhiza scrophulariiflora* (Pennell) Hong. *Biological Conservation*. **124**(4): 463-475

Ghimire SK (2008) Medicinal plants in the Nepal Himalaya: Current issues, sustainable harvesting, knowledge gaps and research priorities. In: *Medicinal plants in Nepal: Anthology of contemporary research.* (Eds.) Jha PK, Karmacharya SB, Chhetri MK, Thapa CB and Shrestha BB. Ecological Society (ECOS). Kathmandu, Nepal**:** 25-42.

Ghimire SK, Sapkota IB, Oli VR and Parajuli RR (2008) *NTFPs of Nepal Himalaya*. WWF Nepal Program, Kathmandu, Nepal.

Gogarten JP, Kibak H, Dittrich P, Taiz L, Bowman EJ, Bowman BJ, Manolson MF, Poole RJ, Date T and Oshima T (1989) *Evolution of the vacuolar H+-ATPase: implications for the origin of eukaryotes. *Proc Natl Acad Sci USA*. **86**(17): 6661-6665

Goldman N (1993) Statistical tests of models of DNA substitution. *J Mol Evol*. **36**(2): 182-98

Gonzalez MA, Baraloto C, J. E, Mori SA, Petronelli P and al. e (2009) Identification of Amazonian trees with DNA barcodes. *PloS one*. **4**: e7483

Govindarajan R, Vijayakumar M, Rawat AK and Mehrotra S (2003) Free radical scavenging potential of *Picrorhiza kurrooa* Royle ex Benth. *Indian J Exp Biol*. **41**(8): 875-9

Graham G, Henry R and Redden R (1994) Identification of navy bean varieties usind random amplification of polymorphic DNA. *Aust J Exp Agr*. **34**: 1173-1176

Groom MJ, Meffe GK and Carroll CR (2006) *Principles of Conservation Biology (3rd ed.)*. Sinauer Associates, Sunderland, MA.

Grunenwald H (2003) Optimization of Polymerase Chain Reactions. In: *PCR Protocols*. (Eds.) Barlett JMS and Stirling D. 226, Humana Press. 89-100.

Gupta A, Khajuria A, Singh J, Bedi KL, Satti NK, Dutt P, Suri KA, Suri OP and Qazi GN (2006) Immunomodulatory activity of biopolymeric fraction RLJ-NE-205 from Picrorhiza kurroa. *Int Immunopharmacol*. **6**(10): 1543-9

Gupta M, Chyi YS, Romero-Severson J and Owen JL (1994) Amplification of DNA markers from evolutionarily diverse genomes using single primers of simple-sequence repeats. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*. **89**(7-8): 998-1006

Gupta PK and Varshney RK (2000) The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. *Euphytica*. **113**(3): 163-185

Gupta PP (2001) *Picroliv: hepatoprotective, immunomodulator. *Drugs of the future*. **26**(1): 25-31

Gyawali MB (2008) *Aspects of Traditional Medicine in Nepal*. Institute of Natural Medicine, University of Toyama, Japan.

Hamrick JL and Godt MJW (1990) Allozyme diversity in plant species. In: **Plant Population Genetics, Breeding, and Genetic Resources*. (Eds.) Brown AHD, Clegg MT, Kahler AL and Weir BS. Sinauer. Sunderland, Massachusetts, USA**:** 43-63.

Han J, Zhang W, Cao H, Chen S and Wang Y (2007) *Genetic diversity and biogeography of the traditional Chinese medicine, *Gardenia jasminoides*, based on AFLP markers. *Biochemical Systematics and Ecology*. **35**(3): 138-145

Hara H, Chater AO and Williams LHJ (1982) *An enumeration of the flowering plants ofNepal*. Trustees of British Museum, London.

Harrison, Laverty and Sterling. (2013) Genetic Diversity [Online] <http://cnx.org/content/m12158/1.2/>

Harrison CJ and Langdale JA (2006) A step by step guide to phylogeny reconstruction. *Plant J*. **45**(4): 561-72

Harsulkar A, Kurane J and Shinde V (2009) **Application of ISSR marker in pharmacognosy: Current update*. 3.

Hart B, Elferink JG and PH N (1992) *Effect of apocynin on the induction of ulcerative lesions in rat skin injected with tubercle bacteria. *Int J Immunopharmacol*. **14**(6)

Hasebe M, Omori T, Nakazawa M, Sano T, Kato M and Iwatsuki K (1994) *rbcL* gene sequences provide evidence for the evolutionary lineages of leptosporangiate ferns. *Proceedings of the National Academy of Sciences*. **91**(12): 5730-5734

He T, Rao G, You R, Ge S and Zhang D (2000) *Genetic diversity of widespread *Ophiopogon intermedius* (Liliaceae s.l.): a comparison with endangered *O. xylorrhizus*. *Biological Conservation*. **96**(2): 253-257

Hebert PD, Cywinska A, Ball SL and deWaard JR (2003) Biological identifications through DNA barcodes. *Proc Biol Sci*. **270**(1512): 313-21

Hebert PDN, Stoeckle MY, Zemlak TS and Francis CM (2004) Identification of Birds through DNA Barcodes. *PLoS Biol*. **2**(10): e312

Heinz DJ (1987) *Sugarcane improvement through breeding*. Elsevier, Amsterdam.

Helentjaris T, Slocum M, Wright S, Schaefer A and Nienhuis J (1986) Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphisms. *Theoret. Appl. Genetics*. **72**(6): 761-769

Hennig W (1966) *Phylogenetic Systematics*. University of Illinois Press, Urbana, Illinois.

Hilu K and Liang H (1997) The *matK* gene: sequence variation and application in plant systematics. *American journal of botany*. **84**(6): 830

Hilu KW, Borsch T, Muller K, Soltis DE, Soltis PS, Savolainen V, Chase MW, Powell MP, Alice LA, Evans R, Sauquet H, Neinhuis C, Slotta TA, Rohwer JG, Campbell CS and Chatrou LW (2003) Angiosperm phylogeny based on *matK* sequence information. *American journal of botany*. **90**(12): 1758-76

Hogbin PM and Peakall R (1999) Evaluation of the conservation of genetic research to the management of endangered plant *Zieria prostrata*. *Conserv Biol*. **13**: 514-522

Hollingsworth PM, Graham SW and Little DP (2011) Choosing and Using a Plant DNA Barcode. *PloS one*. **6**(5): e19254

Hong D, Yang H, Jin C-l, Fischer MA, Holmgren NH and Mill RR. (2014) Scrophulariaceae [Online] <http://www.efloras.org/florataxon.aspx?flora_id=2&taxon_id=10811>

Hong DY (1984) *Neopicrorhiza*. *Opera Bot.* **75**: 56

Hooker JD (1885) *The Flora of British India*. vol 4 Bishen Singh Mahendra Pal Singh, Dehradun.

Hu H and Yang P (2005) *Studies on chemical constituents of *Picrorhiza scrophulariiflora* Pennell. *Chinese Journal of Pharmaceuticals*. **36**(6): 336-339

Huang K, He L, Qu Y, Gao H, Deng X and Wu L (2009) *Isolation and identification of chemical constituents from roots of *Picrorhiza scrophulariiflora* Pennell. . *Shenyang Yaoke Daxue Xuebao*. **26**(2): 112-15

Huang S-X, Liao X, Nie Q-J, Ding L-S and Peng S-L (2004) *Phenyl and Phenylethyl Glycosides from *Picrorhiza scrophulariiflora*. *Helvetica Chimica Acta*. **87**(3): 598-604

Hussain A, Gupta V, Ali J, Ranjekar P and Siddiq E (2000) *Physiological characterization, genetics and molecular mapping of a new source of temperature sensitive genetic male sterility in rice. Fourth International Rice Genetics Symposium, IRRI, Philippines: 95

Indian Herbal Pharmacopoeia (2002) *Indian Herbal Pharmacopoeia, a Revised New Edition*. Indian Drugs Manufacturers Association, Mumbai.

Innis MA and Gelfand DH (1990) *PCR Protocols: A Guide to Methods and Applications*. Academic Press, San Diego, CA.

Innis MA, Myambo KB, Gelfand DH and Brow MA (1988) *DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA. *Proceedings of the National Academy of Sciences*. **85**(24): 9436-9440

Internet visit 1. (2013) Scrophulariaceae [Online] <http://en.wikipedia.org/wiki/Scrophulariaceae>

Internet visit 2. (2013) Vascular Plant Families and Genera - List of genera in family SCROPHULARIACEAE [Online] <http://data.kew.org/cgi-bin/vpfg1992/genlist.pl?SCROPHULARIACEAE>

Internet visit 3. (2014) Multigene families with identical genes [Online] <http://www.eplantscience.com/index/genetics/multigene_families_in_eukaryotes/multigene_families_with_identical_genes.php>

Irshad S, Mahmood M and Perveen F (2012) In-Vitro Anti-Bacterial Activities of Three Medicinal Plants Using Agar Well Diffusion Method. *Research Journal of Biology*. **2**(1): 1-8

Irshad S, Mannan A and Mirza B (2011) *Antimalarial activity of three Pakistani medicinal plants. *Pak J Pharm Sci.* **24**(4): 589-91

IUCN. (2013) The IUCN Red List of Threatened Speceis [Online] <http://www.iucnredlist.org/>

Jaakola L, Suokas M and Häggman H (2010) Novel approaches based on DNA barcoding and high-resolution melting of amplicons for authenticity analyses of berry species. *Food Chemistry*. **123**(2): 494-500

Jaccard P (1908) Nouvelles recherche sur la distribution florale. *Bulletin. Societe Vaudoise Sciences Naturelles*. **44**: 223-270*

Jagetia GC and Baliga MS (2004) The Evaluation of Nitric Oxide Scavenging Activity of Certain Indian Medicinal Plants In Vitro: A Preliminary Study. *Journal of Medicinal Food*. **7**(3): 343-348

Jeena KJ, Joy KL and Kuttan R (1999) *Effect of *Emblica officinalis*, *Phyllanthus amarus* and *Picrorrhiza kurroa* on N-nitrosodiethylamine induced hepatocarcinogenesis. *Cancer Letters*. **136**(1): 11-16

Jonah PM, Bello LL, Lucky O, Midau A and Moruppa SM (2011) Review : The Importance of Molecular Markers in Plant Breeding Programmes. *Global Journal of Science Frontier Research*. **11**(5)

Joshi K, Chavan P, Warude D and Patwardhan B (2004) Molecular markers in herbal drug technology. *Current Science*. **87**(2)

Joshi P and Dhawan V (2007) *Assessment of genetic fidelity of micropropagated *Swertia chirayita* plantlets by ISSR marker assay. *Biol Plant*. **51**(1): 22-26

Joshi SP, Gupta VS, Aggarwal RK, Ranjekar PK and Brar DS (2000) *Genetic diversity and phylogenetic relationship as revealed by inter simple sequence repeat (ISSR) polymorphism in the genus Oryza. *Theoret. Appl. Genetics*. **100**(8): 1311-1320

Joy KL and Kuttan R (1999) *Anti-diabetic activity of *Picrorrhiza kurroa* extract. *Journal of Ethnopharmacology*. **67**(2): 143-148

Joy KL, Rajeshkumar NV, Kuttan G and Kuttan R (2000) Effect of Picrorrhiza kurroa extract on transplanted tumours and chemical carcinogenesis in mice. *J Ethnopharmacol*. **71**(1-2): 261-6

Jurado-Rivera JA, Vogler AP, Reid CA, Petitpierre E and Gomez-Zurita J (2009) DNA barcoding insect-host plant associations. *Proc Biol Sci*. **276**(1657): 639-48

Kant K, Walia M, Agnihotri VK, Pathania V and Singh B (2013) Evaluation of Antioxidant Activity of *Picrorhiza kurroa* (Leaves) Extracts. *Indian J Pharm Sci*. **75**(3): 324-9

Katiyar C, Gupta A, Kanjilal S and Katiyar S (2012) *Drug discovery from plant sources: An integrated approach. *Ayu*. **33**(1): 10-9

Katoch M, Fazli IS, Suri KA, Ahuja A and Qazi GN (2011) Effect of altitude on picroside content in core collections of *Picrorhiza kurrooa* from the north western Himalayas. *Journal of natural medicines*. **65**: 5

Kesanakurti PR, Fazekas AJ, Burgess KS, Percy DM, Newmaster SG, Graham SW, Barrett SC, Hajibabaei M and Husband BC (2011) Spatial patterns of plant diversity below-ground as revealed by DNA barcoding. *Molecular ecology*. **20**(6): 1289-302

Kihara H (1955) *Fauna and flora of Nepal Himalaya*. Kyoto University, Kyoto, Japan.

Kim D-S, Woo E-R, Chae S-W, Ha K-C, Lee G-H, Hong S-T, Kwon D-Y, Kim M-S, Jung Y-K, Kim H-M, Kim H-K, Kim H-R and Chae H-J (2007) *Plantainoside D protects adriamycin-induced apoptosis in H9c2 cardiac muscle cells via the inhibition of ROS generation and NF-κB activation. *Life Sciences*. **80**(4): 314-323

Kim I, Kawahara N and Uchiyama N (2006) *Iridoid glycosides from roots of *Neopicrorhiza scrophulariiflora Phytochemistry*. **67**(24): 2691-2696

Kim IH, Kaneko N, Uchiyama N, Lee JE, Takeya K, Kawahara N and Goda Y (2006) Two phenylpropanoid glycosides from Neopicrorhiza scrophulariiflora. *Chem Pharm Bull*. **54**(2): 275-7

Kim IH, Uchiyama N, Kawahara N and Goda Y (2006) Iridoid glycosides and cucurbitacin glycoside from *Neopicrorhiza scrophulariiflora*. *Phytochemistry*. **67**(24): 2691-6

Kimura M (1981) *Estimation of evolutionary distances between homologous nucleotide sequences. *Proceedings of the National Academy of Sciences of the United States of America*. **78**(1): 454-8

Kitching IJ, Forey PL, Humphries CJ and Williams DM (1998) **Cladistics: The Theory and Practice of Parsimony Analysis (2nd edn)*. Oxford University Press, Oxford, UK.

