GENETIC DIVERSITY ASSESSMENT OF *CITRUS* SPP. OF NEPAL USING PCR-BASED MOLECULAR MARKER TECHNIQUES

M.Sc. Thesis

(2013)

Submitted to

CENTRAL DEPARTMENT OF BIOTECHNOLOGY

Tribhuvan University

Kirtipur, Kathmandu, Nepal

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

Nabin Narayan Munankarmi

Exam Roll no: BT 037-067

T.U. Registration No: 5-2-37-407-2006

GENETIC DIVERSITY ASSESSMENT OF *CITRUS* SPP. OF NEPAL USING PCR-BASED MOLECULAR MARKER TECHNIQUES

M.Sc. Thesis

(2013)

Submitted to

CENTRAL DEPARTMENT OF BIOTECHNOLOGY

Tribhuvan University

Kirtipur, Kathmandu, Nepal

By

Nabin Narayan Munankarmi

Supervisors

Dr. Tribikram Bhattarai

Mrs. Neesha Rana (Basnyat)

Dr. Sangita Shrestha

T.U. Registration No: 5-2-37-407-2006

DEDICATION

Dedicated to my beloved parents

Late Mr. Dev Narayan Munankarmi

and

Mrs. Devaki Munankarmi

ACKNOWLEDGEMENT

I wish to express my heartfelt and sincere gratitude to my respected supervisor **Prof. Dr. Tribikram Bhattarai**, Central Department of Biotechnology, Tribhuvan University, Kirtipur for his invaluable guidance, tones of encouragement, valuable suggestions, critical comments, inspiring endurance and constructive supervision throughout this research work and thesis preparation.

I am deeply indebted to my highly devoted supervisors **Dr. Sangita Shrestha**, Unit chief and **Mrs. Neesha Rana** (**Basnyat**), Senior Scientific Officer, Molecular Biotechnology Unit, Nepal Academy of Science and technology (NAST), Lalitpur for their painstaking effort, meticulous and exemplary guidance, incessant and steadfast inspiration, timely help, moral support, kind constructive supervision and bestowed throughout the course of investigation and during thesis preparation.

I am extremely grateful to the Vice Chancellor of NAST **Prof. Dr. Surendra Raj Kafle**, Secretary **Prof. Dr. Prakash Chandra Adhikari** and Science Faculty Chief **Mr. Ishwor Prasad Khanal** for providing me an opportunity to work in the Molecular Biotechnology Laboratory at NAST.

I owe special debt of gratitude to Senior Scientific officers **Dr. Chirangivi Regmi**, **Dr. Jyoti Maharjan**, **Dr. Deepa Dhital**, **Ms. Jaishree Sijapati**, for their constant help, consistent and constructive suggestions, special guidance and valuable comments from the time of sample collections throughout the lab work and thesis preparation. I would like to extend my sincere thanks to Mr. Pawan Parajuli, Mr. Bikash Baral, Mr. Sushil Bhandari, Mr. Ranjan Koirala, Mr. Jagat Shrestha, Mr. Om Basukala, Ms. Smita Shrestha, Ms. Rita KC, Mrs. Poonam Yadav and all the scientific staffs of NAST for their constant supports, co-operation, enthusiastic suggestion and maintaining productive lab environment. In a meantime, I would like to express my sincere gratitude to NAST for providing me M.Sc. dissertation fellowship.

My heartily thanks goes to **Mr. Hari Prasad Subedi**, **Mr. Ram Lal Shrestha**, **Mr. Kishor Bhandari** and all other staffs at NCRP, Dhankuta for their supports, co-operation, guidance during my one week stay at NCRP for sample collections. Similarly, I would like to express my gratitude to **Mr. Dhan Bahadur Thapa** and other staffs of Central Horticulture Centre (CHC) and National Citrus Development Program (NCDP), Kirtipur for their supports and help during sample collections.

I would like to express my heartfelt gratitude to **Dr. Jwala Bajracharya** (**Chitrakar**), Senior Scientist, Head of Department, Seed Division, NARC and **Dr. Bal Krishna Joshi**, Scientist, Genebank-NARC, Khumaltar for helping me in data analysis and interpretation. I am also indebted to library staffs of different institution like Tribhuvan University Central Library, Central Department of Biotechnology, Central Department of Botany, Central Horticulture Centre, NAST, NCRP, and NARC.

I'm very much delightful to **Associate Prof. Dr. Behrouz Golein**, Iran Citrus Research Institute, Iran, **Dr. Mukunda Ranjit**, SANN International College and Research Center, Gairidhara, Kathmandu and **Dr. Krishna Prasad Poudel**, Horticulture Research Division, Khumaltar, Lalitpur for providing me relevant documents and valuable suggestions.

In a meantime, I'm immensely thankful to University Grant Commission (UGC) for providing me financial support for printing, report preparation and editing, purchase of related books, journals and documentation of findings.

I offer my regards and appreciations to honorable Head of Department, **Associate Prof. Dr. Rajani Malla**, Central Department of Biotechnology, honorable **Prof. Dr. Lakshmaiah Sreerama**, Fulbright Fellow-2010-Tribhuvan University, Dept. of chemistry, St. Cloud University, USA and all the respected faculty members, **Prof. Dr. Mohan Kharel**, **Prof. Dr. Krishna Das Manandhar**, **Dr. Deepak Pant**, **Dr. Tilak Ram Shrestha**, **Dr. Sampoornananda Jha**, **Dr. Giri Raj Tripathi**, **Late Mr. Janardan Pandey**, **Mr. Bal Hari Poudel**, **Mrs. Jarina Joshi**, **Mrs. Smeeta Shrestha** for their sincere advice, tremendous supports, encouragement and vital suggestions.

A special thanks goes to my dear friends namely Jyoti Buju, Prashanna Maharjan, Prabin Shrestha, Sunita Khanal, Rajeev Panthi, Sunita Karki, Usha Adhikari, Santosh Panthi, Nawneet Mishra, Rajesh Lamichhane, Surendra Raj Sharma, Surendra Neupane, Gaurav Chandra Gyawali, Pradeep Shrestha, Rejeena Shrestha, Meena Kusi, Gorakha Raj Giri, Suman K.C, Pratap Khadka, Kushal Shrestha and all my colleagues, seniors, juniors and staffs of CDBT, TU.

Last but not least, I want to pay my special regards and thanks to my family members for their support, patient, guidance, encouragement and deep affection during accomplishing the whole course of my study and this thesis work.

Nabin Narayan Munankarmi

GLOSSARY ACRONYMS

AFLP	Amplified Fragment Length
AMOVA	Analysis of Molecular Variance
AP-PCR	Arbitrarily PCR
m/asl	Meter/Above Sea Level
CAPS	Cleaved Amplified Polymorphic Sequence
CBD	Convention on Biological Diversity
CBOL	Consortium for the Barcode of Life
cDNA	Complementary DNA
CIc	Consensus fork Index
CITES	Convention on International Trade of Endangered Species of Wild Fauna and Flora
CNS	Central Nervous System
CONSEN	Consensus trees and indices
COPH	Cophenetic value
cpDNA	Chloroplast DNA
CTAB	Ceyltrimethyl Ammonium Bromide
DAF	DNA Amplification Fingerprinting
d. f	Degree of Freedom
dNTPs	Deoxynucleotide Phosphate
DNA	Deoxyribonucleic Acid
EDTA	Ethylene-Diaminetetraacetic Acid
EST	Expressed Sequence Tag
EtBr	Ethidium Bromide
FAO	Food and Agriculture Organization
GenAlEx	Genetic Analysis in Excel
GLB	Gel Loading Buffer
G_{ST}	Degree of genetic differentiation among the population
H	Nei's gene diversity
He	Expected Heterozygosity
Ho	Observed Heterozygosity
HWE	Hardy Weinberg Equilibrium
H_S	Mean Heterozygosity within population
H_T	Total Heterozygosity (in overall population)
I	Shannon's Diversity Index
ICIMOD	International Centre for Integrated Mountain Development
I_B	Band Informativeness

IFN	Interferon
IGS	Intergenic spacer
IL	Interleukin
IPR	Intellectual Property Right
IR	Infra-Red
ISH	*In situ* hybridization
ISSR	Inter Simple Sequence Repeat
ITS	Internal Transcribed Spacers of the nuclear rDNA Repeat
IUCN	International Union for Conservation of Nature
Kb	Kilo base
L	Liter
LNA	Locked Nucleic Acid
MAAP	Multiple Arbitrary Amplicon Profiling
MAPs	Medicinal and Aromatic Plants
MAS	Marker Assisted Selection
MatK	MaturaseK gene of chloroplast
ml	milliliter
mM	millimolar
MOEST	Ministry of Environment, Science and Technology
MP-PCR	Microsatellite Primed PCR
MS	Mean Square
mtDNA	Mitochondrial DNA
MVSP	Multi-Variate Statistical Package
MXCOMP	Matrix Comparison
NAST	Nepal Academy of Science and Technology
NARC	Nepal Agriculture Research Council
Na	Observed number of allele
NBS	Nepal Biodiversity Strategy
NCDP	National Citrus Development Program
NCRP	National Citrus Research Program
Ne	Effective number of alleles
NPB	Number of Polymorphic Band
NRTU	Nuclear Ribosomal Transcription Unit
NTSYS	Numerical Taxonomical System (Statistical Package)
OTUs	Operational Taxonomic Units
PCO	Principle Co-ordinate Analysis
PCA	Principle Component Analysis
PCR	Polymerase Chain Reaction

PGR	Plant Genetic Resources
PIC	Polymorphic Information Content
POPGENE	Population Genetic Analysis (Statistical Package)
PP	Percent Polymorphism
PVP	Polyvinyl pyrolidone
QTL	Quantitative Trait Loci
r	Correlation Coefficient
RAPD-PCR	Random Amplified Polymorphic DNA-PCR Technique
rbcL	ribulose-bisphosphate carboxylase gene of chloroplast
REs	Restriction Enzymes
RFLP	Restriction Fragment Length Polymorphism
RNA	Ribonucleic Acid
R_P	Resolving Power
RT-PCR	Reverse Transcription-Polymerase Chain Reaction
SAGE	Serial Analysis of Gene Expression
SAHN	Sequential Agglomerative Hierarchical and Nested (Clustering Method)
SCAR	Sequence Characterized Amplified Region
SDS-PAGE	Sodium Dodecyl Sulphate-Poly-Acrylamide Gel Electrophoresis
SIMGEND	Similarity for Genetic Data
SIMQUAL	Similarity for Qualitative Data
SM	Simple Matching (Coefficient of Similarity)
SNPs	Single Nucleotide Polymorphism
SS	Sum of Square
SSLP	Simple Sequence Length Polymorphism
SSRs	Simple Sequence Repeats
STMs	Sequence Tagged Microsatellite
STRs	Short Tandem Repeats
STS	Sequence Tagged Sites
Taq DNA Pol.	DNA Polymerase Enzyme isolated from *Thermus aquaticus* bacteria
TAE	Tris-Acetate-EDTA Buffer
TE buffer	Tris-EDTA Buffer
TIGR	Institute for Genome Research
TLC	Thin Layer Chromatography
TNB	Total Number of Bands
TOGA	TIGR Orthologous Gene Alignments
TRIPs	Trade Related Aspect of Intellectual Property Right
Tris	Tris (Hydroxymwthyl) Aminomethane
UBC	University of British Columbia

UK	United Kingdom
UPGMA	Un-weighted Pair Group Method of Arithmetic Averages
UPOV	International Convention for the Protection of New varieties of Plants
USA	United State of America
UV	Ultra-Violet
VNTR	Variable Number of Tandem Repeats
WHO	World Health Organization
WPGMA	Weighted Pair Group Method of Arithmetic Averages
WTO	World Trade Organization
ϕ_{PT}	Indicator of Genetic Differentiation

LIST OF TABLES, FIGURES, PHOTO PLATES AND APPENDICES

A. TABLES

Table 1.1 Region-wise total area, production and productivity of Citrus fruits in 2010/11.

Table 1.2 Total area, productive area, production and productivity of commonly grown Citrus.

Table1.3 Area coverage (ha) of different Citrus spp. in 2010/11.

Table 1.4 List of *Citrus* spp. found in Nepal.

Table 2.1 Name of the *Citrus* spp. and their place of origin.

Table 2.2 Top five citrus producing countries of the world with their production and productivity, 2010.

Table 3.1 Altitudinal range, accessions number and locality details of sample collection sites of acid lime landraces.

Table 3.2 Sequences and GC content of screened primers used in ISSR-PCR analysis.

Table 3.3 Characteristics of SSR primers used in present study.

Table 4.1 Altitudinal range, Latitude, Longitude, Accession numbers, Sample collections sites of *Citrus* spp.

Table 4.2 Tested different ISSR PCR-cycling conditions with selected one in bold font.

Table 4.3 Tested different ISSR-PCR parameter and selection of optimized parameters.

Table 4.4 Tested different SSR PCR-cycling conditions with selected one in bold font.

Table 4.5 Testing of different SSR-PCR parameter and selection of optimized parameters.

Table 4.6 Primer sequences, total number of amplified band (TNB), number of polymorphic bands (NPB), percent polymorphism and amplicon size range of the 21 primers used to generate ISSR-PCR profiles in 60 *Citrus aurantifolia* (Acid lime) accessions.

Table 4.7 Assessment of Polymorphic Information Content (PIC), Band informativeness (I_B) and resolving power (R_P) of 21 ISSR primers used to generate ISSR profiles in acid lime accessions.

Table 4.8 Correlation Coefficient values from Mantel test of original similarity matrices.

Table 4.9 Consensus fork index among the UPGMA based phenograms using similarity coefficients among *Citrus aurantifolia* (Acid lime) landraces by ISSR markers.

Table 4.10 Correlation coefficients value (r) obtained from cophenetic values of similarity matrices (Simple Matching, Jaccard's and Dice coefficient) and clusters computed by UPGMA module using MXCOMP module of NTSYS.

Table 4.11 Percentage similarity observed for different acid lime landraces agro-ecological zones based on ISSR similarity matrix (Dice) within and between agro-ecological zones.

Table 4.12 Eigenvalues and the percentage for the co-ordinates of PCO.

Table 4.13 Direct assessments of polymorphism based on monomorphic and polymorphic bands amplified by 21 ISSR primers.

Table 4.14 Analysis of Molecular Variance (AMOVA) for 60 acid lime accessions from three agro-ecological regions.

Table 4.15 Assessment of Alleles per locus, Allele size range, Polymorphic Information Content (PIC) and Probability of Identity (PI) of 12 SSR primers used to generate SSR profiles in 45 citrus accessions belonging to different species.

Table 4.16 Assessment of Number of Different and Effective alleles, Shannon's Information Index, Observed heterozygosity, Expected heterozygosity, Unbiased expected heterozygosity and fixation Index of 12 SSR primers used to generate SSR profiles in citrus accessions belonging to different species.

Table 4.17 Allele Frequency of each of the alleles of individual locus.

Table 4.18 Correlation Coefficient values from Mantel test of original similarity matrices.

Table 4.19 Consensus fork index among the UPGMA based phenograms using similarity coefficients among *Citrus* spp. by SSR markers.

Table 4.20 Correlation coefficient value (r) obtained for cophenetic values of similarity matrices (Simple Matching, Jaccard's and Dice coefficient) and clusters computed by UPGMA module using MXCOMP (matrix comparison) of NTSYS.

Table 4.21 Eigenvalues and the percentage for the co-ordinates of PCO.

B. FIGURES

Figure 3.1 Map of Nepal showing acid lime samples collection sites.

Figure 3.2 Map of Nepal showing *Citrus* spp. samples collection sites.

Figure 4.1 Phenogram generated for 60 *Citrus aurantifolia* (Acid lime) landraces (refer to table 3.1 for sample details) by UPGMA cluster analysis using Simple Matching coefficient of similarity calculated from 234 ISSR loci generated by 21 primers. The clusters are labeled as I, II, III, IV and V.

Figure 4.2 Phenogram generated for 60 *Citrus aurantifolia* (Acid lime) landraces (refer to table 3.1 for sample details) by UPGMA cluster analysis using Jaccard's coefficient of similarity calculated from 234 ISSR loci generated by 21 primers. The clusters are labeled as I, II, III, IV and V.

Figure 4.3 Phenogram generated for 60 *Citrus aurantifolia* (Acid lime) landraces (refer to table 3.1 for sample details) by UPGMA cluster analysis using Dice coefficient of similarity calculated from 234 ISSR loci generated by 21 primers. The clusters are labeled as I, II, III, IV and V.

Figure 4.4 3D – plot of PCO constructed using NTSYS-PC 2.21i to assess dispersion of all acid lime (*C. aurantifolia*) accessions.

Figure 4.5 Assessment of Principal Co-ordinate Analysis (PCO) of Dice similarity matrix carried out with the MVSP 3.21.

Figure 4.6 Phenogram generated for 45 citrus accessions belonging to different species (refer to table 4.1 for sample details) by UPGMA cluster analysis using Simple Matching coefficient of similarity generated by 12 SSR primers. The clusters are labeled as I, II, and III.

Figure 4.7 Phenogram generated for 45 citrus accessions belonging to different species (refer to table 4.1 for sample details) by UPGMA cluster analysis using Jaccard's coefficient of similarity generated by 12 SSR primers. The clusters are labeled as I, II, III and IV.

Figure 4.8 Phenogram generated for 45 citrus accessions belonging to different species (refer to table 4.1 for sample details) by UPGMA cluster analysis using Dice coefficient of similarity generated by 12 SSR primers. The clusters are labeled as I, II, III and IV.

Figure 4.11 Assessment of Principal Co-ordinate Analysis (PCO) of Jaccard's similarity matrix carried out with the MVSP 3.21.

C. PHOTO PLATES

Plate 4.1 Gel Picture of ISSR-PCR for the selection of best cycling condition using primer C4 and the genomic DNA of samples LT-1, LT-2, LT-3,LT- 4, LT-5 and LT-6 in lanes 1-6 respectively (Table 3.1). Lane 'M' are 100 bp plus molecular weight marker; A = Cycling condition as of Fang et al. (1998); B = Cycling condition as of Lui et al. (2011); C = Cycling condition as of Kumar et al. (2010) and D = Cycling condition as of Shahsavar et al. (2007). An arrow indicated the selected one.

Plate 4.2 Gel pictures of ISSR-PCR for the selection of DNA and $MgCl_2$ concentration using primer UBC 809 and the genomic DNA of sample LS-56. Lanes marked 'M' are 100 bp plus molecular weight marker; (A) Each lane is marked with the respective concentrations of template DNA (12.5 to 100 ng) (B) Each lane is marked with the respective concentrations of $MgCl_2$ (1.5mM to 4.0mM) used during ISSR-PCR. An arrow indicated the selected one.

Plate 4.3 Gel pictures of ISSR-PCR for the selection of primer concentration using the genomic DNA sample LS-56 with varying concentration of primer UBC 809. Lane labeled 'M' is 100 bp plus molecular weight marker. Each lane is marked with the respective concentrations of primer UBC 809 (0.1µM to 1.6µM) used during ISSR-PCR. An arrow indicated the selected one.

Plate 4.4 Gel pictures of ISSR-PCR for the selection of dNTPs and *Taq* polymerase concentration using primer UBC 809 genomic DNA of sample LS-56. Lanes marked 'M' are ladder; (A) Each lane is marked with the respective concentrations of dNTPs (0.1mM to 0.5mM) (B) Each lane is marked with the respective concentrations of *Taq* polymerase (0.5U to 2.5U) used during ISSR-PCR. An arrow indicated the selected one.

Plate 4.5 Gel Picture of SSR-PCR for the selection of best cycling condition using primer TC 26 and the genomic DNA of sample 3, 5, 9, 10, 16, 24 and 25 (Table 4.1). Lane 'M' are 100 bp plus molecular weight marker; A = Cycling condition as of Barkley et al. (2006); B = Cycling condition as of Hvarleva et al., 2008; C = Cycling condition as of Kijas et al. (1995) and D = Cycling condition as of Snoussi et al. (2012). An arrow indicated the selected one.

Plate 4.6 Gel pictures of SSR-PCR for the selection of DNA and $MgCl_2$ concentration using primer TAA 27 and genomic DNA of sample 10. Lanes marked 'M' are 100 bp plus molecular weight marker; (A) Each lane is marked with the respective concentrations of template DNA (5 to 40 ng) (B) Each lane is marked with the respective concentrations of $MgCl_2$ (1.5mM to 4.0mM) used during SSR-PCR. An arrow indicated the selected one.

Plate 4.7 Gel pictures of SSR-PCR for the selection of primer concentration using the genomic DNA sample 10 with varying concentration of primer TAA 27. Lane labeled 'M' is 100 bp plus molecular weight marker. Each lane is marked with the respective concentrations of primer UBC 809 (0.1µM to 1.6µM) used during SSR-PCR. An arrow indicated the selected one.

Plate 4.8 Gel pictures of SSR-PCR for the selection of dNTPs and *Taq* polymerase concentration using primer TAA 27 and genomic DNA of sample 10. Lanes marked 'M' are ladder; (A) Each lane is marked with the respective concentrations of dNTPs (0.1mM to 0.5mM) (B) Each lane is marked with the respective concentrations of *Taq* polymerase (0.5U to 2.5U) used during SSR-PCR. An arrow indicated the selected one.

Plate 4.9 ISSR gel pictures for the Primer Screening using template DNA of Sample LS-15. Each lane in [A] [B] and [C] corresponds to different ISSR primers used in the experiment. Lanes marked 'M' are 100 bp plus DNA ladder. An arrow indicated the selected one.

Plate 4.10 ISSR profile amplified with **Primer C 7.** Lanes marked with M are 100bp plus molecular weight markers. [A] Lanes 1-15 represents acid lime samples 1-15; [B] Lanes 16-30 represents acid lime samples 16-30; [C] Lanes 31-45 represents acid lime samples 31-45; [D] Lanes 46-62k represents acid lime samples 46-62k.

Plate 4.11 ISSR profile amplified with **Primer C 8.** Lanes marked with M are 100bp plus molecular weight markers. [A] Lanes 1-15 represents acid lime samples 1-15; [B] Lanes 16-30 represents acid lime samples 16-30; [C] Lanes 31-45 represents acid lime samples 31-45; [D] Lanes 46-62k represents acid lime samples 46-62k.

Plate 4.12 ISSR profile amplified with **Primer UBC 842.** Lanes marked with M are 100bp plus molecular weight markers. [A] Lanes 1-15 represents acid lime samples 1-15; [B] Lanes 16-30 represents acid lime samples 16-30; [C] Lanes 31-45 represents acid lime samples 31-45; [D] Lanes 46-62k represents acid lime samples 46-62k.

Plate 4.13 SSR profile amplified with **Primer AG 14.** Lanes marked with M are 100bp plus molecular weight markers. [A] Lanes 3-52 represents *Citrus* spp.

samples 3-52; [B] Lanes 54-77 represents acid lime samples 54-77; [C] Lanes 80-113 represents acid lime samples 80-113.

Plate 4.14 SSR profile amplified with **Primer CAT 01.** Lanes marked with M are 100bp plus molecular weight markers. [A] Lanes 3-52 represents *Citrus* spp. samples 3-52; [B] Lanes 54-77 represents acid lime samples 54-77; [C] Lanes 80-113 represents acid lime samples 80-113.

Plate 4.15 SSR profile amplified with **Primer TAA 41.** Lanes marked with M are 100bp plus molecular weight markers. [A] Lanes 3-52 represents *Citrus* spp. samples 3-52; [B] Lanes 54-77 represents acid lime samples 54-77; [C] Lanes 80-113 represents acid lime samples 80-113.

D. APPENDICES

Appendix 1 Comparison of different characteristics of most frequently used molecular marker systems in plants.

Appendix 2 List of Citrus spp. collected from Paripatle, Dhankuta and Kathmandu valley, Nepal.

Appendix 3 Quantification of DNA extracted from Acid lime landraces of different agro-ecological region.

Appendix 4 Quantification of DNA extracted from *Citrus* spp. collected from Paripatle, Dhankuta and Kathmandu valley, Nepal.

Appendix 5 Binary data matrix of ISSR-PCR profiles.

Appendix 6 Binary data matrix of SSR-PCR profiles.

Appendix 7 Similarity Coefficient matrix for Simple Matching Coefficient for 60 acid lime samples.

Appendix 8 Similarity Coefficient matrix for Jaccard's Coefficient for 60 acid lime samples.

Appendix 9 Similarity Coefficient matrix for Dice Coefficient for 60 acid lime samples.

Appendix 10 Similarity Coefficient matrix for Simple Matching Coefficient for 45 citrus accessions belonging to different species.

Appendix 11 Similarity Coefficient matrix for Jaccard's Coefficient for 45 citrus accessions belonging to different species.

Appendix 12 Similarity Coefficient matrix for Dice Coefficient for citrus accessions belonging to different species.

Appendix 13 Examples of ISSR profiles generated by ISSR primers.

Appendix 14 Examples of SSR profiles generated by SSR primers.

Appendix 15 Photos Display.

TABLE OF CONTENTS

ACKNOWLEDGEMENT…………………………………………………………...i
GLOSSARY ACRONYMS……………………………………………………….iii
LIST OF TABLES, FIGURES, PHOTO PLATES AND APPENDICES…………vii
TABLE OF CONTENTS………………………………………………………..xv
ABSTRACT xx
CHAPTER 1: INTRODUCTION 1

1.1. Background 2

1.1.1. Citrus fruits and its cultivation in Nepal.. 2
1.1.2. Citrus cultivation area and production in Nepal 3
1.1.3. *Citrus* spp. found in Nepal 5
1.1.4. Nutritional and Economic Importance of Citrus 6
1.1.5. Problems in Citrus Diversity Assessment.. 6
1.1.6. Importance of Diversity Assessment 7
1.1.7. Methods of Diversity Assessment... 7

1.2. Justification 9
1.3. Objectives... 11
1.4. Scope 11

CHAPTER 2: LITERATURE REVIEW 13

2.1. World Production, cultivation and origin of *Citrus* 14
2.2. Citrus and its cultivars... 15
2.3. Problems in the Classification *Citrus* spp. 18
2.4. Genetic Diversity Assessments by Genetic Markers......................... 20

2.4.1. Morphological marker.. 20
2.4.2. Biochemical marker... 20
2.4.3. Molecular marker.. 21

2.4.3.1. Hybridization-Based marker.. 23
2.4.3.2. PCR-Based marker... 23
1) Arbitrary or semi-arbitrary primed PCR techniques.. 23
a) Random Amplified Polymorphic DNA (RAPD) marker.................................... 23
b) Arbitrarily Primed-PCR (AP-PCR) marker.. 24
c) DNA Amplification Fringerprinting (DAF) marker.. 25
d) Amplified Fragment Length Polymorphism (AFLP) marker 25
e) Inter-simple Sequence Repeat (ISSR) marker 26
2) Site-targeted PCR techniques.. 29
a) Simple Sequence Repeats (Microsatellites) marker 29
b) Cleaved Amplified Polymorphic Sequences (CAPS) marker 31
c) Sequence Characterized Amplified Region (SCAR) marker 31
3) DNA Sequencing Based Markers 32

a) Internal Transcribed Spacer (ITS) ... 32
b) Single Nucleotide Polymorphism (SNPs) ... 33
2.5. Application of ISSR marker Technique ... 33
2.5.1. Genomic fingerprinting ... 33
2.5.2. Genetic diversity and phylogenetic analysis ... 34
2.5.3. Genome mapping ... 34
2.5.4. Studies on natural population/speciation ... 35
2.5.5. Determining SSR motif frequency ... 35
2.5.6. Gene tagging and marker assisted selection ... 35
2.6. Application SSR marker Technique ... 36
2.6.1. Genome mapping ... 36
2.6.2. Gender determination ... 37
2.6.3. Cultivar identification and marker-assisted selection ... 37
2.6.4. Genetic diversity and phylogenetic relationships ... 38
2.6.5. Population and evolutionary studies ... 38
2.7. Genetic Diversity ... 38
2.7.1. Significance of Genetic Diversity ... 39
2.7.1.1. Adaptation of crops to biotic and abiotic factors ... 40
2.7.1.2. Food Security ... 40
2.7.1.3. Livelihood of rural people ... 40
2.8. Statistical parameters and estimators for genetic analysis ... 41
2.8.1. Polymorphic Information Content (PIC) ... 41
2.8.2. Primer Resolving Power (Rp) ... 41
2.8.3. Similarity and Dissimilarity indices for molecular marker data ... 41
2.8.4. Correlation Analysis using COPH and MXCOMP of NTSYS ... 42
2.8.5. Correlation Analysis using CONSEN (Consensus Tree and Indices) of NTSYS ... 43
2.8.6. Cluster Analysis ... 43
2.8.7. Diversity Indices for Dominant and Codominant Data ... 43
2.8.7.1. Shannon Index of Diversity (Nei, 1987) ... 44
2.8.7.2. Observed Heterozygosity ... 44
2.8.7.3. Expected Heterozygosity ... 44
2.8.7.4. Nei's Index ... 44
2.8.8. AMOVA and Indicator of Genetic Differentiation ϕ_{PT} ... 45
2.8.9. Principal co-ordinate Analysis (PCO) ... 45
2.9. Diversity study of *Citrus* spp. ... 45
2.9.1. Diversity study of *Citrus* spp. in Nepal ... 47
2.9.2. Diversity study of lime ... 48

CHAPTER 3: MATERIALS AND METHODS 49

3.1. Materials 50

3.1.1. Reagents and Extraction buffers 50

TAE (Tris-Acetate-EDTA) buffer 50
Gel Loading Buffer (GLB) 50
Agarose Gel 50
Tris Buffer 50
EDTA (0.5M, pH 8.0) 51
NaCl 51
Graham's CTAB (Cetyl Trimethyl Ammonium Bromide) extraction buffer 51
TE buffer (Tris-EDTA; 10mM Tris HCl, 1mM EDTA) with RNAse 51
3.1.2. Primers, dNTPs and DNA Dilution 51
3.1.3. Plant Materials 52

3.2. Methodology 54

3.2.1. DNA extraction methods 54

3.2.1.1. CTAB Extraction Method (Graham et al., 1994a) 54
3.2.1.2. DNeasy Mini Kits (Qiagen Ltd, Crawley, UK) 54
3.2.2. DNA Quantification 55
3.2.3. Gel Electrophoresis 55
3.2.4. PCR Optimization 56
3.2.5. Primer screening 56
3.2.6. PCR Profiling 59

3.2.6.1. Genetic Diversity Assessment based on PCR profiling 59
3.2.6.2. Statistical Analysis 59
Estimation of Percent Polymorphism, Polymorphic Information Content, Band Informativeness and Resolving Power 60
Shannon's Information Index, Observed Heterozygosity, Expected Heterozygosity, Probability of Identity and Fixation Index 60
Genetic Diversity Analysis using Similarity matrices and Phenograms 60
Principal Cooridinates Analysis (PCO) 61
Analysis of Molecular Variance (AMOVA) 62

CHAPTER 4: RESULTS 63

4.1. Collection of samples 64

4.1.1. Acid lime (*Citrus aurantifolia*) 64
4.1.2. Different *Citrus* spp 64

4.2. DNA Quantification 65
4.3. ISSR-PCR Optimization 65

4.3.1. ISSR-PCR Cycling conditions optimization 65

4.3.2. ISSR-PCR reaction conditions optimization ... 67

4.4. SSR-PCR Optimization ... 69

4.4.1. SSR-PCR Cycling conditions optimization ... 69
4.4.2. SSR-PCR reaction conditions optimization ... 71

4.5. Primer Screening for ISSR Profiling ... 73
4.6. ISSR Profiling and Genetic Diversity Assessment of acid lime ... 74

4.6.1. Polymorphic Information Content, Band Informativeness and Resolving Power ... 77
4.6.2. Analysis of genetic diversity using the similarity coefficient and cluster analysis ... 78

4.6.2.1. Comparison of similarity coefficients and phenograms of ISSR-PCR profiling ... 78

3D-Plot ... 84
4.6.3. Principal Co-ordinate Analysis (PCO) ... 85
4.6.4. Estimation of within agro-ecological zone genetic similarity of *C. aurantifolia* ... 86
4.6.5. Analysis of Molecular Variance (AMOVA) ... 87

4.7. SSR Profiling and Genetic Diversity Assessment of *Citrus* spp ... 88

4.7.1. Alleles per locus, Polymorphic Information Content and Probability of Identity ... 89
4.7.2. Allele frequency, Shannon's Information Index, Observed and Expected Heterozygosity and Fixation Index ... 91
4.7.3. Genetic diversity analysis using similarity coefficients and cluster analysis ... 93

4.7.3.1. Comparison of Similarity coefficients and construction of phenograms ... 93

4.7.4. Principal Co-ordinates Analysis (PCO) ... 98

CHAPTER 5: DISCUSSION ... 100

5.1. Optimization of PCR conditions of ISSR and SSR ... 101

5.1.1. Optimization of PCR Reaction conditions for ISSR and SSR ... 102
5.1.2. Optimization of PCR cycling conditions of ISSR and SSR ... 104

5.2. Primer Screening for ISSR ... 105
5.3. ISSR Profiling and Genetic Diversity Assessment of Acid lime ... 105

5.3.1. Analysis of Genetic Diversity using NTSYS and MVSP Statistical Packages for Acid lime ... 107

5.3.1.1. Genetic Differentiation Using Analysis of Molecular Variance (AMOVA) for Acid lime of agro-ecological zones ... 111

5.4. SSR Profiling and Genetic Diversity Assessment of *Citrus* spp ... 111

5.4.1. Alleles per locus, Polymorphic Information Content and Probability of Identity 112
5.4.2. Allele frequency, Observed and Expected Heterozygosity, Shannon's Information Index and Fixation Index 114
5.4.3. Analysis of Genetic Diversity using NTSYS and MVSP Statistical Packages for *Citrus* spp. 116

CHAPTER 6: CONCLUSION 119
CHAPTER 7: REFERENCES 123
APPENDICES 150

ABSTRACT

Citrus is one of the most important fruit crops widely cultivated in tropical and subtropical regions of Nepal. Citrus crops cover about 30% of the total area under fruit cultivation. Its cultivation must be promoted to increase yield to fulfill the national and international demand. A large amount of genetic variation exists within and among *Citrus* spp. as a result of long history of cultivation, frequent bud mutation, interspecific and intergeneric hybridization, apomixes and human selection. As morphological traits are inadequate in differentiation of closely related cultivars and biochemical markers doesn't generally distinguish between cultivars differentiated by mutations, DNA based molecular markers adds a new and reliable dimension to identification, phylogeny construction and genetic diversity assessment. Genetic diversity assessment based on molecular marker tools is essential for the selection of parental genotypes for breeding of elite cultivars having desirable qualitative (disease resistant, drought tolerant, off season production etc.) and quantitative (Total Soluble Solid, Tritable Acidity, Vitamin C content, juice percent etc.) traits via Marker Assisted Selection (MAS) and Marker Assisted Introgression in variety development programs. Present investigation consists of two parts *viz*. diversity assessment of acid lime (*C. aurantifolia* Swingle) landraces of different agro-ecological zone of eastern Nepal using ISSR (Inter-Simple Sequence Repeat) markers and diversity assessment of *Citrus* spp. of Nepal using SSR (Simple Sequence Repeat) markers.