Kokubun T and Harborne JB (1995) *Phytoalexin induction in the sapwood of plants of the maloideae (Rosaceae): biphenyls or dibenzofurans. *Phytochemistry*. **40**: 1649-1654

Kress WJ and Erickson DL (2007) A two-locus global DNA barcode for land plants: the coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. *PloS one*. **2**(6)

Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA and Janzen DH (2005) Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences of the United States of America*. **102**(23): 8369-8374

Kullback S and Leibler RA (1951) *On Information and Sufficiency. *The Annals of Mathematical Statistics*. **22**(1): 79-86

Kumar VP, Sivaraj A, Madhumitha G, Saral AM and Senthil BK (2010) *In-vitro* antibacterial activities of *Picrorhiza kurroa* rhizome extract using agar well diffusion method. *International Journal of Current Pharmaceutical Research*. **2**(1): 30-33

Kunwar R, Shrestha K, Dhungana S, Shrestha P and Shrestha K (2010) Floral Biodiversity of Nepal: an update. *Journal of Natural History Museum*. **25**: 295-311

Kuo LY, Li FW, Chiou WL and Wang CN (2011) First insights into fern *matK* phylogeny. *Molecular phylogenetics and evolution*. **59**(3): 556-66

Lahaye R, van der Bank M, Bogarin D, Warner J, Pupulin F, Gigot G, Maurin O, Duthoit S, Barraclough TG and Savolainen V (2008) DNA barcoding the floras of biodiversity hotspots. *Proceedings of the National Academy of Sciences*. **105**(8): 2923-2928

Lavie D and Glotter E (1971) *The cucurbitanes, a group of tetracyclic triterpenes. *Fortschr Chem Org Naturst*. **29**: 307-62

Lawyer FC, Stoffel S, Saiki RK, Myambo K, Drummond R and Gelfand DH (1989) Isolation, characterization, and expression in *Escherichia coli* of the DNA polymerase gene from *Thermus aquaticus*. *J Biol Chem*. **264**(11): 6427-37

Levin RA, Wagner WL, Hoch PC, Nepokroeff M, Pires JC, Zimmer EA and Sytsma KJ (2003) Family-level relationships of Onagraceae based on chloroplast *rbcL* and *ndhF* data. *American journal of botany*. **90**(1): 107-15

Levinson G and Gutman GA (1987) Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Molecular biology and evolution*. **4**(3): 203-21

Lewontin RC (1995) *The Apportionment of Human Diversity. In: *Evolutionary Biology*. (Eds.) Dobzhansky T, Hecht M and Steere W. Springer US. 381-398.

Li G-D and Sun H (2012) *Molecular phylogeny and biogeography of the arctic-alpine genus, *Lagotis* (Plantaginaceae). *Unpublished data*

Li G and Park YJ (2012) *SCAR markers for discriminating species of two genera of medicinal plants, *Liriope* and *Ophiopogon*. *Genetics and molecular research : GMR*. **11**(3): 2987-96

Li JX, Li P, Tezuka Y, Namba T and Kadota S (1998) Three phenylethanoid glycosides and an iridoid glycoside from *Picrorhiza scrophulariiflora*. *Phytochemistry*. **48**(3): 537-542

Li P, Matsunaga K, Yamakuni T and Ohizumi Y (2000) Potentiation of nerve growth factor-action by picrosides I and II, natural iridoids, in PC12D cells. *Eur J Pharmacol*. **406**(2): 203-8

Liu X-L, Qian Z-G, Liu F-H, Yang Y-W and Pu C-X (2011) Genetic diversity within and among populations of *Neopicrorhiza scrophulariiflora* (Scrophulariaceae) in China, an endangered medicinal plant. *Biochemical Systematics and Ecology*. **39**(4-6): 297-301

Long DG, Paton JA, Squirrell J, Woodhead M and Hollingsworth PM (2006) Morphological, ecological and genetic evidence for distinguishing *Anastrophyllum joergensenii* Schiffn. and *A. alpinum* Steph. (Hepaticae; Lophoziaceae). *Journal of Bryology*. **28**: 108-117

Lowe A, Stephen H and Ashton P (2004) *Ecological Genetics: Design, Analysis, and Application*. Blackwell Publishing, Victoria.

Lübberstedt T and Varshney R (2013) **Diagnostics in Plant Breeding*. Springer Science.

Luper S (1998) A Review of Plants Used in the Treatment of Liver Disease: Part 1. *Alternative Medicine Review*. **3**(6): 410-21

Lynch M and Milligan BG (1994) Analysis of population genetic structure with RAPD markers. *Molecular ecology*. **3**(2): 91-9*

Maddison WP, Donoghue MJ and Maddison DR (1984) Outgroup analysis and Parsimony. *Syst. Zool.* **33**: 83-103

Malla SB, Shrestha AB, Rajbhandari SB, Shrestha TB, Adhikari PM and Adhikari SR (1976) *Flora of Langtang and cross section vegetation survey (central zone)*. His Majesty's Ministry of Forests, Department of Medicinal Plants, Kathmandu, Nepal.

Manandhar NP (1986) Ethnobotany of Jumla district, Nepal. *Int. J. Crude Drug Res*. **24**: 81-89

Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Res*. **27**(2): 209-20

Mayr E (1969) *Principles of systematic zoology*. McGraw Hill, New York.

McDermott J and McDonald B (1993) Gene flow in plant pathosystems. *Annu Rev Phytopathol*. **31**: 353-373

Mehrotra R, Rawat S, Kulshrestha D, Patnaik G and Dhawan B (1990) *In vitro studies on the effect of certain natural products against hepatitis B virus. *Indian Journal Medicinal Research*. **92**: 133-138

Meier R, Shiyang K, Vaidya G and Ng PK (2006) DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. *Syst Biol*. **55**(5): 715-28

Meier R, Shiyang K, Vaidya G and Ng PKC (2006) DNA barcoding and taxonomy in Diptera: A tale of high intraspecific variability and low identification success. *Systematic Biology*. **55**: 715-28

Meng F, Hou Z, Li Y, Yang Y and Yu B (2012) *The protective effect of picroside II against hypoxia-reoxygenation injury in neonatal rat cardiomyocytes. *Pharm Biol.* **50**(10): 1226-32

Meyer CP and Paulay G (2005) DNA Barcoding: Error Rates Based on Comprehensive Sampling. *PLoS Biol*. **3**(12): e422

Meyer W, Mitchell TG, Freedman EZ and Vilgalys R (1993) *Hybridization probes for conventional DNA fingerprinting used as single primers in the polymerase chain reaction to distinguish strains of *Cryptococcus neoformans*. *J Clin Microbiol*. **31**(9): 2274-80

Mishra K, Ojha H and Chaudhury NK (2012) Estimation of antiradical properties of antioxidants using DPPH assay: A critical review and results. *Food Chemistry*. **130**(4): 1036-1043

Mittal N, Gupta N, Saksena S, Goyal N, Roy U and Rastogi AK (1998) *Protective effect of picroliv from *Picrorhiza kurroa* against *Leishmania donovani* infections in *Mesocricetus auratus*. *Life Sciences*. **63**(20): 1823-1834

Miwa H, OdrzykoskiI J, Matsui A, Hasegawa M, Akiyama H. and al. e (2009) Adaptive evolution of *rbcL* in *Conocephalum* (Hepaticae, bryophytes). *Gene*. **441**: 169-175

Mondini L, Noorani A and Pagnotta M (2009) Assessing Plant Genetic Diversity by Molecular Tools. *Diversity*. **1**(1): 19-35

Moniz MB and Kaczmarska I (2010) Barcoding of diatoms: nuclear encoded ITS revisited. *Protist*. **161**(1): 7-34

Morgan DR and Soltis DE (1993) *Phylogenetic relationships among members of Saxifragaceae *sensu lato* based on *rbcL* sequence data. *Ann. Mo. Bot. Garden*. **80**: 631-660

Muller KF, Borsch T and Hilu KW (2006) Phylogenetic utility of rapidly evolving DNA at high taxonomical levels: contrasting *matK*, *trnT-F*, and *rbcL* in basal angiosperms. *Molecular phylogenetics and evolution*. **41**(1): 99-117

Mulliken TA (2000) Implementing CITES for Himalayan medicinal plants *Nardostachys grandiflora* and *Picrorhiza kurrooa*. *Traffic Bulletin*. **18**: 63-72

Nagaoka T and Ogihara Y (1997) *Applicability of inter-simple sequence repeat polymorphisms in wheat for use as DNA markers in comparison to RFLP and RAPD markers. *Theoret. Appl. Genetics*. **94**(5): 597-602

Nagpal A and Karki M (2004) *A study on marketing opportunities for Medicinal Aromatic and Dye plants in South Asia*. Medicinal and Aromatic plants program in Asia (MAPPA), International Development Research Centre, New Delhi, India.

Naik PK, Alam MA, Singh H, Goyal V, Parida S, Kalia S and Mohapatra T (2010) *Assessment of genetic diversity through RAPD, ISSR and AFLP markers in *Podophyllum*

hexandrum: a medicinal herb from the Northwestern Himalayan region. *Physiol Mol Biol Plants*. **16**(2): 135-48

Ndiribe C, Pellissier L, Antonelli S, Dubuis A, Pottier J, Vittoz P, Guisan A and Salamin N (2013) Phylogenetic plant community structure along elevation is lineage specific. *Ecology and Evolution*: n/a-n/a

Nei M (1972) Genetic distance between populations. *Am. Nat.* **106**: 283-292

Nei M (1973) Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sci USA*. **70**: 3321-3323

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*. **89**(3): 583-90

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press.

Nei M and Li W (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences of the United States of America*. **76**: 685-686*

Newmaster S, Fazekas AJ and S R (2006) DNA barcoding in land plants: evaluation of *rbcL* in a multigene tiered approach. *Canadian Journal of Botany*. **84**: 335-341

Newmaster SG, Fazekas AJ, Steeves RA and Janovec J (2008) Testing candidate plant barcode regions in the Myristicaceae. *Molecular ecology resources*. **8**(3): 480-90

Nicolalde-Morejón F, Vergara-Silva F, González-Astorga J, Stevenson DW, Vovides AP and Sosa V (2011) A character-based approach in the Mexican cycads supports diverse multigene combinations for DNA barcoding. *Cladistics*. **27**(2): 150-164

Nitta JH (2008) Exploring the utility of three plastid loci for biocoding the filmy ferns (Hymenophyllaceae) of Moorea. *Taxon*. **57**: 725-736

Ogden R, McGough H, Cowan R, Chua L, Groves M and McEwing R (2008) SNP-based method for the genetic identification of ramin *Gonystylus* spp. timber and products: applied research meeting CITES enforcement needs. *Endangered Species Research*. **9**(3): 255-261

Ohba H and Akiyama S (1992) *The alpine flora of the Jaljale Himal, East Nepal*. University of Tokyo, Tokyo, Japan.

Ohler LD and Rose EA (1992) Optimization of long-distance PCR using a transposon-based model system. *PCR Methods Appl*. **2**(1): 51-9

Olsen CS (1998) The trade in medicinal and aromatic plants from central Nepal to northern India. *Econ. Bot.* **52**: 279-92

Olsen CS (2001) Trade in the Himalayan medicinal plant product Kutki – New data. *Medicinal Plant Conservation*. **7**: 11-13

Olsen CS (2005) Trade and conservation of Himalayan medicinal plants: *Nardostachys grandiflora* DC. and *Neopicrorhiza scrophulariiflora* (Pennell) Hong. *Biological Conservation*. **125**(4): 505-514

Pandey BL and Das PK (1989) *Immunopharmacological studies on *Picrorhiza kurroa* Royle-ex-Benth. Part IV: Cellular mechanisms of anti-inflammatory action. *Indian J Physiol Pharmacol*. **33**(1): 28-30

Pandey MR (2006) Use of medicinal plants in Traditional Tibetan therapy system in Upper Mustang, Nepal. *Our Nature*. **4**: 69-82

Peakall R and Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics*. **28**(19): 2537-2539

Peakall ROD and Smouse PE (2006) GenAlex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*. **6**(1): 288-295

Pennell FW (1943) *Scrophulariaceae Of The Western Himalayas*. Academy of Natural Sciences, Lancaster, Pennsylvania.

Pennisi E (2007) Wanted: A Barcode for Plants. *Science*. **318**(5848): 190-191

Phippen WB, Kresovich S, Candelas FG and McFerson JR (1997) Molecular characterization can quantify and partition variation among genebank holdings: a case study with phenotypically similar accessions of *Brassica oleracea* var. capitata L. (cabbage) `Golden Acre'. *Theoret. Appl. Genetics*. **94**(2): 227-234

Pillai PP, Sajan JS, Menon KM, Panicker Jayakumar KS and Subramoniam A (2012) *ISSR analysis reveals high intraspecific variation in *Rauvolfia serpentina* L. – A high value medicinal plant. *Biochemical Systematics and Ecology*. **40**(0): 192-197

Polunin O and Stainton A (1990) **Flowers of the Himalaya* Oxford University Press, Delhi.

Posada D (2003) *Using MODELTEST and PAUP* to select a model of nucleotide substitution. *Curr Protoc Bioinformatics*. **6**(6)

Posada D (2009) Selection of Models of DNA Evolution with jModelTest. In: *Bioinformatics for DNA Sequence Analysis*. (Eds.) Posada D. 537, Humana Press. 93-112.

Posada D and Buckley TR (2004) Model Selection and Model Averaging in Phylogenetics: Advantages of Akaike Information Criterion and Bayesian Approaches Over Likelihood Ratio Tests. *Systematic Biology*. **53**(5): 793-808

Potter RH and Jones MGK (1991) *Molecular analysis of genetic stability. In: Dodds JH (eds) In vitro methods for conservation of plant genetic resources*. Chapman and Hall, London.

Prakash S, Biji K, Gomez S, Murthy K and Babu R (2006) *Genetic diversity analysis of *Sorghum* (*Sorghum bicolor* L. Moench) accessions using RAPD Markers. *Indian Journal of Crop Science*. **1**: 109-112

Press JR, Shrestha KK and Sutton DA (2000) *Annotated Checklist of the Flowering Plants of Nepal*. Natural History Museum, London.

Prevost A and Wilkinson MJ (1999) *A new system of comparing PCR primers applied to ISSR fingerprinting of potato cultivars. *Theoret. Appl. Genetics*. **98**(1): 107-112

Pullin AS (2002) *Conservation biology (1st Edition)*. Cambridge University Press, Cambridge.

Puri A, Saxena RP, Guru PY, Kulshreshtha DK, Saxena KC and Dhawan BN (1992) *Immunostimulant Activity of Picroliv, the Iridoid Glycoside Fraction of *Picrorhiza kurroa*, and its Protective Action against *Leishmania donovani* Infection in Hamsters. *Planta medica*. **58**(6): 528-32

Qiu Y-X, Zhou X-W, Fu C-X and Chan Y-S (2005) A preliminary study of genetic variation in the endangered, Chinese endemic species *Dysosma versipellis* (Berberidaceae). *Bot. Bull. Acad. Sin.* **46**: 61-69

Ramanatha V and Hodgkin T (2002) *Genetic diversity and conservation and utilization of plant genetic resources. *Plant Cell Tiss Organ Culture*. **68**: 1-19

Rani P and Khullar N (2004) Antimicrobial evaluation of some medicinal plants for their anti-enteric potential against multi-drug resistant *Salmonella typhi*. *Phytother Res*. **18**(8): 670-3

Ratnaparkhe MB, Tekeoglu M and Muehlbauer FJ (1998) Inter-simple-sequence-repeat (ISSR) polymorphisms are useful for finding markers associated with disease resistance gene clusters. *Theoret. Appl. Genetics*. **97**(4): 515-519

Ratnasingham S and Hebert PD (2007) bold: The Barcode of Life Data System (http://www.barcodinglife.org). *Mol Ecol Notes*. **7**(3): 355-364

Rawat JM, Rawat B and Mehrotra S (2013) *Plant regeneration, genetic fidelity, and active ingredient content of encapsulated hairy roots of *Picrorhiza kurrooa* Royle ex Benth. *Biotechnology letters*. **35**(6): 961-8

Ray A, Chaudhuri S, Majumdar B and Bandyopadhyay S (2002) *Antioxidant activity of ethanol extract of rhizome of *Picrorhiza kurrooa* on indomethacin induced gastric ulcer during healing. *Indian Journal Clinical Biochem*. **17**(2): 44-51

Reddy MP, Sarla N and Siddiq EA (2002) Inter simple sequence repeat (ISSR) polymorphism and its application in plant breeding. *Euphytica*. **128**: 9-17

Ren BQ, Xiang XG and Chen ZD (2010) Species identification of *Alnus* (Betulaceae) using nrDNA and cpDNA genetic markers. *Molecular ecology resources*. **10**(4): 594-605

Riaz S, Sadia B, Awan FS, Khan IA and Sadaqat HA (2012) *Development of a species-specific sequence-characterized amplified region marker for roses. *Genetics and molecular research : GMR*. **11**(1): 440-447

Rogers SO and Bendich AJ (1987) Ribosomal RNA genes in plants: Variability in copy number and in the intergenic spacer. *Plant molecular biology*. **9**: 509-20

Rohlf FJ (1982) Consensus indices for comparing classifications. *Mathematical Biosciences*. **59**(1): 131-144*

Rohlf FJ (2009) *NTSYSpc: numerical Taxonomy System ver. 2.21i*. Exeter Software, Setauket, New York.

Rohlf FJ and Fisher DR (1968) *Tests for Hierarchical Structure in Random Data Sets. *Systematic Biology*. **17**(4): 407-412

Roy A, Frascaria N, MacKay J and Bousquet J (1992) Segregating random amplified polymorphic DNAs (RAPDs) in *Betula alleghaniensis*. *Theoret. Appl. Genetics*. **85**(2-3): 173-180

Sah JN and Varshney VK (2013) Chemical constituents of *Picrorhiza* genus: a review *American Journal of Essentail Oils and Natural Products*. **1**(2): 22-37

Sambrook J, Williams J, Sharp PA and Grodzicker T (1975) Physical mapping of temperature-sensitive mutations of adenoviruses. *Journal of Molecular Biology*. **97**(3): 369-390

Sanderson MJ (1989) Confidence limits on phylogenies: The bootstrap revisited. *Cladistics*. **5**(2): 113-129*

Sane SA, Shakya N and Gupta S (2011) Immunomodulatory effect of picroliv on the efficacy of paromomycin and miltefosine in combination in experimental visceral leishmaniasis. *Exp Parasitol*. **127**(2): 376-81

Sang T, Crawford D and Stuessy T (1997) Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). *American journal of botany*. **84**(8)

Sanger F, Nicklen S and Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*. **74**(12): 5463-7

Sankar AA and Moore GA (2001) Evaluation of inter-simple sequence repeat analysis for mapping in Citrus and extension of the genetic linkage map. *Theoret. Appl. Genetics*. **102**(2-3): 206-214

Sarwat M (2012) ISSR: a reliable and cost-effective technique for detection of DNA polymorphism. *Methods Mol Biol*. **862**: 103-21

Sass C, Little DP, Stevenson DW and Specht CD (2007) DNA barcoding in the cycadales: testing the potential of proposed barcoding markers for species identification of cycads. *PloS one*. **2**(11)

Schneider H and Schuettpelz E (2006) Identifying fern gametophytes using DNA sequences. *Molecular Ecology Notes*. **6**: 989-991

Sebastian VA, Cruz LD, Subramanian RB and Braganza VJ (2010) Assessment of genetic diversity within and among populations of *Tylophora rotundifolia* using RAPD markers *Gene Conserve*. **9**(37): 94-117

Semagn K, Bjørnstad Å and Ndjiondjop MN (2006) An overview of molecular marker methods for plants. *African Journal of Biotechnology*. **5**(25): 2540-2568

Shakya PR, Adhikari MK, Rajbnandari KR, Chaudhary RP and and Shrestha KK (1997) **Flora of Nepal: Country paper, presented at international seminar in flora of Nepal*. Kathmandu.

Sharma SK and Kumar N (2012) Antimicrobial Screening of *Picrorhiza kurroa* Royle Ex Benth rhizome. *International Journal of Current Pharmaceutical Review and Research*. **3**(3): 60-65

Shaw J, Lickey EB, Schilling EE and Small RL (2007) Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *American journal of botany*. **94**(3): 275-88

Shaw RK, Acharya L and Mukherjee AK (2009) *Assessment of genetic diversity in a highly valuable medicinal plant *Catharanthus roseus* using molecular markers. *Crop Breeding and Applied Biotechnology*. **9**: 52-59

Shrestha S (2001) *Molecular Systematics of Weedy Sporobolus species of Australia*. PhD Thesis submitted to The Univeristy of Queensland, Australia.

Shrestha S (2013) Molecular Marker Techniques. In: *Biological Diversity and Conservation*. (Eds.) Jha PK, Neupane FP, Shrestha ML and Khanal IP. Nepal Academy of Science and Technology. Kathmandu**:** 409-417.

Shrestha TB and Joshi RM (1996) **Rare, Endemic and Endangered Plants of Nepal*. WWF Nepal Program, Kathmandu, Nepal.

Simmons MP, Donovan Bailey C and Nixon KC (2000) Phylogeny reconstruction using duplicate genes. *Molecular biology and evolution*. **17**(4): 469-73

Simons JM, t Hart BA, Ip Vai Ching TRAM, Van Dijk H and Labadie RP (1990) *Metabolic activation of natural phenols into selective oxidative burst agonists by activated human neutrophils. *Free Radical Biology and Medicine*. **8**(3): 251-258

Singh G, Bani S, Singh S, Khjuria A, Sharma L, Gupta B and Banerjee S (1993) *Anti-inflammatory activity of the Iridoids Kutkin, Picroside-I and Kutkoside from *Picrorhiza kurrooa*. *Phytotherapy Research*. **7**: 402-407

Singh N, Gupta A, Singh B and Kaul V (2005) *Quantification of Picroside-I and Picroside-II in *Picrorhiza kurrooa* by HPTLC. *Journal of Liquid Chromatography & Related Technologies*. **28**: 1679-1691

Singh S, Panda MK and Nayak S (2012) *Evaluation of genetic diversity in turmeric (*Curcuma longa* L.) using RAPD and ISSR markers. *Industrial Crops and Products*. **37**(1): 284-291

Singh V and Banyal HS (2011) Antimalarial Effects of *Picrorhiza kurrooa* Royle Ex Benth Extracts on *Plasmodium berghei*. *Asian J. Exp. Biol. Sci.* **2**(3): 529-532

Slatkin M (1987) *Gene Flow and Geographic Structure of Natural Populations. *Science*. **236**: 787-792

Smit H, Berg A, Kroes B, Beukelman C, Ufford H, Quarles V, Dijk H and Labadie R (2000) *Inhibition of T-lymphocyte proliferation by cucurbitacines from *Picrorhiza scrophulariiflora*. *Journal of Natural Products*. **63**(1300-1302)

Smit HF (1968) *Picrorhiza scrophulariiflora, from traditional use to immunomodulatory activity*. PhD Thesis submitted to Universiteit Utrecht, Utrecht.

Smit HF, Kroes BH, van den Berg AJJ, van der Wal D, van den Worm E, Beukelman CJ, van Dijk H and Labadie RP (2000) *Immunomodulatory and anti-inflammatory activity of *Picrorhiza scrophulariiflora*. *Journal of Ethnopharmacology*. **73**(1–2): 101-109

Smith JF, Kress WJ and Zimmer EA (1993) Phylogenetic Analysis of the Zingiberales based on *rbcL s*equences. *Annals of the Missouri Botanical Garden*. **80**: 620-30

Sneath P and Sokal R (1973) *Taxonomy, the principle and practice of numerical classification*. W.H freeman and company, San Francisco, CA.

Sokal R and Michener C (1958) *A statistical method for evaluating systematic relationships. *University of Kansas Science Bulletin*. **38**: 1409-1438

Sokal RR and Rohlf FJ (1962) *The comparison of dendrograms by objective methods. *Taxon*. **11**: 33-40

Srirama R, Senthilkumar U, Sreejayan N, Ravikanth G, Gurumurthy BR, Shivanna MB, Sanjappa M, Ganeshaiah KN and Shaanker RU (2010) Assessing species admixtures in raw drug trade of *Phyllanthus*, a hepato-protective plant using molecular tools. *J Ethnopharmacol*. **130**(2): 208-15

Staub J, Serquen F and Gupta M (1996) Genetic markers, map construction, and their application in plant breeding. *Hort Sci*. **31**: 729-741

Staudacher K, Wallinger C, Schallhart N and Traugott M (2009) DNA barcoding insect-host plant associations. *Proceedings of the Royal Society, series B*. **276**: 639-48

Stech M, Kolvoort E, Loonen MJ, Vrieling K and Kruijer JD (2011) *Bryophyte DNA sequences from faeces of an arctic herbivore, barnacle goose (*Branta leucopsis*). *Molecular ecology resources*. **11**(2): 404-8

Steele KA, Price AH, Witcombe JR, Shrestha R, Singh BN, Gibbons JM and Virk DS (2013) *QTLs associated with root traits increase yield in upland rice when transferred through marker-assisted selection. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*. **126**(1): 101-8

Steinegger E and Hänsel R (1988) **Lehrbuch der Pharmakognosie und Phytopharmazie*. Spriner Verlag, Berlin.