Sixty accessions of acid lime landraces from the different agro-ecological zone of eastern Nepal were analyzed using 21 ISSR markers. Out of 234 scored amplified bands for acid lime accessions, 204 were found to be polymorphic (87.18%). Binary data matrix was created based on presence and absence of ISSR bands. Cluster analysis using ISSR (qualitative) data was performed in the NTSYS-PC (Numerical Taxonomical System) ver. 2.21i statistical package for studying genetic diversity and relationship among various acid lime accessions. Similarity matrices and phenograms were generated by using Simple Matching, Jaccard's and Dice coefficient of similarity and these were compared based on correlation coefficient and consensus indices values. Dice coefficient of similarity in constructing phenogram using UPGMA (Unweighted Pair Group Method of Arithmetic Averages) module was found to be the best for deducing genetic relationship among acid lime accessions. It revealed two major clusters and three minor clusters showing overlapping of acid lime accessions from different agro-ecological zones. The acid lime landraces from High-hill were found to be genetically closest (86.02%, similar) while from Terai were genetically most distant (79.33%, similar). Principal Co-ordinate Analysis (PCO) computed using MVSP (Multi-Variate Statistical Package) ver. 3.21 substantiated the results of phenogram. AMOVA (Analysis of Molecular Variance)

analysis performed using GenAlEx (Genetic Analysis in Excel) ver. 6.5 showed high variations within agro-ecological zones (86%) and low among agro-ecological zone (14%).

Similarly, 45 citrus accessions belonging to different species from Paripatle, Dhankuta and Kathmandu Valley were analyzed using 12 SSR markers. Genetic diversity and relationship among the *Citrus* spp. was studied in the similar manner as that of acid lime using NTSYS-PC ver. 2.21i statistical package. In this study, Jaccard's coefficient of similarity in constructing phenogram using UPGMA module was found to be the best for deducing genetic relationship among *Citrus* spp. accessions under study. It revealed four clusters with kumquat (*Fortunella margarita* and *F. hindsii*) and rootstock citrus (*Poncirus trifoliata* and Old Citranze) out of major clusters. Principal Co-ordinate Analysis (PCO) computed using MVSP ver. 3.21 substantiated the results of phenogram. In present investigation, 60 putative alleles were detected at 12 SSR loci in 45 samples varying from 2 (TAA 27) to 8 (TAA 52) with an average of 5 per locus. Loci TAA 41 (6 alleles, PI value 0.075 and PIC value 0.802) and TAA 52 (8 alleles, PI value 0.089 and PIC value 0.785) were found to be the most informative loci for the investigation of diversity of *Citrus* spp. in this investigation on the basis of number of alleles detected, PI values and PIC values. The Shannon's Information Index, Allele frequency, Observed and Expected Heterozygosity and Fixation Index were computed using GenAlEx ver. 6.5 and their values showed high level of genetic variation among the *Citrus* spp. The result provides genetic information about intraspeicfic diversity and help in identification of similarity group that could be useful for the selection of parental plants to be used in future breeding programs. The results will be helpful for the plant breeders in elite cultivar development purposes. It also provides information about interspecific, intraspecific and even intergeneric genetic similarity and distances of *Citrus* at molecular level using SSR markers and paves the way to tackle the long standing problem of *Citrus* germplasm evaluation, classification and identification in Nepal.

Keywords: *Citrus* spp., acid lime, ISSR, SSR, genetic diversity, molecular marker

CHAPTER 1: INTRODUCTION

1.1. Background

Citrus is a Genus, a flowering plant of Rutaceae family which falls under subfamily *Aurantioideae* and tribe *Citreae*. Citrus was originated in Asian countries extending from the Himalayan foot-hill of Nepal, Bhutan, North Eastern India to North Central China and the Philippines, East Burma, Thailand, Indonesia and New Caledonia in the South East (Lama and kayastha, 1999). Nepal is one of the countries in Asia where Citrus is thought to have originated (Ranjit and Gokarna, 1997). The *Citrus* spp. grows best in subtropical environment, although they have been originated in the tropical region (Larry, 1999). The cultivated species might have been derived from as few as three ancestral species, Pummelo (Citrus *grandis* (L) Osb.), Citron (*C. medica* L.) and Mandarin (*C. reticulata* Blanco) and it has been suggest that all other *Citrus* spp. originated by crosses between these main species or between them and other related genera (Scora, 1975; Barett and Rhodes, 1976), and *C. halimii* has been purposed as fourth basic species of Citrus (Lynch and Milligan, 1994; Kovach, 2007). The taxonomy and systematics of Citrus is complex, and their precise numbers of natural species are unclear as most of the named species are clonally propagated hybrids. Evidence suggests that some of the wild true breeding species are of hybrid origin (Federici et al., 1998). The natural and commercially cultivated Citrus include oranges, grapefruits, lemons, lime and tangerines (Dorji and Yapwattanapun, 2011).

1.1.1. Citrus fruits and its cultivation in Nepal

Citrus fruits are cultivated all over the world in tropical and subtropical regions having suitable soil and climatic conditions. The history of Citrus fruits cultivation in Nepal is not well documented, however the description of the fruits in old scriptures on religious ceremonies and medicinal values indicate that Citrus farming might have started in ancient times, but commercial cultivation of Citrus fruits in Nepal started only after 1970. (NCRP, 2012). In Nepal, citrus fruits are reported to be cultivated in 66 out of 75 districts and citrus crops cover about 30% of the total area under fruit cultivation (Shrestha et al., 2011/12; NCRP, 2012). It is one of the major fruits of Nepal in terms of area coverage, production and export potential, and it has been recognized as high value cash crop by Agriculture Perspective Plan (APP, 1995). At present, major Citrus producing districts of Nepal having more than 1,000 ha area are: Taplejung, Tehrathum, Dhankuta, Ramechahap, Sindhuli, Kavrepalanchowk, Lamjung, Syangja, Salyan and Dailekh. The pocket areas with deep sandy loam soil and soil pH range of 5.0 to 6.5 are most suitable for cultivation of Citrus fruits (Ranjit and Gokarna, 1997). In Nepal, majority of citrus is grown from seed. However, this trend is being replaced by grafted saplings. In Dhading district, a 300 year old seedling mandarin was observed and its fruit

tested out to have excellent qualities. Also, in the Sankhuwasabha district, a 125 year old mandarin seedling was observed and its fruit also had excellent qualities (Shrestha et al., 2011/12). The mid-hills of Nepal are suitable for cultivation of Citrus, particularly for mandarin (*C. reticulata*), sweet orange (*C. sinensis*), lime (*C. aurantifolia*) and lemon (*C. limon*). Citrus is one of the major cash crops of Nepalese farmers in these regions. Mandarin, Sweet orange, Lime, Lemon, Rough lemon, Pummelo, Citron and Sour orange are the main species cultivated in these regions, and limited *Citrus* spp. like pummelo, lemon, rough lemon and lime are grown in Terai belt (Subedi et al., 2010). At present, Mandarin (*C. reticulata*), Junar (*C. sinensis*), Acid lime (*C. aurantifolia*), and Hill lemon (*C. pseudolemon*) are cultivated in commercial scale. The other Citrus fruits like Pummelo (*C. grandis*), Rough lemon (C. *jambhiri*), Sour orange (*C. aurantium*), Sweet lime (C. *limettioides*), Grapefruit (*C. paradise*), Calamondin (*C. madurensis*), Kumquat (*Fortunella japonica*) are found growing in homestead gardens. There is a vast scope of growing Citrus crops in the country (ICIMOD, 2003).

1.1.2. Citrus cultivation area and production in Nepal

Citrus is one of the priority crop of mid-hills from east to west in Nepal (Shrestha et al., 2011/12). The three most important species on citriculture in Nepal include: Suntala (Citrus *reticulata* Blanco), Junar (*C. sinensis* Osbeck) and Kagati (*C. aurantifolia* Swingle) (MoAC, 2011). Of the total area under citrus cultivation, 63.2%, 17.3%, 11.56%, 0.046% and 0.033% are covered by Mandarin, Junar, Acid lime, lemon and other Citrus, respectively (MoAC, 2011). They constitute the key species of commercial Citrus industry of Nepal (Roistacher, 1991). Mandarin is the most important species all across the region as area and production of Citrus fruits are primarily dominated by mandarin. So, the citrus as a whole called as “sutajajaat crop” in Nepal (Subedi, 2010; NCRP, 2012). The second most important species is the sweet orange, and its main coverage is concentrated (about 50%) in Central development region. Acid lime appeared to be the third position particularly concentrated in the Eastern Nepal (Table 1.3). The region wise disintegration data of area, production and productivity of Citrus fruits shows that far-west has the lowest area coverage along with low productivity in 2010/11 (Table 1.1). In fiscal year 2010/11, the total area under Mandarin is estimated to be 22,872 ha having productivity of 179,494 mt. Similarly, sweet orange, lime, lemon and others have their total area, and the productivity values are as 5,365ha/50,679mt, 5,276ha/22,571mt, 1206ha/6432mt and 859ha/4534mt respectively (Table 1.2) (MoAC, 2011). Besides above other Citrus crops and their varieties are also cultivated in Nepal.

Table 1.1 Region-wise total area, production and productivity of Citrus fruits in 2010/11.

Developmental region	Total area (ha)	Productive area (ha)	Total production (ha)	Productivity (mt/ha)
Eastern	9,008	5,546	61,265	11.05
Central	7,324	5,214	63,083	13.00
Western	10,816	7,387	83,610	11.32
Mid-western	4,899	3,059	33,893	11.08
Far-western	3,532	2,403	21,859	9.09
Nepal	35,578	23,609	263,710	11.17

Source: Ministry of Agriculture and Co-operatives, 2011

Table 1.2 Total area, productive area, production and productivity of commonly grown Citrus.

SN	Fruit crops	Area (ha)	Productive area (ha)	Production (mt)	Productivity (mt/ha)
1	Mandarin	22872	14913	179494	12.0
2	Sweet Orange	5365	4089	50679	12.73
3	Lime	5276	2731	22571	8.3
4	Lemon	1206	1087	6432	5.9
5	Others	859	789	4534	5.7
	Citrus total	35578	23609	263710	11.2
	All fruit total	117932	79184	794164	10.0
	Citrus % of all fruits	30.2		33.2	

Source: Ministry of Agriculture and Co-operatives, 2011

Table1.3 Area coverage (ha) of different Citrus spp. in 2010/11.

Developmental region	Citrus spp.					
	Mandarin	Sweet orange	Acid lime	Lemon	Other	Total
Eastern	3510	720	1141	112	697	5546
Central	2330	2131	468	134	151	5214
Western	5452	373	530	633	399	7387
Mid-western	2374	280	252	102	51	3059
Far-western	1247	585	340	106	121	2403
Nepal	14913	4089	2731	1087	789	23609

Source: Ministry of Agriculture and Co-operatives, 2011

1.1.3. Citrus spp. found in Nepal

Nepal is rich in Citrus diversity and many of these species might have their origin in Nepal itself or have been introduced at various times in the history. Lama and kayastha (1999) reported 14 species of Citrus from Pokhara and its surrounding areas. The major *Citrus* spp. in this area were Citrus *reticulata, C. sinensis, C. pseudolemon and C. aurantifolia*. Twelve species of Citrus are available in National Citrus Development Program (Kirtipur), namely, *Citrus reticulata*, *C. unshiu*, *C. sinensis*, *C. grandis*, *C. paradise*, C. lemon, *C. aurantifolia*, *Fortunella japonica*, *Poncirus trifoliate*, *Citrange* etc. Mandarin, sweet orange and acid limes are produced in commercial scale, of which mandarins grown in Nepal are excellent in term of size and quality (NCRP, 2012). Different species of Citrus have been reported from Nepal (including exotic spp.), *viz*, Citrus *reticulata* Blanco, *C. sinensis* Osbeck. *C. aurantium* Linn, *C. aurantifolia* Swingle, *C. pseudolemon* Tanaka, *C. limon* L, *C. grandis* Osbeck, *C. maxima* (Burm.) Merrill, *C. jambhiri* Lush, *C. medica* Linn, *C. limettoides* Tanaka, *C. nobilis* Xc, *C. unshiu* M, *Poncirus trifoliate* L, *Fortunella japonicum* Swingle and possible hybrids (locally known as Chaku and Narayani) (Gmitter and Hu, 1990). *Citrus* spp. mainly found in Nepal along with their local and common names are shown in Table 1.4.

Table 1.4 List of *Citrus* spp. found in Nepal.

S.N	Nepali name	English name	Scientific name
1	सुन्तला	Mandarin	Citrus *reticulate* Blanco.
2	जुनार	Sweet Orange	Citrus *sinensis* Osbeck.
3	कागती	Lime	Citrus *aurantifolia* Swingle
4	मुन्तला	Kumquate	*Fortunella margarita* Swingle (oval type) *Fortunella japonica* Swingle (round type)
5	भोगते	Pummelo	Citrus *grandis* Osbeck.
6	निबुवा	Hill Lemon	Citrus *limon*
7	बिमिरो	Citron	Citrus *medica* Linn.
8	चाक्सी	Sweet Lime	Citrus *limettioides* Tanaka
9	काली ज्यामिर	Sour Orange	Citrus *aurantium* Linn.
10	नाईटे ज्यामिर	Rough Lemon	Citrus *jambhiri* Lush.
11	संकत्रा	Possible hybrid	
12	काठे	Possible hybrid	
13	कमला	Possible hybrid	
14	नरायणि	Possible hybrid	
15	चकुपांउ	Possible hybrid	

16	उन्सु	Satsuma mandarin	Citrus *unshiu*
17	कर्नाखट्टा	Karna Khatta	Citrus *Karna* Raff
18	रङ्गपुर कागती	Rangapur Lime	Citrus *limonia* Osbeck.
19	ग्रेप फ्रूट	Grape fruit	Citrus *paradise* Macf.
20	क्यालामण्डिन	Calamondin	Citrus *medurensis* Lou.

Source: Nepal Citrus Development Programme, Kirtipur, Kathmandu (2010/11)

1.1.4. Nutritional and Economic Importance of Citrus

Citrus fruits contribute augmenting food, improvement in nutrition, generation of employment and income, and it also help in maintaining healthy environment. Citrus fruits and fruit juice are being refreshing and good source of vitamins, especially Vitamin-C and Vitamin B complex play an important role in human nutrition, and these are useful for the treatment of diseases arising from vitamin deficiencies including cough and cold (Cox, 1974). Citrus fruits are well-known for their dietary, nutritional, medicinal and cosmetic properties and are also good sources of citric acid, flavonoids, phenolics, pectins, limonoids, ascorbic acid, etc. (Dugo and Di Giacomo, 2002). Fruit juice contains sugars (glucose and sucrose) and acid (mainly citric acid and a little of malic acid) (Aubert et al., 1990). Besides nutritive value of fruit pulp, rind is rich in pectin, glucosides (hesperidin in oranges and lemons, naringins in grapefruit and pummelo) and essential oils. Major constituents of Citrus are carbohydrates, vitamins, acids, nitrogen compounds, enzymes, pigments, lipids and volatile compounds (Shrestha, 1999). In developed country, per capita consumption of Citrus fruit is about 10 kg/year where as in Asian countries, it is only about 4 kg/year (Aubert et al., 1990).

Various citrus components are currently being studied with respect to health, including phenolics, carotenoids, limionoids, vitamin C, water soluble B-complex vitamins, specifically folic acid and thiamine (vitamin B), Citrus pectins as a dietary fiber, therapeutic value of the Citrus flavonoids etc (Lallan and Singh, 2006). Iron found in the citrus constitutes about 1.3mg/450g of edible part (Shah, 1992). Citrus fruit is given to sick people with jaundice and high fever, and for curing diseases like dysentery and beriberi, while bitter glucoside “Naringin” provides prevention against malaria (Radha and Mathew, 2007). The carotenoids might involve as first line of defense against ROS (Reactive Oxygen Species), anthocyanin had ability to modulate capillary permeability and resistance (Lallan and Singh, 2006).

1.1.5. Problems in Citrus Diversity Assessment

In Citrus, there are frequent incidences of hybridization, apomixis (parthenogenesis or embryo developed from an unfertilized egg), polyploidy and bud mutation. This makes

the taxonomy of the genus Citrus chaotic and creates problems in the genetic diversity assessment (Kumar et al., 2010). Morphological markers alone have serious limitation for diversity assessment as they are influenced by the environment. Biochemical markers are powerful than morphological but these are limited in numbers for the assessment. On the other hand, molecular markers are abundant, superior and give unbiased estimation but different molecular markers also have their own limitations (Semagn et al., 2006).

1.1.6. Importance of Diversity Assessment

In particular, an adequate knowledge of existing genetic diversity, where in plant population it is found and how to best utilize it, is of fundamental interest for basic science and applied aspects like the efficient management of crop genetic resources. The improvement of crop genetic resources is dependent on continuous infusions of wild relatives, traditional varieties and the use of modern breeding techniques. These processes require an assessment of diversity at some level, to select resistant, highly productive varieties (Mondini et al., 2009). In general, improvement in the yield of a crop species is accompanied with the reduction in variability among the cultivated varieties of that species leading to the genetic erosion which in turn limits further improvement in the species. Germplasm collections aim at minimizing the detrimental effects of genetic erosion by collecting and preserving the variability in crops and their related species. For this, genetic diversity assessment plays an important role. The narrow genetic base in improved crops due to commonness of one or more parents in their ancestry leads to genetic vulnerability. Such worst conditions can be avoided by using genetically appropriate diverse and unrelated parents in breeding programs (Singh, 2005). Genetic diversity assessment is required for understanding the distribution and extent of genetic variation within and between species. Knowledge of crop species evolution and of the relationship between crop species and their wild relatives has been of considerable value for the conservation and use of plant genetic resources (Bhattacharya and Dutta, 1996). Understanding genetic variability in citrus is critical for characterizing germplasm, controlling genetic erosion and the registration of new cultivars (Herrero et al., 1996; Barkley et al., 2006).

1.1.7. Methods of Diversity Assessment

Analysis of the genetic diversity of Citrus fruit is crucial. Assessment of genetic diversity is the basis for breeding, variety development and conservation of the genetic resources. Traditionally, evaluation of genetic diversity and characterization of varieties have been based on morpho-physiological traits such as color, size and maturity period,

which are adapted by most of the Citrus breeders. In recent years, development of DNA marker technology has provided efficient tools to evaluate and measure the genetic diversity and cultivar identification (Ni et al., 2002).

Morphological Markers

They generally correspond to the visually scorable qualitative traits and have been found in nature or as a result of mutagenesis experiments. They are usually dominant or recessive (Chawla, 2002). Morphological traits controlled by single locus can be used as genetic markers provided their expression is reproducible over a range of environment. Their number is very limited and can't distinguish the heterozygous from homozygous individual (Lalitha, 1999). The morpho-agronomic marker may not reflect true genetic identities and diversities (Ferdinandez et al., 2001).

Molecular Markers

Introduction of Polymerase chain reaction technology has clearly changed the way in which molecular data are gathered for applications ranging from nucleotide sequencing and phylogenetic reconstruction to population level studies. They have great potential for studies of hybridization and introgression, conservation biology and phylogenetic (Soltis et al., 1998). Understanding taxonomy, phylogenetic relationships, and genetic variability in citrus is critical for determining genetic relationships, characterizing germplasm, controlling genetic erosion, designing sampling strategies or core collections, establishing breeding programs, and the registration of new cultivars (Herrero et al., 1996). As DNA based markers are phenotypically neutral, abundant, less subject to environmental effects, and developmental stages of plant, they are preferred over conventional morphological and biochemical markers for genetic diversity studies. A range of (PCR) - based molecular marker techniques are now available for the assessment of genetic diversity. ISSR and SSR markers are commonly used molecular marker in diversity assessment.

ISSR (Inter Simple Sequence Repeat) marker

ISSR marker is a simple and quick method that combines most of the advantages of microsatellites (SSRs) and amplified fragment length polymorphism (AFLP) to the universality of random amplified polymorphic DNA (RAPD). It involves amplification of DNA segments present at an amplifiable distance in between two identical microsatellite repeat regions oriented in opposite direction (Semagn et al., 2006). Its attraction is due to their high frequency of polymorphism, technical simplicity, cost effectiveness, requirement of few nanogram of DNA, no requirement of prior DNA sequence

information and feasibility of automation (Reddy et al., 2002; Weising et al., 2005; Kurane et al., 2009).

SSR (Simple Sequence Repeat) marker

Microsatellites or SSRs are mono, di, tri, tetra and penta nucleotide repeat units abundantly found in eukaryotic genomes. They have gained considerable importance in plant genetics and breeding owing to many desirable genetic attributes including hypervariability, multiallelic nature, codominant inheritance, reproducibility, relative abundance, extensive genome coverage including organelles genomes, chromosome specific location and amenability to automation and high throughput genotyping (Kalia et al., 2011).

1.2. Justification

Citrus is an economically important crop and cultivated in an ever-widening area since ancient times that has long been the object of intensive breeding programs aimed at improvement of fruit quality, and obviously it seems that a systematic program of genetic analysis is not only desirable but required if progress is to be made in solving taxonomic and breeding programs in Citrus (Torress et al., 1978). Assessment of genetic diversity using morphological markers alone has serious limitations, especially in species of a complex genus like Citrus, whose taxonomy is otherwise chaotic due to frequent incidences of inter-specific and inter-generic hybridization, apomixes, long history of cultivation, polyploidy and bud mutations (Shahsavar et al., 2007; Kumar et al., 2010). Besides, the approach of molecular fingerprinting is complementary to phenotypic measures in quantifying genetic changes as it shows variations in DNA that may not be phenotypically expressed (Shehata et al., 2009).

Nepal is the home land of Citrus spp. because of wide availability and cultivation of wild *Citrus spp.*, but there are limited work carried out on evaluation and characterization of citrus genetic materials (Verma, 1994). Despite the commercial and nutritional importance of citrus, very scarce studies employing DNA based marker systems to assess genetic diversity of citrus have been attempted in Nepal. In Nepal, Citrus production period is confined between Kartik and Poush (Oct and Jan), but the demand is throughout the year. Significant national requirement of Citrus fruits are fulfilled from India even in seasonal period because of low production, narrow production period and poor infrastructure facilities for storage (NCRP, 2012). Study on genetic diversity using molecular marker can efficiently provide information about history and biology of cultivars and clue for identification of genetically diverse cultivars which will help in

production of the elite varieties for off-season production, so that production period of Citrus fruits can be prolonged and made available to market for longer periods (Metais et al., 2000). Knowledge of relative genetic relationships among genotypes is useful in breeding programs to organize germplasm resources and will be potential source for selection of elite cultivars for the breeding and variety development programs. Therefore, assessment of genetic diversity within and between species of citrus using SSR (Simple Sequence Repeats) marker may provide guidelines to select parental genotypes for breeding of elite cultivars and variety development programs, germplasm conservation, solving classification problems as well as sustain genetic diversity of *Citrus spp.* in Nepal.

Acid lime (Citrus *aurantifolia* Swingle) is an important commercial fruit crop, cultivated in Terai to high hills of Nepal. Production and quality are very low and decreasing in every generation due to low quality planting materials, poor orchard management and high level of genetic erosion (NCRP, 2012; Shrestha et al., 2012a). Till now there is no appropriate high yielding variety of acid lime. A survey conducted in 14 major cities of Nepal showed that 94.5% (1875.0 tons) of lime sold from Kalimati market, one of wholesale market at Kathmandu was imported from India and likewise 68% of the lime sold in rest of 13 cities was imported from India (Dhakal and Bhattarai, 2002). Since it is a perennial crop, variety development through breeding is more tedious and time consuming. High variation was observed within the acid lime genotypes in Nepal at different altitudinal range (Sapkota, 2006; Shrestha et al., 2012b). Study of various characters at molecular level using molecular marker techniques is more reliable and appropriate for selection of superior parental genotypes for breeding of elite cultivars and variety development programs. The genetic diversity of acid lime at different agro-ecological zones in Nepal has been carried out at molecular level using co-dominant SSR (Simple Sequence Repeats) marker (Shrestha et al., 2012a). In this context, assessment of genetic diversity of acid lime landraces using the dominant ISSR (Inter-Simple Sequence Repeat) marker would be desirable to compare the results with that of SSR marker and for the selection of genotypes for future breeding programs.

A range of PCR-based molecular marker techniques are now available for the assessment of genetic diversity. Microsatellites or SSRs are mono, di, tri, tetra and penta nucleotide repeat units abundantly found in eukaryotic genomes. Variation in number of repeat motifs of each of these repeat units varies in different cultivars and species and forms the basis of genetic diversity assessment. They are codominant so able to distinguish heterozygotes and homozygotes. ISSR is based on sequence variation in the genomic regions between Simple Sequence Repeats (SSR) or Microsatellites. However,

they are inherited in dominant Mendalian fashion. These methods have wide range of applications including the characterization of genetic relatedness among populations, genetic fingerprinting, gene tagging, Marker Assisted Selection (MAS) and breeding, detection of clonal variation, cultivar identification, phylogenetic analysis, detection of genomic instability of hybridization etc.

1.3. Objectives

The overall objective of this study is to assess genetic diversity of acid lime landraces found in different altitudinal gradients of Eastern Nepal and Citrus spp. collected from NCRP (National Citrus Research Program), Pripatle, Dhankuta and from Kathmandu valley using PCR-based ISSR and SSR molecular marker techniques respectively. Specific objectives of this project are:

a. Genetic study of Acid lime using ISSR markers
   i. Extraction and quantification of DNA from all collected samples.
   ii. Optimization of PCR reaction and cycling conditions for ISSR-PCR amplification.
   iii. Genetic diversity assessment using NTSYS-PC ver. 2.21i and MVSP ver. 3.21.
   iv. Analysis of Molecular Variance (AMOVA) using GenAlEx ver. 6.5.

b. Genetic study of *Citrus* spp. using SSR markers
   i. Extraction and quantification of DNA from all collected samples.
   ii. Optimization of PCR reaction and cycling conditions for SSR-PCR amplification.
   iii. Genetic diversity assessment using NTSYS-PC ver. 2.21i and MVSP ver. 3.21.
   iv. Genetic analysis using GenAlEx ver. 6.5 and PowerMarker ver. 3.25.

1.4. Scope

Citrus spp. in Nepal exhibits a wide range of morphological diversity. The present research is helpful in providing valuable information on the genetic diversity of different *Citrus* spp. as well as acid lime landraces of eastern Nepal. Data on the distribution of genetic variability and relationships among species could be used to design sampling strategies or core collections in which agronomically important genes would be searched and investigated. It is believed to be useful for individual researchers, governmental and non-governmental agencies and institution working in this field. The mapping of microsatellite markers and quantitative traits would provide a tool for marker assisted

breeding. This study will provide important information for choosing parental lines for breeding programs in a germplasm collection, classification of plant germplasm accessions, and further curation and acquisition of new plant germplasm accessions. Furthermore, the information will also useful for the characterization of accessions in the plant germplasm collections and taxonomic studies.

CHAPTER 2: LITERATURE REVIEW

2.1. World Production, cultivation and origin of *Citrus*

Citrus is the most produced fruit in the world with over the 116 million tons of production (Isabel et al., 1999). The commercially important *Citrus* spp. and many related genera of the subfamily Aurantioidae, family Rutaceae, are indigenous to Sourthern Asia bordered by Northeast India, Myanmar (Burma) and the Yunnan province of China (Gmitter and Hu, 1990; Kochhar, 1998). About 78 species of citrus coming under this family have their origin in India (Hooker, 1872; Bhattacharya and Dutta, 1996). North West India is the place of origin of citron (*Citrus medica* L.), the lime (*C. aurantifolia* Swingle) originated thousands of years ago in Malaysia and then enters the Thailand, India and other countries while lemon (*C. limon* (L.) Burm. f.) were originated in Nepal/India through natural crossing between citron and lime. The genera *Pocirus* and *Fortunella* are native to China (Subedi, 2010). Sweet orange (*C. sinesis* Osbeck) originated in the Sourthern China from where it was spread and introduced in India. Pommelo (*C. grandis* Osbeck) originated in Fijji Island and in China. Grapefruit (*C. paradise* Macf) is a native of West Indies (Reuther et al., 1967-1989). Thus, except grapefruit, most of the important species in the genus *Citrus* and the related genera originated in the old world. Introduction and spread of these species to new world started during the beginning of Christian-era (Gmitter et al., 1992; Radha and Mathew, 2007). The table 2.1 shows the origin of different *Citrus* spp.

Table 2.1 Name of the *Citrus* spp. and their place of origin

Citrus spp.	Place of Origin	Inferences
Citrus medica	India	Natural/True Species
C. aurantium	China	Hybrid (mandarin x maxima)
C. sinensis	China	Hybrid (mandarin x maxima)
C. maxima	China, Malaysia	Natural/True Species
C. reticulata	China	Natural/True Species
C. aurantifolia	Malaysia	Natural
C. limon	India, Nepal	Hybrid (citron X lime)
C. paradise	West Indies	Mutant of maxima
Poncirus trifoliata	Northern China	Natural
Fortunella species	West-East China	Natural

Source: Subedi (2010)

Among the most common citrus fruits, sweet orange occupy nearly 2/3rd of world's total area. They are grown in about 114 countries of the world (FAO, 2002). The most significant production areas are found in the Americas (led by Brazil, the United States, Mexico, and Argentina), the Mediterranean basin (led by Spain, Italy, Egypt, and Turkey),

and the south and east Asian regions (led by China, India, and Japan) (Talon and Fred, 2008). The leading centers producing citrus are Italy, Sicily, Spain, Greece, Argentina, Brazil, Mexico, Japan, China, Israel, India and Australia. About 90% of grapefruits of the world are produced in Florida. Italy leads in lemon, Mexico and India are main producer of acid lime, while Japan mainly grows mandarin (FAO, 1998). The top five citrus producing of the world in 2010 are shown in table 2.2.

Table 2.2 Top five citrus producing countries of the world with their production and productivity, 2010

Country	Productive Area (ha)	Total Production (mt)	Productivity (ha/mt)
China	174750	4886900	27.96
India	55500	764300	13.77
Mexico	21200	103600	4.86
Iran	6400	87000	13.59
USA	2100	43700	20.80
World	1245710	11763628	9.44

Source: FAO, 2010

2.2. Citrus and its cultivars

Citrus known as “sour fruits” are members of the family Rutaceae. These plants are large shrubs or small trees, reaching 5-15 m tall, with spiny shoots and alternately arranged evergreen leaves with an entire margin. The flowers are solitary or in small corymbs, each flower 2-4 cm in diameter, with five (rarely four) white petals and numerous stamens. The fruit is a hesperidium, a specialized berry, globose to elongated, filled with pulp vesicles. Citrus fruits are refreshing, delicious to eat and provide minerals and vitamins (i.e rich source of vitamin C). The fruits can be made to make cold drinks, squashes, jam, jelly and marmalade (Helgi et al., 2005).

Citrus *sinensis* Osbeck (Sweet Orange)

The sweet orange is the fruit of the Citrus spp., *Citrus sinensis* in the family Rutaceae. The fruit of the Citrus *sinensis* is called sweet orange to distinguish it from that of the Citrus *aurantium*, the bitter orange. The orange is a hybrid, possibly between pomelo (Citrus *maxima*) and mandarin (Citrus *reticulata*), cultivated since ancient times (Nicolosi et al., 2000). The sweet orange (tight skin orange) requires dry and semi-arid conditions coupled with distinct summer and winter with low annual precipitation. In general, low humidity and severe winter result in good color development and external appearance in fruits whereas high humidity favors thin skin and plentiful juice (ICIMOD, 2003). Orange trees are widely grown in tropical and subtropical climates for its sweet fruit, which can be eaten fresh or processed to obtain juice, and for its fragrant peel. In Nepal,

it ranks second after mandarin in terms of area and production, and cultivated in more than two dozen districts where Ramechhap and Sindhuli together accounts more than 50 % of total production in Nepal (NCRP, 2012).

Citrus *reticulata* Blanco (Mandarin)

The Mandarin (Citrus *reticulata*) is an indigenous fruit and known by its local name in different places, most important Citrus spp. of Nepal, covers nearly 58 % of Citrus area in the country (Paudyal and Chalise, 2007). Mandarin group includes all types of loose skin oranges locally called as Suntala. Flowering starts from March, which lasts for about a month and need of pollinators is critical for setting good fruits and higher productivity (ICIMOD, 2003). The mandarin is easily peeled with the fingers, starting at the thick rind covering the depression at the top of the fruit, and can be easily split into even segments without squirting juice. Unlike sweet orange, the genetic variation in the mandarin group is associated with sexual hybridization among a great number of species and interspecific hybrids (Cameron and Frost, 1968).

Citrus *aurantifolia* Swingle (Acid lime)

Limes comprise a varied group of types – both acid and sweet varieties. There are two kinds of acid limes, the small-fruited Mexican (West Indian, Key) type, *C. aurantifolia* Swingle, and the large-fruited Tahiti (Persian, Bears) lime, *C. lantifolia* Tan., which is triploid and therefore seedless. *C. limettioides* Tan., the sweet lime, is commonly referred to as the Indian or Palestine sweet lime, is native to north-eastern Indian. *C. aurantifolia* is native to Malaysian region of South-Western Asia, while *C. lantifolia* probably originated in the East and then spread to Persia, and then to Tahiti, probably via Brazil and Australia, and finally to California (Khan, 2007).

The Key lime (*C. aurantifolia*) is a Citrus spp. with a globose fruit, 2.5–5 cm in diameter (1–2 in) that is yellow when ripe but usually picked green commercially. It is smaller and seedier, with a higher acidity, a stronger aroma, and a thinner rind. It is valued for its unique flavor compared to other limes, with the Key lime usually having a more tart and bitter flavor. The Key lime (Citrus *aurantifolia* Swingle) is also known as West Indian lime, bartender's lime, Omani lime, acid lime or Mexican lime. *C. aurantifolia* is a shrubby tree, to 5 m (16 ft), with many thorns. Dwarf varieties exist which can be grown indoors during winter months and in colder climates. Its trunk rarely grows straight, with many branches, often originating quite far down on the trunk. The leaves are ovate, 2.5–9 cm (1–3.5 in) long, resembling orange leaves (the scientific name *aurantifolia* refers to this resemblance to the leaves of the orange, *C. aurantium*) (Morton, 1987).

It is one of the important commercial fruit crops of Nepal, which ranks third after mandarin and sweet orange in terms of area under coverage and production (NCRP, 2012). It is cultivated in 60 out of 75 districts of Nepal. The cultivation range of acid lime fruits are scattered at different altitudinal range in Nepal from Terai to high hills and east to west (Shrestha et al., 2012a). It is commonly known as "Kagati". Cultivation range of acid lime in Nepal is 800 m to 1400 m als in the mid hills stretching from east to west, but potentiality of cultivation range could be much wider from 125 m asl Terai to 1800 m asl in high hills of Nepal and normal production period is limited between Kartik and Poush (Dhakal and Bhattarai, 2002). Three cultivars of lime have been grown in Terai area, i.e. acid lime (Pahade Kagati or Sun Kagati), eureka (Chasme Kagati) and natural hybrid types (Paudyal and Shrestha, 2004). Among them acid lime has high commercial value in the market due to better aroma, appropriate size and medicinal values. It is considered as a high value commodity and has been given the number one priority by Master Plan for Horticulture Development (Bhattacharya and Dutta, 1996). In Nepal, Eastern Development Region (EDR) occupies the highest position in terms of area and production of acid lime (MoAC, 2011). It is used as juice, dessert, pickle and other medicinal purpose and contains important limiting vitamin like vitamin-C and minerals like calcium and iron, so it can overcome the malnutrition problem. It works against asthma and also as antidepressant, stress relief, colds, flu, fever, nosebleeds, mouth ulcers, throat infections and boils (Reuther et al., 1967; Tisserland, 1995; Shrestha et al., 2012a).

C. limon Burm. (Lemon)

It is assumed that lemons derive from citrons (Khan, 2007). Lemon fruit quality is excellent in semi-arid irrigated areas and coastal areas. In humid tropics, lemon trees produce fruit with coarser peels. The fruit size and shape of lemon varies greatly from spherical in Baramasi lemon or Eureka to oblong in Assam lemons and fruits develop nipples or mammila at the stylar end, which is almost absent in some cultivars and prominent in others (Ladaniya, 2008). It was reported that lemons were thought to be natural hybrids of a citron and a lime (Scora, 1975; Barett and Rhodes, 1976).