Steven GN and Subramanyam R (2009) Testing plant barcoding in a sister species complex of pantropical Acacia (Mimosoideae, Fabaceae). *Molecular ecology resources*. **9**: 172-80

Stewart C and Excoffier L (1996) Assessing population genetic structure and variability with RAPD data application to *Vaccinium macrocarpon* (American Cranberry). *Journal of evolutionary biology*. **9**: 153-171

Stolk J, Hiltermann TJ, Dijkman JH and Verhoeven AJ (1994) *Characteristics of the inhibition of NADPH oxidase activation in neutrophils by apocynin, a methoxy-substituted catechol. *American Journal of Respiratory Cell and Molecular Biology*. **11**(1): 95-102

Stolk J, Rossie W and Dijkman JH (1994) *Apocynin improves the efficacy of secretory leukocyte protease inhibitor in experimental emphysema. *American Journal of Respiratory and Critical Care Medicine*. **150**(6): 1628-1631

Stuppner H, Kahling H, Seligmann O and Wagner H (1990) *Minor cucurbitacin glycosides from *Picrorhiza kurrooa*. *Phytochemistry*. **29**(5): 1633-37

Stuppner H and Wagner H (1989a) *Minor Iridoid and Phenol Glycosides of *Picrorhiza kurrooa*. *Planta medica*. **55**(5): 467-9

Stuppner H and Wagner H (1989b) *New cucurbitacin glycosides from *Picrorhiza kurrooa*. *Planta medica*. **55**(6): 559-63

Sturm S and Stuppner H (2001) Analysis of iridoid glycosides from *Picrorhiza kurroa* by capillary electrophoresis and high performance liquid chromatography-mass spectrometry. *Chromatographia*. **53**(11-12): 612-618

Suh JP, Jeung JU, Lee JI, Choi YH, Yea JD, Virk PS, Mackill DJ and Jena KK (2010) *Identification and analysis of QTLs controlling cold tolerance at the reproductive stage and validation of effective QTLs in cold-tolerant genotypes of rice (*Oryza sativa* L.). *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*. **120**(5): 985-95

Sullivan J and Joyce P (2005) *Model Selection In Phylogenetics. *Annual Review of Ecology, Evolution, and Systematics*. **36**(1): 445-466

Suwal PN (1993) *Medicinal plants of Nepal*. His Majesty's Government of Nepal, Ministry of Forests and Soil Conservation, Department of Medcinal Plants, Kathmandu, Nepal.

Tamhankar S, Ghate V, Raut A and Rajput B (2009) *Molecular profiling of "Chirayat" complex using Inter Simple Sequence Repeat (ISSR) markers. *Planta medica*. **75**(11): 1266-70

Tamura K, Peterson D, Peterson N, Stecher G, Nei M and Kumar S (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular biology and evolution*. **28**: 2731-39

Tate JA and Simpson BB (2003) Paraphyly of *Tarasa* (Malvaceae) and Diverse Origins of the Polyploid Species. *Systematic Botany*. **28**(4): 723-737

Tavaré S (1986) *Some Probabilistic and Statistical Problems in the Analysis of DNA Sequences. In: *American Mathematical Society: Lectures on Mathematics in the Life Sciences*. (Eds.). 17, Amer Mathematical Society. 57-86.

Teklewood A and Becker H (2006) Geographic pattern of genetic diversity among 43 Ethiopian mustard (*Brassica carinata* A. Braun) accession as revealed by RAPD analysis. *Genet Res Crop Evol.* **53**: 1173-1185

Thompson JD, Higgins DG and Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. **22**(22): 4673-80

Tilman D (2000) Causes, Consequences and Ethics of Biodiversity. *Nature*. **405**: 208-211

Transue D, Fairbanks D, Robison L and Anderson W (1994) Species Identification by RAPD analysis of grain amaranth genetic resources. *Crop Science*. **34**: 1385-89

Tsuda Y, Goto S and Y. I (2004) RAPD analysis of genetic variation within and among four natural populations of *Betula maximowicziana*. *Silvae Genetica*. **53**: 234-239

Tsumura Y, Ohba K and Strauss SH (1996) Diversity and inheritance of inter-simple sequence repeat polymorphisms in Douglas-fir (*Pseudotsuga menziesii*) and Sugi (*Cryptomeria japonica*). *Theoret. Appl. Genetics*. **92**(1): 40-45

Usman MRM, Yamgar S, Gadgoli C and Salunkhe D (2012) Preliminary screening and antimicrobial activity of *Picrorhiza kurroa* royle ethanolic extracts. *Int. J. Pharm. Sci. Rev. Res.* **14**(1): 73-76

Vaidya AB, Antarkar DS, Doshi JC, Bhatt AD, Ramesh V, Vora PV, Perissond D, Baxi AJ and Kale PM (1996) *Picrorhiza kurroa* (Kutaki) Royle ex Benth as a hepatoprotective agent - experimental & clinical studies. *Journal of postgraduate medicine*. **42**(4): 105-8

Valentini A, Miquel C, Nawaz MA, Bellemain E, Coissac E, Pompanon F, Gielly L, Cruaud C, Nascetti G, Wincker P, Swenson JE and Taberlet P (2009) New perspectives in diet analysis based on DNA barcoding and parallel pyrosequencing: the *trnL* approach. *Molecular ecology resources*. **9**(1): 51-60

Verma PC, Basu V, Gupta V, Saxena G and Rahman LU (2009) Pharmacology and chemistry of a potent hepatoprotective compound Picroliv isolated from the roots and rhizomes of *Picrorhiza kurroa* Royle ex. Benth. (Kutki). *Current pharmaceutical biotechnology*. **10**(6): 641-9

Visen PK, Saraswat B and Dhawan BN (1998) Curative effect of picroliv on primary cultured rat hepatocytes against different hepatotoxins: an in vitro study. *J Pharmacol Toxicol Methods*. **40**(3): 173-9

Wang D, He Z, Feng B and Yang C (1993) *Chemical constituents from *Picrorhiza scrophulariiflora*. *Acta Botanica Yunnanica*. **15**(1): 83-88

Wang G, Mahalingam R and Knap HT (1998) (C-A) and (G-A) anchored simple sequence repeats (ASSRs) generated polymorphism in soybean, Glycine max (L.) Merr. *Theoret. Appl. Genetics*. **96**(8): 1086-1096

Wang H, Sun Y, Ye WC, Xiong F, Wu JJ, Yang CH and Zhao SX (2004) *Antioxidative phenylethanoid and phenolic glycosides from *Picrorhiza scrophulariiflora*. *Chemical & pharmaceutical bulletin*. **52**(5): 615-7

Wang H, Wu FH, Xiong F, Wu JJ, Zhang LY, Ye WC, Li P and Zhao SX (2006) Iridoids from *Neopicrorhiza scrophulariiflora* and their hepatoprotective activities *in vitro*. *Chem Pharm Bull*. **54**(8): 1144-9

Wang H, Ye WC, Jiang RW, Wu JJ, Mak TC, Zhao SX and Yao XS (2004) *Three new cyclopentanoid monoterpenes from *Picrorhiza scrophulariiflora*. *Planta medica*. **70**(4): 382-4

Wang H, Ye WC, Xiong F and Zhao SX (2004) *Phenylethanoid glycosides from root of *Picrorhiza scrophulariiflora*. *Zhongguo Zhong yao za zhi = Zhongguo zhongyao zazhi = China journal of Chinese materia medica*. **29**(6): 531-4

Wang X, Yang R, Feng S, Hou X, Zhang Y, Li Y and Ren Y (2012a) Genetic variation in *Rheum palmatum* and *Rheum tanguticum* (Polygonaceae), two medicinally and endemic species in China using ISSR markers. *PloS one*. **7**(12): e51667

Wang XM, Hou XQ, Zhang YQ, Yang R, Feng SF, Li Y and Ren Y (2012b) *Genetic Diversity of the Endemic and Medicinally Important Plant *Rheum officinale* as Revealed by Inter-Simpe Sequence Repeat (ISSR) Markers. *International journal of molecular sciences*. **13**(3): 3900-15

Weber D and Helentjaris T (1989) Mapping RFLP loci in maize using B-A translocations. *Genetics*. **121**(3): 583-90

Weinges K, Kloss P and Henkels W (1972) *Natural Products from Medicinal Plants. XVII. Picroside-II, A New 6-Vanilloylcatapol from *Picrorhiza kurrooa* Royle ex. Benth. *Justus Liebigs Ann. Chem*. **759**(173-182)

White T, Bruns T, Lee S and Taylor J (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: *PCR Protocols: A Guide to Methods and Applications*. (Eds.) Innis M, Gelfand D, Shinsky J and White T. Academic Press. 315-322.

WHO. (2013) Traditional Medicine [Online] <http://www.who.int/topics/traditional_medicine/en/>

Wiemers M and Fiedler K (2007) Does the DNA barcoding gap exist? - a case study in blue butterflies (Lepidoptera: Lycaenidae). *Front Zool*. **4**: 8

Wolfe K (1991) Protein-coding genes in chloroplast DNA compilation of nucleotide sequences, database entries, and rates of molecular evolution In: *Cell culture and somatic cell genetics of plants*. (Eds.) Bogorad L and Vasil I. 7B, Academic Press. SanDiego**:** 467-482.

Wolfe KH, Li WH and Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences of the United States of America*. **84**(24): 9054-8

Wu KS, Jones R, Danneberger L and Scolnik PA (1994) Detection of microsatellite polymorphisms without cloning. *Nucleic Acids Res*. **22**(15): 3257-8

Xiao M, Li Q, Wang L, Guo L, Li J, Tang L and Chen F (2006) *ISSR analysis of the genetic diversity of the endangered species *Sinopodophyllum hexandrum* (Royle) Ying from Western Sichuan Province, China. *Journal of Integrative Plant Biology*. **48**(10): 1140-46

Xie Z, Hu H, Kong D and Yang P (2007) *Two new compounds from *Picrorhiza scrophulariiflora*. *Chinese Journal of Pharmaceutical*. **38**(3): 221-226

Xiong J (2006) *Essential Bioinformatics*. Cambridge University Press, New York, USA.

Yang LC, Deng H, Yi Y, Zhang XM, Wang YZ and Lin JQ (2010) *Identification of medical *Dendrobium* herbs by ISSR marker. *Zhong yao cai = Zhongyaocai = Journal of Chinese medicinal materials*. **33**(12): 1841-4

Yeh FC, Yang RC, Boyle T, Ye ZH and Mao JX (1997) POPGENE: the user-friendly shareware for population genetic analysis. Molecular Biology and Biotechnology Centre, University of Alberta, Canada

Young ND, Menancio-Hautea D, Fatokun CA and Danesh D (1992) RFLP technology, crop improvement, and international agriculture. In: *Biotechnology: Enhancing research on tropical crops in Africa*. (Eds.) Thottappilly G, Monti L, Moham D and Moore A. Technical Center for Agriculture and Rural Cooperation, International Institute of Tropical Agriculture.**:** 221-230.

Yu J, Xue J-H and Zhou S-L (2011) New universal *matK* primers for DNA barcoding angiosperms. *Journal of Systematics and Evolution*. **49**(3): 176-181

Yu KF, A.V. D and Pauls KP (1993) Random amplified polymorphic DNA (RAPD) analysis. In: *Methods in plant molecular biology and biotechnology*. (Eds.) Glick B and Thompson J. CRC Press, Boca Raton-Ann Arbor-London-Tokyo. 380.

Zhang D, Chen S and Gao Q (2007) *Patterns of genetic variation in *Swertia przewalskii*, an endangered endemic species of the Qinghai-Tibet Plateau. *Biochemical genetics*. **45**(1-2): 33-50

Zhang F, Lv Y, Dong H and Guo S (2010) *Analysis of genetic stability through intersimple sequence repeats molecular markers in micropropagated plantlets of *Anoectochilus formosanus* Hayata, a medicinal plant. *Biological & pharmaceutical bulletin*. **33**(3): 384-8

Zhang HL, Hsia KC and Lou TC (1965) Botanical and pharmacognostical studies of the Chinese drug hu-huang-lien, *Rhizome picrorrhizae*. *Yao xue xue bao = Acta pharmaceutica Sinica*. **12**(12): 808-16

Zhou YQ, Jing JZ, Li ZY, Hao J, Jia JF, Zhang BH and Hao JG (2005) Genetic diversity of yam (*Dioscorea opposita* Thunb) detected by ISSR markers. *Shi yan sheng wu xue bao*. **38**(4): 324-30

Zhu T, Huang K, Deng X, Zhang Y, Xiang H, Gao H and Wang D (2008) Three New Caffeoyl Glycosides from the roots of *Picrorhiza Scrophulariiflora*. *Molecules*. **13**(4): 729-735

Zietkiewicz E, Rafalski A and Labuda D (1994) Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. *Genomics*. **20**(2): 176-83

Zou LC, Zhu TF, Xiang H, Yu L, Yan ZH, Gan SC, Wang DC, Zeng S and Deng XM (2008) New secoiridoid glycosides from the roots of *Picrorhiza scrophulariiflora*. *Molecules*. **13**(9): 2049-57

Zou X, Liao X, Ding L and Peng S (2007) *Phenyl and phenylethyl glycosides from *Picrorhiza scrophulariiflora*. *Journal of Asian Natural Products*. **9**(3-5): 443-448

Zurawski G and Clegg MT (1984) The barley chloroplast DNA *atpBE*, *trnM2*, and *trnV1* loci. *Nucleic Acids Res*. **12**(5): 2549-59

Note: References marked with an asterisk (*) are abstract of original references cited.

APPENDICES

APPENDIX I

Preparation of Reagents and Extraction Buffers

1. Preparation of 50X TAE (Tris-Acetate-EDTA) Stock buffer

To prepare 50 X TAE stock buffers, Tris base 242 gm (Qualigens Fine Chemicals, Mumbai) was dissolved in approximately 750 ml deionized water. To this solution, 57.1 ml glacial acetic acid (Qualigens Fine Chemicals, Mumbai) was added followed by 100 ml of 0.5 M EDTA pH 8.0 (Promega Corporation, USA). The final volume was made to 1L. This stock solution was stored at Room Temperature.

2. Preparation of 1X TAE from 50X Stock solution

20 ml of stock solution was diluted using 980 ml of double distilled water to make working 1X TAE solution. The final solute concentration of Tris-acetate at this stage was 40 mM while that of EDTA was 1mM.

3. Preparation of 5X Gel Loading Buffer (GLB)

2.5 gm Sucrose was dissolved in 7 ml deionized water. 25 mg Bromophenol Blue (Fermentas Life Sciences, Canada) was added to the solution and final volume of 10 ml was made. Thus prepared buffer was used in 1:1 ratio (by volume) with DNA sample while 1:4 ratio (by volume) with PCR product during electrophoresis.

4. Preparation of 1.5 % Agarose Gel

1.5 gm Agarose (Promega, Product of Spain) was dissolved in 100 ml 1X TAE buffer in microwave oven. When the solution cooled down to approximately 55^{o}C, 5 µl of Ethidium Bromide solution was added. The solution was poured onto gel casting tray with appropriate comb (12-25 toothed) for gel loading well.

5. Preparation of Tris Buffer (1M, pH 8.0 and pH 7.5)

60.55 gm (Qualigens Fine Chemicals, Mumbai) Tris Base was added to 400 ml of deionized water. The final pH was adjusted to either 8.0 or 7.5, as needed, by adding concentrated HCl. The final volume of 500 ml was then made, autoclaved and stored at Room Temperature until needed.

6. Preparation of EDTA (0.5M, pH 8.0)

93.05 gm EDTA (Disodium Ethylene diamine tetra acetate.2H_2O) was added to 400 ml deionized water in Schott bottle and mixed on a magnetic stirrer. The pH was adjusted to 8.0 by adding sodium hydroxide pellets (10 gm approx.). The final volume was adjusted to 500 ml using deionized water and autoclaved and stored at RT until needed.

7. Preparation of 4M NaCl

117 gm NaCl (Qualigens Fine Chemicals, Mumbai) was dissolved in 400 ml deionized water in Schott bottle using magnetic stirrer. The final volume was adjusted to 500 ml with deionized water. The solution was autoclaved and stored at room temperature.

8. Preparation of Graham's CTAB (Hexadecyl Trimethyl Ammonium Bromide) extraction buffer (2% CTAB, 1.4M NaCl, 0.1M EDTA, 0.1M Tris HCl, pH 8.0) (Graham et. al. 1994)

Tris Buffer (100ml of 1M solution, pH 8.0), EDTA (200 ml of 0.5M solution, pH 8.0), NaCl (350 ml of 4M solution) and 20 gm CTAB (Loba Chemie, India) were dispensed and mixed in a sterile 1L Schott bottle. The final volume was adjusted to 1L with deionized water.

9. Preparation of Doyle and Doyle extraction buffer (100mM Tris Hcl, 20 mM Na-EDTA, 1.4 M NaCl, 2% CTAB, 0.2% β – mercaptoethanol, 1.5% polyvinyl pyrrolidone, Doyle and Doyle, 1990)

Tris Buffer (1ml of 1M stock, pH 8.0), Na-EDTA (0.4 ml of 0.5M stock), NaCl (3.5 ml of 4M stock), 0.2 gm CTAB (Loba Chemie, India), PVP (1.5 ml) were dispensed and mixed in a bottle. The final volume was adjusted to 10 mL with deionized water. β – mercaptoethanol (20 µl) was added to freshly prepared buffer inside a fumehood.

10. Preparation of TE buffer (Tris-EDTA; 10 mM Tris HCl, 1 mM EDTA) with RNAse

EDTA (1 ml of 0.5M stock, pH 8.0) and Tris buffer (5 ml of 1M stock) were mixed and dissolved and the final volume was adjusted to 500 ml. This solution was autoclaved and stored at room temperature RNAse A (200 µl of 5 mg/ml) was added to the 200 ml of buffer so that the final concentration is 10 µg/ml for fresh use.

11. Primers, dNTPs and DNA dilution

Primers, available in lyophilized form, were diluted to required working concentration of 10 µM using sterile distilled water.Commercially supplied dNTP mix (10 mM each, Fermentas Life Sciences) was used in PCR reaction which was stored at -20^{o}C until used.

Initial concentration of extracted DNA solution was estimated using Biophotometer (Eppendorf, Germany) and as such the working concentration of 25 ng was prepared by diluting the initial concentration. The working stock solution of DNA was stored at -20^{o}C until used.

APPENDIX II

Preparation of Media and Reagents

1. Preparation of Nutrient Broth

22.5 grams of Nutrient Broth powder (MAST Group Ltd., Merseyside, UK) was dissolved in 150 ml of distilled water in a conical flask. The broth solution was autoclaved at 121^{o}C at 15 psi after aliquoting 10 ml each in labeled test tubes.

2. Preparation of Mueller Hilton Agar (MHA)

38.0 grams of Mueller Hilton Agar powder (MAST Group Ltd., Merseyside, UK) was weighed and dissolved in 1000 ml of distilled water. Around 20 ml was poured into sterile petriplates in triplicates for each organism.

3. Preparation of 0.5 McFarland solutions

First, a 1% (w/v) solution of anhydrous Barium chloride ($BaCl_2$) was prepared. Similarly, a 1% (v/v) solution of sulfuric acid (H_2SO_4) was also prepared. These solutions were mixed in proportions in a test tube as described in the table below to prepare desired MacFarland Turbidity Standard Solution.

Table 3.10 Macfarland scale and respective concentrations of $BaCl_2$ and H_2SO_4

MacFarland Scale Number	Amount of 1% $BaCl_2$ (mL)	Amount of 1% H_2SO_4 (mL)
0.5	0.05	9.95
1	0.1	9.9
2	0.2	9.8
3	0.3	9.7
4	0.4	9.6
5	0.5	9.5
6	0.6	9.4
7	0.7	9.3
8	0.8	9.2
9	0.9	9.1
10	1.0	9.0

4. Preparation of Extract Solution

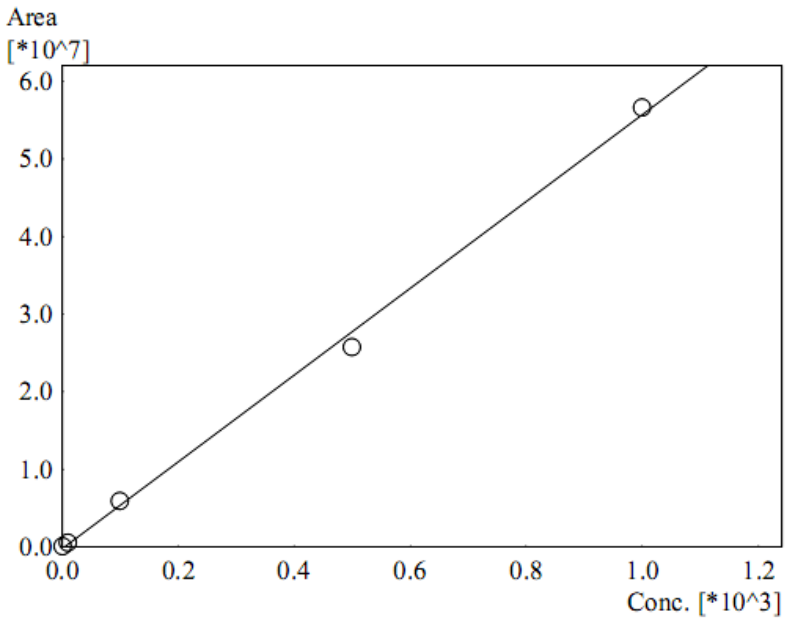
50 mg each of the methanolic extract from Manaslu Conservation Area (MCA), Langtang National Park (LNP), and Kanchanjunga Conservation Area (KCA) was weighed and dissolved in 1 ml Methanol to prepare a 50 mg/ml extract solution in eppendorf tubes (Eppendorf, Germany).

APPENDIX IV

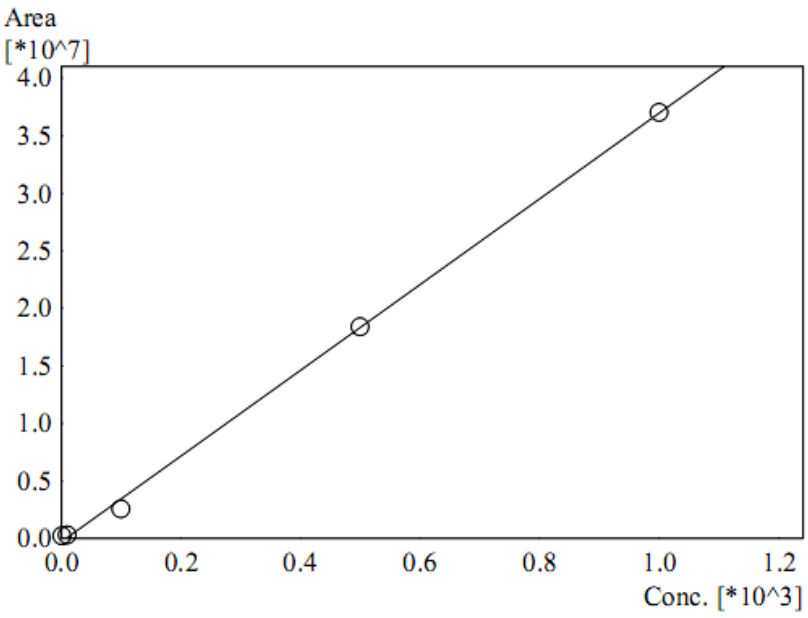
(Binary Matrix of ISSR profiles, Simple Matching Coefficient Similarity Matrix, Jaccard Coefficient Similarity matrix, Dice Coefficient Similarity Matrix)

APPENDIX V

Calibration curve of Picroside I using various concentrations (1 µg/ml to 1000 µg/ml)



Calibration curve of Picroside II using various concentrations (1 µg/ml to 1000 µg/ml)



APPENDIX VI

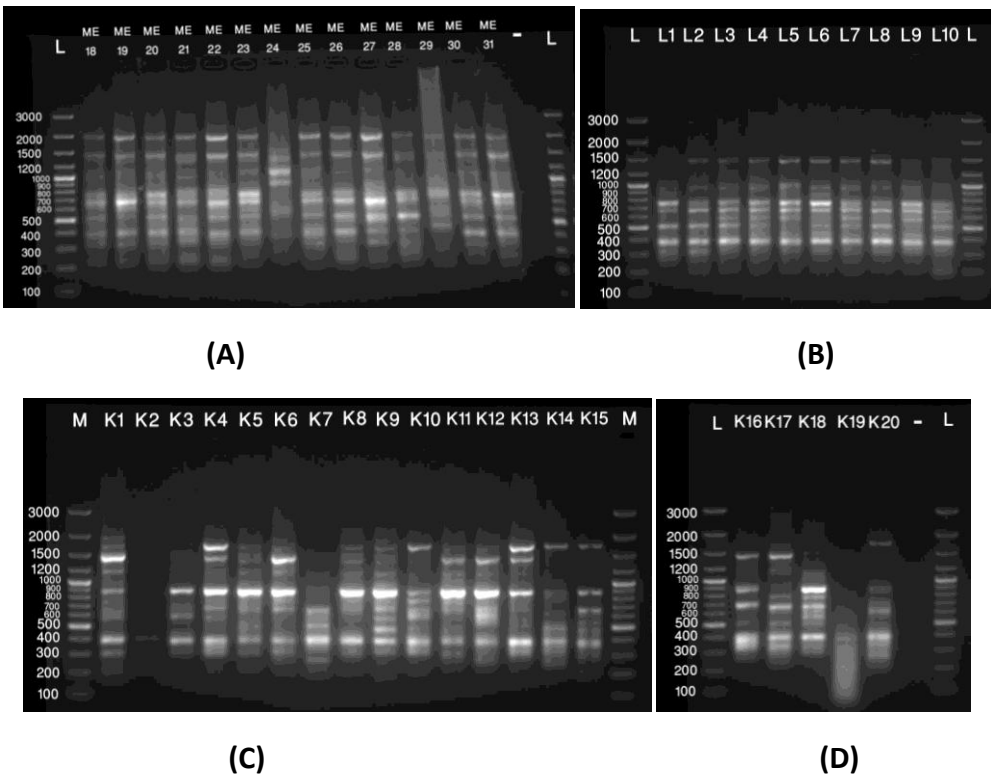


Plate 1 ISSR profile amplification using UBC 834. Lanes marked with M are 100 bp+ DNA ladder. (A) ISSR profile of accessions from MCA; (B) ISSR profile of accessions from LNP; (C) ISSR profile of accessions from KCA (K1-K15) and (D) accessions from K16-K20.