C. paradise Macfadyen (Grapefruit)

Grapefruit is appears to have originated as a mutation or hybrid of the Shaddock in the West Indies perhaps Barbados (Khan, 2007). It derived its name as fruits are borne in clusters like grapes. This is one of the large citrus fruits (although smaller than pummelos) with diameters of 10–15 cm. This is a typically tropical fruit with a high heat-summation requirement. People in North America, Europe, and Japan prefer fresh

grapefruit. There are red-fleshed and white fleshed cvs. Red flesh is due to lycopene and carotenoid pigments (Ladaniya, 2008).

C. grandis or *C. maxima* Osbeck (Pummelo or Shaddock)

It is thought to be originated in southern China and found growing wildly in China, mainly north-eastern states of India, and also in the Malay region of Southeast Asia. Fruits and flowers of pummelo are the largest among commercially grown citrus in the world. Fruits that are 20–25 cm in diameter is common, even up to 30 cm in diameter can be seen. Flesh of the pummelo is firm with crisp carpellary membranes and juice sacs (Ladaniya, 2008).

C. medica Lin. (Citron)

Citron fruits are oblong in shape like some of the lemons (long-oval or ellipsoidal to obovate in shape), but fruits are very seeded, quite large with yellow colored very thick, rough and bumpy peel at maturity. These are grown since ancient times in India and considered to be native to India. In Assam state, it is widely available and known as Bira-Jira or Bakel-Khowa-Tenga. The Etrog or Ethrog citron is widely cultivated in Israel due to its importance in Jewish religious ceremonies (Ladaniya, 2008).

2.3. Problems in the Classification *Citrus* spp.

Citrus and its close relatives are represented by 28 genera in the tribe Citreae of the subfamily Aurantioideae in the family Rutaceae (Swingle and Reece, 1967). There are currently two commonly used classifications of Citrus *viz*. 1) Swingle (Swingle and Reece, 1967) and 2) Tanaka (Tanaka, 1977) who recognized 16 and 162 species of citrus respectively. However, there are only three basic true species of Citrus within the subgenus Citrus as defined by Swingle: citron (*C. medica*), mandarin (*C. reticulata*), and pummelo (*C. maxima*) and other cultivated *Citrus* spp. within the subgenus Citrus are believed to be hybrids derived from these true species, species of the subgenus Papeda, or closely related genera (Scora, 1975; Barett and Rhodes, 1976). In recent years, this idea has gained support from data derived from molecular markers (Federici et al., 1998; Nicolosi et al., 2000). The geographic distribution of *C. halimii* in Thailand and Malaysia (Karp et al., 1996) is not coincident with the distribution of the other species of Citrus studied, and this species has been proposed as the fourth basic species of Citrus (Lynch and Milligan, 1994; Kovach, 2007). The difficulty in classifying Citrus taxa is mainly due to repeated cross-pollination and to adventitious nucellar embryony, which stabilizes and perpetuates hybrid taxa (Scora, 1975). Selection and propagation of many natural or man-made hybrids and many mutants during a long history of cultivation makes Citrus

classification even more complex (Fang et al., 1998). Another problem in Citrus taxonomy is disagreement on what degree of divergence justifies species status, and whether apparent hybrids among naturally occurring forms should be assigned species rank (Roose et al., 1995). The research on germplasm characterization, genetic variation and breeding in oranges and other Citrus spp. have been hampered because of characteristics related to the reproductive biology of these species, i.e. high inter-specific fertility, apomictic reproduction, polyembryony, a long juvenile phase and a paucity of polymorphic DNA markers (Bretó et al., 2001; Corazza-Nunes et al., 2002).

Botanical classification of Citrus:

Order	-	Geraniales (21 families)
Sub-order	-	Garaniinea (12 families)
Family	-	Rutaceae (7 sub-families)
Sub-family	-	Aurantoidae (2 tribes)
Tribe	-	Citrae (3 sub-tribe)
Group C	-	Citrus fruit trees (6 genera)
Genus	-	1. Poncirus (1 species)
		2. Fortunella (4 species)
		3. Citrus

Classification of Genus Citrus:

Classification of Citrus fruits is a very problematic one, which awaits still perfect solution. The genus Citrus contains many kinds or types that differ as to their fruits, flowers, leaves, and twigs. Currently, there are 2 outstanding systems of classification for Citrus. They are those of Walter T. Swingle, a USDA scientist who did much of his work in Florida, and Tyosaburo Tanaka of Japan. Swingle’s system is relatively simple, containing 16 species. He is commonly referred to as a “lumper” because he lumps a large number of kinds into a relatively small number of groups. Tanaka’s system initially included 145 species and he is known as a “splitter” because he has split the genus Citrus into many small groups. He has continued to add to this list (Krezdorn). Swingle divides the genus Citrus into two sub-genus i.e. the sub-genus *Papeda* containing six species and the sub-genus Citrus (*Eu*Citrus) with ten species while Tanaka classify the genus Citrus into two sub-genus *Archi*Citrus in 5 sections and *Meta*Citrus in 3 sections. These sections were further divided in 13 sub-sections, eight groups, and two sub-groups. Two micro-groups and total 145 species were described. In 1961, he added two new sub-sections, one new group and 12 new species, and becomes 157 species. Further refinement continued and he purposed 162 species to genus Citrus (Swingle and Reece, 1967; Tanaka, 1977; Singh and Naqui, 2001).

2.4. Genetic Diversity Assessments by Genetic Markers

Understanding the molecular basis of the essential biological phenomena in plants is crucial for the effective conservation, management, and efficient utilization of plant genetic resources (PGR). Within the last twenty years, molecular biology has revolutionized conventional breeding techniques in all areas. Biochemical and Molecular techniques have shortened the duration of breeding programs from years to months, weeks, or eliminated the need for them all together. The use of molecular markers in conventional breeding techniques has also improved the accuracy of crosses and allowed breeders to produce strains with combined traits that were impossible before the advent of DNA technology (Stuber et al., 1999). Various genetic markers have been used for the assessment of genetic diversity of plant species. Genetic markers have been considered as the specific locations on the chromosome which serve as landmarks of genome analysis. A marker must be polymorphic, that is, it must exist in different form. This polymorphism in the marker can be detected at three levels: phenotype (morphological), differences in proteins (biochemical), or differences in the nucleotide sequence of DNA (molecular) (Chawla, 2002).

2.4.1. Morphological marker

Morphological markers are generally corresponds to the qualitative traits that can be scored visually. They are usually dominant or recessive. The genetic markers used in plants are those affecting morphological characters including genes for dwarfism, albinism, and altered leaf morphology (Chawla, 2002). Unfortunately, there are several undesirable factors that are associated with morphological markers. First, they are highly influenced by the environmental factors. Second, these mutant traits often have undesirable features such as dwarfism or albinism. And lastly, performing breeding experiments with these markers is time consuming, labor intensive and the large populations of plants required need large plots of land and/or greenhouse space in which to be grown (Stuber et al., 1999). Use of morphological traits may be helpful but often inadequate in differentiation of closely related cultivars. On the other hand, certain morphologically different variants may be phylogeneticly closely related. Therefore they are not so reliable in genetic diversity assessment (Chawla, 2002).

2.4.2. Biochemical marker

Markers that reveal polymorphisms at the protein level are known as biochemical markers. Biochemical markers are proteins produced as a result of gene expression which can be separated by electrophoresis to identify the alleles. The most commonly used protein markers are isozymes which are variant forms of the same enzyme

(Vodenicharova, 1989). The enzymes are extracted, and run on denaturing electrophoresis gels. The denaturing component in the gels (usually SDS, sodium dodeacylsulphate) unravels the secondary and tertiary structure of the enzymes and they are then separated on the basis of net charge and mass. Polymorphic differences occur on the amino acid level allowing singular peptide polymorphism to be detected and utilized as a polymorphic biochemical marker. Biochemical markers are superior to morphological markers in that they are generally independent of environmental growth conditions (Stuber et al., 1999). Protein markers reveal differences in the gene sequence and function as co-dominant markers. However, their use is limited due to their limited number in any crop species and also because they are subject to post-translational modifications (Staub et al., 1982). Isozymes are convenient genetic markers for plants breeders and plant taxonomists. They have been used as biochemical markers to study taxonomic relationships between *Citrus* spp. (Rahman and Nito, 1994; Protopapadakis and Papanikolaus, 1999). This method is useful in distinguishing cultivars derived by sexual reproduction but does not generally distinguish between cultivars differentiated by mutations (McDermott and Mcdonald, 1993; Karp et al., 1997a; Rahman et al., 2001). Its application to Citrus genetics proven reliable tools for monitoring *in vitro* culture tissue, distinguishing nucellar from zygotic plants, and helpful for assessing Citrus genetic organization, cultivar identification, heterozygosity, and phylogeny (Ollitrault, 1990).

2.4.3. Molecular marker

A molecular marker is a DNA sequence readily detected and whose inheritance can be easily monitored. It can be defined as a genomic locus, detected through probe or specific starters (primer) which, in virtue of its presence, distinguishes unequivocally the chromosomic trait which it represents as well as the flanking regions at the 3' and 5' extremity (Barcaccia et al., 2000). The use of molecular marker is based on naturally occurring DNA polymorphism (i.e.: base pair deletions, substitutions, additions or patterns) (Gupta et al., 1994). Genetic polymorphism is defined as the simultaneous occurrence of a trait in the same population of two or more discontinuous variants or genotype (Chawla, 2002). Molecular markers are superior to other forms of MAS (Marker Assisted Selection) because they are relatively simple to detect, abundant throughout the genome even in highly bred cultivars, completely independent of environmental conditions and can be detected at virtually any stage of plant development. Molecular markers are imperative in the assessment of genetic diversity, ranging from nucleotide level (SNPs) to gene and allele frequencies (genotype information), and devising various germplasm conservation programs. Molecular

markers are also helpful in studying population structure, authentication of germplasms, solving taxonomic problems, and assigning plants to their correct taxonomic hierarchies which are important for phylogenetic studies (Slatkin, 1987). Molecular markers may or may not correlate with phenotypic expression of a genomic trait. They offer numerous advantages over conventional, phenotype-based alternatives as they are stable and detectable in all tissues regardless of growth, differentiation, development, or defense status of the cell. Additionally, they are not confounded by environmental, pleiotropic and epistatic effects. An ideal molecular marker should possesses the following features: (1) be polymorphic and evenly distributed throughout the genome; (2) provide adequate resolution of genetic differences; (3) generate multiple, independent and reliable markers; (4) be simple, quick and inexpensive; (5) need small amounts of tissue and DNA samples; (6) link to distinct phenotypes; and, (7) require no prior information about the genome of an organism. Nevertheless, no molecular marker presents all the listed advantages. The different methods of molecular assessment differ from each other with respect to important features such as genomic abundance, level of polymorphism detected, locus specificity, reproducibility, technical requirements and cost. Depending on the need, modifications in the techniques have been made, leading to a second generation of advanced molecular markers (Barcaccia et al., 2000). Recently, various DNA markers have been used extensively to study phylogenetic relationships among plants (Whitkus et al., 1994). They include the use of restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), microsatellite or simple sequence repeat (SSR) and inter-simple sequence repeat (ISSR) markers. Some of these methods are labor intensive and time consuming while some others may be less reproducible or require a previous knowledge of the sequence of the target DNA (MacPherson et al., 1993; Fang et al., 1997). Nevertheless, some of these methods have been used in phylogenetic studies of Citrus variants (Green et al., 1986; Kijas et al., 1995; Fang et al., 1997; Kijas et al., 1997). The different molecular techniques employed are either non-PCR (hybridization based) based or PCR based or both or sequencing based techniques (Joshi et al., 1999; Spooner et al., 2005; Weising et al., 2005; Semagn et al., 2006).

Hybridization-Based marker

e.g., RFLP (Restriction Fragment Length Polymorphism), VNTR (Variable Number Tandem Repeats)

PCR-Based marker

Various PCR-based techniques are of two types depending upon the primer used for amplification:

i) Arbitrary or semi-arbitrary primed PCR techniques that developed without prior sequence information (e.g., AP-PCR, DAF, RAPD, AFLP, ISSR).
ii) Site-targeted PCR techniques that developed from known DNA sequences (e.g., EST, CAPS, SSR, SCAR, STS).

DNA Sequencing Based Markers

e.g., ITS (Internal Transcribed Spacer) and SNP (Single Nucleotide Polymorphism)

2.4.3.1. Hybridization-Based marker

This include RFLPs (restriction fragment length polymorphism) and VNTR (variable number tandem repeats) loci where probes such as random genomic clones, cDNA clones, and probes for microsatellite and mini-satellite sequences are hybridized to filters containing DNA which has been digested with restriction enzymes .The polymorphisms in the case of VNTR loci is due to a difference in the number of repeats, while RFLPs are generated due to events such as point mutations, inversions, deletions or translocations (Botstein et al., 1980; Jeffreys et al., 1985; Neale and Williams, 1991; Sarwat, 2012).

2.4.3.2. PCR-Based marker

PCR is a molecular biology technique for enzymatically replicating (amplifying) small quantities of DNA without using a living organism. It is used to amplify a short (usually up to 10 kb), well-defined part of a DNA strand from a single gene or just a part of a gene (Semagn et al., 2006). The use of this kind of marker has been exponential, following the development by Mullis et al. of the Polymerase Chain Reaction (PCR). This technique consists of the amplification of several discrete DNA products, deriving from regions of DNA which are flanked by regions of high homology with the primers. These regions must be close enough to one another to permit the elongation phase (Mullis et al., 1986).

1) Arbitrary or semi-arbitrary primed PCR techniques

a) Random Amplified Polymorphic DNA (RAPD) marker

RAPD was the first to become available and is most common and frequently used PCR-based techniques, developed and applied in plants (Welsh and McClelland, 1990; Wiliams et al., 1990). RAPD is a PCR-based method which employs single primer of arbitrary nucleotide sequence with 10 nucleotides to amplify anonymous PCR fragments from genomic template DNA and the resulting fragments are size separated on an ethidium bromide stained agarose gel. Homology between the primer and the template

DNA results in the existence of certain band. Any mutation that prevents the hybridization of a primer to the template DNA at a certain locus results in the absence of this band. The number of RAPD loci used for the analysis of plant genomes varies between several tens to several hundreds. The use of short primers is necessary to increase the probability that, although the sequences are random, they are able to find homologous sequences suitable for annealing. DNA polymorphisms are then produced by "rearrangements or deletions at or between oligonucleotide primer binding sites in the genome" (Wiliams et al., 1990).

RAPD are multilocus markers and their mode of inheritance is dominant. The major advantages of this system are very simple genotyping technology and no prior sequence information is required, therefore, RAPD has become very popular in many laboratories. The major disadvantages of this system are low level of polymorphism, dominant mode of inheritance and the short (10 nucleotide) primer result in very sensitive PCRs and as a result repeated PCR results are not always identical (low reproducibility) (Paran et al., 1991; Adamblondon et al., 1994; Xue et al., 2010). The RAPD assays produce fragments from homozygous dominant or heterozygous alleles. No fragment is produced from homozygous recessive alleles because amplification is disrupted in both alleles (Semagn et al., 2006). Among the different types of molecular markers available, RAPD has been widely used for the assessment of genetic diversity among rare species (Wiliams et al., 1990). RAPD markers have been used in the detection of DNA polymorphism, cultivar fingerprinting (Xu et al., 1995) and in identification of markers for specific genes (Paran et al., 1991).

b) Arbitrarily Primed-PCR (AP-PCR) marker

To overcome the limitation of reproducibility associated with RAPD, AFLP technology was developed by the Dutch company, Keygene. This is a special case of RAPD wherein discreet amplification patterns are generated by employing single primers of 10-50 bases in length in PCR of genomic DNA (Chawla, 2002). Simple and reproducible fingerprints of complex genomes can be generated using single arbitrarily chosen primers and the polymerase chain reaction (PCR). No prior sequence information is required. This method involves two cycles of low stringency amplification followed by PCR at higher stringency (Welsh and McClelland, 1990). It was used for fingerprinting of the genomes (Welsh and McClelland, 1990; Joshi et al., 1999).

c) DNA Amplification Fringerprinting (DAF) marker

A single arbitrary primer, as short as 5 bases, is used to amplify DNA using PCR. DAF requires careful optimization of parameters; however, it is extremely amenable to automation and fluorescent tagging of primers for early and easy determination of amplified products (Chawla, 2002). The original DAF protocol is mainly different from RAPD in that it uses short primers (at least 5 bp), higher primer concentrations, two-temperature cycles instead of three-temperature cycles, and detection of amplification product on AgNO3 stained polyacrylamide gel (Semagn et al., 2006). This technique has been useful in genetic typing and mapping based on its complex and simple patterns of bands it produce (Joshi et al., 1999).

d) Amplified Fragment Length Polymorphism (AFLP) marker

AFLP technique combines the power of RFLP with the flexibility of PCR-based technology by ligating primer-recognition sequences (adaptors) to the restricted DNA (Lynch and Walsh, 1998). The first step in AFLP analysis involves restriction digestion of genomic DNA (about 500 ng) with a combination of rare cutter (EcoRI or PstI) and frequent cutter (MseI or TaqI) restriction enzymes. Double-stranded oligonucleotide adaptors are then designed in such a way that the initial restriction site is not restored after ligation. Such adaptors are ligated to both ends of the fragments to provide known sequences for PCR amplification (Semagn et al., 2006). PCR amplification will only occur where the primers are able to anneal to fragments which have the adaptor sequence plus the complementary base pairs to the additional nucleotides called selective nucleotides. An aliquot is then subjected to two subsequent PCR amplifications under highly stringent conditions with primers complementary to the adaptors, and possessing 3` selective nucleotides of 1 – 3 bases. The first PCR (pre-amplification) is performed with primer combinations containing a single bp extension while final (selective) amplification is performed using primer pairs with up to 3-bp extension. Because of the high selectivity, primers differing by only a single base in the AFLP extension amplify a different subset of fragments. A primer extension of one, two or three bases reduces the number of amplified fragments by factors of 4, 16 and 64, respectively. Ideal primer extension lengths will vary with genome size of the species and will result in an optimal number of bands: not too many bands to cause smears or high levels of band co-migration during electrophoresis, but sufficient to provide adequate polymorphism (Vos et al., 1995). The AFLP banding profiles are the result of variation in the restriction sites ot in the intervening region (Spooner et al., 2005). AFLP fragments are visualized either on agarose gel or on denaturing polyacrylamide gels with autoradiography, AgNO3 staining or automatic DNA sequencers. AFLP is applied in population genetics, molecular

taxonomy, genetic linkage map construction, germplasm construction, identification of cultivars, clones (Joshi et al., 1999; Weising et al., 2005). AFLP has been used successfully for variation analysis of number of plants; *Jatropha curcas* (Tatikonda et al., 2009); *Olea europaea* L. (Sensi et al., 2003); *Swertia* species (Misra et al., 2010); Tea (Paul et al., 1997).

e) Inter-simple Sequence Repeat (ISSR) marker

Inter simple sequence repeat (ISSR)-PCR is a technique, which involves the use of microsatellite sequences as primers in a polymerase chain reaction to generate multilocus markers (Reddy et al., 2002). ISSRs exhibit the specificity of microsatellite markers, but need no sequence information for primer synthesis enjoying the advantage of random markers (Joshi et al., 2000). The primers are not proprietary and can be synthesized by anyone. It is a simple and quick method that combines most of the advantages of microsatellites (SSRs) and amplified fragment length polymorphism (AFLP) to the universality of random amplified polymorphic DNA (RAPD). ISSR markers are highly polymorphic and are useful in studies on genetic diversity, phylogeny, gene tagging, genome mapping and evolutionary biology. To date, PCR with anchored or unanchored microsatellite-complementary primers is probably the most widely applied variant of inter-repeat PCR. The technique uses microsatellites, usually 16-25bp long, as primers in a single primer PCR reaction targeting multiple genomic loci to amplify mainly inter simple sequence repeats of different sizes (Reddy et al., 2002). The approach relies on the existence of inversely oriented tandem repeat arrays made up of the same or a very similar motif. When such inverted repeat blocks are located close to each other (i.e., within about 5kb), the inter-repeat region may be amplified by PCR with one tandem repeat-specific primer (Weising et al., 2005). The basic premise is that primer annealing sites are distributed evenly throughout the genome such that the primer will anneal to two sites oriented on opposing DNA strands. If these are within an appropriate distance of one another, the region between the two primers will be amplified through PCR. The region would not be amplified if there were divergence at the primer binding sites, if a binding site was lost or if structural rearrangements of the chromosomes had occurred (Kurane et al., 2009). The microsatellite repeats used as primers for ISSRs can be di-nucleotide, tri-nucleotide, tetra-nucleotide or penta-nucleotide. Tri and tetra-nucleotides are less frequent and their use in ISSRs is lesser than the di-nucleotides (Semagn et al., 2006). The primers used can be either unanchored (Meyer et al., 1993; Gupta et al., 1994; Wu et al., 1994) or more usually anchored at 3` or 5` end with 1 to 4 degenerate bases extended into the flanking sequences (Zietkiewicz et al., 1994). When unanchored i.e. only the SSRs are used as primers, the primer tends to slip within the

repeat units during amplification leading to smears instead of clear bands. Extending the primer (anchoring) with 1 to 4 degenerate nucleotides at the 3' end or 5' end assures annealing only to the ends of a microsatellite in template DNA, thus obviating internal priming and smear formation. Secondly, the anchor allows only a subset of the microsatellites to serve as priming sites. When 5' anchored primers are used, the amplified products include the microsatellite sequences and their length variations across a genome and therefore give more number of bands and a higher degree of polymorphism. Usually di-nucleotide repeats, anchored either at 3' or 5' end reveal high polymorphism. The primers anchored at 3' end give clearer banding pattern as compared to those anchored at 5' end (Tsumura et al., 1996; Nagaoka and Ogihara, 1997; Blair et al., 1999; Joshi et al., 2000). In general, primers with (AG), (GA), (CT), (TC), (AC), (CA) repeats show higher polymorphism than primers with other di-, tri- or tetra-nucleotide repeats. (AT) repeats are the most abundant di-nucleotides in plants but the primers based on (AT) would self- anneal and not amplify. Tri-and tetra-nucleotides are less frequent and their use in ISSRs is lesser than the di-nucleotides. ISSRs use longer primers (15–30 mers) as compared to RAPD primers (10 mers), which permit the subsequent use of high annealing temperature leading to higher stringency. The annealing temperature depends on the GC content of the primer used and ranges from 45 to 65°C (Reddy et al., 2002). The amplified products are usually 200–2000 bp long and possible to detection by both agarose and polyacrylamide gel electrophoresis. About 92–95% of the scored fragments could be repeated across DNA samples of the same cultivar and across separate PCR runs when detected using polyacrylamide (Fang and Roose, 1997; Moreno et al., 1998).

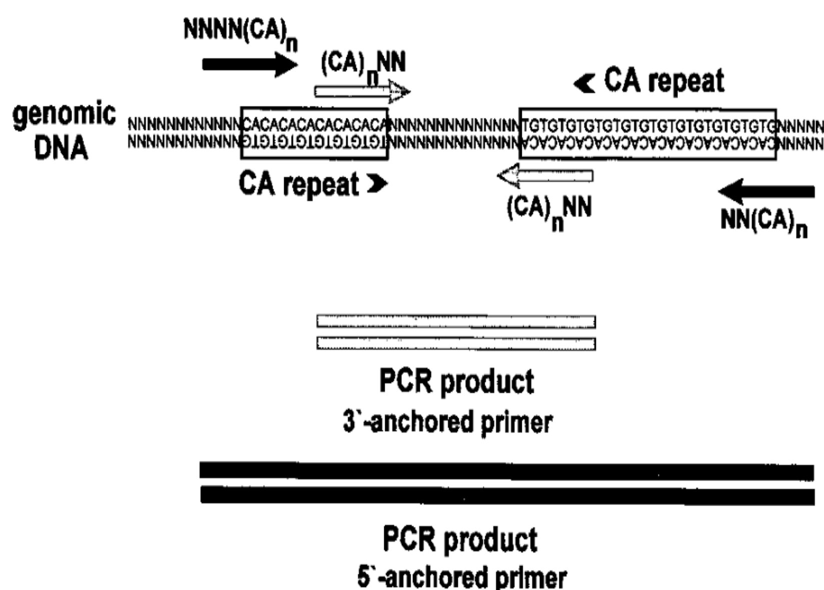


Figure 2.1 Inter-SSR PCR: A single primer targeting a $(CA)_n$ repeat, anchored either at the 3′ (light arrow) or the 5′ end (dark arrow) of the repeat, is used to amplify genomic sequence flanked by two inversely oriented $(CA)_n$ elements.

ISSRs segregate mostly as dominant markers following simple Mendelian inheritance so, can't resolve heterozygous genotype and the resulting fragment may either be homozygous (AA) or heterozygous (Aa) and absence of fragment means genotype (aa) (Gupta et al., 1994; Tsumura et al., 1996; Moodie et al., 1997; Ratnaparkhe et al., 1998; Wang et al., 1998). However, they have also been shown to segregate as co-dominant markers in some cases thus enabling distinction between homozygotes and heterozygotes (Wu et al., 1994; Akagi et al., 1996; Wang et al., 1998; Sanker and Moore, 2001). ISSR markers overcome the shortcomings of the low reproducibility of RAPD, the high cost of AFLP, the complexity of SSR, and represent a fast and a cost efficient technique (Kurane et al., 2009). The evolutionary rate of change within microsatellites is considerably higher than most other types of DNA, so the likelihood of polymorphism in these sequences is greater. The source of variability in the ISSRs can be attributed to any combination of template DNA, nature of Primer used or Detection Method (Reddy et al., 2002). The technique is simple, quick, and the use of radioactivity is not essential. Polyacrylamide gel electrophoresis (PAGE) in combination with radioactivity was shown to be most sensitive, followed by PAGE with AgNO3 staining and then agarose gel with EtBr system of detection (Semagn et al., 2006). Markedly, higher number of bands were resolved per primer when polyacrylamide gel was used compared to agarose (Moreno et al., 1998). However, high levels of polymorphism have been detected even when products of ISSR amplification are resolved on agarose gels (Tsumura et al., 1996; Wolff and Morgan-Richards, 1998; Arcade et al., 2000; Sanker and Moore, 2001).

The ISSR primers are used to generate the variation in a given DNA sample include one of the highly variable microsatellite sequences and an arbitrary pair of bases at the 3′ end (anchored). One samples for variation among DNA samples in small PCR reactions using one primer at a time. Where the primer successfully locates two microsatellite regions within an amplifiable distance away on the DNA strands of the DNA sample (template DNA) the PCR reaction will generate a band of a particular size (=molecular weight) for that "locus" representing the intervening stretch of DNA between the microsatellites. Usually, several too many such "paired" microsatellite areas exist in a particular DNA sample, so one gets that many bands generated in the reaction for that sample (Kurane et al., 2009).

Several earlier molecular studies in Citrus have utilized ISSR markers for assessing genetic diversity in clementine (Bretó et al., 2001) and in lemon (Gulsen and Roose, 2001), cultivar identification in Citrus and Poncirus (Fang and Roose, 1997; Capparelli et al., 2004), constructing genetic linkage map in Citrus *grandis* and *Poncirus trifoliata* (Shankar and Moore, 2001), in management of Citrus germplasm resources (Krueger and Roose, 2003), discrimination of clones of sour oranges (Pasquale et al., 2006), and molecular differentiation in Citrus germplasm resources of Iran (Shahsavar et al., 2007).

2) Site-targeted PCR techniques

a) Simple Sequence Repeats (Microsatellites) marker

The genomes of higher organisms contain three types of multiple copies of simple repetitive DNA sequences (satellite DNAs, minisatellites, and microsatellites) arranged in arrays of vastly differing size (Armour et al., 1999; Hancock, 1999). Simple sequence repeats (SSRs) were discovered and developed in humans (Litt and Luty, 1989; Edwards et al., 1991) and were first applied to plants (Akkaya et al., 1992). Some authors (Armour et al., 1999) define microsatellites as 2–8 bp repeats, others (Goldstein and Pollock, 1997) as 1–6 or even 1–5 bp repeats (Schlotterer, 1998). A strict definition of 2–6 bp repeats, in line with the descriptions of the original authors has been suggested (Chambers and MacAvoy, 2000). SSRs are characterized by a low degree of repetition per locus (5–100), random dispersed distribution of about (10^4–10^5) per genome (Tautz, 1993) and high degree of length polymorphism (Zane et al., 2002). They are the fastest evolving pieces of DNA and can be used for population studies (Ayad et al., 1997). The repeats of longer units form minisatellites or, in the extreme case, satellite DNA. SSR allelic differences are the results of variable numbers of repeat units within the microsatellite structure. The repeated sequence is often simple, consisting of two, three or four nucleotides (di-, tri-, and tetranucleotide repeats, respectively). One common example of a microsatellite is a dinucleotide repeat (CA)n, where n refers to the total number of repeats that ranges between 10 and 100. These markers often present high levels of inter and intra-specific polymorphism, particularly when tandem repeats number is ten or greater (Queller et al., 1993). Dinucleotide repeats dominate, followed by mono and tetranucleotide repeats, and trinucleotide repeats are least dominant. Among dinucleotides, (CA)n repeats are most frequent, followed by (AT)n, (GA)n and (GC)n, the last type of repeat being rare. There are only four possible types of dinucleotide repeat, because CA = AC = GT = TG, GA = AG = CT = TC, AT = TA, and GC = CG (Ellegren, 2004). Poly G and Poly A are the simplest of the microsatellites. The number and composition of microsatellite repeat differ in plant and animals. The frequency of repeat greater than 20 bp has been estimated to occur every 33kb in plants, whereas in mammals it is found to occur every 6kb. In human, AC or TC are very

common repeats and in plants, AT is more common, followed by AG or TC. In general, plants have about ten times less SSRs than do humans (Chawla, 2002). Microsatellites are born from regions in which variants of simple repetitive DNA sequence motifs are already over represented (Tautz et al., 1986). The bulk of simple repeats are embedded in non-coding DNA, either in the intergenic sequence or in the introns. Microsatellites that are used as genetic markers are usually of this type and are generally assumed to evolve neutrally. In contrast to unique DNA, microsatellite polymorphisms derive mainly from variability in length rather than in the primary sequence. Moreover, genetic variation at many microsatellite loci is characterized by high heterozygosity and the presence of multiple alleles (Ellegren, 2004). It is now well established that the predominant mutation mechanism in microsatellite tracts is 'slipped-strand mispairing' (Levinson and Gutman, 1987).

SSR markers have gained considerable importance in plant genetics and breeding owing to many desirable attributes including hypervariability, multiallelic nature, codominant inheritance, reproducibility, relative abundance, extensive genome coverage (including organellar genomes), chromosome specific location, amenability to automation and high throughput genotyping (Parida et al., 2009). Replication slippage also occurs during PCR amplification of simple sequence repeats in vitro leading to formation of stutter bands i.e. minor products that differ in size from the main product by multiples of the length of the repeat unit (Hauge and Litt, 1993; Murray et al., 1993). Quantitative experiments show that the *Taq* polymerase slippage rate increases with the number of repeat units and is inversely correlated with repeat unit length (Shinde et al., 2003). PCR-induced stutter bands have been observed by many microsatellite users; tetra-nucleotide repeat markers typically give fewer stutter bands than di-nucleotide and, in particular, mononucleotide repeats and they generally appear as products that are shorter than the size of allele being amplified (Ellegren, 2004). SSR have several advantages over other molecular markers. For example, (i) microsatellites allow the identification of many alleles at a single locus, (ii) they are evenly distributed all over the genome, (iii) they are co-dominant, (iv) little DNA is required and (v) the analysis can be semi-automated and performed without the need of radioactivity. During the last decade, microsatellite markers (SSRs) were extensively exploited for identification of cultivars, assessment of genetic diversity, phylogenetic studies and management of germplasm resources/collections due to their high level of polymorphism, even distribution in the genome, codominant Mendelian inheritance, transferability between laboratories and conservation across taxa (Hvarleva et al., 2008).

The genetic variability in some Iranian sweet oranges and Mandarins were assessed using SSR markers (Golein et al., 2005). ISSR, SSR, and isozymes were utilized for assessment of diversity, phylogenetic relationships, and parentage in lemon (*C. limon* (L.) Burm. f.) accessions and related taxa was done and found little genetic variation among lemon accessions (Gulsen and Roose, 2001). Genetic diversity and population structure in a Citrus germplasm collection located at the University of California, Riverside were assessed using 24 SSR markers and conclude that there are only a few naturally occurring species of Citrus and the most other types of Citrus arose through various hybridization events between these naturally occurring forms (Barkley et al., 2006). Characterization Citrus cultivar and clones in Cyprus were carried out through microsatellite and RAPD analysis (Hvarleva et al., 2008).

b) Cleaved Amplified Polymorphic Sequences (CAPS) marker

CAPS marker is a combination of PCR and RFLP, and it was originally a named PCR-RFLP (Maeda et al., 1990). CAPS markers are inherited mainly in a codominant manner (Matsumoto and Tsumura, 2004). This technique involves PCR amplified DNA fragments to be digested with a restriction enzymes to reveal restriction site polymorphisms. The digested restriction fragments are subjected to agarose gel electrophoresis followed by staining with ethidium bromide for bands visualization (Chawla, 2002). Hence, CAPS markers rely on differences in restriction enzyme digestion patterns of PCR fragments caused by nucleotide polymorphism between samples. Critical steps in the CAPS marker approach include DNA extraction, PCR conditions, and the number or distribution of polymorphic sites. However, the ability of CAPS to detect DNA polymorphism is not as high as SSRs and AFLPs because nucleotide changes affecting restriction sites are essential for the detection of DNA polymorphism by CAPS (Semagn et al., 2006).

c) Sequence Characterized Amplified Region (SCAR) marker

A SCAR marker is a genomic DNA fragment that is identified by PCR amplification using a pair of specific oligonucleotide primers (Paran et al., 1991; McDermott et al., 1994). SCARs are derived by cloning and sequencing the two ends of RAPD markers that appeared to be diagnostic for specific purposes. Two 24-base oligonucleotide primers corresponding to ends of the fragment (the 5' ten bases are the same as original 10-mer used in the RAPD reaction and the 14 internal bases from the end) have been synthesized. These primers with their increased specificity generally amplify a single highly repeatable band, not the 5-10 bands for the progenitor 10 base primers (Chawla, 2002). SCARs are advantageous over RAPD markers as they detect only a single locus,

their amplification is less sensitive to reaction conditions, and they can potentially be converted into codominant markers (Paran et al., 1991).