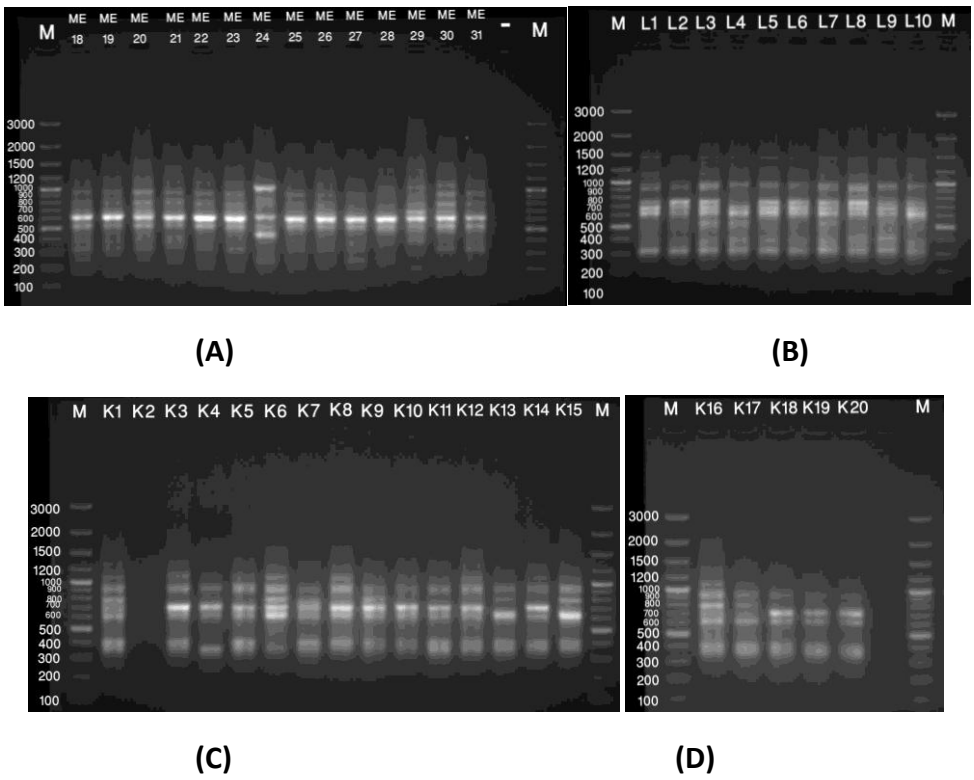


Plate 2 ISSR profile amplification using UBC 835. Lanes marked with M are 100 bp+ DNA marker. (A) ISSR profile of accessions from MCA; (B) ISSR profile of accessions from LNP; (C) ISSR profile of accessions from KCA (K1-K15) and (D) accessions from K16-K20.

APPENDIX VII – Photo Display

[A]

[B]

[C]

[D]

[E]

Photo: [A], [B], [C] and [D] *N. scrophulariiflora* plant as found in natural habitat. Note the rhizome, leaves, flower and fruit. [E] Research team, Molecular Biotechnology Lab, NAST.

Appendix 1

Binary Matrix Scores of ISSR profiles

	ME 18	ME 19	ME 20	ME 21	ME 22	ME 23	ME 25	ME 26	ME 27	ME 28	ME 29	ME 30	ME 31	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11	K12	K13	K14	K15	K16	K17	K18	K19	K20
1300	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1250	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	9	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1190	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1000	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	9	0	0	0	0	0	0	0	0	0	0	1	0	0	0
950	1	1	1	1	1	1	1	0	1	1	0	1	1	1	0	0	1	0	0	0	1	0	0	0	9	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0
900	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0
800	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	9	1	0	1	0	0	0	0	0	0	0	0	1	0	0
710	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0
700	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	9	0	0	0	9	0	0	0	1	0	0	0	0	0	0	0	0	0	0
650	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	1	0
610	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0
600	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	9	0	0	0	9	0	0	1	0	0	0	0	1	0	0	0	1	0	0
550	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	9	0	0	0	0	0	0	0	1	0	0	0	0	0	0
510	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	1	1	1	1	1	1	1	1	0	9	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0
500	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	9	0	0	1	9	1	1	1	0	1	1	1	1	1	1	1	1	0	0
490	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0
450	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0
400	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	1	0	0	9	0	0	0	1	1	0	0	0	0	0	0	0	0	0
390	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	1	9	0	0	0	0	0	0	0	0	1	0	0	0	0	0
350	0	0	0	1	0	0	0	0	0	0	1	0	0	1	1	1	1	1	1	1	1	1	1	0	9	0	0	0	9	0	1	0	0	0	0	0	0	0	0	0	0	1	0
300	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	9	1	0	0	9	0	1	0	0	0	0	0	0	0	0	0	0	1	0
290	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	1	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0
200	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	9	1	1	0	9	0	0	1	0	0	0	1	1	0	1	0	0	0	0
100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	1	1	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1700	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	1	0	1	0	0	1	1	1	0	1	1	1	0	0	0
1600	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	9	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1400	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
1200	0	0	1	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1190	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1
1100	0	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1050	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	9	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1010	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
900	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	0	1	0	9	1	0	0	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1
890	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
850	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	9	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
800	0	0	1	1	1	1	1	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
750	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
700	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
650	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
600	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
590	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	9	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
550	1	0	1	1	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
500	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
450	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
410	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
400	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
390	1	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
300	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
290	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0
1900	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0
1500	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	0	0
1450	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1300	0	0	1	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1200	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	0	1	1	0	1	0	1	0	0	1	1	0	0	0
1000	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	0	1	0	1	1	1	1	1	1	0	1	1	0	0
850	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	1	1	1	0	0	1
800	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0	1	0	1	0	0	0	0	0	1	0
900	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
750	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
700	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	1	1	1	1	1	1	0	0	1
650	0	0	0	1	0	1	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	1
550	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	1	0
500	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1	0	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
400	1	1	1	1	1	0	0	1	1	1	1	1	1	1	0	0	0	1	1	0	0	1	1	1	0	1	1	1	1	0	1	1	1	1	1	0	1	1	1	0	0	0	0
300	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
200	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
250	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2000	1	1	1	1	1	1	1	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0
1750	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1700	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	1	0	1	1	1	1	1	1	1	1	0	0	0	0	0
1600	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1500	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	0	1	1	0	0	1	1	1	0	1	1	0	1	1	1	0	0	1	1	0	0	0
1350	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1200	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	1	1	1	0	1	1	0	0	0	0	1	0	0	0
1000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
900	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	0	0	0	0	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	0	1
950	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
800	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
750	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0
700	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	0	1	1	0	1	0	0	1	1	1	1	0	0
600	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0	1	1	1	0	1	1	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0
500	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	0	1	0	0	0	0	0	0	1	0	1	1	1	0	0	0	0	0	1	1	0	0
400	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1
300	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	0	0	1	1	0	1	1	1	1	0	1	1	1	1	0	1	1	1	1	0	1	1	1	0	1	0	0	1
200	1	1	1	1	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1500	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	1	1	1	1	0	9	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	9
1300	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	9	1	0	1	1	0	0	0	0	0	1	0	1	1	0	0	0	0	9
1200	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	1	1	0	1	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9
1000	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9
950	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9
900	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	9	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	9
800	1	1	1	1	1	1	0	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	1	9	0	0	0	1	0	0	1	0	0	0	1	0	1	0	0	0	0	9
750	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	1	1	1	1	1	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9
700	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	9	1	1	1	1	1	1	1	1	1	1	0	1	0	0	0	0	0	9
600	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	1	9	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	9
500	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9
400	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	9	1	0	1	0	1	1	0	0	1	1	1	1	1	1	1	1	1	9
450	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9

350	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	1	1	0	1	1	0	0	9	9
300	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	9
200	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	9
100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	9

Appendix 2

Simple Matching Coefficient Similarity Matrix

	ME18	ME19	ME20	ME21	ME22	ME23	ME25	ME26	ME27	ME28	ME29	ME30	ME31	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11	K12	K13	K14	K15	K16	K17	K18	K19	K20
ME18	1.00																																										
ME19	0.94	1.00																																									
ME20	0.90	0.93	1.00																																								
ME21	0.90	0.91	0.90	1.00																																							
ME22	0.93	0.96	0.95	0.93	1.00																																						
ME23	0.91	0.94	0.92	0.90	0.94	1.00																																					
ME25	0.90	0.92	0.91	0.89	0.93	0.92	1.00																																				
ME26	0.91	0.92	0.90	0.88	0.93	0.91	0.95	1.00																																			
ME27	0.89	0.92	0.89	0.88	0.91	0.88	0.92	0.91	1.00																																		
ME28	0.88	0.89	0.86	0.85	0.89	0.88	0.89	0.90	0.87	1.00																																	
ME29	0.81	0.82	0.79	0.80	0.83	0.80	0.79	0.81	0.77	0.84	1.00																																
ME30	0.90	0.91	0.88	0.87	0.91	0.89	0.90	0.89	0.86	0.89	0.81	1.00																															
ME31	0.89	0.90	0.89	0.86	0.90	0.88	0.89	0.88	0.85	0.88	0.80	0.99	1.00																														
L1	0.67	0.69	0.65	0.67	0.68	0.69	0.69	0.68	0.66	0.72	0.68	0.68	0.68	1.00																													
L2	0.66	0.69	0.68	0.67	0.69	0.69	0.70	0.69	0.67	0.69	0.65	0.66	0.66	0.90	1.00																												
L3	0.65	0.69	0.66	0.65	0.67	0.69	0.69	0.68	0.64	0.67	0.66	0.67	0.67	0.86	0.90	1.00																											
L4	0.65	0.68	0.65	0.66	0.66	0.69	0.70	0.68	0.66	0.68	0.64	0.66	0.66	0.88	0.90	0.92	1.00																										
L5	0.62	0.66	0.64	0.63	0.65	0.66	0.67	0.66	0.63	0.66	0.64	0.65	0.65	0.89	0.91	0.90	0.91	1.00																									
L6	0.68	0.70	0.67	0.68	0.68	0.69	0.69	0.69	0.66	0.71	0.67	0.68	0.68	0.90	0.92	0.90	0.88	0.91	1.00																								
L7	0.62	0.66	0.64	0.64	0.65	0.68	0.67	0.67	0.63	0.66	0.64	0.64	0.64	0.88	0.95	0.92	0.92	0.93	0.93	1.00																							
L8	0.66	0.67	0.65	0.66	0.66	0.68	0.70	0.68	0.66	0.69	0.62	0.66	0.66	0.90	0.92	0.90	0.92	0.92	0.90	0.91	1.00																						
L9	0.69	0.72	0.68	0.70	0.70	0.71	0.71	0.71	0.69	0.70	0.69	0.69	0.69	0.89	0.90	0.89	0.90	0.91	0.92	0.92	0.91	1.00																					
L10	0.68	0.70	0.67	0.65	0.68	0.68	0.68	0.68	0.66	0.69	0.65	0.66	0.66	0.85	0.88	0.86	0.85	0.87	0.87	0.87	0.85	0.89	1.00																				
K1	0.72	0.75	0.72	0.71	0.73	0.74	0.71	0.72	0.70	0.73	0.72	0.74	0.74	0.65	0.64	0.63	0.62	0.65	0.66	0.63	0.62	0.69	0.65	1.00																			
K2	0.79	0.77	0.73	0.78	0.78	0.75	0.76	0.78	0.76	0.80	0.84	0.78	0.76	0.69	0.64	0.67	0.65	0.63	0.69	0.64	0.68	0.73	0.66	0.78	1.00																		
K3	0.69	0.72	0.71	0.67	0.71	0.70	0.70	0.70	0.68	0.72	0.71	0.72	0.73	0.69	0.66	0.66	0.63	0.69	0.68	0.65	0.64	0.70	0.68	0.89	0.77	1.00																	
K4	0.69	0.72	0.71	0.68	0.71	0.71	0.71	0.70	0.70	0.73	0.69	0.72	0.72	0.65	0.61	0.63	0.62	0.64	0.64	0.61	0.60	0.66	0.62	0.89	0.74	0.86	1.00																
K5	0.71	0.73	0.71	0.69	0.73	0.72	0.72	0.73	0.70	0.74	0.71	0.74	0.73	0.65	0.64	0.63	0.61	0.65	0.67	0.63	0.62	0.68	0.64	0.88	0.78	0.87	0.89	1.00															
K6	0.71	0.73	0.70	0.71	0.72	0.71	0.69	0.70	0.70	0.72	0.73	0.73	0.73	0.63	0.58	0.58	0.58	0.60	0.62	0.56	0.57	0.63	0.61	0.85	0.75	0.82	0.80	0.82	1.00														
K7	0.77	0.79	0.76	0.76	0.79	0.77	0.78	0.79	0.76	0.80	0.79	0.79	0.78	0.68	0.68	0.67	0.66	0.66	0.69	0.68	0.66	0.70	0.66	0.82	0.87	0.83	0.80	0.85	0.82	1.00													
K8	0.71	0.73	0.73	0.70	0.73	0.72	0.71	0.72	0.70	0.74	0.72	0.74	0.74	0.63	0.62	0.61	0.59	0.62	0.65	0.60	0.60	0.66	0.64	0.85	0.77	0.84	0.82	0.87	0.88	0.84	1.00												
K9	0.75	0.77	0.74	0.73	0.75	0.75	0.72	0.73	0.72	0.75	0.71	0.76	0.75	0.65	0.62	0.62	0.62	0.63	0.65	0.61	0.61	0.65	0.68	0.82	0.75	0.83	0.79	0.82	0.86	0.85	0.87	1.00											
K10	0.73	0.76	0.73	0.73	0.75	0.74	0.73	0.74	0.73	0.74	0.72	0.73	0.73	0.66	0.64	0.63	0.63	0.64	0.66	0.63	0.63	0.69	0.65	0.82	0.81	0.82	0.78	0.81	0.82	0.88	0.83	0.84	1.00										
K11	0.70	0.72	0.72	0.69	0.71	0.70	0.69	0.70	0.69	0.71	0.69	0.70	0.70	0.64	0.62	0.61	0.60	0.61	0.65	0.60	0.62	0.64	0.62	0.82	0.75	0.83	0.77	0.83	0.83	0.86	0.87	0.87	0.88	1.00									
K12	0.70	0.72	0.70	0.69	0.71	0.70	0.69	0.71	0.69	0.71	0.70	0.70	0.70	0.65	0.62	0.63	0.63	0.64	0.66	0.62	0.62	0.66	0.63	0.82	0.78	0.81	0.78	0.84	0.87	0.87	0.86	0.85	0.87	0.94	1.00								
K13	0.72	0.74	0.74	0.72	0.75	0.73	0.73	0.73	0.71	0.74	0.74	0.74	0.74	0.64	0.64	0.64	0.62	0.64	0.66	0.63	0.63	0.67	0.65	0.82	0.81	0.83	0.77	0.83	0.85	0.87	0.87	0.88	0.87	0.91	0.90	1.00							
K14	0.74	0.75	0.74	0.73	0.75	0.74	0.73	0.73	0.73	0.74	0.74	0.73	0.73	0.65	0.64	0.62	0.63	0.64	0.66	0.62	0.63	0.69	0.65	0.83	0.81	0.83	0.78	0.84	0.84	0.88	0.87	0.86	0.89	0.89	0.91	0.91	1.00						
K15	0.71	0.73	0.71	0.71	0.73	0.71	0.70	0.72	0.70	0.72	0.72	0.72	0.72	0.63	0.62	0.60	0.60	0.62	0.64	0.60	0.61	0.66	0.62	0.80	0.80	0.80	0.74	0.83	0.86	0.87	0.85	0.86	0.89	0.88	0.91	0.94	0.91	1.00					
K16	0.76	0.76	0.75	0.72	0.76	0.74	0.74	0.76	0.74	0.78	0.75	0.75	0.74	0.65	0.64	0.65	0.62	0.64	0.68	0.63	0.64	0.68	0.66	0.77	0.83	0.77	0.76	0.81	0.83	0.86	0.84	0.81	0.82	0.84	0.86	0.87	0.86	0.87	1.00				
K17	0.76	0.77	0.76	0.73	0.77	0.75	0.76	0.77	0.74	0.78	0.75	0.76	0.76	0.64	0.66	0.65	0.63	0.63	0.67	0.64	0.65	0.67	0.66	0.76	0.80	0.77	0.74	0.79	0.84	0.88	0.84	0.81	0.82	0.84	0.84	0.86	0.85	0.86	0.93	1.00			
K18	0.77	0.77	0.75	0.73	0.76	0.76	0.75	0.77	0.73	0.77	0.74	0.75	0.75	0.66	0.64	0.66	0.65	0.65	0.68	0.65	0.65	0.69	0.68	0.74	0.85	0.77	0.72	0.77	0.76	0.87	0.80	0.83	0.83	0.82	0.84	0.85	0.86	0.86	0.89	0.89	1.00		
K19	0.72	0.72	0.71	0.72	0.74	0.72	0.74	0.75	0.73	0.72	0.75	0.72	0.70	0.64	0.65	0.67	0.66	0.64	0.67	0.65	0.67	0.70	0.66	0.73	0.86	0.77	0.69	0.75	0.73	0.84	0.80	0.77	0.82	0.81	0.80	0.84	0.83	0.83	0.86	0.84	0.89	1.00	
K20	0.73	0.74	0.71	0.73	0.76	0.76	0.75	0.76	0.74	0.73	0.75	0.77	0.75	0.69	0.69	0.70	0.68	0.69	0.73	0.69	0.69	0.72	0.67	0.78	0.82	0.79	0.73	0.80	0.81	0.87	0.81	0.81	0.88	0.84	0.85	0.89	0.89	0.90	0.88	0.85	0.87	0.90	1.00