3) DNA Sequencing Based Markers

Now-a-days, DNA sequencing has become popular for genetic studies. Among the numerous approaches for study of polymorphism at DNA level, the most direct strategy is determination of the nucleotide sequence of a defined region. Two basic strategies for DNA sequencing were developed during mid 1970: a) Chemical based method (Maxam and Gilbert, 1997) and b) Enzyme based method (Sanger et al., 1977). Of these two, Sanger method (Chain Termination Method) is the most commonly used method. DNA sequencing provides highly robust, reproducible, and informative datasets, and can be adapted to different levels of discriminatory potential by choosing appropriate genomic target regions. However, this method is very tedious and expensive when very large numbers of individual have to be assayed such as in population genetics, phylogeography and marker assisted breeding programs (Weising et al., 2005). DNA sequencing of different genomes (nuclear, chloroplast and mitochondrial) and different genes (matK, rbcL, ndhF of chloroplast genome and ribosomal DNA sequences such as 18S, 26S, ITS and 5S of nuclear genomes) are widely applied in the phylogenetic reconstruction and diagonostic development for DNA based studies (Shrestha et al., 2003).

a) Internal Transcribed Spacer (ITS)

The nuclear rDNA region has frequent insertions/delitions which can be phylogenetically informative (Baldwin et al., 1995). The nuclear ribosomal transcriptional unit (NRTU) is comprised of 18s, 5.8s and 25s genes with two Internal Transcribed Spacer (ITS) (ITS-1 and ITS-2). Genes encoding ribosomal RNA and spacers occurs in tandem repeats that are hundreds to thousands of copies long; each separated by regions of non-transcribed DNA termed intergenic spacer (IGS). ITS and IGS are used for genetic variation studies at the genus, species and population levels (Soltis et al., 1998; Alvarez and Wendel, 2003).
Beside, sequence comparison of ITS region is widely used in taxonomy and molecular phylogeny because it a) is easy to amplify even from small quantities of DNA and b) has a high degree of variation even between closely related species. ITS of nrDNA have been widely used for resolving phylogenetic relationships among closely related species (Lanoue et al., 1996; Shrestha et al., 2003; Li et al., 2009; Joshi, 2011); molecular authentifaication of herbal materials (Zhang et al., 2007); genetic diversity assessment (Mondini et al., 2009); inra-specific variation study (Haque et al., 2009) and DNA barcoding (Zuo et al., 2010).

b) Single Nucleotide Polymorphism (SNPs)

SNP is characterized by a single base substitution at a particular position, represents the most abundant source of DNA polymorphism in organisms (Weising et al., 2005). They represent a definite position at a chromosomal site at which the DNA sequences of two individuals differs by a single base (A, T, G or C). Direct sequencing of DNA fragments (amplified by PCR) from several individuals and aligning is direct way to indentify SNP polymorphisms (Chawla, 2002). In addition, the dramatic increase in the number of DNA sequences submitted to the database has made it possible to identify SNPs for several crops by electronic mining (e-mining) without the need for sequencing (Taillon-Miller et al., 1998; Somers et al., 2003). The SNPs are usually more prevalent in the non-coding regions of the genome. Within the coding regions, when an SNP is present, it can generate either non-synonymous mutations that result in an amino acid sequence change (Sunyaev et al., 1999), or synonymous mutations that not alter the amino acid sequence. Synonymous changes can, however, modify mRNA splicing, resulting in phenotypic differences (Richard and Beckman, 1995). This approach consists of the identification and alignment of sequences from the same locus from different sources (genotypes) allowing the detection of SNPs along these DNA sequences. The availability of Expressed Sequence Tag (EST) databases makes it possible to target polymorphism in functional region of the genomes and even to specific gene (Chawla, 2002). They are used for a wide range of purposes, including rapid identification of crop cultivar and construction of ultra high density genetic maps (Spooner et al., 2005; Mondini et al., 2009).

2.5. Application of ISSR marker Technique

2.5.1. Genomic fingerprinting

DNA fingerprinting is an important tool for characterization of germplasm and establishment of the identity of varieties/hybrids/parental sources etc. in plant breeding and germplasm management (Reddy et al., 2002). Di-nucleotide based ISSR primers anchored at 5′ or 3′ end have been used in fingerprinting studies with high reproducibility for maintenance of cocoa collection (Charters and Wilkinson, 2000). ISSRs showed sufficient polymorphism to distinguish between various cultivars of chrysanthemum (Wolff et al., 1995). Microspore de-rived plants could be distinguished from those derived from somatic tissues in anther culture of flax at an early seedling stage (Chen et al., 1998).

2.5.2. Genetic diversity and phylogenetic analysis

ISSRs have been used to estimate the extent of genetic diversity at inter- and intra-specific level in a wide range of crop species which include rice (Joshi et al., 2000), wheat (Nagaoka and Ogihara, 1997), fingermillet (Salimath et al., 1995), *Vigna* (Ajibade et al., 2000), sweet potato (Huang and Sun, 2000) and *Plantago* (Wolff and Morgan-Richards, 1998). Anchored SSR primers for instance, have been found to be more useful and reproducible than isozymes, RFLPs and RAPDs in the diversity analysis of trifoliate orange germplasm (Fang et al., 1997). ISSR is the marker of choice for assessment of genetic diversity in cocoa (Charters and Wilkinson, 2000), gymnosperms such as Douglas fir and sugi (Tsumura et al., 1996) and even fungi (Hantula et al., 1996). In a study on white lupin, among 10 primers used any two were sufficient to distinguish all the 37 accessions studied (Gilbert et al., 1999). Similarly, 4 primers were sufficient to distinguish 34 cultivars of potato (Prevost and Wilkinson, 1999) and 3 primers could distinguish 16 genotypes of redcurrant (Lanham and Brennan, 1998). The use of such highly informative primers lowers the cost, time and labor for diversity analysis (Reddy et al., 2002).

It has been found effective in resolving problems relating to the phylogeny of Asian cultivated rice *Oryza sativa* (Joshi et al., 2000), wheat (Nagaoka and Ogihara, 1997), finger millet (Salimath et al., 1995), *Vigna* (Ajibade et al., 2000) and Diplotaxis species (Martin and Sanchez-Yelamo, 2000). Significant genome/species specific ISSR markers have been reported in four genera *Oryza* (Joshi et al., 2000), *Lolium* and *Festuca* (Pasakinskiene et al., 2000) and *Diplotaxis* (Martin and Sanchez-Yelamo, 2000) which are useful in delineating species.

2.5.3. Genome mapping

Unmapped ISSR can be used to saturate RFLP and SSR linkage maps. The RFLP map of barley was saturated with 60 ISSRs (referred as RAMPs in the study) which mapped to all chromosomes (Becker and Heun, 1995). In Einkorn wheat’s, the nine ISSR markers mapped at or close to the RFLP marker positions (Kojima et al., 1998). ISSRs have been used along with AFLP and RAPD markers in the mapping of Japanese and European larch genomes (Arcade et al., 2000). The genetic linkage map of Citrus was further saturated using 75 ISSR markers, which were dispersed among all the linkage groups (Shankar and Moore, 2001). In soybean, 58 ISSR markers were mapped onto 18 RAPD/RFLP linkage groups (Wang et al., 1998). CA polymorphisms had a biased distribution and GA polymorphisms were randomly dispersed (Reddy et al., 2002).

2.5.4. Studies on natural population/speciation

The hypervariable nuclear ISSR markers have proved useful in testing hypotheses of speciation, introgression and systematics (Wolff and Morgan-Richards, 1998). The hybrid origin of *Penstemon clevelandi* was cleared by the use of just 8 ISSR markers. Population of *P. clevelandi* has been found to have an additive profile of bands of the two proposed progenitor species *viz. P. centranthifolius* and *P. spectabilis*. As the population of *P. spectabilis* lacked the additive profile of bands of its proposed putative parents. The hybrid origin of *P. spectabilis* was thus negated and its origin was attributed instead to introgression of genes and not the genome of a related species (Wolfe et al., 1998). Variation within and between population can be compared using dispersed multilocus markers such as ISSR. The amount of variation between *O. granulate* populations from different regions (49.2%) was found to be higher than that between populations within a region (38%) or within a population (12%) using ISSR markers (Qian et al., 2001).

2.5.5. Determining SSR motif frequency

Analysis of ISSR provides insights into the organization (clustered or not), frequency and levels of polymorphism of different simple sequence repeats in a genome. In rice and wheat, di-nucleotide simple sequence repeats used as primers gave the maximum number of bands and are, therefore, more common than any SSRs with larger units (Nagaoka and Ogihara, 1997; Blair et al., 1999). Poly (GA) based 3'-anchored primers produced 5 times as many bands as those with poly (GT) motif indicating low frequency or lack of clustering of (GT) motif (Blair et al., 1999). Using ISSRs it has been found that tetra-nucleotide repeats were abundant across eukaryotic genomes (Gupta et al., 1994) and that tetramers of tetra-nucleotides AGAC and GACA are scattered within the genome of grasses (Pasakinskiene et al., 2000). The demonstrated in Brassica has shown that enhanced recovery of microsatellite markers is possible using ISSR primers (Varghese et al., 2000).

2.5.6. Gene tagging and marker assisted selection

There are great contributions of DNA markers closely linked to important agronomic traits for practical crop improvement programs. In rice, an ISSR marker generated by primer $(AG)_8YC$ was converted to a sequence tagged site (STS) marker to identify the fertility restoration gene, Rf-1 (Akagi et al., 1996). Such co-dominant marker can be used in management of genetic purity of hybrid seed. In chickpea, ISSR markers UBC 855_{500} generated by primer $(AG)_8YT$ and UBC 825_{1200} using primer $(AG)_8T$ were linked to the gene conferring resistance to race 4 of Fusarium wilt (Ratnaparkhe et al., 1998). Recently, two allelic dominant DNA markers were identified using ISSTR-PCR, one linked

in coupling and the other in repulsion phase to a major locus Fgr, which modulates fructose to glucose ratio in tomatoes (Levin et al., 2000). Another trait of value in hybrid seed production *viz*., temperature-sensitive genic male sterility has been tagged with an ISSR marker UBC 855_{1060} in rice (Hussain et al., 2000).

ISSRs also used to generate species specific, gene specific and trait specific markers. 87 putative genome/species specific markers were identified during delineating the phylogenetic relationship among different species of the genus *Oryza* (Joshi et al., 2000). The 582 bp inter-SSR *Festuca* specific sequence and 1350 bp *F. arundinacea* specific sequence have potential as markers to confirm presence of closely linked *Festuca* genes (Pasakinskiene et al., 2000). Likewise, race specific markers have been developed in various fungi groups using ISSRs (Hantula et al., 1996).

2.6. Application SSR marker Technique

Biotechnology is integral to the application of robust, high through-put detection of species-specific and species or genus-transferred microsatellites, or simple sequence repeat (SSR) markers (Wang et al., 2009). Microsatellites as DNA markers have been widely used in many living organisms including *E. coli* (Gur-Arie et al., 2000); (Schlotterer, 2000), humans (Beckmann and Soller, 1990), mice (Love et al., 1990), cows and sheep (Moore et al., 1991), plants (Condit and Hubbell, 1991), and insects (Hughes and Queller, 1993; Goldstein and Clark, 1995; de Rosas et al., 2008; Exeler et al., 2008; Kim et al., 2008; Pol et al., 2008) for different kinds of basic genetics research.

2.6.1. Genome mapping

Genome mapping consists of genetic mapping, comparative mapping, physical mapping, and association mapping. Genetic mapping with microsatellite markers in plants were first reported in tropical trees and then reported in soybean, rice, etc. Comparative mapping has been successfully conducted in many plant species, including the Solanaceae family, grasses, crucifers, legumes and other species (Wang et al., 2009). Comparative mapping to *Brassica* identified genomic blocks that have been maintained since the divergence of the Arabidopsis and *Brassica* lineages (Schranz et al., 2007). Microsatellite markers have been used for comparative mapping between *Quercus robur* (L.) and *Castanea sativa* (Mill.) (Barreneche et al., 2004). A physical map of a *2 Mb BAC contig in the region around 80 cM of Arabidopsis thaliana chromosome 2 was constructed using SSR markers and BAC end-sequences (Wang et al., 1997).

2.6.2. Gender determination

A small group of flowering plants are sexually dimorphic where the female plants are commercially valued for production of fruits (papaya, kiwi fruit, dates, seabuckthorn, etc.) and seeds (pistachio, nutmeg, black pepper, jojoba, etc.). However, sex of most of the dioecious plants is not revealed morphologically and the male and female plants cannot be distinguished at seedling stage. This problem is exacerbated in species where the sex of an individual is revealed only after flowering which may take few months (papaya, Coccinia) to several years (date palm, nutmeg and jojoba). Sex linked microsatellite markers have been described in several species like hemp (Rode et al., 2005), wild strawberry (Spigler et al., 2008), hop (Jakse et al., 2008), *Carica* papaya (Parasnis et al., 1999) and *Actinidia chinensis* (Fraser et al., 2009). In papaya microsatellite probe $(GATA)_4$ was used as a diagnostic marker and demonstrated the sex-specific DNA variation at any stage of the plant development (Parasnis et al., 1999). The constructed of gene-rich female, male and consensus linkage maps of the diploid species *A. chinensis* were done using 644 microsatellite markers (Fraser et al., 2009).

2.6.3. Cultivar identification and marker-assisted selection

A specific set of SSR markers should be developed and selected for each economically important agricultural species, so that they can be employed to protect the intellectual property rights of plant commercial companies (Wang et al., 2009). In tomato, a set of 65 SSR markers has been selected for distinguishing 19 diverse tomato cultivars (He et al., 2003). In soybean, a set of SSR markers has been selected for the identification of soybean cultivars (Song et al., 1999).

SSR marker-assisted selection (MAS) can greatly enhance the efficiency of plant breeding programs. SSR markers used for selection can be classified into flanking SSR markers (closely linked to the locus for a trait) and targeted gene SSR markers (developed within the targeted gene itself). The gene for resistance to the Hessian fly (H32), the gene for adult-plant leaf rust resistance and yellow pigment content (YP) were successfully selected by flanking SSR markers of Xgwm3 and Xcfd223 (Sardesai et al., 2005), GWM296 (Hiebert et al., 2007), and Xwmc809 (He et al., 2008), respectively, in wheat breeding programs. A major QTL for submergence stress (Sub1) was selected by employing flanking SSR markers in a rice backcross breeding program (Neeraja et al., 2007). The polymorphic SSR marker QLB1, co-segregated with the locus for resistance to the barley yellow mosaic virus, was used for resistance selection in barley breeding (Tyrka et al., 2008). Microsatellite markers have also been used to study the association between host race of Acyrthosiphon pisum and their symbiants (Simon et al., 2003). As

more whole genome sequence projects are completed, more gene-targeted SSR markers should be available for marker-assisted selection.

2.6.4. Genetic diversity and phylogenetic relationships

Genetic diversity assessment and phylogenetic relationship construction will provide important information for choosing parental lines for breeding programs, classification of plant germplasm accessions, and further curation and acquisition of new plant germplasm accessions **(**Wang et al., 2009)**.** The genetic diversity and phylogenetic relationships from germplasm collections such as a temperate bamboo collection (Barkley et al., 2005), a Citrus variety collection (Barkley et al., 2006) and a cultivated and wild peanut collection (Barkley et al., 2007; Cuc et al., 2008) have been assessed by SSR or transferable SSR markers. Recently, SSR markers have been developed in eggplant (*Solanum melongena* L.) and the genetic diversity of 38 cultivated eggplant varieties and phylogenetic relationships of its related species have been assessed (Stagel et al., 2008).

2.6.5. Population and evolutionary studies

Plants cannot migrate large distance like animals, except in cases of plant introduction and domestication by humans, seed dispersal, and pollen flow, all of which affect the genetic structure of natural populations. Traditionally, plant evolution study is based taxonomic and phenotypic data (such as morphology and karyotype). Researchers can now integrate data collected from molecular assays such as microsatellites, karyotyping, and phenotyping to provide a more robust analysis of their species of interest (Wang et al., 2009). Chloroplast SSR markers have been used to survey the variation among barley populations (including wild *Hordeum* species, wild progenitors of cultivated barley, barley landraces, and barley cultivars) (Provan et al., 1999). The genotyping results based on 50 SSR markers confirmed the reclassification of the cultivated potato into four species: *S. tuberosum*, *S. ajanhuiri*, *S. juzepczukii*, and *S. curtilobum* (Spooner et al., 2007). The tomato (*Lycopersicon esculentum*) has been renamed *Solanum lycopersicon* (Spooner et al., 1993; Olmstead et al., 1999) and 'Dicots' are being considered as 'Eudicots' (Slotis and Slotis, 2003).

2.7. Genetic Diversity

Biological diversity occurs at three different levels: genetic diversity (variation in genes and genotype), species diversity (species richness) and ecosystem diversity (communities of species and their environment) (Rao et al., 2002). Genetic diversity is important for the survival of wild species and as a source of genetic variation for

cultivated plants (Weising et al., 2005). Genetic diversity is commonly used expression to describe the heritable variation found within biological entities and can be measured at the individual, population, and species level. At any particular locus, diversity may be present within an individual. Diversity may also be present within a population, when the alleles present at a variable locus are found in different individuals. In addition, diversity may be present within species and allelic variation may change between populations across the species where some populations exhibit alleles that other lack and vice versa. The distinction between the proportion of heterozygosity and level of polymorphism, at the individual population, and species levels, are the two of the values that diversity statistics attempt to quantify (Lowe et al., 2004). Genetic diversity studies have two objectives, the analysis of level of polymorphism at a given hierarchical level (most usually populations) and study of distribution of polymorphisms among the different hierarchical levels (among the populations) investigated (Angela et al., 1998). The ability of a species to respond to varied environment depends in the level of genetic diversity it contains (Ayala and Kiger, 1984). In all sexually reproducing organisms, each individual plant or animal contains a different mixture of genes and is known as Genetic diversity (Miller and Tangley, 1991). This incredible variation within species is what allows populations to adapt to changes in climate and other environmental conditions so genetic diversity plays an important role in survival and adaptability of organism to both biotic and abiotic factors. The diversity is the result of natural evolutionary processes such as hybridization, natural selection, spontaneous mutation and genetic drift; this has led to the enormous range of today’s adapted genotypes. Therefore, genetic diversity is essential for the long term survival and continuing evolution of plants and animals (Templeton, 1994). The more the variations better the chance that at least some of the individuals would have an allelic variant that are suited for new environment and would produce offspring with the variant that would in turn reproduce and continue the population into subsequent generations (Cooke and Reeves, 1998).

2.7.1. Significance of Genetic Diversity

Agriculture began when farmers started to gather seeds from wild plant and began sowing them to grow food since 12.000 years ago (Shand, 1997). Genetic diversity and agriculture are interrelated as genetic diversity is critical for agriculture since it promotes sustainable agriculture and agriculture contributes to conservation and sustainable use of genetic diversity (Oluwaponmile, 2009). The followings are the significance of genetic diversity:

2.7.1.1. Adaptation of crops to biotic and abiotic factors

The survival and long term viability of species are dependent on its ability to recover from population decline which is ultimately linked to the variation in its genetic make-up (Vrigenhoek, 1994). Those species with narrow and overall lack of genetic diversity in specific traits may have less ability to survive and adapt to evolving biotic and changing environmental conditions. Genetic diversity provides species with the ability to adapt to changing environment and evolve, by increasing their tolerance to abiotic factors like frost, high temperature, drought and water logging, as well as increasing their resistance to particular biotic factors like disease, pest and parasites (Oluwaponmile, 2009). Genetic diversity also helps in soil and water conservation, maintenance of soil fertility and biota, and pollination, all of which are essential for human survival (Anonymous, 2008).

2.7.1.2. Food Security

Food security is a basic right that must be upheld (Anonymous, 2008). In Agriculture, genetic diversity has pivotal support in ensuring the food security. More than 400 species of crop pests developed resistance to one or more of the pesticides used to control them (Miller and Tangley, 1991). However, genetic diversity ensures the constant infusion of new hardy genes into our crops and this ensures that pest and disease are put to check. In 1970, the United States of America lost 15 percent of its corn; worth $ 1 billion, when a fungus spread rapidly across the Midwest (Miller and Tangley, 1991). In such case, genetic uniformity allowed a disease that is always present at low levels, even in the best of the years, to run wild. Only introducing new corn varieties containing new genes will stopped the fungus in track (Miller and Tangley, 1991). A very similar occurrence is the cause of the infamous potato (*Solanum tuberosum*) famine in Ireland. Since potatoes are produced by vegetative propagation so entire crop is essentially a clone of one plant, no genetic diversity is developed, that's why susceptible to epidemics. In the 1840s, much of Ireland's population depended on potatoes for food; they plant "lumper" variety which is susceptible to a rot causing plasmodiophorid, *Phythophthera infestans*. As a result vast majority of potato crops were destroyed and left tens of thousands of people to starve to death (Anonymous, 2008a).

2.7.1.3. Livelihood of rural people

Genetic diversity and other components of biodiversity are crucial to the alleviation of poverty, due to the basic goods and ecosystem services it provides. Seventy percent of world's poor live in rural areas and depend directly on biological diversity for their

livelihoods (Anonymous, 2008). More than 3 billion people depend on marine and coastal biodiversity for their live hoods, while over 1.6 billion people, including 1 billion living in poverty, rely on forests and non-timber forest products (Anonymous, 2008). Habitat degradation and loss of biodiversity are threatening the livelihoods of more than 1 billion people living in dry and sub-humid lands (Anonymous, 2008).

2.8. Statistical parameters and estimators for genetic analysis

2.8.1. Polymorphic Information Content (PIC)

PIC estimates the degree of polymorphism of marker which is proportion of individuals that are heterozygous for a marker (Shrestha, 2012). High PIC value indicates rich heterozygosity which in turn is associated with a high degree of polymorphism (Arya et al., 2011). Genetic diversity is based on the number of alleles per locus and the frequency of alleles per locus. A locus with low information has few alleles at low frequency (Koirala, 2010). A PIC value of less than 0.25 indicate low polymorphism, a value between 0.25 and 0.5 indicate average polymorphism and a value higher than 0.5 indicates a highly polymorphic locus (Botstein et al., 1980). The marker with highest values is best used to distinguish variety and markers with lowest values indicate the rare allele and are useful in structuring the varieties (Luce et al., 2001).

2.8.2. Primer Resolving Power (Rp)

The varieties distinguishing power of each primer can be studied by calculating resolving power (Rp), which is based on allocation of alleles within the sampled genotypes (Prevost and Wilkinson, 1999). It is based on the distribution of bands detected within the sampled accessions i.e. band informativeness. It is used as an index for good primers (Azizi et al., 2009).

2.8.3. Similarity and Dissimilarity indices for molecular marker data

The generated binary data matrix from presence and absence of fragment is the qualitative (nominal) one. So for analysis of such data, similarity indices should be calculated. This can be done by using SimQual computational algorithm in NTSYS Statistical package. However, proper choice of similarity, s or dissimilarity coefficient $d = 1 - s$, is important and depends on factors such as the properties of the marker system employed, the genealogy of the germplasm, the operational taxonomic unit (OTU) under consideration (e.g., lines, populations), the objectives of the study, and necessary preconditions for subsequent multivariate analyses (Reif et al., 2005).

According to the examination done for genetical and mathematical properties of 10 di/similarity coefficient widely used in germplasm surveys, when molecular marker data are 'allelic informative' (i.e. allele frequency can be determined from the marker data), the estimates of coefficient between OTUs can be calculated from the difference in the allele frequencies but when marker data are 'allelic non-informative' the estimates of coefficient between OTUs under consideration based on the presence or absence of observed bands as described by (Sneath and Sokal, 1973) can be calculated by one of the three coefficients: Simple Matching similarity coefficient (SM) (Sokal and Michener, 1958), Jaccard's similarity coefficient (J) (Jaccard, 1908) and Dice similarity coefficient (D) equivalent to Nei – Li coefficient. Among these three coefficients, both J and D don't involve shared absence of bands (Nei and Li, 1979; Reif et al., 2005).

2.8.4. Correlation Analysis using COPH and MXCOMP of NTSYS for the selection of best fitted cluster analysis

COPH module performs three types of computations. First, it can take a hierarchical system of clusters and produce a symmetrical matrix of "cophenetic" (ultrametric) similarity or dissimilarity values. These can be used to test for the goodness of fit of a clustering to a set of data. Second, it can take a rooted or unrooted tree and produce a matrix of path length distances between all pairs of OTUs. This could be used to test the goodness of fit of an additive tree to a dissimilarity matrix. Third, it can take a rooted tree and produce a matrix giving the expected variances and covariances based on a Brownian motion model of evolution (Rohlf, 2009).

The program MXCOMP performs two kinds of matrix comparisons i.e. can take two symmetric similarity or dissimilarity matrices and plot one matrix against the other element by element (but with the diagonal value ignored) and can linearly adjust two matrices for the effects of a third matrix and then plot them against each other for a 3-way Mantel test. In both cases, it computes the product-moment correlation, r, and the Mantel test statistics, Z, to measure the degree of relationship between the matrices (Rohlf, 2009).

Cophenetic value matrix (COPH matrix of ultrameric values) from a tree matrix can be used by the MXCOMP program to measure the goodness of fit of a cluster analysis to the similarity or dissimilarity matrix on which it was based (Sokal, 1979). The module can be used, for example, to compare trees produced by different clustering methods or to construct a tree that is a summary of many trees. The input to this module consists of sets of nested clusters in the form of a tree matrix - such as produced by the SAHN clustering module.

2.8.5. Correlation Analysis using CONSEN (Consensus Tree and Indices) of NTSYS

This module produces a consensus tree and computes consensus indices from two or more rooted labeled trees. A consensus tree is a tree (actually just a dendrogram since it consists of nested sets and their levels, the term "tree" is used in a more technical sense in graph theory) that represents the consensus topology (subset relationships) of two or more trees. A consensus index is a numerical value that indicates the degree to which the consensus tree is resolved (fully bifurcating). The trees being compared must all have the same number of OTUs (Rohlf, 2009).

2.8.6. Cluster Analysis

Cluster analysis refers a group of multivariate techniques whose primary purpose is to group individuals or objects based on the characteristics they possess, so that individuals with similar descriptions are mathematically gathered into the same cluster (Hair et al., 1995). There are two main ways of analyzing the resulting similarity or dissimilarity matrix and displaying the results (Karp et al., 1997b). One is a dendrogram (or tree diagram) linking together in clusters samples that are more genetically similar to each other than to samples in other clusters. More widely used algorithm for such cluster analysis are Unweighted Pair Group Method with Arithmetic Averages (UPGMA), Neighbour-joining Method and Ward's Method (Karp et al., 1997b). Another approach is the use of Principal Coordinate Analysis (PCO) to produce 2 or 3-dimentional scatter plot of the samples such that the geometrical distances among samples in the plot reflect the genetic distances among them with a minimum of distortion (Karp et al., 1997b). 'Eigen value' defined the amount of total variation displayed on the principal co-ordinate (PC) axes, e.g. the first PC summarizes most of the variability present in the original data relative to all the remaining PCs (Kovach, 2007). Both the tree diagram and the PCO are called 'Phenetic' method as they are based on measures of overall distance or similarity among samples.

2.8.7. Diversity Indices for Dominant and Codominant Data

The main problem with using the dominant data derived from multilocus screening methods, to estimate diversity statistics is that the frequency or heterozygous is unknown, as they are indistinguishable from homozygotes. Hence, it is not possible to assess directly whether a particular population is in Hardy-Weinberg equilibrium. There are two approaches to deal with this problem. The first uses methods that ignore this problem and thus limitation must be borne in mind when the results are considered (eg. Nei's and Shannon's diversity indices). The second address the shortcomings of the data

and uses specific statistics to offset these drawbacks (Lynch and Milligan, 1994; Lowe et al., 2004).

2.8.7.1. Shannon Index of Diversity (Nei, 1987)

Shannon's index of diversity (I) is the diversity measure widely used in ecology but applied to genetics (Lewontin, 1972). Shannon's index, I (Shannon, 1948), originally developed as a measure of entropy in information theory, has become widely used statistic for quantifying levels of diversity (Lewontin, 1972). It is generally preferred, as the resultant values are normally distributed which allows more refined statistical tests to be applied to the data, and for each locus Shannon's index produce values from 0 – 0.73 when natural log is used (Lowe et al., 2004).

2.8.7.2. Observed Heterozygosity

This value is the number of heterozygotes recorded at a particular locus expressed as a proportion of the total number of loci surveyed. Number of heterozygotes divided by the total number of individuals typed is observed heterozygosity. This measure is widely applied to co-dominant data for diploid organisms (Lowe et al., 2004).

2.8.7.3. Expected Heterozygosity

It is the probability that, at a single locus, any two alleles, chosen at random from the population, are different to each other. Expected heterozygosity is calculated using an unbiased formula from allele frequencies assuming Hardy-Weinberg equilibrium. It is a useful measure of informativeness of a locus: loci with expected heterozygosity of 0.5 or less are in general not very useful for large-scale parentage analysis. In general when summarising the properties of a locus, it is customary to quote expected heterozygosity rather than observed heterozygosity. (Nei, 1987).

2.8.7.4. Nei's Index

Nei's method was originally derived for use with (co-dominant) proteins, and is based on a measure of expected heterozygosity. For dominant data, the concept of heterozygosity is not applicable and the estimate becomes 'gene diversity' simply a measure of variability but still of statistical value (Nei, 1987). The average expected heterozygosity is Nei's gene divesity. The average gene diversity is the average of this quantity across all loci. The value is adjusted for variation in sample sizes (Nei, 1978) through multiplication by 2n/ (2n – 1), where n = sample size. For each locus Nei's index produces values from 0 – 0.5 when the natural log is used (Lowe et al., 2004).

2.8.8. AMOVA and Indicator of Genetic Differentiation ϕ_{PT}

AMOVA (Analysis of Molecular Variance) is a method of estimating population differentiation using the molecular data. This method can analyze a variety of molecular data like molecular marker data (RAPD, ISSR or AFLP), direct sequence data, or phylogenetic trees based on such molecular data (Excoffier et al., 1992). AMOVA can be performed in GenAlEx (genetic analysis in excel), Arlequin, WINAMOVA in RAPD, ISSR and AFLP to partition the total molecular variance between and within populations. AMOVA is recent statistical procedure for the hierarchical partitioning of genetic variation among populations and regions, and the estimation of widely used F-statistic and/or their analogues. ϕ_{PT} is the estimation of population genetic differentiation provided by GenAlEx when binary or haploid data are analyzed for comparative studies (Peakall and Smouse, 2012).

2.8.9. Principal co-ordinate Analysis (PCO)

Principal coordinates analysis (PCO) is a multivariate technique that allows one to find and plot the major patterns within a multivariate data set. The mathematics is complex, but in essence, PCO is a process by which the major axes of variation are located within a multidimensional data set and can be viewed as a more general form of PCA. It then performs an eigenanalysis of the matrix, giving eigenvalues and eigenvectors. It gives a direct ordination of the cases and is useful in situations where there are more variables than cases (PCA is not recommended under this circumstance). The disadvantage is that no results are given for the variables, unlike PCA, which provides component loadings. The main advantage of PCO is that many different kinds of similarity or distance measures can be used (Kovach, 2007).

2.9. Diversity study of Citrus spp.

Large diversity existing now-a-days within the genus *Citrus* is related to the high number of taxonomic units (species and hybrid), apomixis, widely sexual compatibility between *Citrus* and related genera, the high frequency of bud mutations and the long history of cultivation and wide dispersion including natural and human selection (Scora, 1988; Hvarleva et al., 2008). Genetic improvement of Citrus spp. through conventional breeding methods has been hampered by the long juvenile period, high heterozygosity, large plant size and nucellar embryony (Gmitter et al., 1992). Genetic relationships in a germplasm collection of 198 cultivars and accessions of 54 species belonging to Citrus and 13 related genera by means of ten isoenzymatic systems were studied and found that the Citrus spp. were distributed in two main groups: the orange-mandarin group and the lime-lemon-citron-pummelo group (Karp et al., 1997a). Several previous studies

have utilized various molecular markers (ISSR, RAPD, AFLP, and SSR) to fingerprint accessions, evaluate phylogenetic relationships among accessions, and examine the level of genetic diversity in Citrus. Many of these studies have targeted specific Citrus groups or sampled a few individuals of each taxon. The variability of 24 Clementine (*C. reticulate* Blanco) accessions by utilizing ISSR, RAPD, and AFLP markers examined and found that only two varieties of 24 could be distinguished (Bretó et al., 2001). Assessment of diversity, phylogenetic relationships, and parentage in lemon accessions and related taxa were carried out utilizing ISSR, SSR, and isozymes, and found little genetic variation among lemon accessions (Gulsen and Roose, 2001). For the facilitation of a representative collection of lemon cultivars, an early classification of lemon genotypes are required. So in Spain, a study of molecular discrimination of lemon cultivars were carried out using RAPD, ISSR and IRAP (Inter-retrotransposon amplified polymorphism) markers (Bernet et al., 2004). In another study, 48 trifoliate orange (Poncirus trifoliata) accessions were classified into four groups employing isozymes, RFLP, and ISSR markers to (Fang et al., 1997). Similarly, ISSR markers were utilized to distinguish closely related cultivars, many of which had arisen by selection of spontaneous mutations. This study showed that ISSR markers could distinguish some (but not all) of these closely related accessions (Fang and Roose, 1997). RAPD, SCAR, and cpDNA markers were used to elucidate phylogenetic relationships and genetic origins of hybrids in 36 accessions of Citrus and one accession from each of four related genera (Nicolosi et al., 2000). The phylogenetic relations of 88 accessions representing 45 Citrus spp. and six related genera were examined by utilizing RFLP and RAPD markers (Federici et al., 1998). The genetic diversity and population structure in Citrus germplasm collection, University of California, Riverside was assessed employing SSR markers and found that *Fortunella* clusters within the genus Citrus but *Poncirus* is a sister genus to Citrus and supports the hypothesis of few naturally occurring species of Citrus and most other types are arose through various hybridization events between naturally occurring forms (Barkley et al., 2006). Employing microsatellite markers in Iran, genetic variability assessment in some Iranian sweet orange and mandarins were carried out and found that majority of sweet orange accessions showed a narrow genetic base suggesting that the observed morphological polymorphism within the group must be associated with somatic mutations and the markers sufficiently discriminated variation within the mandarins (Golein et al., 2005). In Cyprus, Citrus cultivars and clones were characterized using SSR and RAPD markers which reveal clone specific RAPD markers and both markers allowed discrimination of studied accessions at species level (Hvarleva et al., 2008). Now-a-days, grafted saplings are commonly used than seedlings, so Tunisian Citrus rootstocks germplasm genetic diversity was assessed using SSR markers and found diversity due to high heterozygosity (Snoussi et al., 2012). In the phenetic relationship

analysis between Citrus *indica* Tanaka and few commercially important Citrus spp. by ISSR marker, revealed diverse genetic relationship between Citrus *indica* and other five Citrus spp.: *C. reticulate*, *C. sinensis*, *C. aurantifolia*, *C, macroptera* and *C. maxima* (Chibame et al., 2010).

2.9.1. Diversity study of Citrus spp. in Nepal

Nepal is one of the countries where Citrus was believed to be originated and is rich in Citrus diversity. *Citrus* spp. are the cash generating fruit crops of mid-hill region of Nepal (Lama and kayastha, 1999). The diversity study of Citrus in Nepal is chiefly based on the morphological qualitative (tree shape, fruit shape, fruit base shape, pulp color, seed surface and color etc) and quantitative characters (tree height, fruit weight, juice content, total soluble solids, citric acid content, number of seeds per fruit etc). Most of the study included single individual *Citrus* spp. or few *Citrus* spp. The study of *Citrus* spp. diversity in Pokhara Valley and its surrounding areas (Lama and kayastha, 1999) was done in terms of morphology and found large genetic diversity. They realized many local landraces and cultivars of Citrus may have several desirable traits such as disease resistance, good fruit quality, tolerant to climatic conditions, etc. Farmer's participatory selection of superior pummelo (Citrus *grandis* L. Osbeck) trees was carried out in Terai and mid-hills of Nepal to assess the diversity, consumer's preference for quality fruit (Paudyal, 2000). Similarly, study of pummelo at unexplored area of Sankhuwasabha and Bhojpur district was conducted to assess the phenotypic variation in order to identify the superior phenotypes in respect to fruit quality and yield (Rijal, 2002). The possibility of early production of two Satsuma mandarin (Citrus *unshiu* Marcvitch) namely Okitsu Wase and Miygawa Wase was explored by evaluating for their fruit characters. The study indicated that warmer areas of 900-1000 m altitude could be the suitable locations for early production of the varieties (Paudyal and Chalise, 2007). Harvesting period of sweet orange in Nepal is very narrow (January-February) as most plantations are composed of mid-season variety. Evaluation study of fourteen variety of sweet orange on the basis of fruit characters and maturity period was carried out at National Citrus Research Program, Dhankuta to select early, mid and late maturing varieties for the mid-hill ecological region of the country to extend harvesting duration (Paudyal and Subedi, 2008). Due to limited resources and lab facilities only few high level of diversity study (biochemical and molecular) had been done in Nepal. The genetic diversity of Citrus *reticulate* Blanco and Citrus *limon* (L) Burm. F. of Nepal was analyzed through isozyme analysis. They used 19 population (10 of lemon and 9 of mandarin) and four enzyme system namely peroxidase (PRX), malate dehydrogenase (MDH), shikimate dehydrogenase (SKD) and malic acid (ME) for this study. Fourteen different phenotypes

identified in this study suggesting considerable genetic diversity of Citrus population in Nepal. Populations of lemon clustered in one group where that of mandarin scattered at various level of similarity ranging from 95% to 100%. This study indicates lemon population of Nepal is very much similar whereas mandarin showed amount of inter-population as well as intra-population diversity (Pahari and Bimb, 2004).

2.9.2. Diversity study of lime

In situ germplasm evaluation was undertaken in Terathum, Jhapa, Morang, Sunsari, and Chitwan districts in 2002 and 2003 to study diversity and identify superior lime genotypes for off-season production on the basis of tree and fruit characteristics. Maturity period of all the acid lime genotypes from hill was October to January but genotypes that can be harvested from July to November were identified from Terai of which two accessions maturing during rainy season with better fruit quality have been selected for commercial production (Paudyal and Shrestha, 2004). Participatory varietal selection study on acid lime was carried out by NCRP Dhankuta in farmer's fields of Jhapa, Morang, Sunsari and Chitwan during 2062/63 to 2066/67. On the basis of growth, yield, fruit quality, maturity period, tolerant to gummosis and citrus canker disease, NCRP-49, NCRP-53 and NCRP-55 were identified as superior types while NCRP-49 and NCRP-55 can be harvested from the first week of Asar to and end of Bhadra (Hamrick and Godt, 1996). Eighteen landraces of acid lime (Citrus *aurantifolia*) were characterized and evaluated using IPGRI descriptors for Citrus in the year 2004/05 to find superior landraces in Horticulture farm at IAAS (Institute of Agriculture and Animal Science), Rampur, Chitwan, Nepal. Landraces 010P12, 37P3, 24P3 and 01P3 were found suitable, both for the main and off-season crop after winter (Sapkota, 2006).

Till now, a single study of Citrus (acid lime) genetic diversity employing molecular marker has been done. The genetic diversity of acid lime (Citrus *aurantifolia* Swingle) Landraces at different agro-ecological zone in Nepal were assessed using SSR (Simple Sequence Repeat) markers. The average number of allele was detected 2.91 in Terai accessions followed by 2.82 and 2.73 in high hill and mid hill accessions respectively. Proportion of polymorphic allele was found 100% in Terai and high hill. The average polymorphic information content (PIC) value was observed higher 0.53 in Terai as compared to 0.50 in mid and 0.49 in high hills. Average genetic similarity index among the acid lime found to be 0.80 where as lowest similarity index was 0.77 observed in Terai and highest of 0.82 in the mid hills. The result of this study shows that, moderate genetic diversity of acid lime accessions was found in the Terai as compared to mid hill and high hill of Nepal (Shrestha et al., 2012a).

CHAPTER 3: MATERIALS AND METHODS

3.1. Materials

3.1.1. Reagents and Extraction buffers

TAE (Tris-Acetate-EDTA) buffer

Stock solution of TAE (50X) buffer was prepared by dissolving 242 gm Tris base (Qualigens Fine Chemicals, Mumbai) 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 up to 1L. The final solution is (50X) TAE stock buffer. This stock solution was stored at room temperature (Sambrook and Russell, 2001).
The working solution of 1X TAE buffer was made by diluting the stock solution (50X) in deionised water or ddH_2O. Final solute concentrations were 40mM Tris-acetate and 1mM EDTA.

Gel Loading Buffer (GLB)

5X GLB (Gel Loading Buffer) was prepared by dissolving 2.5g of sucrose in 7ml deionized water in which 25mg bromophenol blue (Fermentas Life Sciences, Canada) was added and the final volume made up to 10ml. This prepared gel loading buffer (GLB) was added to the sample in proportions as 1 (GLB) to 1 (DNA sample) and 1 (GLB) to 4 (PCR product) by volume, during electrophoresis.

Agarose Gel

Agarose (2 gm for ISSR analysis and 4 gm for SSR analysis; Promega, Spain) was dissolved in TAE buffer (100ml, 1X) by heating in microwave. It was then cooled to approximately 55°C and poured on to the gel casting tray with an appropriate comb (8-17 toothed) fixed in place for well formation. The gel was cooled to room temperature before further use.

Tris Buffer (1M)

1 M Tris Buffer of PH 8 or 7.5 as stock solutions were prepared by adding tris base (60.55gm, Qualigens Fine Chemicals, Mumbai) to deionized water (400ml). The pH was adjusted to 8.5 or 7.5 by the addition of concentrated HCL. The final volume was then made up to 500ml, autoclaved and stored at room temperature until needed (Sambrook and Russell, 2001).

EDTA (0.5M, pH 8.0)

93.05gm Disodium Ethylene Diamine Tetra acetate.$2H_2O$ (EDTA, Promega, Product of Spain) was added to a Schott bottle containing deionized water (400ml), mixed by stirring and the pH was adjusted to 8.0 by adding NaOH pellets (approximately, 10gm). The volume was adjusted to 500ml with deionized water, autoclaved and stored at room temperature until needed.

NaCl (4M)

4M Sodium Chloride was prepared by dissolving, 117gm NaCl, (Qualigens Fine Chemicals, Mumbai) in a Schott bottle containing deionized water (400l), mixed by stirring. The final volume was adjusted up to 500ml with deionized water, autoclaved and stored at room temperature.

Graham's CTAB (Cetyl Trimethyl Ammonium Bromide) extraction buffer (2% CTAB, 1.4m NaCl, 0.1M EDTA, 0.1M Tris HCl, pH 8.0) (Graham *et al.*, 1994)

CTAB extraction buffer was prepared by mixing 100 ml of 1M Tris buffer (p^H 8.0) and 200ml of 0.5M EDTA solution, (P^H 8.0) and 350ml of NaCl solution (4M) and 20gm CTAB (Loba Chemie, India) in a 1L sterile Schott bottle, the final volume was made up to 1L with deionized water.

TE buffer (Tris-EDTA; 10mM Tris HCl, 1mM EDTA) with RNAse

1ml of EDTA (0.5M stock; pH 8.0) was added to a bottle containing 5ml of Tris buffer (1M solution pH 8) and the final volume made up to 500ml. The solution was autoclaved and stored at room temperature. 200µl RNAse solution (5mg/ml RNAse A) was added to 200ml of Tris-EDTA buffer to make final concentration of 10µg/ml in sterile Schott bottle for fresh use.

3.1.2. Primers, dNTPs and DNA Dilution

Primers were diluted to the working concentration of 10µM by sterile distilled water.
Commercially supplied dNTP mix (10mM each, Fermentas Life Sciences) was used for PCR reaction. It was stored at -20°C until use.
Dilution of genomic DNA to required concentration was carried out by initial estimation of concentration of DNA by using Bio-photometer (Eppendorff, Germany).

3.1.3. Plant Materials

About 19-21 lime trees based on morphological and framer's preferences were selected for study from each of the three agro-ecosites such as Terai (< 600 masl), mid-hill (600-1200 masl) and High-hill (> 1200 masl) of Eastern Development Region of Nepal (fig 3.1). A total of 60 young leaf (6-8 weeks after sprout) samples of acid lime (*C. aurantifolia*) landraces were collected randomly from selected trees at different altitudinal range by the Phd student Mr. Ram Lal Shrestha (Table 3.1). The samples were used in part I of the study i.e. ISSR-based genetic diversity assessment of acid lime.

Table 3.1 Altitudinal range, accessions number and locality details of sample collection sites of acid lime landraces

Above 1200 masl			600-1200 masl			Less than 600 masl		
Acc no	**Altitude, m**	**VDC-Ward no**	**Acc no**	**Altitude, m**	**VDC-Ward no**	**Acc no**	**Altitude, m**	**VDC-Ward no**
LT-1	1605	Okhre-8	LD-49	1185	Bodhe-1	LM-43	135	Sunpur-2
LT-17	1750	Fachmara-7	LKv-60	1285	Balara-1	LM-44	135	Sunpur-2
LT-18	1710	Fachmara-9	LKm-61	1285	Balara-1	LD-45	135	Sunpur-2
LT-15	1655	Fachmara-9	LKr-62	1285	Balara-1	LD-58	135	Sunpur-2
LD-50	1638	Rajarani-9	LD-48	1181	Bodhe-1	LS-34	128	Narsing-2
LT-8	1505	Okhre-8	LD-25	1180	Balara-1	LS-35	128	Narsing-4
LT-22	1505	Sudap-1	LD-26	1175	Balara-1	LS-36	128	Narsing-4
LT-9	1500	Okhre-5	LD-27	1175	Balara-1	LS-37	128	Narsing-4
LT-21	1485	Fachamara-1	LD-28	1175	Balara-1	LS-38	128	Narsing-4
LT-20	1410	Fachamara-8	LD-29	1175	Balara-1	LS-39	128	Narsing-4
LT-16	1405	Fachamara-7	LD-30	1175	Balara-1	LS-40	128	Narsing-4
LT-19	1350	Fachamara-7	LD-59	1175	Balara-1	LS-41	128	Narsing-4
LT-13	1315	Fachamara-7	LT-4	1155	Okhre-1	LS-42	128	Narsing-4
LT-12	1310	Fachamara-7	LT-5	1155	Okhre-3	LS-56	128	Narsing-4
LT-14	1308	Fachamara-7	LT-6	1150	Okhre-3	LS-57	128	Narsing-4
LT-23	1308	Sudap-7	LD-31	1150	Dhnk -3	LM-51	125	Pathari-2
LT-3	1305	Okhre-8	LT-7	1145	Okhre-2	LM-52	125	Pathari-2
LD-24	1290	Balehara-8	LT-10	1135	Okhre-3			
LT-2	1285	Okhre-1	LT-11	1130	Okhre-3	LM-54	125	Pathari-2
LD-46	1278	Bodhe-2	LD-32	1130	Balhra-3	LM-55	125	Pathari-2
			LD-33	1130	Balhra-1	-	-	-

LT, Lime Terhathum. LD, Lime Dhankuta. LM, Lime Morang. LS, Lime Sunsari. LKm, Lime madrasi. LKr, Lime Rampur. LKv, Lime Vanarasi, VDC, Village Development Committee. M, meter. asl, Above sea level.

Source: (Shrestha et al., 2012a)

Eighty four and 29 young leaf samples of citrus accessions belonging to different species (total 113 samples) were collected from National Citrus Research Program (NCRP), Paripatle, Dhankuta and the Kathmandu valley respectively (fig 3.2). Among them only

45 samples were used for the SSR based study i.e. part II of this investigation (Appendix 2 and Table 4.1).

Each acid lime plant was designated with accession code with respect to the district. The *Citrus* spp. samples were designated by numbers. The samples were packed in small Whatman lens cleaning tissue (lens paper) pouches after washing with clean water and placed in a bag with silica gel and brought to Molecular Biotechnology Laboratory, Nepal Academy of Science and Technology (NAST), Khumaltar for DNA extraction and subsequent study.

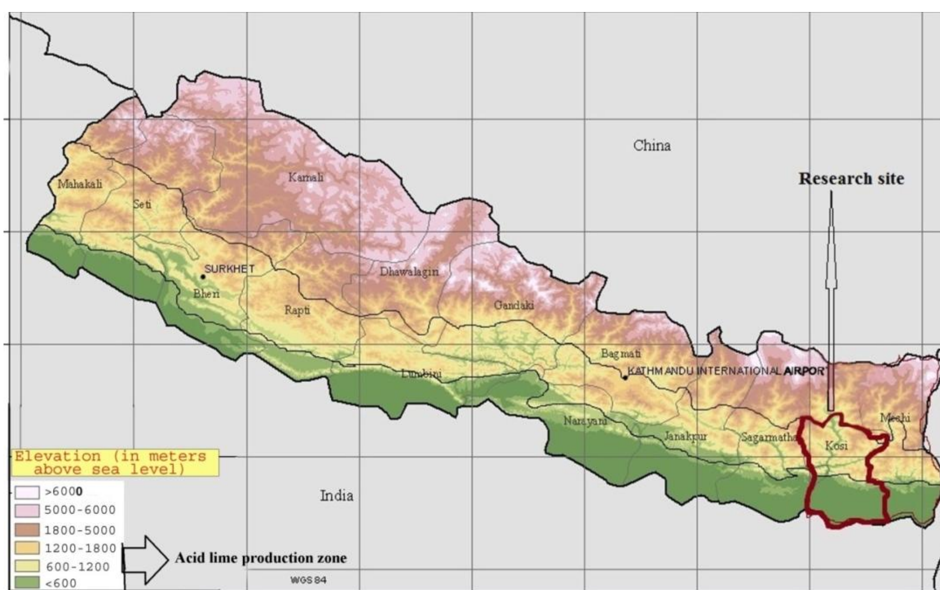


Figure 3.1 Map of Nepal showing acid lime samples collection sites

Figure 3.2 Map of Nepal showing *Citrus* spp. samples collection sites

3.2. Methodology

3.2.1. DNA extraction methods

Two major DNA extraction techniques were employed for DNA extraction.

3.2.1.1. CTAB Extraction Method (Graham et al., 1994a)

Approximately 200 mg of preserved leaf of various *Citrus* spp. was ground in sterilized mortar and pestle (sterilization done with 2% sodium hypochloride and autoclaved at 121°C and 15 lbs pressure) using liquid Nitrogen. The ground samples were treated with 1000µl CTAB buffer and transferred to sterile Eppendroff tubes (1.5 ml). The mixtures were then incubated at 65°C for 15 minutes. Following this, the tubes were centrifuged at 11000 rpm (∼13000g) for 15 minutes and the supernatant were transferred to a clean sterile microfuge tubes and equal volumes (∼500µl) of chloroform isoamyl alcohol (24:1) were added. The solutions were mixed gently by inversions several times. Tubes were then centrifuged at 11000 rpm for 2 minutes at 25°C and the upper aqueous phase was transferred to clean Eppendroff tubes and re-extracted with equal volume of chloroform isoamyalchol (24:1). Tubes were centrifuged for 2 minutes at 11000 rpm and upper aqueous layer was transferred to new tubes. Each sample was then treated with 1/10th volume (approximately 50 µl) of Ammonium acetate (7.5 M) followed by addition of equal volume (i.e 500 µl) of ice-cold (-20°C) absolute ethanol. Then the tubes were placed overnight at -20°C to allow precipitation of DNA.

Next day, the tubes were centrifuged at 11000 rpm for 20 minutes at 4°C. The supernatants were discarded and the pellets were quickly washed twice with ice cold 70% ethanol (stored at -20°C; ∼300µl), washing by adding alcohol around wall of Eppendroff tube gently and pipetting without touching the DNA. In order to get rid of unwanted salts, samples were briefly spinned at 11000 rpm for 1 minute and excess alcohol pipetted off. Pellets were then dried in laminar air flow for 5 minutes after centrifugation. Finally, pellets were resuspended in TE buffer (with RNase), centrifuged for few seconds for touchdown and stored at -20°C until use.

3.2.1.2. DNeasy Mini Kits (Qiagen Ltd, Crawley, UK)

The DNA extraction of acid lime samples using DNeasy mini kits according to Qiagen catalogue provided.

A water bath was turned on and heated to 65°C and the supplied AE buffer was incubated. Approximately, 100 mg collected preserved leaves of acid lime was ground to a fine powder in a sterilized mortar and pestle in presence of liquid nitrogen. Then, the

grounded tissue was put into labeled microfuge tube and 400 µl AP1 buffer was added. To this, 4 µl Rnase A (100 mg/ml stock) was added and vortexed to remove clumps. This was then incubated at 65°C for 10 minutes to remove the lysed cells and the contents were mixed by inverting the tubes 2-3 times. Then, 130 µl AP2 buffer was added, mixed and incubated on ice for 5 minutes and centrifuged for 5 minutes at 20,000g (14000 rpm). After that the supernatant was transferred to liliac QIAshredder spin column in a collection tube and spinned for 2 minutes at full speed of 13000 rpm. Flow through but not the pellet was transferred to a clean microgfuge tube labeled with genotype name. 225 µl AP3 buffer and 450 µl 100% ethanol was added to cleared lysate very gently and mixed with the tip and then 650 µl of the mixture applied into the Dneasy column (clear) in a 2 ml collection tube and spinned for 1 minute at 8000 rpm. The remaining sample was repeated and reused the collection tube to spin again for 2 minutes at 8000 rpm and the flow through was discarded. Dneasy column was kept in a new 2 ml tube and 500 µl AW buffer was added and spinned for 1 minute at 8000rpm. Flow through butt was discarded and a collection tube with pellet was kept. Again 500 µl AW buffer was added to Dneasy column and spinned for 2 minutes at maximum speed of 13000 rpm to dry column membrane. The column was carefully removed and the collections tubes and contents were discarded. Finally, DNeasy column was transferred to a microfuge tube and 100 µl preheated AE buffer was pipetted out directly onto the column membrane. Then it was incubated at room temperature for 5 minutes and spinned for 1 minute at 8000 rpm. The last two steps were repeated again with the addition of 100 µl of AE buffer for elution. The eluted DNA samples were properly labeled and stored at -20°C for further use.

3.2.2. DNA Quantification

The yield of DNA extracted from leaf tissue was measured using spectrophotometric method (Biophotometer, Eppendrof-AG22331, Germany). Each DNA sample was quantified for its concentration along with its purity assessment using the Biophotometer.

3.2.3. Gel Electrophoresis

The extracted DNA was analyzed on a 1.5% Agarose gel in TAE buffer (1X) at 50V (8.33 V/cm) for half an hour, where as amplified products of ISSR-PCR were analyzed on a 2% Agarose gel at 50V (8.33 V/cm) for two hours and amplified products of SSR-PCR on a 4% Agarose agarose (Sambrook and Russell, 2001; Semagn et al., 2006) at 70V (11.67 V/cm) for one and half hour using Major Science gel tank (Taiwan). Total volume loaded in well was 4µl DNA and 1µl GLB (6X) for DNA analysis and 10µl PCR product and 2µl GLB (6X)

for product analysis. The Ethidium Bromide (EtBr) (10mg/ml; Promega) was added during gel preparation at a concentration of 5µl per 100ml gel. After gel run, the gels were visualized on a gel documentation system (IN GENIUS, Syngene Bio-imaging, UK).

3.2.4. PCR Optimization

All the experiments of PCR were carried out in thermal cycler (BIOER, China). PCR reaction conditions were optimized by varying several PCR parameters. These parameters includes: PCR Cycling conditions, DNA concentration, $MgCl_2$ concentration, Primer concentration, dNTPs concentration and Taq polymerase concentration and cycling conditions. The four different ISSR-PCR programs were assessed for the selection of best program. Similarly, four different SSR-PCR programs were assessed for the selection of best program.

Different ranges of concentration of PCR-reaction parameters *viz*. $MgCl_2$, Taq polemerase, dNTPs, template DNA concentration and primer concentrations were assessed including PCR cycling conditions for the selection of best conditions for *Citrus aurantifolia* (acid lime) and different *Citrus* spp. For the selection of best and productive results, a range of DNA concentrations (1.2, 25.0, 37.5, 50.0, 62.5, 75.0, 87.5 and 100.0 ng) of sample were assessed. Similarly, different $MgCl_2$ concentrations (1.5, 2.0, 2.5, 3.0, 3.5 and 4.0 mM), dNPrimer concentrations (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 and 1.6 µM), dNTPs concentrations (0.1, 0.2, 0.3, 0.4 and 0.5 mM) and Taq polymerase concentrations (0.5, 1.0, 1.5, 2.0 and 2.5 Unit) were assessed for the selection of best and optimum concentrations for PCR amplification sample DNA templates. For controlling false positive results, a negative control reaction was run in each optimization experiments.

3.2.5. Primer screening

Using the optimized ISSR-PCR reaction and cycling conditions, 49 different ISSR primers; one (C7 or $HVH(GA)_7T$) from Shahsavar et al. (2007), four (C1, C2, C4 and C5) from Chibame et al. (2010) and three (C8, C9 and C10) from Fang et al. (1998) and remaining from UBC set 9, University of British Columbia, Oligonucleotides Synthesis Laboratory, Vancouver, British Columbia, Canada were screened against the *Citrus aurantifolia* (acid lime) genomic DNA (Table 3.2).

Twelve different SSR primers were selected and used. Five were already used and designed by Golein et al. (2005), next five by Jannati et al. (2009) and remaining two by

Barkley et al. (2006). The primer pairs were synthesized by Eurofins Genomic Test Pvt Ltd, Banglor, India (Table 3.3).

Table 3.2 Sequences and GC content of screened primers used in ISSR-PCR analysis

SN	Primer Code	Primer Sequence (5' - 3')	Primer Length	GC content (%)
1	C1	TCTCTCTCTCTCTCTCTCCC	20 mer	55
2	C2	AGCAGCAGCAGCGT	14 mer	64.3
3	C4	CTCCTCCTCGC	11 mer	72.7
4	C5	CACCACCACGC	11 mer	72.7
5	C7	HVHGAGAGAGAGAGAGAT	18 mer	47.2
6	C8	TCCTCCTCCTCCTCCRY	17 mer	64.7
7	C9	BDBTCCTCCTCCTCCTCC	18 mer	63.9
8	C10	HVHTCCTCCTCCTCCTCC	18 mer	63.9
9	UBC 807	AGAGAGAGAGAGAGAGT	17 mer	47.1
10	UBC 809	AGAGAGAGAGAGAGAGG	17 mer	52.9
11	UBC 810	GAGAGAGAGAGAGAGAT	17 mer	47.1
12	UBC 811	GAGAGAGAGAGAGAGAC	17 mer	52.9
13	UBC 812	GAGAGAGAGAGAGAGAA	17 mer	47.1
14	UBC 817	CACACACACACACACAA	17 mer	47.1
15	UBC 818	CACACACACACACACAG	17 mer	52.9
16	UBC 819	GTGTGTGTGTGTGTGTA	17 mer	47.1
17	UBC 820	GTGTGTGTGTGTGTGTC	17 mer	52.9
18	UBC 821	GTGTGTGTGTGTGTGTT	17 mer	52.9
19	UBC 822	TCTCTCTCTCTCTCTCA	17 mer	47.1
20	UBC 824	TCTCTCTCTCTCTCTCG	17 mer	52.9
21	UBC 825	ACACACACACACACACT	17 mer	47.1
22	UBC 826	ACACACACACACACACC	17 mer	52.9
23	UBC 830	TGTGTGTGTGTGTGTGG	17 mer	52.9
24	UBC 834	AGAGAGAGAGAGAGAGYT	18 mer	47.2
25	UBC 835	AGAGAGAGAGAGAGAGYC	18 mer	52.8
26	UBC 836	AGAGAGAGAGAGAGAGYA	18 mer	47.2
27	UBC 840	GAGAGAGAGAGAGAGAYT	18 mer	47.2
28	UBC 841	GAGAGAGAGAGAGAGAYC	18 mer	52.8
29	UBC 842	GAGAGAGAGAGAGAGAYG	18 mer	52.8
30	UBC 844	CTCTCTCTCTCTCTCTAGC	19 mer	52.6
31	UBC 845	CTCTCTCTCTCTCTCTAGG	19 mer	52.6
32	UBC 848	CACACACACACACACARG	18 mer	52.8
33	UBC 850	GTGTGTGTGTGTGTGTYC	18 mer	52.8
34	UBC 853	CTCTCTCTCTCTCTCTAGT	19 mer	47.4

35	UBC 856	ACACACACACACACACYA	18 mer	47.2
36	UBC 857	ACACACACACACACACYG	18 mer	52.8
37	UBC 860	TGTGTGTGTGTGTGTGRA	18 mer	47.2
38	UBC 864	ATGATGATGATGATGATG	18 mer	33.3
39	UBC 867	GGCGGCGGCGGCGGCGGC	18 mer	100
40	UBC 868	GAAGAAGAAGAAGAAGAA	18 mer	33.3
41	UBC 873	GACAGACAGACAGACA	16 mer	50.0
42	UBC 880	GAAGAGGAGAGGAGA	15 mer	60.0
43	UBC 886	VDVCTCTCTCTCTCTCT	17 mer	50.0
44	UBC 888	BDBCACACACACAGACA	17 mer	50.0
45	UBC 889	DBDACACACACACACACA	18 mer	47.2
46	UBC 890	VHVGTGTGTGTGTGTGT	17 mer	50.0
47	UBC 891	HVHTGTGTGTGTGTGTG	17 mer	50.0
48	UBC 895	AGAGTTGGTAGCTCTTGATC	20 mer	45.0
49	UBC 900	ACTTCCCCACAGGTTAACACA	21 mer	47.6

Where, H = non-G, Y = Pyrimidine, B = non-A, D = non-C, V = non-T

Table 3.3 Characteristics of SSR primers used in present study.

Primer code	**Primer sequence 5'-3' (F) and 5' – 3' (R)**	**Repeat motif**	**Allele no.**	**Allele size**	**PIC value**	**Reference**
TAA45	GCACCTTTTATACCTGACTCGG (F) TTCAGCATTTGAGTTGGTTACG (R)	TAA	10	132-171	0.95	(Golein et al., 2005)
TAA52	GATCTTGACTGAACTTAAAG (F) ATGTATTGTGTTGATAACG (R)	TAA	8	132-174	0.94	,,
TAA41	AGGTCTACATTGGCATTGTC (F) ACATGCAGTGCTATAATGAATG (R)	TAA	9	122-185	0.91	,,
TAA15	GAAAGGGTTACTTGACCAGGC (F) CTTCCCAGCTGCACAAGC (R)	TAA	9	141-204	0.89	,,
TAA3	AGAGAAGAAACATTTGCGGAGC (F) GAGATGGGACTTGGTTCACACG (R)	TAA	10	133-172	0.87	,,
CAT01	GCTTTCGATCCCTCCACATA (F) GATCCCTACAATCCTTGGTCC (R)	CAT / CTT	12	138-172	0.89	(Jannati et al., 2009)
CAC15	TAAATCTCCACTCTGCAAAAGC (F) GATAGGAAGCGTCGTAGACCC (R)	CAC	4	148-163	0.36	,,
TAA27	GGATGAAAAATGCTCAAAATG (F)	TAA	10	158-230	0.85	,,

	TAGTACCCACAGGGAAGAGAGC (R)					
CT19	CGCCAAGCTTACCACTCACTAC (F) GCCACGATTTGTAGGGGATAG (R)	CT	14	117-171	0.84	,,
TC26	CTTCCTCTTGCGGAGTGTTC (F) GAGGGAAAGCCCTAATCTCA (R)	TC	7	93-119	0.83	,,
AG14	AAAGGGAAAGCCCTAATCTCA (F) CTTCCTCTTGCGGAGTGTTC (R)	GA	20	119-163	0.85	(Barkley et al., 2006)
GT03	GCCTTCTTGATTTACCGGAC (F) TGCTCCGAACTTCATCATTG (R)	GT	19	149-197	0.83	,,

3.2.6. PCR Profiling

DNA from the leaf samples of all 60 acid limes were tested with screened primers for the study of genetic diversity and identification polymorphic markers. Similarly, DNA from the leaf samples of all 113 citrus accessions belonging to different species were tested with screened primers for same purposes.

3.2.6.1. Genetic Diversity Assessment based on PCR profiling

An amplified DNA fragment that was present in DNA samples of at least one accession and was absent in the DNA sample of at least one accession is known as a polymorphic fragment. The polymorphic fragments were identified in amplified DNA produced by different DNA samples. The selected standard accessions and the individual accessions were then screened for the presence or absence of observed polymorphic markers amplified with primers. The markers were scored visually. Molecular size of the PCR products were estimated by comparison with 100bp plus DNA ladder (Gene Ruler^{TM}, Fermentas LIFE SCIENCES, #SM0323), 100bp DNA ladder (Promega, Madison, WI 5371-5399, USA) and 100bp DNA ladder (BIONEER, D-1030). Total number of band (TNB), Number of polymorphic band (NPB) amplified by selected primers were estimated.

3.2.6.2. Statistical Analysis

The observed results from the different statistical processes were documented and analyzed for the study of genetic diversity of acid lime and different *Citrus* spp. of Nepal. Statistical tools used for diversity analysis are as follows:

Estimation of Percent Polymorphism, Polymorphic Information Content, Band Informativeness and Resolving Power

Primer banding characteristics *viz*. Percent Polymorphism (PP), Polymorphic Information Content (PIC), Band Informativeness (I_B) and Resolving Power (R_P) for each of the primers were calculated by the use of Microsoft Office Excel 2007.

Polymorphic Information content (PIC) values were calculated for each of the ISSR primer according to the formula: PIC = $1-\sum (P_{ij})^2$, where P_{ij} is the frequency of the ith pattern revealed by the jth primer summed across all patterns revealed by the primers (Botstein et al., 1980). For each primer percent polymorphism (PP) has been calculated as NPB_s/TNB_s, where NPB is the number of polymorphic bands and TNB is Total number of bands amplified.

For the assessment of primer combination ability to differentiate between accessions Resolving Power was calculated (Prevost and Wilkinson, 1999), $R_p = \sum I_B$, where I_B is the bandinformativeness with $I_B = 1 - [2 X (0.5 – P)]$, where P is the proportion of accessions containing the band.

Shannon's Information Index, Observed Heterozygosity, Expected Heterozygosity, Probability of Identity and Fixation Index

The Shannon's Information Index, observed and expected heterozygosity were calculated using GenAlEx ver. 6.5 (Peakall and Smouse, 2012) and the value tallied with the results obtained from PowerMarker V3.25.

Shannon's Information Index, I = -1 X ∑[pi X ln (pi)]

Observed Heterozygosity, Ho = Number of heterozygotes / N

Expected Heterozygosity, He = $1 - \sum pi^2$

Unbiased Expected Heterozygosity, uHe = (2N / (2N-1)) X He

Probability of Identity, PI = $2 X [\sum(pi^2)^2] - \sum(pi)^4$

Fixation Index, F = (He - Ho) / He = 1 - (Ho / He)

Where, pi = pi is the frequency of the ith allele for the population; N = Sample Size

Genetic Diversity Analysis using Similarity matrices and Phenograms

After fragment sizing and matching, the qualitative binary data matrix were created for bands presence or absence obtained from PCR profiles, generated by primers, from samples and were analyzed using Numerical Taxonomy and Multivariate System (NTSYS-PC, version 2.21i, Exeter software, Setauket, New York, USA). Bands were scored as discrete variables ("1" for presence and "0" for absence). Amplification failure was

scored as "9", which was designated in the analysis procedure as an indicator of missing data (Transue et al., 1994). Similarity indices were calculated using SimQual (Similarity for qualitative data) computational algorithm. Based on similarity matrices, the SAHN (Sequential, Agglomerative, Hierarchial and Nested) clustering was performed using UPGMA (Unweighted Pair Group method with Arithmetic Mean) algorithm (Sneath and Sokal, 1973). From these, phenograms were constructed to show the relationship among the *Citrus aurantifolia* (acid lime) landraces of Eastern Nepal. Estimates of similarity were computed on the basis of following three different measures:

1) Simple Matching Coefficient (SM) (Sokal and Michener, 1958)

S_{ij} = a+d/a+b+c+d

2) Dice's Coefficient of Similarity (D) (Dice, 1945; Nei and Li, 1979)

S_{ij} = 2a/2a+b+c

3) Jaccard's Coefficient (J) (Jaccard, 1908)

S_{ij} = a/a+b+c

Where,

S_{ij} = the similarity between two 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, and

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

Consensus indices (CI) were calculated using Strict consensus method for the each combination of coefficient and UPGMA clustering. Cophenetic value of each tree was generated using COPH algorithm. Mantel test was applied in NTSYS-PC 2.21i program through MXCOMP procedure in the comparison of original matrices and the matrices obtained from cophenetic values.

Using NTSYS-PC, 3D-plot of the distribution of all *Citrus aurantifolia* (acid lime) accessions and *Citrus* spp. were constructed for the illustration of variation as compared to phenogram using suitable similarity matrix with the analysis of Eigen vector.

Principal Cooridinates Analysis (PCO)

Relationship among the *Citrus aurantifolia* (acid lime) accessions as well as *Citrus* spp. was studied using PCO with MVSP version 3.2 (Multivariate Statistical Package) program (Kovach, 2007).

Analysis of Molecular Variance (AMOVA)

Genetic variation was evaluated for the binary data matrix created for all ISSR loci by analysis of molecular variance (AMOVA) (Excoffier et al., 1992) using the statistical program GENEALEx 6.5 (Peakall and Smouse, 2001). The AMOVA based on Euclidean metric of (Excoffier et al., 1992) is given by:

$E = (\epsilon^2_{XY}) = n\ [1-2n_{XY}/2n]$

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 (Peakall and Smouse, 2001).

CHAPTER 4: RESULTS

4.1. Collection of samples

4.1.1. Acid lime (*Citrus aurantifolia*)

Sixty samples of acid limes previously collected from different agro-ecological zones of eastern Nepal were used for ISSR investigation (Table 3.1).

4.1.2. Different *Citrus* spp.

Among the total 113 citrus accessions belonging to different species collected from National Citrus Research Program (NCRP), Paripatle, Dhankuta and Kathmandu Valley (Appendix 2), only 45 samples with at least two cultivars per species were selected for SSR based study, due to resource and time constraints. The samples included 7 mandarins, 8 sweet oranges, 1 sour orange, 2 grapefruits, 10 acid limes, 2 citrons, 2 lemons, 3 pummelos, 2 kumquats, 2 rootstocks i.e. trifoliate and citranze, 3 unknown samples and one each of tangore, tangelo and hokse (natural hybrid) (Table 4.1). The common name of *Citrus aurantifolia* is acid lime, *C. reticulata* is mandarin, *C. sinensis* is sweet orange, *C. aurantium* is sour orange (kalijyamir), *C. limon* is lemon, *C. grandis* or *C. maxima* is pummel, *C. paradisi* is grapefruit and *Fortunella* spp. is kumquat.

Table 4.1 Altitudinal range, Latitude, Longitude, Accession numbers, Sample collections sites of *Citrus* spp.