Appendix 3

Jaccard Coefficient Similarity Matrix

	ME18	ME19	ME20	ME21	ME22	ME23	ME25	ME26	ME27	ME28	ME29	ME30	ME31	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11	K12	K13	K14	K15	K16	K17	K18	K19	K20
ME18	1.00																																										
ME19	0.73	1.00																																									
ME20	0.63	0.73	1.00																																								
ME21	0.62	0.65	0.65	1.00																																							
ME22	0.72	0.84	0.79	0.73	1.00																																						
ME23	0.63	0.74	0.71	0.62	0.76	1.00																																				
ME25	0.62	0.69	0.66	0.61	0.71	0.69	1.00																																			
ME26	0.64	0.67	0.62	0.57	0.70	0.65	0.78	1.00																																		
ME27	0.58	0.68	0.63	0.58	0.67	0.56	0.71	0.66	1.00																																	
ME28	0.53	0.56	0.49	0.47	0.55	0.54	0.56	0.57	0.53	1.00																																
ME29	0.31	0.34	0.30	0.33	0.36	0.30	0.28	0.30	0.26	0.36	1.00																															
ME30	0.61	0.65	0.56	0.55	0.63	0.59	0.61	0.56	0.52	0.55	0.31	1.00																														
ME31	0.59	0.62	0.60	0.53	0.61	0.57	0.58	0.54	0.50	0.53	0.30	0.95	1.00																													
L1	0.17	0.20	0.18	0.20	0.18	0.21	0.21	0.17	0.18	0.22	0.14	0.18	0.19	1.00																												
L2	0.18	0.23	0.24	0.23	0.23	0.24	0.25	0.22	0.23	0.21	0.14	0.20	0.20	0.69	1.00																											
L3	0.17	0.22	0.21	0.20	0.20	0.23	0.24	0.20	0.19	0.18	0.14	0.20	0.21	0.60	0.72	1.00																										
L4	0.18	0.22	0.22	0.22	0.20	0.25	0.26	0.21	0.22	0.20	0.14	0.20	0.21	0.65	0.71	0.77	1.00																									
L5	0.15	0.20	0.20	0.19	0.19	0.21	0.23	0.19	0.19	0.18	0.13	0.19	0.20	0.69	0.75	0.71	0.75	1.00																								
L6	0.18	0.21	0.21	0.21	0.19	0.21	0.22	0.20	0.20	0.22	0.13	0.19	0.20	0.69	0.76	0.69	0.66	0.73	1.00																							
L7	0.14	0.20	0.19	0.20	0.18	0.22	0.22	0.20	0.18	0.17	0.12	0.16	0.17	0.66	0.86	0.76	0.78	0.79	0.78	1.00																						
L8	0.17	0.19	0.20	0.20	0.19	0.21	0.24	0.19	0.20	0.19	0.09	0.19	0.20	0.69	0.75	0.71	0.76	0.77	0.70	0.74	1.00																					
L9	0.16	0.21	0.19	0.21	0.19	0.21	0.22	0.20	0.19	0.17	0.13	0.18	0.19	0.65	0.69	0.65	0.70	0.71	0.72	0.74	0.72	1.00																				
L10	0.21	0.24	0.24	0.21	0.22	0.22	0.23	0.21	0.22	0.22	0.14	0.19	0.20	0.59	0.67	0.61	0.61	0.64	0.62	0.64	0.60	0.66	1.00																			
K1	0.20	0.25	0.25	0.22	0.23	0.25	0.21	0.20	0.20	0.21	0.17	0.25	0.26	0.17	0.18	0.17	0.17	0.22	0.20	0.18	0.15	0.19	0.21	1.00																		
K2	0.11	0.10	0.09	0.10	0.10	0.09	0.10	0.10	0.10	0.14	0.10	0.13	0.13	0.05	0.05	0.05	0.05	0.04	0.05	0.05	0.05	0.06	0.05	0.21	1.00																	
K3	0.18	0.23	0.24	0.18	0.21	0.20	0.20	0.18	0.19	0.20	0.16	0.22	0.25	0.23	0.23	0.22	0.20	0.28	0.23	0.22	0.20	0.23	0.26	0.63	0.21	1.00																
K4	0.19	0.24	0.25	0.20	0.22	0.23	0.23	0.19	0.23	0.23	0.14	0.23	0.24	0.20	0.17	0.19	0.20	0.22	0.18	0.17	0.15	0.18	0.18	0.64	0.18	0.56	1.00															
K5	0.19	0.22	0.22	0.19	0.22	0.21	0.21	0.21	0.20	0.21	0.13	0.24	0.23	0.17	0.19	0.16	0.16	0.21	0.20	0.17	0.15	0.18	0.18	0.59	0.21	0.56	0.62	1.00														
K6	0.23	0.27	0.26	0.25	0.26	0.24	0.22	0.22	0.25	0.24	0.20	0.28	0.29	0.18	0.15	0.15	0.15	0.19	0.18	0.14	0.13	0.16	0.18	0.55	0.17	0.48	0.47	0.47	1.00													
K7	0.20	0.24	0.21	0.22	0.24	0.21	0.23	0.22	0.22	0.23	0.18	0.24	0.23	0.11	0.15	0.13	0.14	0.14	0.14	0.16	0.12	0.12	0.13	0.35	0.29	0.38	0.32	0.41	0.38	1.00												
K8	0.19	0.22	0.25	0.20	0.22	0.21	0.19	0.19	0.20	0.21	0.15	0.24	0.25	0.14	0.16	0.13	0.13	0.17	0.17	0.13	0.12	0.15	0.18	0.51	0.18	0.48	0.45	0.54	0.61	0.38	1.00											
K9	0.29	0.32	0.31	0.28	0.30	0.30	0.26	0.25	0.26	0.28	0.19	0.32	0.31	0.20	0.18	0.19	0.19	0.22	0.21	0.18	0.16	0.18	0.26	0.45	0.17	0.50	0.42	0.46	0.58	0.46	0.58	1.00										
K10	0.17	0.23	0.22	0.21	0.22	0.20	0.20	0.19	0.21	0.18	0.11	0.18	0.20	0.14	0.14	0.13	0.14	0.17	0.15	0.14	0.12	0.15	0.17	0.39	0.18	0.42	0.34	0.37	0.45	0.48	0.42	0.47	1.00									
K11	0.17	0.20	0.25	0.18	0.20	0.18	0.18	0.17	0.18	0.17	0.12	0.18	0.19	0.15	0.16	0.15	0.14	0.16	0.18	0.14	0.15	0.14	0.17	0.43	0.15	0.48	0.36	0.46	0.51	0.46	0.56	0.57	0.55	1.00								
K12	0.15	0.19	0.20	0.17	0.19	0.17	0.16	0.17	0.17	0.16	0.12	0.17	0.18	0.16	0.15	0.16	0.17	0.20	0.19	0.16	0.14	0.15	0.17	0.42	0.16	0.42	0.37	0.47	0.60	0.48	0.53	0.52	0.51	0.76	1.00							
K13	0.19	0.23	0.26	0.23	0.24	0.22	0.22	0.19	0.21	0.20	0.17	0.22	0.24	0.15	0.18	0.16	0.16	0.18	0.17	0.16	0.15	0.16	0.19	0.43	0.21	0.45	0.33	0.43	0.54	0.46	0.54	0.60	0.52	0.67	0.63	1.00					
K14	0.19	0.21	0.24	0.21	0.22	0.20	0.20	0.17	0.21	0.18	0.15	0.18	0.20	0.13	0.14	0.12	0.14	0.17	0.15	0.13	0.12	0.15	0.17	0.41	0.19	0.44	0.34	0.45	0.49	0.48	0.53	0.52	0.54	0.58	0.63	0.65	1.00				
K15	0.17	0.21	0.22	0.21	0.22	0.18	0.18	0.19	0.19	0.18	0.15	0.20	0.21	0.13	0.15	0.12	0.13	0.17	0.16	0.13	0.13	0.15	0.16	0.38	0.20	0.40	0.29	0.45	0.55	0.48	0.50	0.54	0.57	0.57	0.65	0.77	0.63	1.00			
K16	0.19	0.19	0.21	0.15	0.19	0.16	0.16	0.17	0.19	0.20	0.11	0.17	0.16	0.09	0.11	0.12	0.10	0.12	0.14	0.10	0.10	0.10	0.14	0.25	0.20	0.26	0.25	0.32	0.42	0.35	0.40	0.37	0.31	0.42	0.46	0.47	0.43	0.49	1.00		
K17	0.22	0.24	0.25	0.20	0.23	0.21	0.23	0.22	0.22	0.23	0.14	0.21	0.23	0.10	0.16	0.14	0.13	0.13	0.14	0.13	0.14	0.11	0.16	0.25	0.17	0.28	0.24	0.30	0.47	0.44	0.42	0.39	0.34	0.44	0.43	0.46	0.42	0.48	0.64	1.00	
K18	0.21	0.21	0.21	0.17	0.19	0.20	0.18	0.19	0.17	0.18	0.09	0.17	0.18	0.11	0.11	0.13	0.14	0.14	0.14	0.13	0.12	0.12	0.17	0.19	0.22	0.26	0.18	0.23	0.28	0.38	0.30	0.42	0.34	0.36	0.40	0.41	0.43	0.46	0.46	0.49	1.00
K19	0.10	0.12	0.16	0.15	0.14	0.13	0.14	0.15	0.16	0.08	0.09	0.10	0.10	0.07	0.11	0.12	0.11	0.11	0.09	0.11	0.10	0.09	0.11	0.17	0.24	0.26	0.14	0.20	0.22	0.31	0.32	0.27	0.32	0.33	0.29	0.38	0.34	0.36	0.35	0.34	0.44
K20	0.13	0.16	0.18	0.19	0.19	0.21	0.18	0.17	0.20	0.13	0.06	0.17	0.16	0.11	0.16	0.12	0.13	0.17	0.15	0.13	0.14	0.13	0.13	0.25	0.13	0.30	0.21	0.29	0.39	0.36	0.36	0.37	0.48	0.39	0.40	0.50	0.50	0.56	0.42	0.40	0.38