Acc no.	Name of *Citrus* spp.	Variety	Collection Site	Altitude (m)	Latitude (N)	Longitude (E)
3	*C. paradisi*	Marsheedless	Hort. Centre, Kirtipur	1328	27°40.751'	085°16.974'
5	*C. sinensis*	Local selection	Hort. Centre, Kirtipur	1330	27°40.715'	085°16.934'
9	*C. medica*	Local collection	Hort. Centre, Kirtipur	1333	27°40.701'	085°16.932'
10	*C. sinensis*	Washington Navel	Hort. Centre, Kirtipur	1351	27°40.509'	085°16.860'
16	*C. aurantium*	Unknown	Hort. Centre, Kirtipur	1352	27°40.479'	085°16.862'
24	Unknown	Unknown	Chahabil, Kathmandu	1345	27°40.470'	085°16.360'
25	Unknown	Unknown	Chahabil, Kathmandu	1346	27°40.429'	085°16.802'
35	*C. reticulata*	Khoku	NCRP-01, Dhankuta	1362	27°00.178'	087°18.555'
37	Natural hybrid	Unknown	NCRP, Dhankuta	1350	27°00.102'	087°18.492'
38	*C. unshiu*	Okitsuwase	NCRP-06, Dhankuta	1362	27°00.178'	087°18.555'
40	*C. clementina*	Nules	NCRP-101, Dhankuta	1362	27°00.178'	087°18.555'
42	*C. tangerina*	Dancy	NCRP-92, Dhankuta	1358	27°00.197'	087°18.585'
45	Old Citrange	Unknown	NCRP-38, Dhankuta	1345	27°00.154'	087°18.595'
51	*C. sinensis*	Valencia Late (local)	NCRP-13, Dhankuta	1354	27°00.193'	087°18.603'
52	*C. aurantifolia*	Unknown	NCRP-52, Dhankuta	1373	27°00.258'	087°18.642'
54	*Poncirus trifoliata*	Unknown	NCRP-36, Dhankuta	1330	27°00.082'	087°18.590'
56	*C. paradise*	Shamber	NCRP-45, Dhankuta	1358	27°00.264'	087°18.637'
59	*C. aurantifolia*	Unknown	NCRP-47, Dhankuta	1369	27°00.282'	087°18.649'
60	*C. limon*	Unknown	NCRP-53, Dhankuta	1376	27°00.284'	087°18.641'
62	*C. aurantifolia*	Unknown	NCRP-51, Dhankuta	1371	27°00.287'	087°18.644'
63	*C. aurantifolia*	Unknown	NCRP-55, Dhankuta	1373	27°00.282'	087°18.642'

65	*C. aurantifolia*	Unknown	NCRP-50, Dhankuta	1370	27°00.288'	087°18.646'
67	*C. sinensis*	Dhankuta Local	NCRP-33, Dhankuta	1350	27°00.192'	087°18.585'
68	*C. maxima*	Phultrac	NCRP-42, Dhankuta	1362	27°00.266'	087°18.633'
69	*C. sinensis*	Delicious Seedless	NCRP-27, Dhankuta	1345	27°00.175'	087°18.550'
70	*C. medica*	Dhankuta Local	NCRP-62, Dhankuta	1361	27°00.265'	087°18.638'
71	*C. sinensis*	Valencia Late (France)	NCRP-86, Dhankuta	1350	27°00.188'	087°18.594'
74	Natural hybrid	Local Hokse	NCRP-41, Dhankuta	1360	27°00.262'	087°18.636'
75	Tangelo	Seminole	NCRP-75, Dhankuta	1344	27°00.066'	087°18.572'
77	Mandarin	Kinnow	NCRP-02, Dhankuta	1362	27°00.276'	087°18.652'
80	*C. sinensis*	Malta Blood Red	NCRP-16, Dhankuta	1349	27°00.107'	087°18.575'
82	*C. sinensis*	Mosambi	NCRP-18, Dhankuta	1350	27°00.201'	087°18.608'
85	*C. aurantifolia*	Unknown	NCRP-48, Dhankuta	1377	27°00.288'	087°18.648'
87	*C. aurantifolia*	Unknown	NCRP-46, Dhankuta	1378	27°00.291'	087°18.680'
92	Tangor	Murkotte	NCRP-103, Dhankuta	1338	27°00.025'	087°18.571'
94	*C. aurantifolia*	Unknown	NCRP-49, Dhankuta	1374	27°00.287'	087°18.645'
97	*C. maxima*	Henderson	NCRP-76, Dhankuta	1342	27°00.087'	087°18.574'
103	*C. reticulata*	Kamal	NCRP-09, Dhankuta	1362	27°00.275'	087°18.650'
104	*Fortunella margarita*	Unknown	NCRP-105, Dhankuta	1358	27°00.266'	087°18.640'
105	*C. aurantifolia*	Unknown	NCRP-60, Dhankuta	1364	27°00.276'	087°18.636'
107	*Fortunella hindsii*	Unknown	NCRP-106, Dhankuta	1360	27°00.208'	087°18.641'
108	*C. mitis*	Kalamondin	NCRP-12, Dhankuta	1356	27°00.286'	087°18.658'
111	*C. limon*	Unknown	NCRP-57, Dhankuta	1365	27°00.279'	087°18.641'
112	*C. maxima*	Unknown	Bhaktapur	1307	27°40.084'	085°24.912'
113	*C. aurantifolia*	Unknown	Bhaktapur	1309	27°40.039'	085°25.009'

4.2. DNA Quantification

Quantification of DNA was done on Bio-photometer (Eppendorff, Germany). The DNA concentration of acid lime landraces extracted by using Qiagen kit ranged from 4 µg/ml to 126.8 µg/ml and purity range from 0.5 to 2.13 (Appendix 3). While the DNA concentration of *Citrus* spp. extracted by using CTAB method ranged from 61.8 µg/ml to 808.9 µg/ml and purity ranged from 1.17 to 2.07 (Appendix 4).

4.3. ISSR-PCR Optimization

4.3.1. ISSR-PCR Cycling conditions optimization

Among the four cycling conditions assessed (Fang et al., 1998; Shahsavar et al., 2007; Kumar et al., 2010; Lui et al., 2011) (Table 4.2), PCR program (cycling condition) described by Shahsavar et al. (2007) with initial denaturation of 94°C for 2 minutes followed by 40 cycles of denaturation at 94°C for 30 seconds, annealing at 50°C for 45 seconds, elongation at 72°C for 2 minutes and final elongation of 7 minutes at 72°C was

found to be the best for ISSR profiling of Nepalese *Citrus aurantifolia* landraces and hence selected for subsequent ISSR profiling experiments (Plate 4.1).

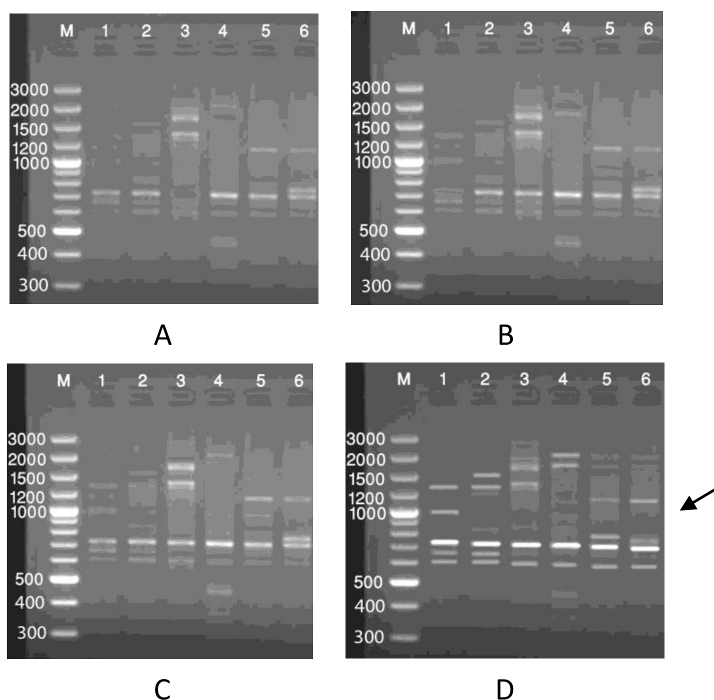


Plate 4.1 Gel Picture of ISSR-PCR for the selection of best cycling condition using primer C4 and the genomic DNA of samples LT-1, LT-2, LT-3,LT- 4, LT-5 and LT-6 in lanes 1-6 respectively (Table 3.1). Lane 'M' are 100 bp plus molecular weight marker; A = Cycling condition as of Fang et al. (1998); B = Cycling condition as of Lui et al. (2011); C = Cycling condition as of Kumar et al. (2010) and D = Cycling condition as of Shahsavar et al. (2007). An arrow indicated the selected one.

Table 4.2 Tested different ISSR PCR-cycling conditions with selected one in bold font.

Fang et al. (1998)			**Shahsavar et al. (2007)**		
Initial Denaturation	94°C for 4 min		**Initial Denaturation**	**94°C for 5 min**	
Denaturation	94°C for 30 s	27 cycles	**Denaturation**	**94°C for 30 s**	**40 cycles**
Annealing	52°C for 45 s		**Annealing**	**50°C for 45 s**	
Extention (Elongation)	72°C for 2 min		**Extention (Elongation)**	**72°C for 2 min**	
Final Extension	72°C for 7 min		**Final Extension**	**72°C for 7 min**	
Hold	4°C for 4 min		**Hold**	**4°C for 5 min**	

Kumar et al. (2010)			**Lui et al. (2011)**		
Initial Denaturation	94°C for 4 min		Initial Denaturation	94°C for 5 min	
Denaturation	94°C for 1 min		Denaturation	94°C for 30 s	

Annealing	52^oC for 1 min	35 cycles	Annealing	52^oC for 45 s	35 cycles
Extention (Elongation)	72^oC for 2 min		Extention (Elongation)	72^oC for 2 min	
Final Extension	72^oC for 7 min		Final Extension	72^oC for 5 min	
Hold	4^oC for 4 min		Hold	4^oC for 4 min	

4.3.2. ISSR-PCR reaction conditions optimization

The ISSR-PCR reaction conditions optimization were done using sample LM-56 acid lime (*C. aurantifolia*) and primer UBC 810.

Table 4.3 Tested different ISSR-PCR parameter and selection of optimized parameters.

S.N	PCR parameters	Range tested	Selected optimized condition	Remarks
1	DNA concentration	12.5, 25, 37.5, 50, 62.5, 75, 87.5, 100	25 ng	Most of the concentration produced the reproducible bands but 25 ng of template DNA was selected as optimum condition since it is minimum concentration that produces discernible bands (Plate 4.2).
2	$MgCl_2$ concentration	1.5, 2.0, 2.5, 3.0, 3.5, 4.0	3.0 mM	Crispy bands observed at 1.5 – 4.0 and 3.0mM $MgCl_2$ was selected as optimum concentration for subsequent experiments as it revealed more number of bands with clear and crispy nature (Plate 4.2).
3	Primer concentration	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	0.4µM	The banding pattern suggested primer concentration 0.2µM – 1.6µM to be good. But 0.4µM concentration was selected for subsequent ISSR-PCR profiling experiments as it produce clear and crispy bands (Plate 4.3)
4	dNTPs concentration	0.1, 0.2, 0.3, 0.4, 0.5	0.4mM	Observed ISSR patterns revealed 0.4mM produces crispy bands so 0.4mM concentration selected as optimum (Plate 4.4).
5	*Taq* polymerase	0.5, 1.0, 1.5, 2.0,	1.5U	Though all the concentration

	concentration	2.5		produces ISSR crispy bands, the best banding pattern was considered for 1.5U *Taq* polymerase (Plate 4.4).

The final optimized ISSR-PCR reaction condition for *Citrus aurantifolia* (Acid lime) included 25ng genomic DNA, 3.0mM $MgCl_2$, 2.5µl (10mM) of 10X PCR reaction buffer (Fermentas Co., 100mM Tris-HCl/p^H 8.8 at 25°C, 500mM KCL, 0.8% Nonidet P_{40}), 0.4µM Primer, 0.4mM dNTPs and 1.5U *Taq* polymerase in a 25µl reaction volume.

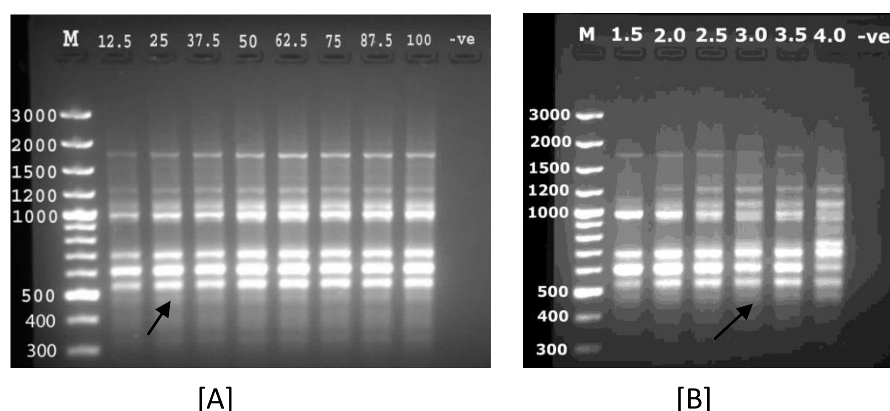


Plate 4.2 Gel pictures of ISSR-PCR for the selection of DNA and $MgCl_2$ concentrration using primer UBC 809 and the genomic DNA of sample LS-56. Lanes marked 'M' are 100 bp plus molecular weight marker; (A) Each lane is marked with the respective concentrations of template DNA (12.5 to 100 ng) (B) Each lane is marked with the respective concentrations of $MgCl_2$ (1.5mM to 4.0mM) used during ISSR-PCR. An arrow indicated the selected one.

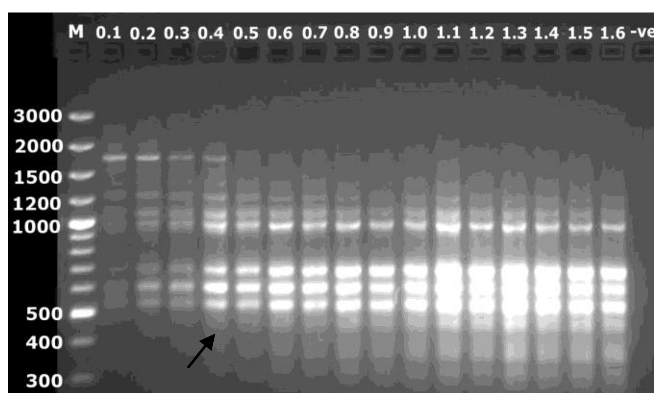


Plate 4.3 Gel pictures of ISSR-PCR for the selection of primer concentration using the genomic DNA sample LS-56 with varying concentration of primer UBC 809. Lane labeled 'M' is 100 bp plus molecular weight marker. Each lane is marked with the respective concentrations of primer UBC 809 (0.1µM to 1.6µM) used during ISSR-PCR. An arrow indicated the selected one.

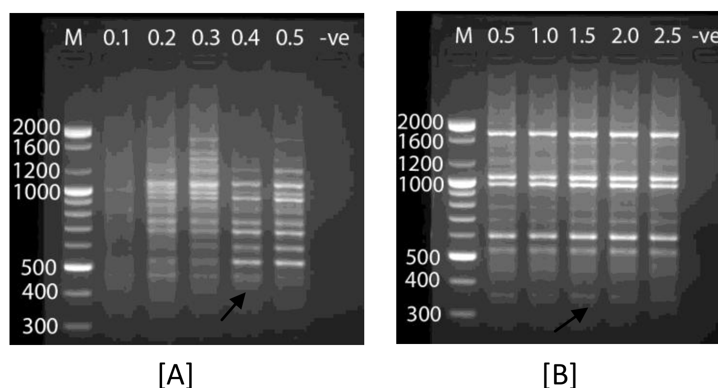
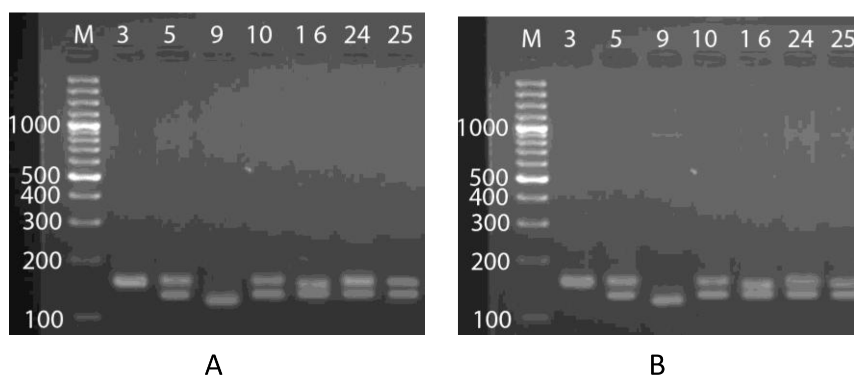


Plate 4.4 Gel pictures of ISSR-PCR for the selection of dNTPs and *Taq* polymerase concentration using primer UBC 809 genomic DNA of sample LS-56. Lanes marked 'M' are ladder; (A) Each lane is marked with the respective concentrations of dNTPs (0.1mM to 0.5mM) (B) Each lane is marked with the respective concentrations of *Taq* polymerase (0.5U to 2.5U) used during ISSR-PCR. An arrow indicated the selected one.

4.4. SSR-PCR Optimization

4.4.1. SSR-PCR Cycling conditions optimization

Among the four cycling conditions assessed (Kijas et al., 1995; Barkley et al., 2006; Hvarleva et al., 2008; Snoussi et al., 2012) (Table 4.4), PCR program described by Snoussi et al. (2012) with 94°C for 5 min followed by 35 cycles of denaturation of 94°C for 30 second, annealing at primer specific temperature (47-61°C) for 1 minute and elongation at 72°C for 30 seconds and final elongation at 72°C for 4 minutes was found to be the best which gave clear, scorable and crispy bands (plate 4.5) and therefore selected for SSR-PCR profilings.



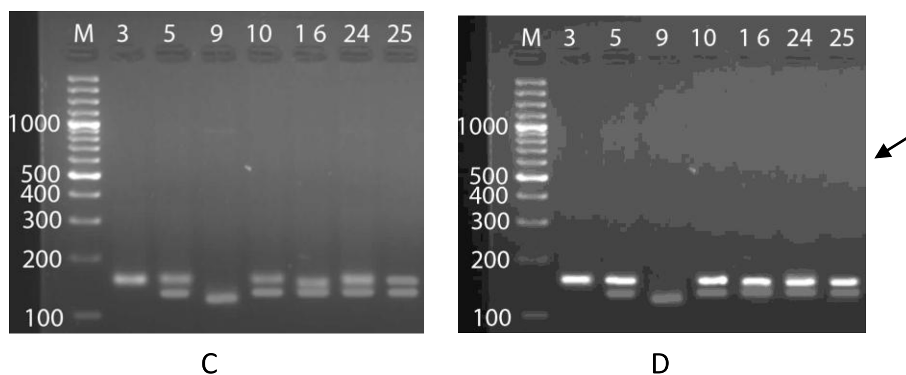


Plate 4.5 Gel Picture of SSR-PCR for the selection of best cycling condition using primer TC 26 and the genomic DNA of sample 3, 5, 9, 10, 16, 24 and 25 (Table 4.1). Lane 'M' are 100 bp plus molecular weight marker; A = Cycling condition as of Barkley et al. (2006); B = Cycling condition as of Hvarleva et al., 2008; C = Cycling condition as of Kijas et al. (1995) and D = Cycling condition as of Snoussi et al. (2012). An arrow indicated the selected one.

Table 4.4 Tested different SSR PCR-cycling conditions with selected one in bold font.

Kijas et al. (1995)			**Barkley et al. (2006)**		
Initial Denaturation	94°C for 5 min		Initial Denaturation	95°C for 5 min	
Denaturation	94°C for 1 min	32 cycles	Denaturation	95°C for 30 s	38 cycles
Annealing	Primer specific temp for 30 s		Annealing	Primer specific temp for 30 s	
Extention (Elongation)	72°C for 1 min		Extention (Elongation)	72°C for 1 min	
Final Extension	72°C for 4 min		Final Extension	72°C for 7 min	
Hold	4°C for 4 min		Hold	4°C for 5 min	

Hvarleva et al. (2008)			**Snoussi et al. (2012)**		
Initial Denaturation	94°C for 5 min		**Initial Denaturation**	**94°C for 5 min**	
Denaturation	94°C for 1 min	35 cycles	**Denaturation**	**94°C for 30 s**	**35 cycles**
Annealing	Primer specific temp for 1 min		**Annealing**	**Primer specific temp for 1 min**	
Extention (Elongation)	72°C for 2 min		**Extention (Elongation)**	**72°C for 30 s**	
Final Extension	72°C for 30 min		**Final Extension**	**72°C for 4 min**	
Hold	4°C for 4 min		**Hold**	**4°C for 5 min**	

4.4.2. SSR-PCR reaction conditions optimization

The SSR-PCR reaction conditions optimization were done using sample 10 (Sweet orange, Washington navel) and primer TAA 27.

Table 4.5 Testing of different SSR-PCR parameter and selection of optimized parameters.

S.N	PCR parameters	Range tested	Selected optimized condition	Remarks
1	DNA concentration	12.5, 25, 37.5, 50, 62.5, 75, 87.5, 100	25 ng	Most of the concentration produced the clear bands but 25 ng of template DNA was selected as optimum condition since it is minimum concentration that produces discernible and clear heterozygous bands (Plate 4.6).
2	$MgCl_2$ concentration	1.5, 2.0, 2.5, 3.0, 3.5, 4.0	3.0 mM	Crispy Bands observed at 1.5 – 4.0 and 3.0mM $MgCl_2$ was taken as optimum concentration for subsequent experiments as it reveal clear and crispy heterozygous bands (Plate 4.6).
3	Primer concentration	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	0.4µM	The banding pattern suggested primer concentration 0.2µM – 1.6µM to be good. But 0.4µM concentration was selected for subsequent ISSR-PCR profiling experiments as it gave clear heterozygous bands (Plate 4.7).
4	dNTPs concentration	0.1, 0.2, 0.3, 0.4, 0.5	0.2mM	Observed SSR patterns revealed 0.2mM produces crispy and intense heterozygous bands so 0.2mM concentration selected as optimum (Plate 4.8).
5	*Taq* polymerase concentration	0.5, 1.0, 1.5, 2.0, 2.5	2U	Though all the concentration produces heterozygous bands but the best banding pattern was considered for 2U *Taq* polymerase as at this concentration heterozygous bands are intense and clears (Plate 4.8).

The final optimized SSR-PCR reaction condition for *Citrus* spp. included 25ng genomic DNA, 3.0mM $MgCl_2$, 2.5µl (10mM) of 10X PCR reaction buffer (Fermentas, 100mM Tris-HCl/p^H 8.8 at 25°C, 500mM KCL, 0.8% Nonidet P_{40}), 0.4µM Primer, 0.2mM dNTPs and 2U *Taq* polymerase in a 25µl reaction volume.

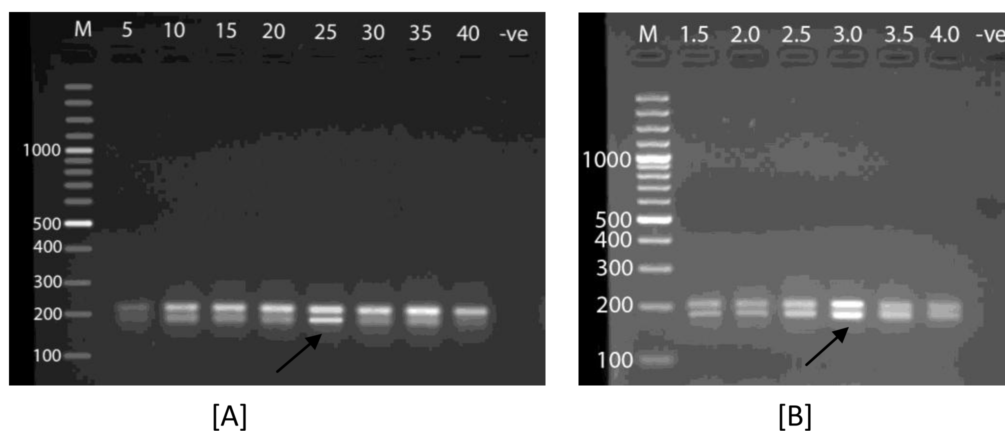


Plate 4.6 Gel pictures of SSR-PCR for the selection of DNA and $MgCl_2$ concentration using primer TAA 27 and genomic DNA of sample 10. Lanes marked 'M' are 100 bp plus molecular weight marker; (A) Each lane is marked with the respective concentrations of template DNA (5 to 40 ng) (B) Each lane is marked with the respective concentrations of $MgCl_2$ (1.5mM to 4.0mM) used during SSR-PCR. An arrow indicated the selected one.

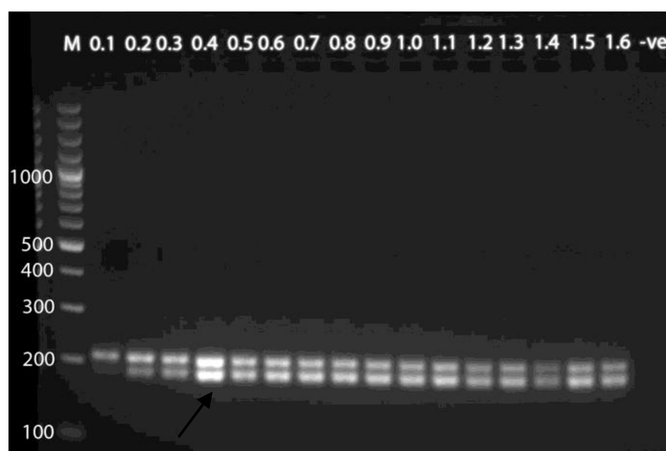


Plate 4.7 Gel pictures of SSR-PCR for the selection of primer concentration using the genomic DNA sample 10 with varying concentration of primer TAA 27. Lane labeled 'M' is 100 bp plus molecular weight marker. Each lane is marked with the respective concentrations of primer UBC 809 (0.1µM to 1.6µM) used during SSR-PCR. An arrow indicated the selected one.

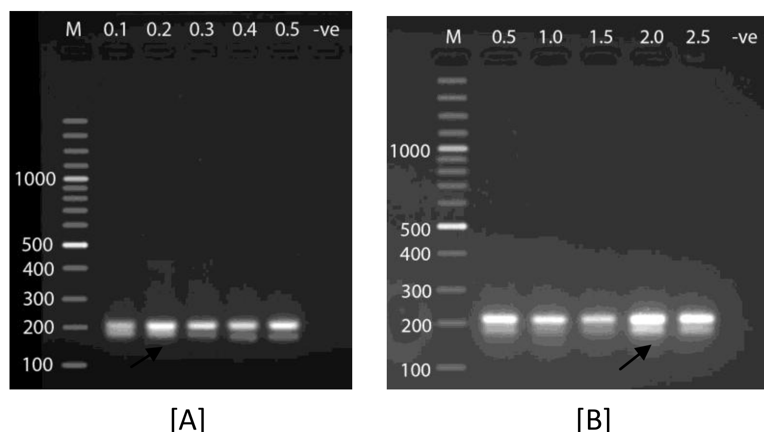
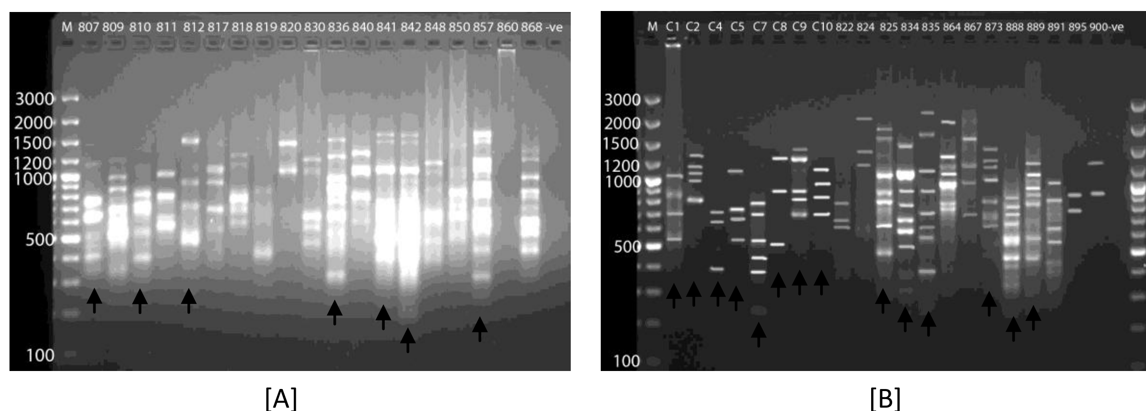


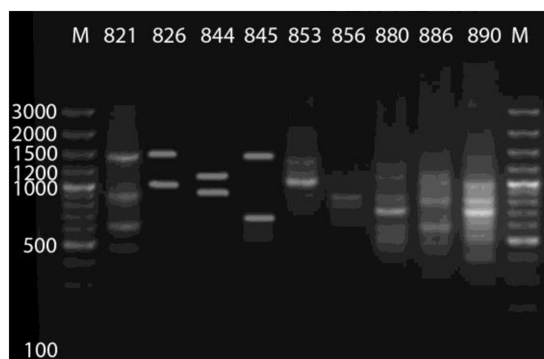
Plate 4.8 Gel pictures of SSR-PCR for the selection of dNTPs and *Taq* polymerase concentration using primer TAA 27 and genomic DNA of sample 10. Lanes marked 'M' are ladder; (A) Each lane is marked with the respective concentrations of dNTPs (0.1mM to 0.5mM) (B) Each lane is marked with the respective concentrations of *Taq* polymerase (0.5U to 2.5U) used during SSR-PCR. An arrow indicated the selected one.

4.5. Primer Screening for ISSR Profiling

The optimized ISSR-PCR reactions and cycling conditions was used for screening of 49 different ISSR primers (Table 3.1) using the fresh genomic DNA of acid lime (*C. aurantifolia*) sample LS-15. Out of 49 primers, 21 primers (Table 4.6 and Plate 4.9) which gave multiple, scorable, crispy and reproducible bands were selected finally to be used in ISSR-PCR analysis involving 60 accessions of *Citrus aurantifolia* (acid lime) under study. The experiments were repeated twice for the conformation of reproducibility of ISSR amplifications.



[A] [B]

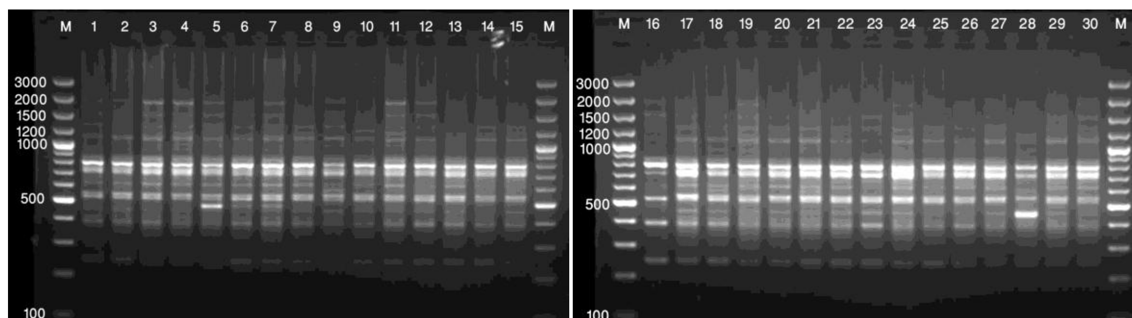


[C]

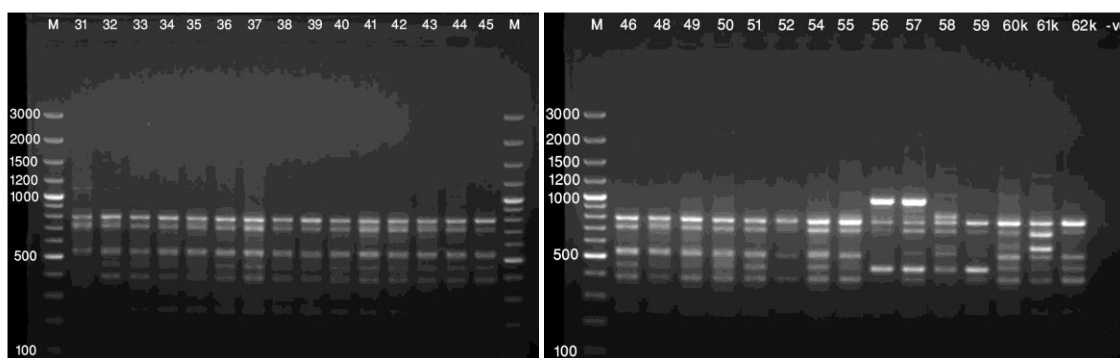
Plate 4.9 ISSR gel pictures for the Primer Screening using template DNA of Sample LS-15. Each lane in [A] [B] and [C] corresponds to different ISSR primers used in the experiment. Lanes marked 'M' are 100 bp plus DNA ladder. An arrow indicated the selected one.

4.6. ISSR Profiling and Genetic Diversity Assessment of acid lime

The ISSR-PCR profiles generated by the selected 21 UBC primers were used for genetic diversity assessment of *Citrus aurantifolia* (Acid lime) landraces under study. Representative ISSR profiles generated by primers C 7, C 8 and UBC 842 are shown (Plates 4.10 – 4.12).



[**A**] [**B**]



[**C**] [**D**]

Plate 4.10 ISSR profiles amplified with **Primer C 7.** Lanes marked with M are 100bp plus molecular weight markers. [A] Lanes 1-15 represents acid lime samples 1-15; [B] Lanes 16-30 represents acid lime samples

16-30; [C] Lanes 31-45 represents acid lime samples 31-45; [D] Lanes 46-62k represents acid lime samples 46-62k.

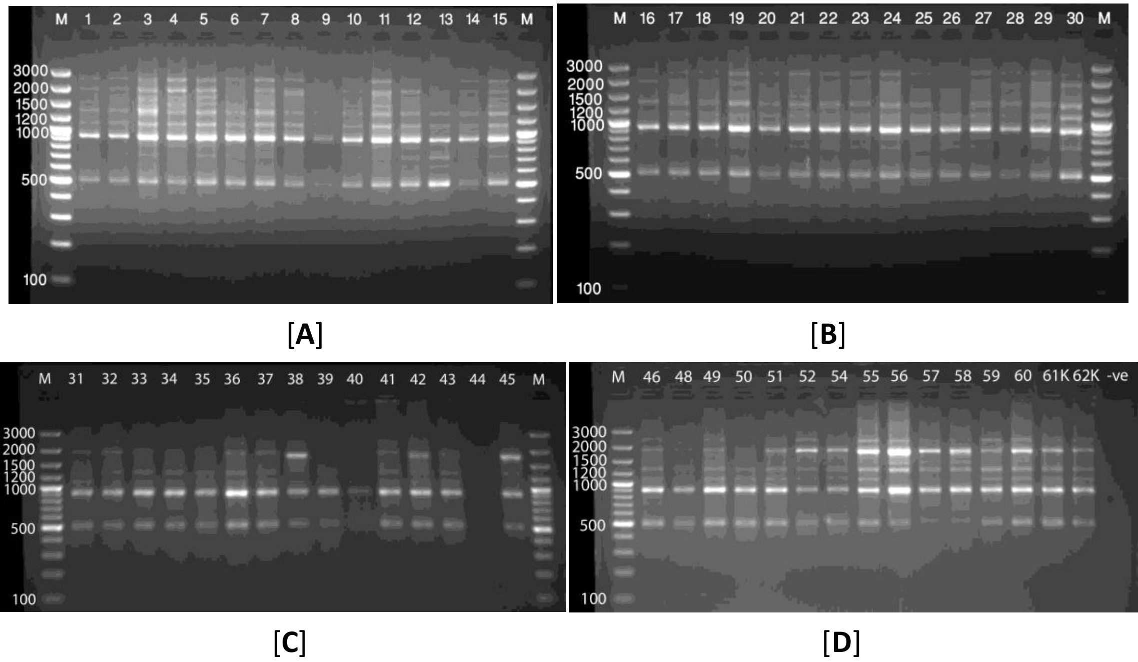


Plate 4.11 ISSR profiles amplified with **Primer C 8.** Lanes marked with M are 100bp plus molecular weight markers. [A] Lanes 1-15 represents acid lime samples 1-15; [B] Lanes 16-30 represents acid lime samples 16-30; [C] Lanes 31-45 represents acid lime samples 31-45; [D] Lanes 46-62k represents acid lime samples 46-62k.

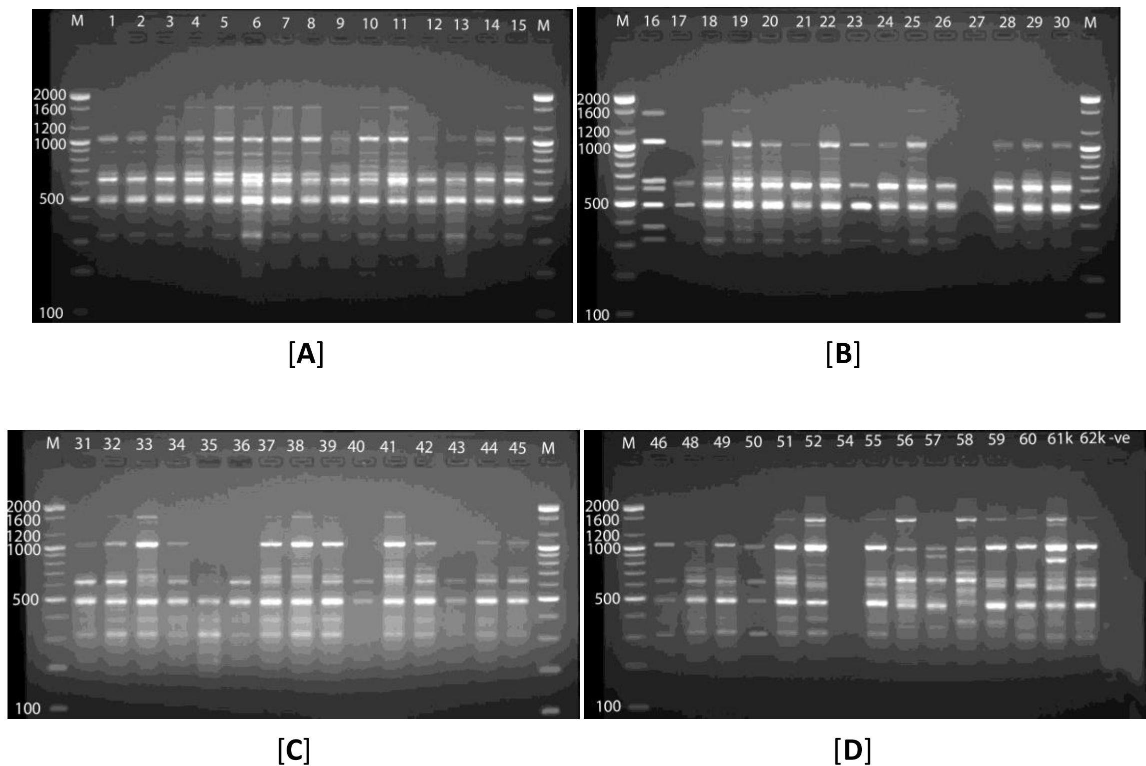


Plate 4.12 ISSR profiles amplified with **Primer UBC 842.** Lanes marked with M are 100bp plus molecular weight markers. [A] Lanes 1-15 represents acid lime samples 1-15; [B] Lanes 16-30 represents acid lime samples 16-30; [C] Lanes 31-45 represents acid lime samples 31-45; [D] Lanes 46-62k represents acid lime samples 46-62k.

A total of 234 loci were amplified across the 60 acid lime accessions with an average amplification of 9.72 bands per primer. Among these total amplified bands, 204 (87.18%) bands were polymorphic while 30 (12.82%) were found to be monomorphic. Polymorphic bands produced by different primers ranged from 55.56% to 100%. Among 21 primers, 8 primers revealed 100% polymorphisms. The number of scorable bands produced per primer ranged from 7 to 18 with the variation in amplicon size ranging from 250 bp to 3200 bp. The highest number of ISSR loci (18) was produced by primer UBC 857 where as the lowest number of ISSR loci (7) was produced by primer C1 and UBC 834 in the total accessions tested (Table 4.6).

Table 4.6 Primer sequences, total number of amplified band (TNB), number of polymorphic bands (NPB), percent polymorphism and amplicon size range of the 21 primers used to generate ISSR-PCR profiles in 60 *Citrus aurantifolia* (Acid lime) accessions.

SN	Primer Code	Primer Sequence (5' - 3')	Primer Length	TNB	NPB	Polymorphisms (%)	Amplicon size range (bp)
1	C1	TCTCTCTCTCTCTCTCTCCC	20 mer	**7**	5	71.43	550-1800
2	C2	AGCAGCAGCAGCGT	14 mer	10	10	100	500-3200
3	C4	CTCCTCCTCGC	11 mer	14	14	100	300-2200
4	C5	CACCACCACGC	11 mer	12	11	91.67	600-2500
5	C7	HVHGAGAGAGAGAGAGAT	18 mer	17	15	88.23	250-2000
6	C8	TCCTCCTCCTCCTCCRY	17 mer	9	7	77.78	520-2800
7	C9	BDBTCCTCCTCCTCCTCC	18 mer	9	6	66.67	520-2000
8	C10	HVHTCCTCCTCCTCCTCC	18 mer	11	8	72.73	500-2000
9	UBC 807	AGAGAGAGAGAGAGAGT	17 mer	9	9	100	450-1300
10	UBC 810	GAGAGAGAGAGAGAGAT	17 mer	16	16	100	390-1980
11	UBC 812	GAGAGAGAGAGAGAGAA	17 mer	12	9	75	450-1500
12	UBC 825	ACACACACACACACACT	17 mer	9	5	**55.56**	500-2000
13	UBC 834	AGAGAGAGAGAGAGAGYT	18 mer	**7**	5	71.43	310-1550
14	UBC 835	AGAGAGAGAGAGAGAGYC	18 mer	10	10	100	400-2900
15	UBC 836	AGAGAGAGAGAGAGAGYA	18 mer	10	10	100	320-1450
16	UBC 841	GAGAGAGAGAGAGAGAYC	18 mer	10	10	100	320-1700
17	UBC 842	GAGAGAGAGAGAGAGAYG	18 mer	12	12	100	320-1650
18	UBC 857	ACACACACACACACACYG	18 mer	**18**	17	94.44	300-3000
19	UBC 873	GACAGACAGACAGACA	16 mer	14	13	92.86	470-3000
20	UBC 888	BDBCACACACACACAGACA	17 mer	9	7	77.78	480-1450

21	UBC 889	DBDACACACACACACACA	18 mer	9	5	55.56	480-1400
	Total			234	204		
Average Polymorphic Bands per primer = 9.72				**Average Polymorphism = 87.18**			

Where, H = non-G, Y = Pyrimidine, B = non-A, D = non-C, V = non-T

4.6.1. Polymorphic Information Content, Band Informativeness and Resolving Power

The Polymorphic Information Content (PIC) score ranged from 0.74 (UBC 807) to 0.93 (UBC 857) with an average of 0.85. The Band informativeness of the 21 ISSR primers ranged from 0.42 (C 4) to 1.77 (C 1) with an average of 1.12. The Resolving Power (R_P) of ISSR primers ranged from 4.63 (UBC 807) to 23.16 (UBC 857) with an average of 12.03. The primers that have highest PIC value also gave highest R_P score (Table 4.7).

Table 4.7 Assessment of Polymorphic Information Content (PIC), Band informativeness (I_B) and Resolving power (R_P) of 21 ISSR primers used to generate ISSR profiles in acid lime accessions.

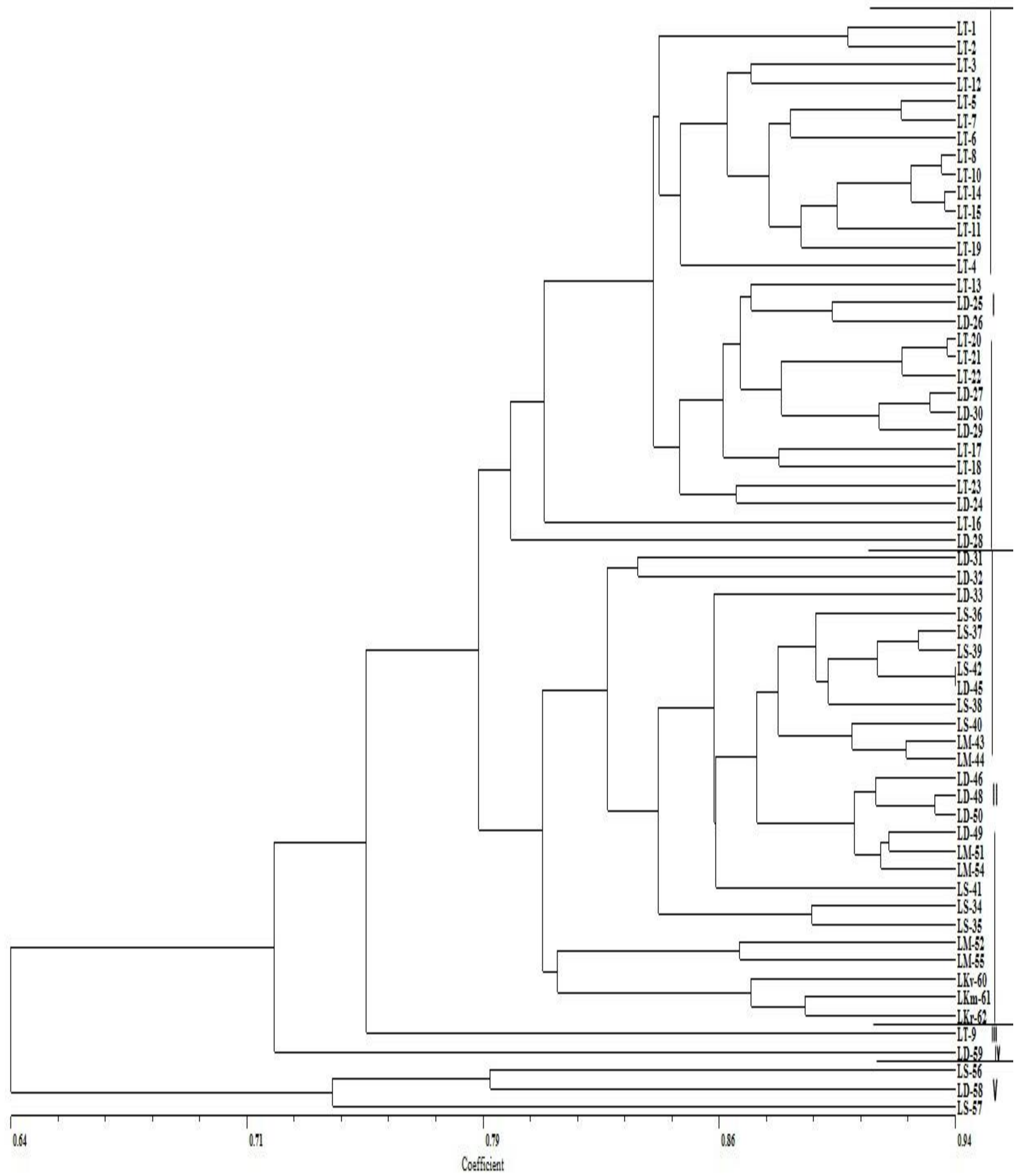
SN	Primer Code	Primer Sequence (5' - 3')	PIC	I_B range	I_B average	R_P
1	C1	TCTCTCTCTCTCTCTCTCCC	0.85	1.4-2.0	**1.77**	12.4
2	C2	AGCAGCAGCAGCGT	0.83	0.06-1.8	0.83	8.33
3	C4	CTCCTCCTCGC	0.82	0.06-1.66	**0.42**	5.87
4	C5	CACCACCACGC	0.83	0.03-1.86	0.60	7.23
5	C7	HVHGAGAGAGAGAGAGAT	0.89	0.03-2.0	0.87	14.87
6	C8	TCCTCCTCCTCCTCCRY	0.87	0.56-1.96	1.47	13.23
7	C9	BDBTCCTCCTCCTCCTCC	0.86	0.1-1.96	1.50	13.56
8	C10	HVHTCCTCCTCCTCCTCC	0.79	0.03-2.0	0.78	8.63
9	UBC 807	AGAGAGAGAGAGAGAGT	**0.74**	0.03-1.86	0.51	**4.63**
10	UBC 810	GAGAGAGAGAGAGAGAT	0.85	0.03-2.0	1.0	12
11	UBC 812	GAGAGAGAGAGAGAGAA	0.91	0.03-1.83	1.03	16.6
12	UBC 825	ACACACACACACACACT	0.87	0.13-1.93	1.51	13.63
13	UBC 834	AGAGAGAGAGAGAGAGYT	0.81	0.06-1.96	1.28	8.96
14	UBC 835	AGAGAGAGAGAGAGAGYC	0.86	0.06-1.93	1.29	12.96
15	UBC 836	AGAGAGAGAGAGAGAGYA	0.87	0.16-1.86	1.13	11.3
16	UBC 841	GAGAGAGAGAGAGAGAYC	0.87	0.16-1.83	1.08	10.83
17	UBC 842	GAGAGAGAGAGAGAGAYG	0.89	0.06-1.86	1.04	12.53
18	UBC 857	ACACACACACACACACYG	**0.93**	0.03-1.96	1.28	**23.16**
19	UBC 873	GACAGACAGACAGACA	0.89	0.16-1.93	0.96	13.43
20	UBC 888	BDBCACACACACAGACA	0.87	0.03-1.96	1.47	13.23
21	UBC 889	DBDACACACACACACACA	0.87	0.1-2.0	1.70	15.33
		Average	**0.85**		**1.12**	**12.03**

4.6.2. Analysis of genetic diversity using the similarity coefficient and cluster analysis

4.6.2.1. Comparison of similarity coefficients and phenograms of ISSR-PCR profiling

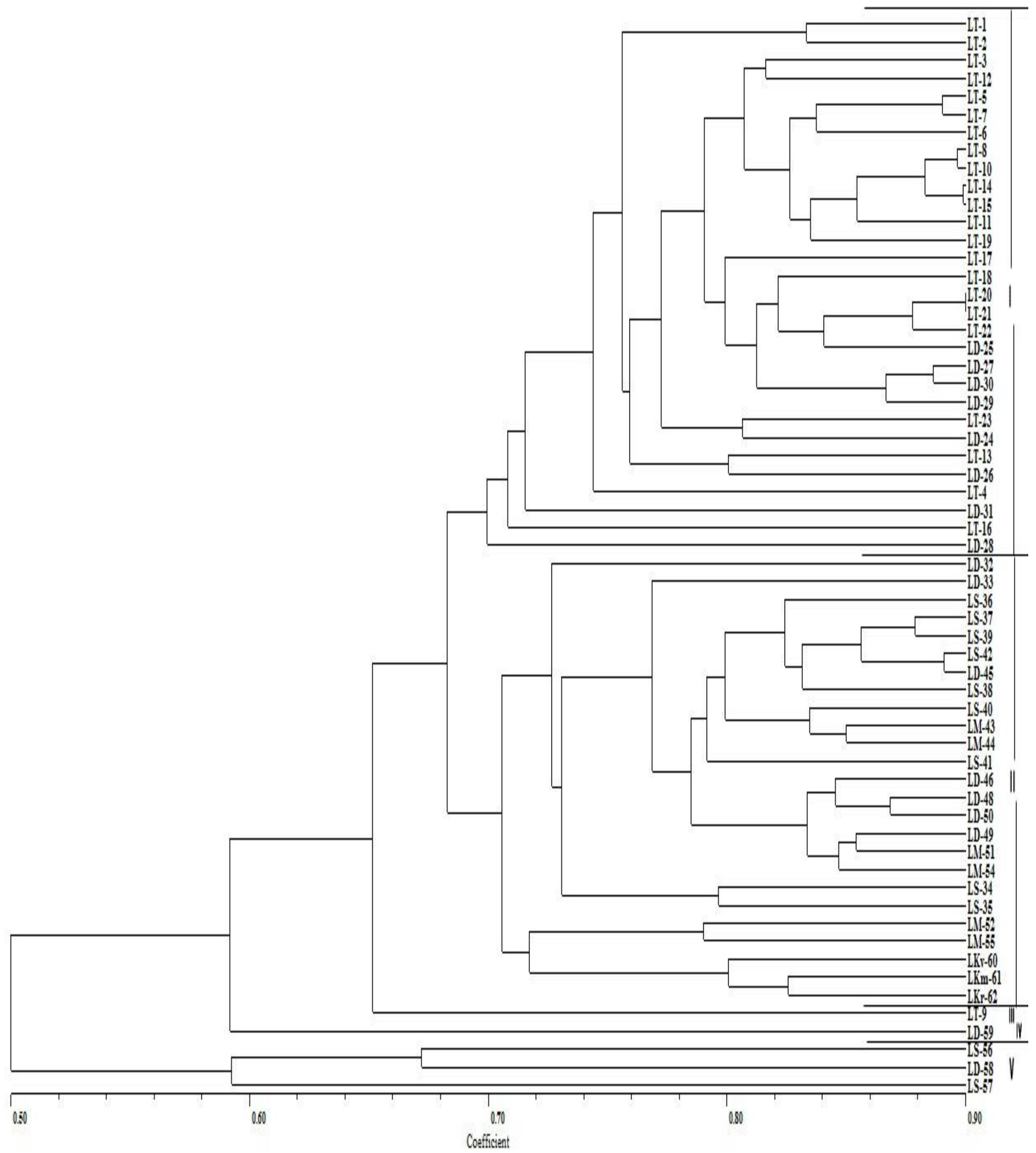
Different similarity matrices were generated using binary data based on presence and absence of ISSR loci from SimQual computational algorithm of NTSYS-PC ver. 2.21i. The varied range of similarity coefficients were obtained, using Simple Matching (SM), Jaccard's (J) and Dice (D) coefficient i.e. SM (0.54 - 0.94), J (0.42 - 0.90) and Dice (0.57 - 0.95) with an average similarity coefficient of 0.79, 0.69 and 0.81 respectively (Appendix 7, 8 and 9).

In order to represent the genetic relationship among all 60 accessions of *Citrus aurantifolia* (Acid lime) landraces belonging to different agro-ecological zone of Eastern Nepal, all the three similarity coefficients were employed for the construction of phenograms. Phenograms topologies generated from three coefficients were comparable except for few variations in range of polymorphism and slight alterations in position of few accessions (cluster I consist of 29 accessions and cluster II consist of 26 accessions) in the phenogram generated from SM coefficient (*viz.* fig 4.1, 4.2 and 4.3).



Coefficient of Similarity (Simple Matching)

Figure 4.1 Phenogram generated for 60 *Citrus aurantifolia* (Acid lime) landraces (refer to table 3.1 for sample details) by UPGMA cluster analysis using Simple Matching coefficient of similarity calculated from 234 ISSR loci generated by 21 primers. The clusters are labeled as I, II, III, IV and V.



Coefficient of Similarity (Jaccard's)

Figure 4.2 Phenogram generated for 60 *Citrus aurantifolia* (Acid lime) landraces (refer to table 3.1 for sample details) by UPGMA cluster analysis using Jaccard's coefficient of similarity calculated from 234 ISSR loci generated by 21 primers. The clusters are labeled as I, II, III, IV and V.

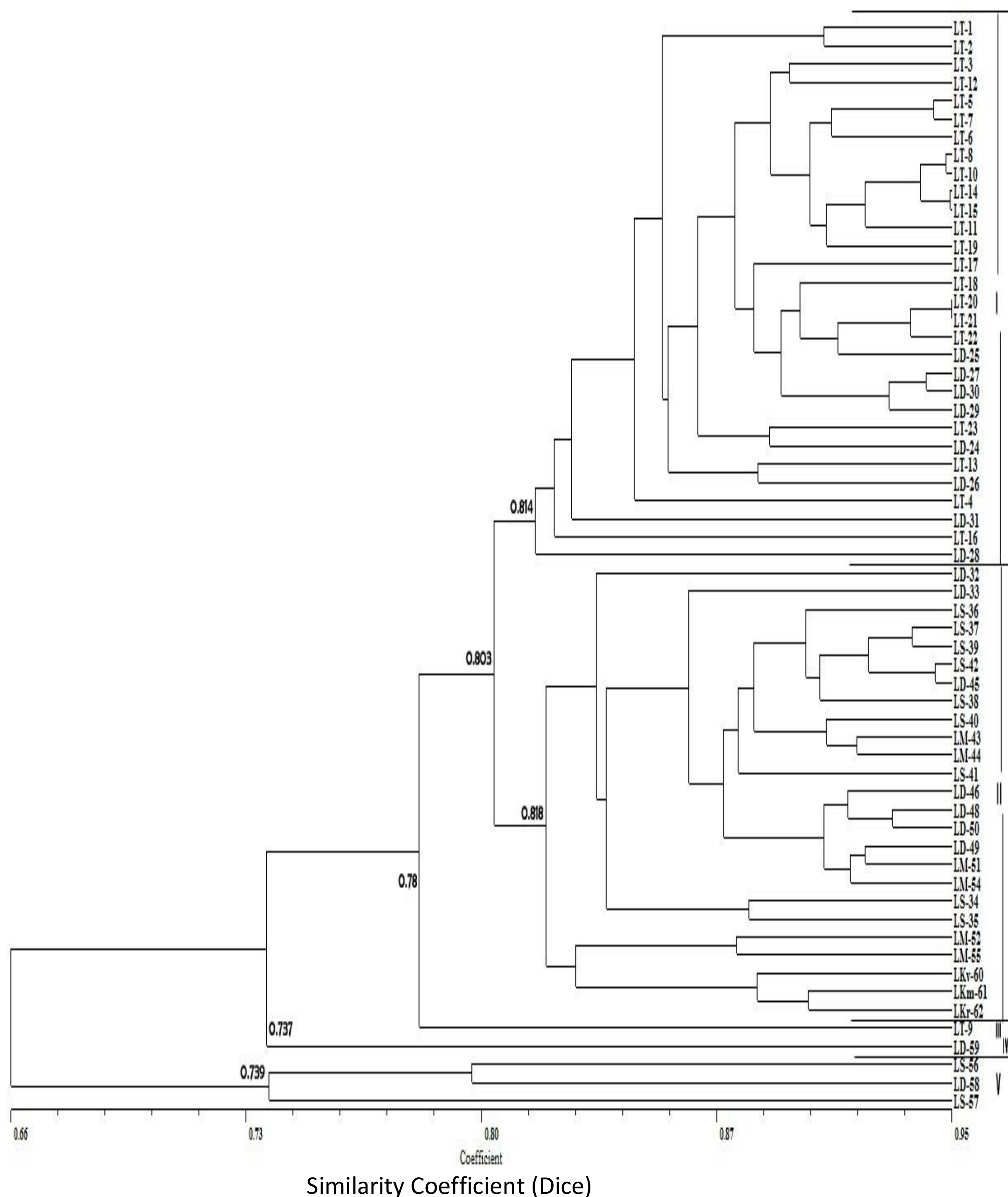


Figure 4.3 Phenogram generated for 60 *Citrus aurantifolia* (Acid lime) landraces (refer to table 3.1 for sample details) by UPGMA cluster analysis using Dice coefficient of similarity calculated from 234 ISSR loci generated by 21 primers. The clusters are labeled as I, II, III, IV and V.

The correlation values were computed from the comparison of original matrices by applying Mantel test (Khan, 2007). The result from the Mantel test (Matrix comparison) of original matrices showed that the correlation value between Jaccard's and Dice was the highest and significant (0.99710) in comparison to SM and J (0.98143) and SM and

Dice (0.98318) (Table 4.8). This suggests those Dice or Jaccard's coefficients are more suitable than Simple Matching coefficient.

Table 4.8 Correlation Coefficient values from Mantel test of original similarity matrices.

	Simple Matching (SM)	Jaccard (J)	Dice (D)
Simple Matching (SM)	******	0.98143	0.98318
Jaccard (J)		******	0.99710
Dice (D)			******

In order to evaluate trees constructed from UPGMA clustering by genetic similarity coefficients, consensus indices (CI) were calculated using Strict consensus method for the each combination of coefficient and UPGMA clustering. The highest Consensus fork index (CI_C = 1.00000) was observed for Jaccard and Dice coefficient. The consensus fork indices for Simple Matching coefficient and Dice coefficient was found to be same as that of the Simple Matching coefficient and Jaccard coefficient (CI_C = 0.74138) (Table 4.9). This result supports that phenograms constructed from Dice or Jaccard's coefficient will be more suitable than that constructed from Simple Matching coefficient. The overall topologies of three phenograms are more or less similar except for the SM tree in which cluster I consist of 29 accessions and cluster II consist of 26 accessions (Fig 4.1-4.3).

Table 4.9 Consensus fork index among the UPGMA based phenograms using similarity coefficients among *Citrus aurantifolia* (Acid lime) landraces by ISSR markers.

	Simple Matching (SM)	Jaccard (J)	Dice (D)
Simple Matching (SM)	******	0.74138	0.74138
Jaccard (J)		******	1.00000
Dice (D)			******

The clustering based on Unweighted Pair Group Method of Arithematic Averages (UPGMA) for Dice coefficient observed to give a highest cophenetic correlation value of 0.90356 and lowest cophenetic correlation value of 0.88396 was observed for UPGMA clustering using Simple Matching coefficient where as intermediate value 0.89800 was observed for UPGMA clustering using Jaccard's coefficient (Table 4.10).

On comparing the correlation value with the standard chart of goodness of fit, Simple Matching coefficient with UPGMA clustering method had a lowest Correlation Coefficient value (r = 0.88396), which indicate good fit for cluster analysis. On the other hand for Dice coefficient, the value was 0.90356, which indicate very good fit. Therefore,

Simple Matching coefficient of similarity with UPGMA clustering method is considered as least suitable where as Dice coefficient of similarity with UPGMA clustering method is considered best for deducing the genetic relationship among different *Citrus aurantifolia* (Acid lime) landraces.

Table 4.10 Correlation coefficients value (r) obtained from cophenetic values of similarity matrices (Simple Matching, Jaccard's and Dice coefficient) and clusters computed by UPGMA module using MXCOMP module of NTSYS.

Clustering modules of similarity	Simple Matching	Jaccard	Dice
UPGMA	0.88396	0.89800	0.90356

In the present study, Dice similarity with UPGMA clustering yielded highest correlation coefficient value of 0.90356 followed by J and SM similarity respectively. From this results, the three coefficient can be judged in the decreasing order as D > J > SM to be used for interpretation of genetic relationship in *Citrus aurantifolia* (Acid lime) landraces.
From all the statistical comparison made above for the similarity coefficients, Dice coefficient was found to be best and chosen finally, for the interpretation of the results on genetic diversity and the relationships among various accessions representing different agro-ecological zone. Now, on the basis of Dice similarity coefficient (Appendix 9), genetic similarity percent (%) was estimated within and between agro-ecological zones (Table 4.11). The individual genetic similarity among various *C. aurantifolia* accessions has been assessed from the pair wise comparison of Dice pair similarity matrix. From the pair wise comparison, acid lime accessions from High-hill (LT-21 and LT-20) were observed to be most similar (0.95) and accessions from Terai (LS-56 and LS-35) revealed to be genetically least similar (0.571). Every individual accession could be compared individually with any of 60 acid lime accessions to check their genetic relatedness and distances (Table 4.11).

Table 4.11 Percentage similarity observed for different acid lime landraces agro-ecological zones based on ISSR similarity matrix (Dice) within and between agro-ecological zones.

Zone/zone	High-hill	Average	Mid-hill	Average	Terai	Average
High-hill	75-95	86.02	68-95	83.6	57-95	82.68
Mid-hill			70-94	82.94	57-94	79.98
Terai					57-94	79.33

Zone = Agro-ecological zone

Phenogram was constructed on the basis of Dice similarity coefficient using SAHN module of NTSYS-PC ver. 2.21i. Phenogram was obtained by distance matrix strategy using UPGMA algorithm. The UPGMA clustering revealed the genetic diversity and relationship among acid lime accessions from geographically diverse agro-ecological zones. Based on the phenogram and Dice similarity coefficient level, genetic relationships (or similarity) within 60 acid lime accessions ranged from 57% to 95% with an average of 81% was observed. The 60 accessions were separated into 2 major (cluster I and II) and 3 minor clusters (cluster III, IV and V). The accessions from High-hill, mid-hill and Terai were clustered with each other. The cluster I clustered majority of accessions i.e. 30 accessions from High-hill and Mid-hill agro-ecological zone. The highest similarity coefficient was observed for the accession LT-14 and LT-15 (0.945) and LT-20 and LT-21 (0.95) i.e. they are in same genetic level, followed by LT-8 and LT-10 (0.943), LT-5 and LT-7 (0.940), LD-27 and LD-30 (0.938), LT-1 and LT-2 (0.907), LT-3 and LT-12 (0.897), LT-23 and LD-24 (0.891) and LT-13 and LD-26 (0.887) under cluster I and the lowest similarity coefficient (i.e. highest genetic distance) was observed between LS-56 and LS-35 (0.571). Cluster II clustered 25 accessions from High-hill, Mid-hill and Terai agro-ecological zone with the exotic varieties of Vanarasi, Madrasi and Rampur (LKv-60, LKm-61 and LKr-62 respectively). Accession LS-42 and LD-45 has similarity coefficient of 0.940 followed by 0.993 for LS-37 and LS-39, 0.927 for LD-48 and LD-50, 0.919 for LD-49 and LM-51, 0.917 for LM-43 and LM-44, 0.902 for LKm-61 and LKr-62, 0.884 for LS-34 and LS-35 and 0.881 for LM-52 and LM-55 in this cluster. Cluster III and cluster IV consist of single accession, LT-9 and LD-59 from High-hill and Mid-hill respectively. Cluster V clustered three accessions (LS-56, LD-58 and LS-57) from Terai agro-ecological zone. The accession LS-56 and LD-58 has 0.80 similarity coefficient. The accessions from High-hill and Mid-hill were observed to be closely related with a high average genetic similarity (86 % and 83 % respectively) in comparison to Terai (79 %). The cluster group I and II has narrow genetic relationship followed by cluster III, V and IV. The cluster group II separated from cluster group I at a similarity coefficient of 0.803 and similarly cluster group IV separated from rest of the group at a similarity coefficient of 0.66. There was small genetic variation between cluster groups I and II (similarity %, 81.4 and 81.8 respectively) and cluster groups IV and V (similarity %, 73.7 and 73.9 respectively), where as wider variation was observed between cluster group IV and II (Fig 4.3). The result shows high level of genetic similarity within the acid lime landraces.

3D-Plot

The construction of 3D-plot of PCO (Principal Component Analysis) was carried out with analysis of Eigen vector using NTSYS-PC 2.21i for all the acid lime (*C. aurantifolia*) accessions, using Dice similarity matrix. It showed the comparable dispersion and

scatterness as shown by the phenogram of the individual accessions from different agro-ecological zones (Fig 4.4).

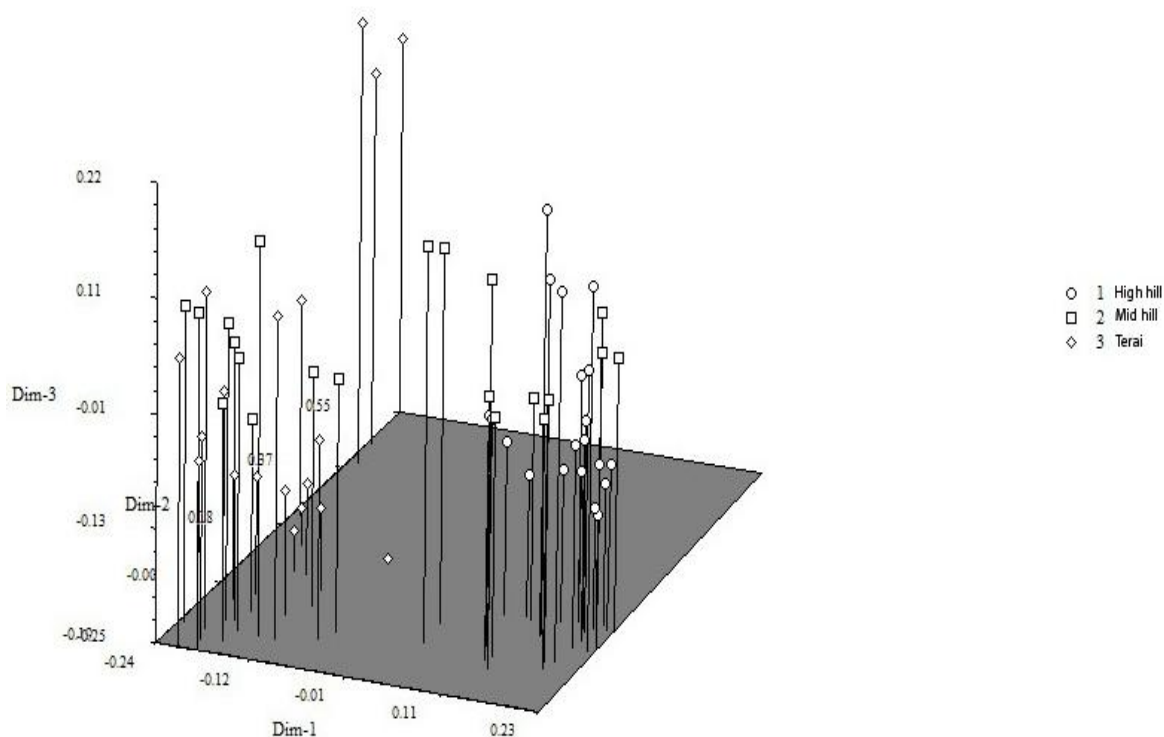


Figure 4.4 3D – plot of PCO constructed using NTSYS-PC 2.21i to assess dispersion of all acid lime (*C. aurantifolia*) accessions.

4.6.3. Principal Co-ordinate Analysis (PCO)

Similarly, Eigen analysis done in Principal Co-ordinate analysis (PCO) by MVSP 3.22 (Multi-Variate Statistical Package) also distinctly clustered the acid lime (*C. aurantifolia*) accessions in scattered plot that is in congruent with the 3D – plots illustration of genetic relationship of the acid lime (*C. aurantifolia*) in the present investigation (Fig 4.5). PCO also revealed the same association as revealed in the phenogram generated by NTSYS. The groups were discriminated, with axes 1 and 2 expanding 22.85% of the total variation. The PCO exhibited genetic variation among the acid lime landraces in the study (Table 4.12).