Appendix 4

Dice Coefficient Similarity Matrix

	ME18	ME19	ME20	ME21	ME22	ME23	ME25	ME26	ME27	ME28	ME29	ME30	ME31	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11	K12	K13	K14	K15	K16	K17	K18	K19	K20
ME18	1.00																																										
ME19	0.85	1.00																																									
ME20	0.78	0.85	1.00																																								
ME21	0.76	0.79	0.79	1.00																																							
ME22	0.84	0.91	0.88	0.85	1.00																																						
ME23	0.78	0.85	0.83	0.77	0.86	1.00																																				
ME25	0.77	0.81	0.80	0.76	0.83	0.82	1.00																																			
ME26	0.78	0.81	0.76	0.72	0.82	0.78	0.88	1.00																																		
ME27	0.74	0.81	0.77	0.73	0.80	0.72	0.83	0.80	1.00																																	
ME28	0.69	0.72	0.66	0.64	0.71	0.70	0.72	0.73	0.69	1.00																																
ME29	0.48	0.51	0.46	0.49	0.53	0.47	0.43	0.46	0.42	0.53	1.00																															
ME30	0.76	0.78	0.72	0.71	0.78	0.74	0.76	0.72	0.68	0.71	0.47	1.00																														
ME31	0.74	0.77	0.75	0.69	0.76	0.72	0.74	0.70	0.67	0.69	0.46	0.98	1.00																													
L1	0.29	0.33	0.31	0.33	0.30	0.34	0.34	0.29	0.31	0.36	0.24	0.30	0.32	1.00																												
L2	0.31	0.37	0.38	0.37	0.37	0.38	0.40	0.35	0.37	0.34	0.24	0.33	0.34	0.82	1.00																											
L3	0.29	0.35	0.35	0.33	0.33	0.37	0.38	0.34	0.31	0.30	0.25	0.33	0.34	0.75	0.83	1.00																										
L4	0.30	0.36	0.36	0.36	0.34	0.40	0.41	0.35	0.36	0.33	0.24	0.34	0.35	0.79	0.83	0.87	1.00																									
L5	0.26	0.34	0.34	0.32	0.32	0.35	0.37	0.32	0.32	0.31	0.24	0.32	0.33	0.81	0.86	0.83	0.86	1.00																								
L6	0.30	0.35	0.34	0.35	0.32	0.34	0.36	0.33	0.33	0.36	0.24	0.32	0.34	0.82	0.86	0.82	0.80	0.84	1.00																							
L7	0.24	0.33	0.32	0.33	0.30	0.36	0.36	0.33	0.31	0.29	0.22	0.28	0.30	0.79	0.92	0.86	0.87	0.88	0.88	1.00																						
L8	0.29	0.32	0.33	0.34	0.31	0.35	0.39	0.32	0.34	0.33	0.16	0.31	0.33	0.81	0.86	0.83	0.86	0.87	0.83	0.85	1.00																					
L9	0.28	0.35	0.32	0.35	0.32	0.34	0.36	0.33	0.33	0.29	0.23	0.30	0.31	0.79	0.82	0.79	0.82	0.83	0.84	0.85	0.83	1.00																				
L10	0.35	0.39	0.38	0.35	0.36	0.36	0.38	0.35	0.37	0.36	0.24	0.32	0.34	0.74	0.80	0.76	0.76	0.78	0.77	0.78	0.75	0.79	1.00																			
K1	0.33	0.40	0.40	0.36	0.38	0.40	0.34	0.34	0.33	0.35	0.29	0.40	0.41	0.29	0.31	0.29	0.29	0.36	0.33	0.31	0.26	0.33	0.35	1.00																		
K2	0.19	0.18	0.16	0.19	0.19	0.17	0.18	0.19	0.18	0.25	0.17	0.24	0.22	0.10	0.09	0.09	0.09	0.08	0.10	0.09	0.10	0.11	0.09	0.35	1.00																	
K3	0.30	0.37	0.38	0.30	0.34	0.34	0.33	0.30	0.32	0.33	0.28	0.36	0.40	0.38	0.38	0.36	0.33	0.44	0.38	0.36	0.33	0.38	0.41	0.77	0.34	1.00																
K4	0.31	0.38	0.40	0.34	0.36	0.37	0.37	0.32	0.38	0.37	0.24	0.38	0.39	0.33	0.30	0.32	0.33	0.36	0.31	0.29	0.26	0.31	0.31	0.78	0.31	0.71	1.00															
K5	0.31	0.36	0.36	0.31	0.36	0.35	0.35	0.34	0.34	0.35	0.24	0.38	0.37	0.29	0.31	0.28	0.27	0.34	0.33	0.29	0.26	0.31	0.31	0.74	0.35	0.72	0.77	1.00														
K6	0.38	0.42	0.41	0.40	0.41	0.39	0.36	0.37	0.40	0.38	0.34	0.44	0.45	0.30	0.26	0.27	0.26	0.31	0.30	0.24	0.23	0.27	0.31	0.71	0.29	0.64	0.64	0.64	1.00												
K7	0.33	0.39	0.35	0.36	0.38	0.35	0.37	0.36	0.36	0.38	0.30	0.38	0.37	0.20	0.26	0.24	0.25	0.25	0.25	0.28	0.21	0.21	0.23	0.52	0.44	0.55	0.49	0.58	0.56	1.00										
K8	0.31	0.36	0.40	0.34	0.36	0.35	0.33	0.32	0.34	0.35	0.26	0.38	0.40	0.25	0.27	0.24	0.23	0.29	0.29	0.23	0.22	0.26	0.31	0.67	0.31	0.65	0.62	0.70	0.76	0.56	1.00									
K9	0.44	0.49	0.47	0.44	0.46	0.46	0.41	0.40	0.42	0.44	0.31	0.48	0.47	0.33	0.31	0.31	0.32	0.36	0.35	0.31	0.28	0.31	0.42	0.63	0.29	0.67	0.59	0.63	0.73	0.63	0.74	1.00								
K10	0.29	0.37	0.36	0.34	0.36	0.33	0.33	0.32	0.34	0.30	0.20	0.31	0.33	0.25	0.25	0.23	0.25	0.29	0.27	0.25	0.22	0.26	0.29	0.56	0.30	0.59	0.51	0.54	0.62	0.65	0.59	0.64	1.00							
K11	0.29	0.33	0.40	0.31	0.33	0.30	0.30	0.29	0.31	0.30	0.21	0.31	0.32	0.27	0.27	0.25	0.25	0.28	0.31	0.25	0.26	0.24	0.29	0.60	0.26	0.65	0.53	0.63	0.67	0.63	0.72	0.73	0.71	1.00						
K12	0.27	0.32	0.34	0.30	0.31	0.29	0.28	0.30	0.30	0.28	0.21	0.29	0.31	0.27	0.26	0.28	0.29	0.33	0.31	0.27	0.24	0.27	0.29	0.59	0.28	0.59	0.54	0.64	0.75	0.65	0.69	0.68	0.68	0.86	1.00					
K13	0.32	0.37	0.41	0.37	0.39	0.36	0.36	0.33	0.34	0.33	0.30	0.37	0.38	0.26	0.30	0.28	0.27	0.31	0.29	0.28	0.27	0.27	0.32	0.60	0.34	0.62	0.50	0.60	0.70	0.63	0.70	0.75	0.68	0.80	0.78	1.00				
K14	0.32	0.34	0.39	0.34	0.36	0.33	0.33	0.29	0.34	0.30	0.26	0.31	0.33	0.22	0.25	0.21	0.25	0.29	0.27	0.23	0.22	0.26	0.29	0.59	0.31	0.61	0.51	0.62	0.66	0.65	0.69	0.68	0.70	0.73	0.78	0.78	1.00			
K15	0.29	0.34	0.36	0.34	0.36	0.31	0.31	0.32	0.32	0.30	0.27	0.34	0.35	0.23	0.26	0.22	0.23	0.29	0.27	0.24	0.22	0.27	0.27	0.55	0.33	0.57	0.45	0.62	0.71	0.65	0.67	0.70	0.73	0.73	0.79	0.87	0.78	1.00		
K16	0.32	0.32	0.34	0.27	0.31	0.28	0.28	0.29	0.32	0.33	0.19	0.29	0.28	0.17	0.20	0.21	0.18	0.22	0.24	0.18	0.19	0.18	0.25	0.40	0.33	0.41	0.40	0.49	0.59	0.52	0.57	0.54	0.48	0.59	0.63	0.64	0.60	0.66	1.00	
K17	0.36	0.38	0.40	0.33	0.38	0.35	0.37	0.36	0.35	0.37	0.24	0.35	0.37	0.19	0.28	0.24	0.23	0.23	0.25	0.24	0.24	0.20	0.28	0.41	0.29	0.44	0.38	0.46	0.64	0.61	0.59	0.56	0.51	0.61	0.60	0.63	0.59	0.65	0.78	1.00
K18	0.35	0.35	0.34	0.29	0.31	0.34	0.31	0.32	0.29	0.30	0.16	0.29	0.31	0.20	0.20	0.23	0.24	0.24	0.24	0.22	0.21	0.21	0.29	0.32	0.36	0.41	0.30	0.38	0.44	0.55	0.46	0.59	0.51	0.53	0.58	0.58	0.60	0.63	0.63	0.66
K19	0.18	0.21	0.28	0.26	0.25	0.24	0.25	0.26	0.27	0.15	0.17	0.18	0.18	0.12	0.20	0.21	0.21	0.19	0.16	0.20	0.18	0.17	0.21	0.30	0.38	0.41	0.25	0.33	0.36	0.47	0.48	0.43	0.48	0.50	0.45	0.55	0.51	0.53	0.52	0.51
K20	0.23	0.27	0.31	0.32	0.32	0.35	0.31	0.29	0.33	0.23	0.11	0.29	0.28	0.20	0.27	0.21	0.23	0.30	0.26	0.24	0.24	0.22	0.23	0.40	0.24	0.46	0.35	0.44	0.56	0.53	0.53	0.54	0.65	0.57	0.57	0.67	0.67	0.71	0.59	0.57

	K19	K20
K19	1.00	
K20	0.61	1.00

APPENDIX III

DNA Quantification Results obtained using Biophotometer (Eppendorff, Germany)

Sample ID	Dilution Factor	Absorbance (nm)				Ratio 260/280	Ratio 260/230	Concentration (µg/ml)
		230	260	280	320			
ME 18	10:90	0.276	0.324	0.170	0.011	1.90	1.17	162.2
ME 19	10:90	0.464	0.494	0.260	0.018	1.90	1.06	247.2
ME 20	10:90	0.525	0.857	0.452	0.029	1.90	1.63	428.6
ME 21	10:90	0.378	0.399	0.216	0.019	1.85	1.05	199.3
ME 22	10:90	0.455	0.572	0.303	0.019	1.89	1.26	149.6
ME 23	10:90	0.192	0.094	0.055	0.011	1.70	0.49	46.9
ME 25	10:90	0.304	0.311	0.166	0.014	1.87	1.02	100.6
ME 26	10:90	0.336	0.382	0.210	0.020	1.82	1.14	190.9
ME 27	10:90	0.493	0.553	0.321	0.042	1.72	1.12	276.6
ME 28	10:90	0.284	0.271	0.166	0.028	1.63	0.95	135.3
ME 29	10:90	0.250	0.222	0.135	0.020	1.64	0.89	111.0
ME 30	10:90	0.217	0.174	0.107	0.020	1.62	0.80	86.8
ME 31	10:90	0.285	0.285	0.168	0.021	1.69	1.00	142.7
L 1	10:90	0.309	0.232	0.151	0.057	1.53	0.75	115.9
L 2	10:90	0.291	0.328	0.174	0.013	1.88	1.13	164.0
L 3	10:90	0.339	0.375	0.228	0.050	1.65	1.11	187.7
L 4	10:90	0.291	0.283	0.155	0.023	1,83	0.97	141.4
L 5	10:90	0.275	0.210	0.122	0.025	1.73	0.76	104.8
L 6	10:90	0.353	0.303	0.183	0.034	1.65	0.86	151.4
L 7	10:90	0.351	0.353	0.217	0.050	1.63	1.01	176.5
L 8	10:90	0.335	0.356	0.198	0.027	1.80	1.06	177.9
L 9	10:90	0.389	0.417	0.248	0.069	1.68	1.07	208.7
L 10	10:90	0.316	0.280	0.164	0.046	1.71	0.88	139.8
K 1	10:90	0.539	0.672	0.405	0.064	1.66	1.25	335.9

K 2	10:90	0.279	0.256	0.142	0.023	1.80	0.92	128.0
K 3	10:90	0.411	0.503	0.317	0.068	1.58	1.22	251.4
K 4	10:90	0.449	0.578	0.344	0.037	1..68	1.29	289.2
K 5	10:90	0.705	0.976	0.633	0.154	1.54	1.38	487.8
K 6	10:90	0.609	0.938	0.486	0.036	1.93	1.54	469.2
K 7	10:90	0.180	0.109	0.064	0.015	1.71	0.60	54.5
K 8	10:90	0.836	1.252	0.874	0.250	1.43	1.50	626.0
K 9	10:90	0.412	0.531	0.354	0.082	1.50	1.29	265.7
K 10	10:90	0.353	0.431	0.279	0.056	1.54	1.22	215.5
K 11	10:90	0.337	0.413	0.246	0.041	1.68	1.23	206.5
K 12	10:90	0.574	0.836	0.492	0.075	1.70	1.46	417.8
K 13	10:90	0.616	0.819	0.464	0.079	1.76	1.33	409.5
K 14	10:90	0.470	0.701	0.415	0.066	1.69	1.49	350.5
K 15	10:90	0.517	0.581	0.326	0.058	1.78	1.12	290.3
K 16	10:90	0.618	0.616	0.428	0.111	1.44	1.00	308.2
K 17	10:90	0.450	0.539	0.332	0.065	1.63	1.20	269.5
K 18	10:90	0.414	0.471	0.325	0.085	1.45	1.14	235.5
K 19	10:90	0.496	0.652	0.417	0.076	1.56	1.31	326.0
K 20	10:90	0.391	0.535	0.305	0.043	1.75	1.37	267.4