Table 4.12 Eigenvalues and the percentage for the co-ordinates of PCO

Eigen Values		
	Axis 1 (Co-ordinate 1)	**Axis 2 (Co-ordinate 2)**
Eigenvalues	323.271	185.777
Percentage	14.514	8.341

Cum. Percentage	14.514	22.855

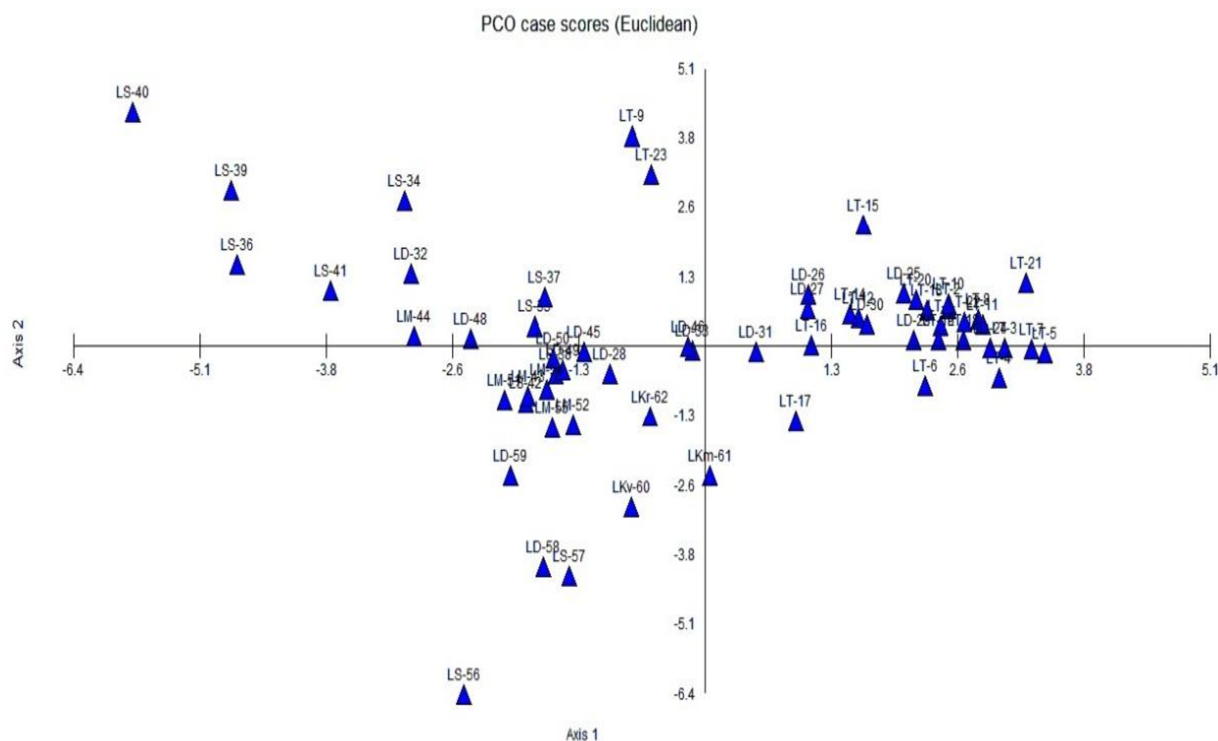


Figure 4.5 Assessment of Principal Co-ordinate Analysis (PCO) of Dice similarity matrix carried out with the MVSP 3.21.

4.6.4. Estimation of within agro-ecological zone genetic similarity of *C. aurantifolia*

The intra-zone diversity of acid lime landraces was shown according to the polymorphism of ISSR bands. On assessing each agro-ecological zone for their polymorphism using each primer, the most diverse accessions of acid lime have been observed for Terai (80.69%) and the lowest for High-hill (65.15%) (Table 4.13).

Table 4.13 Direct assessments of polymorphism based on monomorphic and polymorphic bands amplified by 21 ISSR primers.

Zon/Sample no.	High-hill (20)			Mid-hill (21)			Terai (19)		
Primers	M	P	% Pol	M	P	% Pol	M	P	% Pol
C1	4	3	42.85	4	3	42.86	3	4	57.14
C2	2	7	77.78	1	8	88.89	0	9	100
C4	0	12	100	0	11	100	0	9	100
C5	2	9	81.81	2	9	81.81	1	8	88.89
C7	5	7	58.33	2	13	86.67	3	10	76.92
C8	4	5	55.56	4	5	55.56	2	7	77.78
C9	5	3	37.5	5	4	100	3	6	66.67

C10	3	1	25	3	5	62.5	4	6	60
UBC 807	0	5	100	1	6	85.71	1	7	87.5
UBC 810	4	11	73.33	0	14	100	1	9	90
UBC 812	3	5	62.25	3	7	70	3	8	72.72
UBC 825	7	1	14.3	7	2	22.22	4	5	55.56
UBC 834	3	4	57.14	3	3	50	2	2	50
UBC 835	5	4	44.44	2	7	77.78	1	7	87.5
UBC 836	1	8	88.89	1	9	90	1	7	87.5
UBC 841	3	7	70	1	9	90	0	9	100
UBC 842	1	9	90	1	11	91.67	1	11	91.67
UBC 857	3	12	80	5	13	72.22	2	15	88.23
UBC 873	3	11	78.57	3	10	76.92	1	12	92.30
UBC 888	4	4	50	4	4	50	2	7	77.78
UBC 889	7	1	12.5	5	3	37.5	4	5	55.56
Total	69	129		57	156		39	163	
Average	3.28	6.14		2.71	7.42		1.86	7.76	
SD	1.88	3.48		1.87	3.68		1.28	2.79	
% Pol	**65.15**			**73.23**			**80.69**		

Where, Zon = Agro-ecological zone Sample no. = Number of Samples processed P = Number of Polymorphic bands M = Number of Monomorphic band % Pol = Percent Polymorphism

4.6.5. Analysis of Molecular Variance (AMOVA)

According to AMOVA analysis performed using GenAlEx 6.5, there was significant partitioning of the genetic variation ($p < 0.001$), with 14% occurring among the ecological zone and remaining 86% within the ecological zone. Indicator of the genetic differentiation (ϕ_{PT}) for the three agro-ecological zones was 0.139. This suggests that genetic differentiation of acid lime was comparably lower among than genetic differentiation of acid lime within agro-ecological zone variation accounting 86% variation (Table 4.14).

Table 4.14 Analysis of Molecular Variance (AMOVA) for 60 acid lime accessions from three agro-ecological regions

Source of variation	d.f	SS	MS	Estimated variance	Total variance	P-value	ϕ_{PT}
Among agro-zone	2	185.227	92.613	3.539	14%	< 0.001	0.139
Within agro-zone	57	1248.298	21.900	21.900	86%		
Total	59	1433.525		25.439	100%		

Where, d.f = Degree of freedom; SS = Sum of square; MS = Mean sum of square; ϕ_{PT} = Indicator of the genetic differentiation

4.7. SSR Profiling and Genetic Diversity Assessment of *Citrus* spp.

The SSR-PCR profiles generated by the 12 SSR primer pairs were used for genetic diversity assessment of *Citrus* spp. under study collected from Dhankuta and Kathmandu valley, Nepal. Representative SSR profiles generated by primers AG 14, CAT 01 and TAA 41 are shown (Plates 4.13 – 4.15).

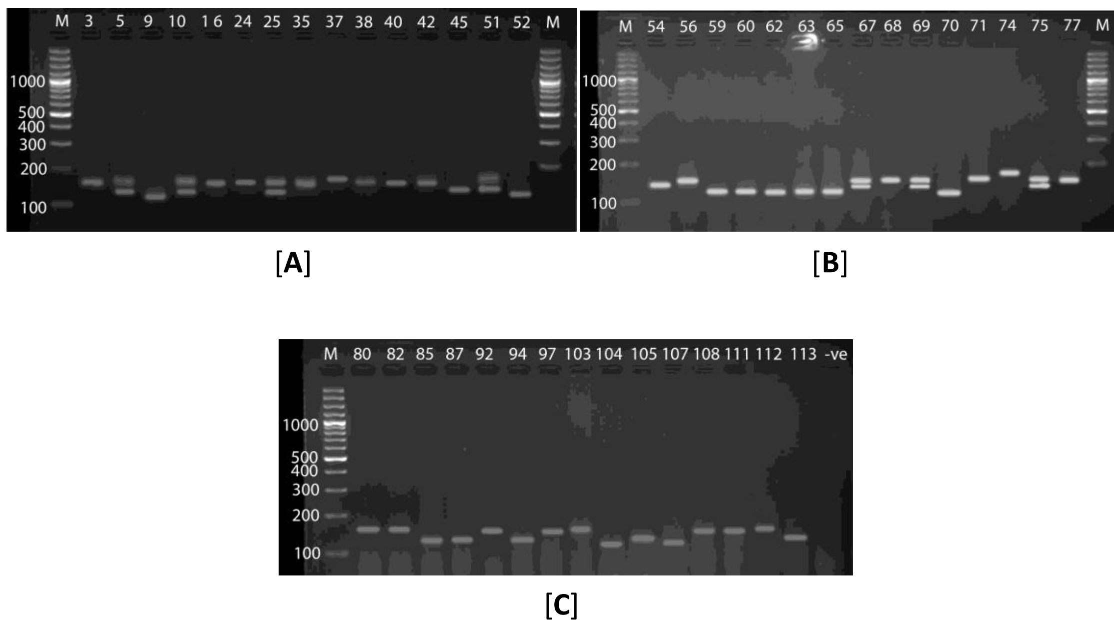
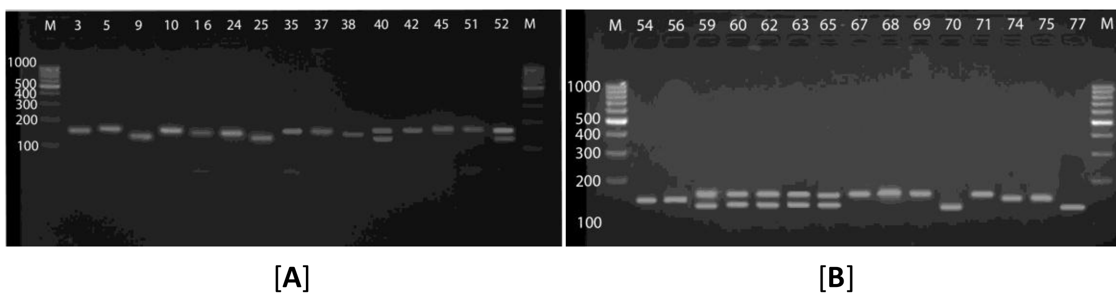
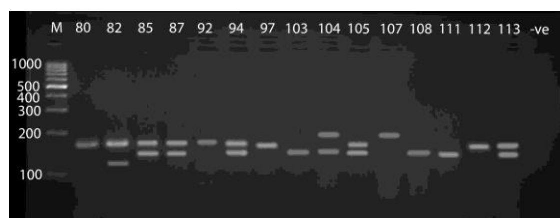


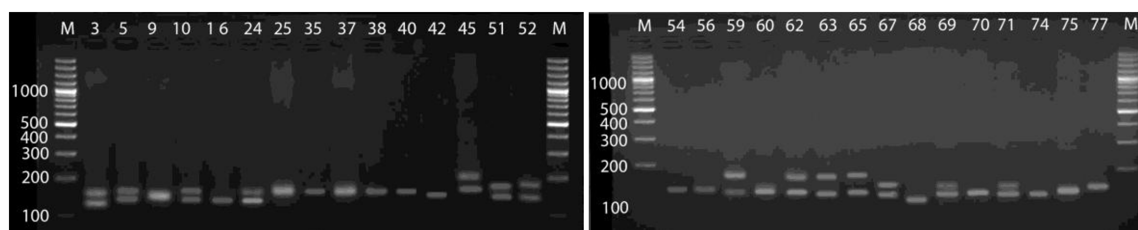
Plate 4.13 SSR profiles amplified with **Primer AG 14.** Lanes marked with M are 100bp plus molecular weight markers. [A] Lanes 3-52 represents *Citrus* spp. samples 3-52; [B] Lanes 54-77 represents acid lime samples 54-77; [C] Lanes 80-113 represents acid lime samples 80-113.



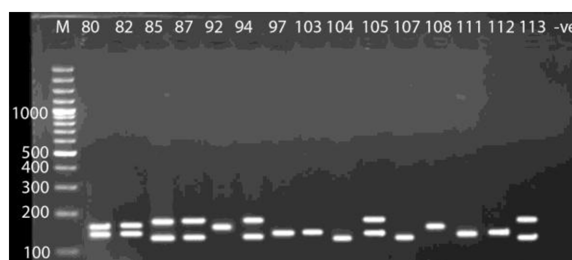


[**C**]

Plate 4.14 SSR profiles amplified with **Primer CAT 01.** Lanes marked with M are 100bp plus molecular weight markers. [A] Lanes 3-52 represents *Citrus* spp. samples 3-52; [B] Lanes 54-77 represents acid lime samples 54-77; [C] Lanes 80-113 represents acid lime samples 80-113.



[**A**] [**B**]



[**C**]

Plate 4.15 SSR profiles amplified with **Primer TAA 41.** Lanes marked with M are 100bp plus molecular weight markers. [A] Lanes 3-52 represents *Citrus* spp. samples 3-52; [B] Lanes 54-77 represents acid lime samples 54-77; [C] Lanes 80-113 represents acid lime samples 80-113.

4.7.1. Alleles per locus, Polymorphic Information Content and Probability of Identity

All the 12 SSR primers pairs used in this study successfully amplified all *Citrus* spp. samples and showed 100% polymorphism i.e. no monomorphic bands were observed for the all 12 SSR primers pairs. A total of 60 putative alleles were amplified at 12 SSR loci in the 45 accessions with an average of 5 alleles per locus. The amplified allele size ranged from 80-270 bp. Out of 12 primer pairs, a minimum of 2 alleles were detected by the primers TAA 27 and a maximum 8 allele was detected for primer TAA 52. The di-nucleotide markers produced on average 4.5 alleles per locus, where as tri-nucleotide

repeat markers had a mean of 5.25 alleles per locus. All the 12 primers amplified clear, well resolved fragments with little stutter (minor product that differ in size from the main product by multiples of the length of the repeat unit). The SSR marker could distinguish between various *Citrus* spp. Among the 12 SSR primers, three primers gave rare alleles i.e. alleles detected in only one individual in the population. Primer CAT 01 gave allele of size 130 bp for sample 82, primer TAA 52 gave allele of size 230 bp for sample 38 and TAA 41 gave allele of size 200 bp for sample 45. Similarly, the PIC score ranged from 0.497 with Primer TAA 27 to 0.802 with Primer TAA 41 with an average value of 0.662. Except the primer TAA 3, TAA 27 and CAC 15 all other primers have high PIC value (>0.6). These PICs were calculated using Microsoft Office Excel 2007. The probability of identity (PI) was calculated using GenAlEx ver. 6.5, and ranged from 0.075 to 0.383 with an average of 0.143 (Table 4.15).

Table 4.15 Assessment of Alleles per locus, Allele size range, Polymorphic Information Content (PIC) and Probability of Identity (PI) of 12 SSR primers used to generate SSR profiles in 45 citrus accessions belonging to different species.

Primer code	Primer sequence	Repeat motif	Alleles /locus	Allele size (approx.)	PIC value	PI
TAA45	GCACCTTTTATACCTGACTCGG (F) TTCAGCATTTGAGTTGGTTACG (R)	TAA	5	90-160	0.705	0.160
GT03	GCCTTCTTGATTTACCGGAC (F) TGCTCCGAACTTCATCATTG (R)	GT	4	150-200	0.675	0.214
CAT01	GCTTTCGATCCCTCCACATA (F) GATCCCTACAATCCTTGGTCC (R)	CAT / CTT	6	130-200	0.745	0.115
TAA15	GAAAGGGTTACTTGACCAGGC (F) CTTCCCAGCTGCACAAGC (R)	TAA	5	140-200	0.690	0.148
TAA3	AGAGAAGAAACATTTGCGGAGC (F) GAGATGGGACTTGGTTCACACG (R)	TAA	6	130-270	0.548	0.312
CT19	CGCCAAGCTTACCACTCACTAC (F) GCCACGATTTGTAGGGGATAG (R)	CT	5	140-180	0.602	0.206
TAA27	GGATGAAAAATGCTCAAAATG (F) TAGTACCCACAGGGAAGAGAGC (R)	TAA	**2**	170-220	**0.497**	**0.383**
TAA52	GATCTTGACTGAACTTAAAG (F) ATGTATTGTGTTGATAACG (R)	TAA	**8**	80-230	0.785	**0.089**
CAC15	TAAATCTCCACTCTGCAAAAGC (F) GATAGGAAGCGTCGTAGACCC (R)	CAC	4	150-180	0.532	0.338
TAA41	AGGTCTACATTGGCATTGTC (F) ACATGCAGTGCTATAATGAATG	TAA	**6**	130-200	**0.802**	**0.075**
TC26	CTTCCTCTTGCGGAGTGTTC (F)	TC	4	130-160	0.692	0.142

	GAGGGAAAGCCCTAATCTCA (R)					
AG14	AAAGGGAAAGCCCTAATCTCA (F)	GA	5	130-170	0.674	0.159
	CTTCCTCTTGCGGAGTGTTC (R)					
Average			**5**		**0.662**	**0.195**
Total	**12 pairs**		**60**			

Where, PIC = Polymorphic Information Content; PI = Probability of Identity

4.7.2. Allele frequency, Shannon's Information Index, Observed and Expected Heterozygosity and Fixation Index

For the calculation of these parameters, GenAlEx ver. 6.5 was employed. The results were compared with the results obtained from PowerMarker ver. 3.25, and the similar results were obtained. The allele frequency table shows the list of allele size in each primer (locus) and allele frequency of each allele (Table 4.17). The effective number of alleles ranged from 1.850 to 4.844 with an average of 3.009. The observed heterozygosity ranged from 0.111 (for primer CAC 15) to 0.778 (for primer TAA 27) and expected heterozygosity ranged from 0.460 (for Primer TAA 3) to 0.794 (for primer TAA 41) (Nei's) or 0.465 to 0.802 (unbiased) and it revealed a medium, average length polymorphism rates for SSR markers of 0.638 (Nei's) or 0.645 (unbiased). Similarly, the Shannon's information index ranged from 0.677 to 1.678 with an average of 1.226 and the Fixation Index ranged from -0.606 to 0.798 with an average of 0.446 (Table 4.16).

Table 4.16 Assessmet of Number of Different and Effective alleles, Shannon's Information Index, Observed heterozygosity, Expected heteroxygosity, Unbiased expected heterozygosity and fixation Index of 12 SSR primers used to generate SSR profiles in 45 citrus accessions belonging to different species.

Primer Code	N	Na	Ne	I	Ho	He	uHe	F
TAA45	45	5	3.103	1.291	0.467	0.678	0.685	0.311
GT03	45	4	2.677	1.097	0.444	0.626	0.633	0.291
CAT01	45	6	3.828	1.441	0.311	0.739	0.747	0.579
TAA15	45	5	3.250	1.316	0.489	0.692	0.700	0.294
TAA3	45	6	1.850	1.000	0.244	**0.460**	**0.465**	0.468
CT19	45	5	2.435	1.172	0.133	0.589	0.596	0.774
TAA27	45	2	1.939	**0.677**	**0.778**	0.484	0.490	-0.606
TAA52	45	8	3.990	**1.678**	0.311	0.749	0.758	0.585
CAC15	45	4	1.942	0.834	**0.111**	0.485	0.491	0.771
TAA41	45	6	**4.844**	**1.624**	0.467	**0.794**	**0.802**	0.412
TC26	45	4	3.306	1.286	0.222	0.698	0.705	0.681
AG14	45	5	2.945	1.294	0.133	0.660	0.668	0.798
Average	**45**	**5.000**	**3.009**	**1.226**	**0.343**	**0.638**	**0.645**	**0.446**

S.E	0.000	0.426	0.264	0.085	0.056	0.032	0.033	0.110
St. Dev	0.000	1.477	0.916	0.295	0.195	0.111	0.112	0.382

Where, N = No. of Samples; Na = No. of Different Alleles; Ne = No. of Effective Alleles; Ho = Observed Heterozygosity; He = Nei's Expected Heterozygosity; uHe = Unbiased Expected Heterozygosity; I = Shannon's Information Index; F = Fixation Index; S.E = Standard Error; St. Dev = Standard Deviation

Table 4.17 Allele Frequency of each of the alleles of individual locus.

Primer TAA 45			**Primer GT 03**			**Primer CAT 01**			**Primer TAA 15**		
S.N	**Allele Size (bp)**	**Frequency**	**S.N**	**Allele Size (bp)**	**Frequency**	**S.N**	**Allele Size (bp)**	**Frequency**	**S.N**	**Allele Size (bp)**	**Frequency**
1	**90**	0.133	1	**150**	0.033	1	**130**	0.011	1	**140**	0.022
2	**100**	0.067	2	**160**	0.122	2	**140**	0.100	2	**150**	0.156
3	**130**	0.333	3	**170**	0.400	3	**150**	0.289	3	**160**	0.422
4	**150**	0.433	4	**200**	0.444	4	**160**	0.278	4	**180**	0.311
5	**160**	0.033				5	**170**	0.300	5	**200**	0.089
						6	**200**	0.022			

Primer TAA 3			**Primer CT 19**			**Primer TAA 27**			**Primer TAA 52**		
S.N	**Allele Size (bp)**	**Frequency**	**S.N**	**Allele Size (bp)**	**Frequency**	**S.N**	**Allele Size (bp)**	**Frequency**	**S.N**	**Allele Size (bp)**	**Frequency**
1	**130**	0.011	1	**140**	0.033	1	**180**	0.411	1	**80**	0.056
2	**140**	0.044	2	**150**	0.178	2	**210**	0.589	2	**90**	0.433
3	**150**	0.722	3	**160**	0.600				3	**100**	0.100
4	**160**	0.067	4	**170**	0.089				4	**130**	0.178
5	**180**	0.089	5	**180**	0.100				5	**160**	0.044
6	**270**	0.067							6	**180**	0.078
									7	**200**	0.100
									8	**230**	0.011

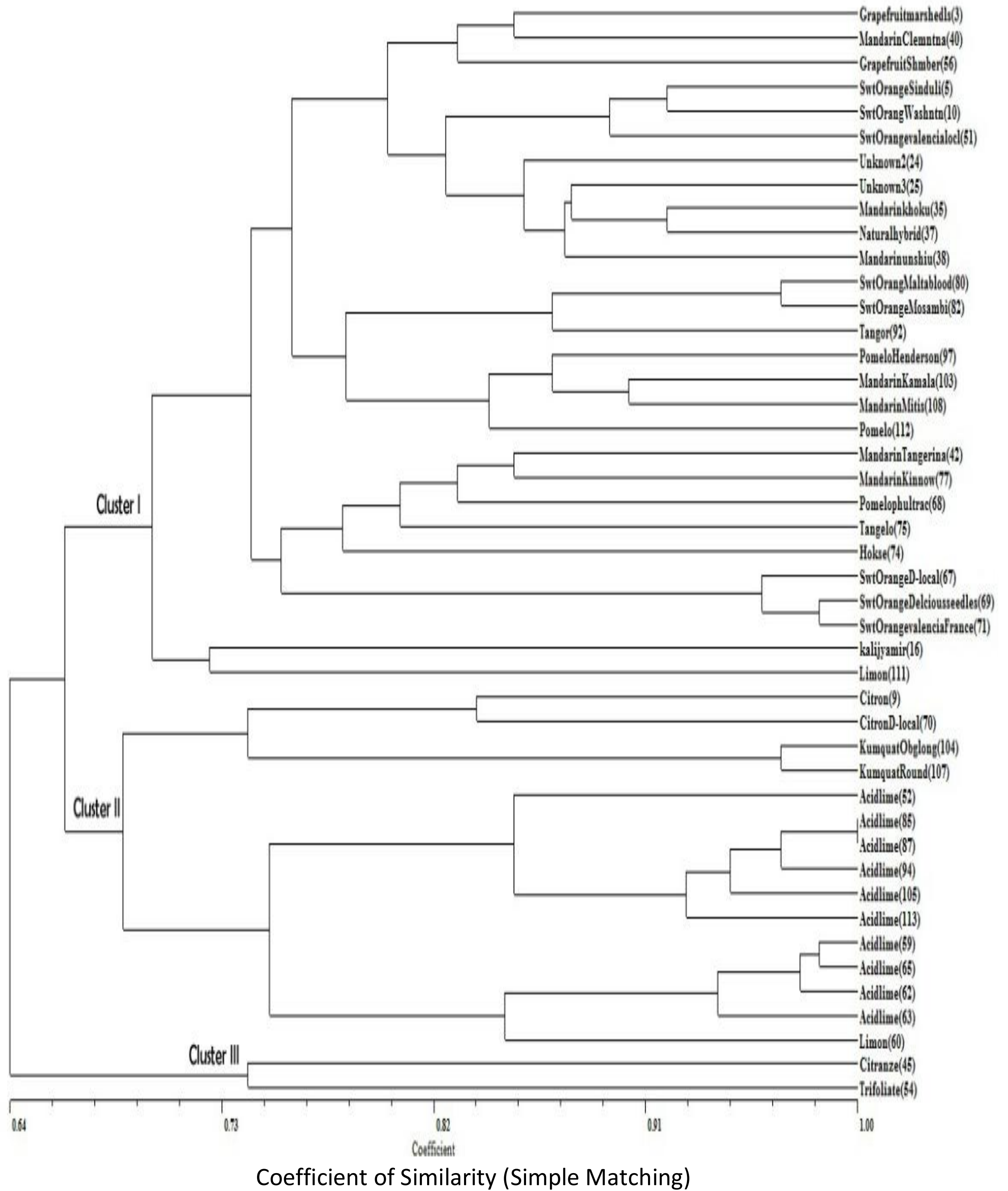
Primer CAC 15			**Primer TAA 41**			**Primer TC 26**			**Primer AG 14**		
S.N	**Allele Size (bp)**	**Frequency**	**S.N**	**Allele Size (bp)**	**Frequency**	**S.N**	**Allele Size (bp)**	**Frequency**	**S.N**	**Allele Size (bp)**	**Frequency**
1	**150**	0.033	1	**130**	0.178	1	**130**	0.200	1	**130**	0.222
2	**160**	0.289	2	**140**	0.256	2	**140**	0.244	2	**140**	0.144
3	**170**	0.656	3	**150**	0.200	3	**150**	0.122	3	**150**	0.078
4	**180**	0.022	4	**160**	0.233	4	**160**	0.433	4	**160**	0.511
			5	**180**	0.122				5	**170**	0.044
			6	**200**	0.011						

4.7.3. Genetic diversity analysis using similarity coefficients and cluster analysis

4.7.3.1. Comparison of Similarity coefficients and construction of phenograms

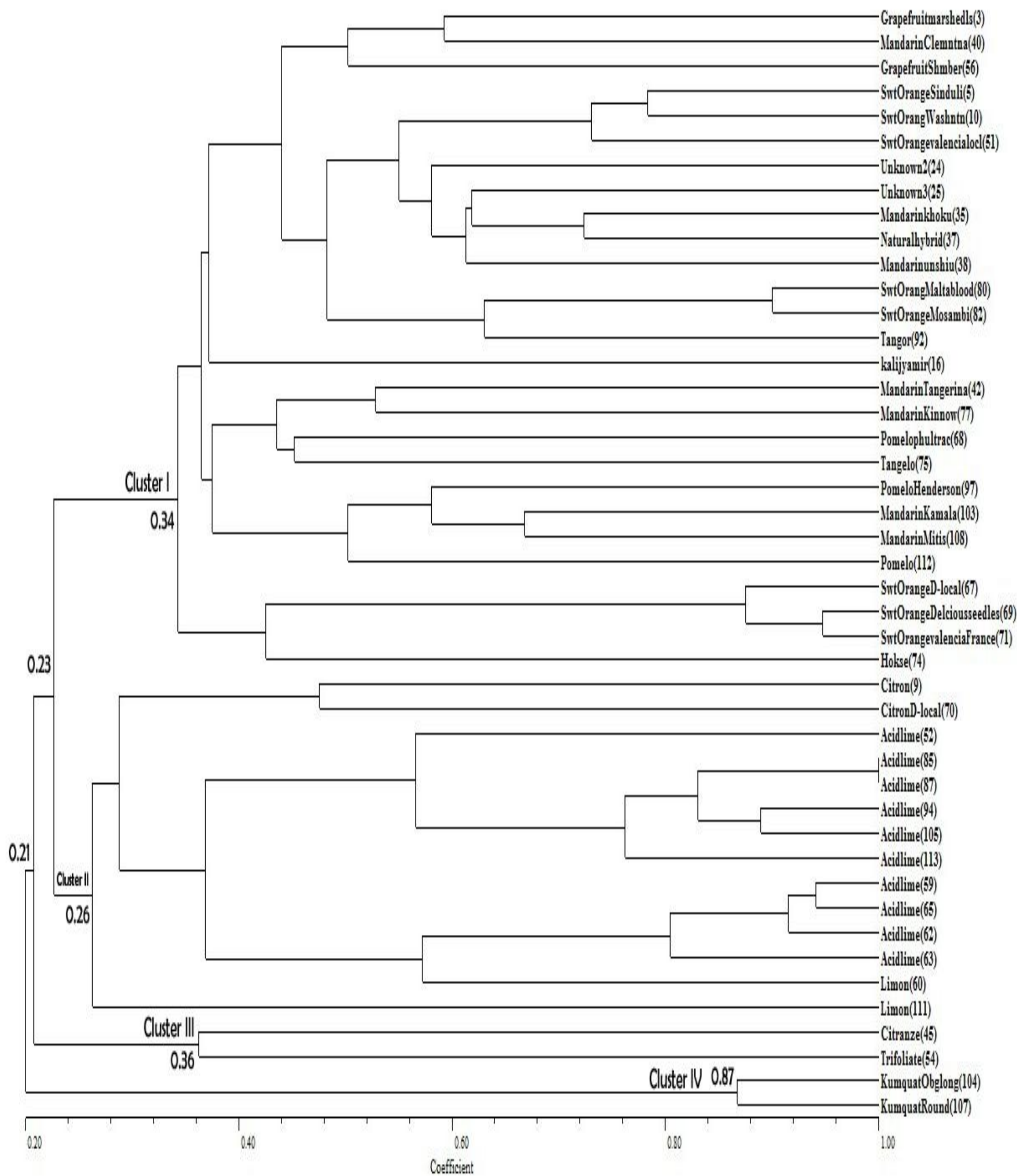
As done in ISSR analysis of acid lime, the similarity coefficients were generated from the binary data based on presence and absence of SSR loci using SimQual computational algorithm of NTSYS-PC ver. 2.21i. The ranges of similarity coefficients were obtained, using Simple Matching (SM), Jaccard's (J) and Dice (D) coefficient i.e. SM (0.53 – 1.0), J (0.10 – 1.0) and Dice (0.12 – 1.0) with an average of 0.716, 0.30 and 0.45 respectively (Appendix 10, 11 and 12).

The above similarity coefficients were employed for the construction of phenograms in order to represent genetic relationship among the 45 accessions belonging to different *Citrus* spp. The topology of phenograms was comparable but little variation was observed in case of phenogram generated by using Simple Matching coefficient. The two kumquats i.e. round and oblong forms separate cluster in case of phenograms generated by using Jaccard's and Dice coefficient (*viz.* fig 4.6, 4.7 and 4.8).



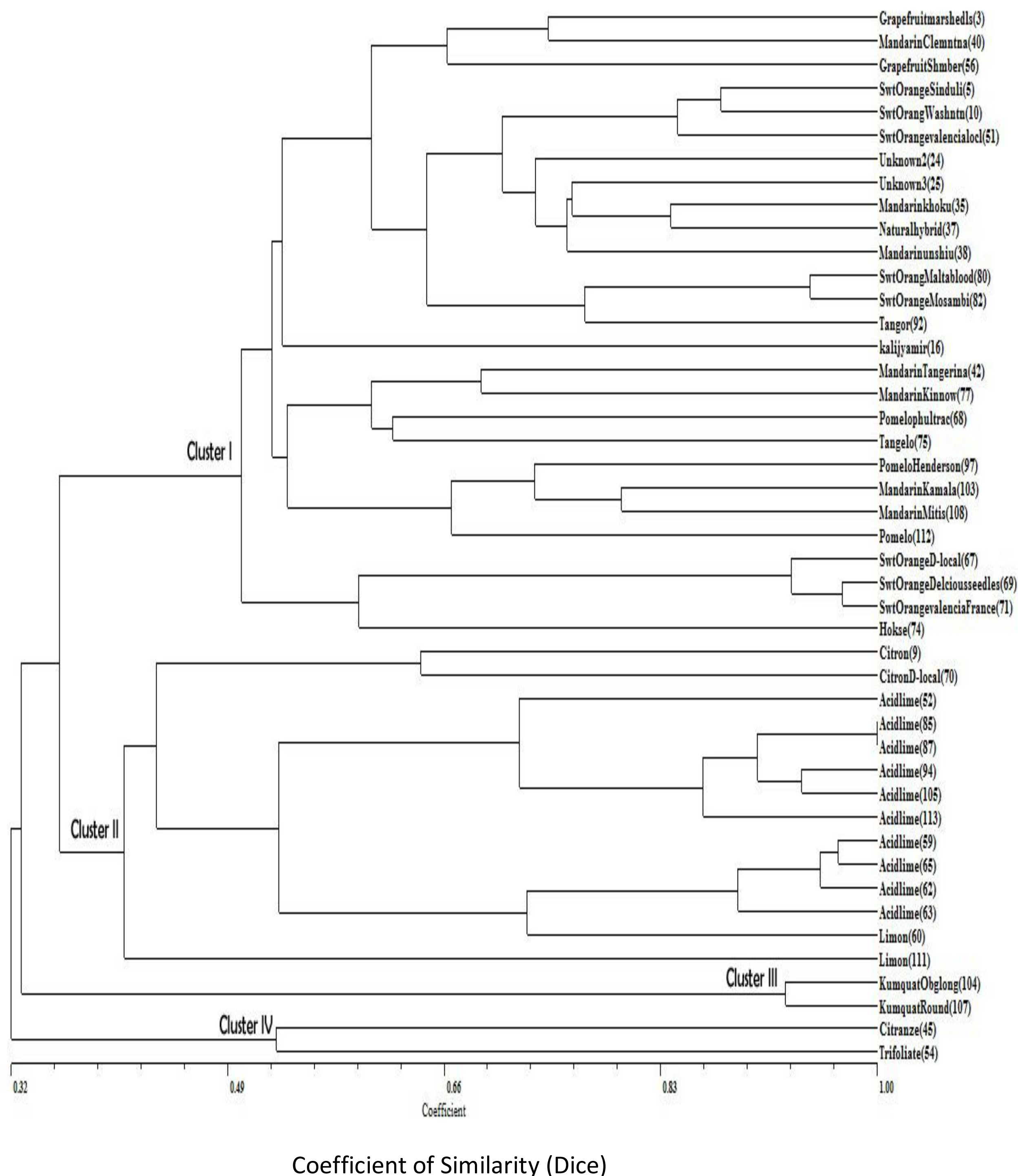
Coefficient of Similarity (Simple Matching)

Figure 4.6 Phenogram generated for 45 citrus accessions belonging to different species (refer to table 4.1 for sample details) by UPGMA cluster analysis using Simple Matching coefficient of similarity generated by 12 SSR primers. The clusters are labeled as I, II, and III.



Coefficient of Similarity (Jaccard's)

Figure 4.7 Phenogram generated for 45 citrus accessions belonging to different species (refer to table 4.1 for sample details) by UPGMA cluster analysis using Jaccard's coefficient of similarity generated by 12 SSR primers. The clusters are labeled as I, II, III and IV.



Coefficient of Similarity (Dice)

Figure 4.8 Phenogram generated for 45 citrus accessions belonging to different species (refer to table 4.1 for sample details) by UPGMA cluster analysis using Dice coefficient of similarity generated by 12 SSR primers. The clusters are labeled as I, II, III and IV.

In order to select best similarity coefficient and the best fitted phenogram, original similarity matrix were compared using Mantel test to calculate correlation values. The correlation values suggest Dice or Jaccard's coefficient is more suitable than Simple Marching coefficient (Table 4.18).

Table 4.18 Correlation Coefficient values from Mantel test of original similarity matrices.

	Simple Matching (SM)	**Jaccard (J)**	**Dice (D)**
Simple Matching (SM)	******	0.95034	0.95463
Jaccard (J)		******	0.98573
Dice (D)			******

Consensus indices (CI) were calculated using Strict consensus method for the each combination of coefficient and UPGMA clustering for the evaluation of trees. The highest consensus fork index (CI_C = 0.97674) was observed for Jaccard's and Dice coefficient on comparison to others (Table 4.19). This result supports phenograms constructed from Dice or Jaccard's coefficient will be more suitable than that constructed from others.

Table 4.19 Consensus fork index among the UPGMA based phenograms using similarity coefficients among *Citrus* spp. by SSR markers.

	Simple Matching (SM)	**Jaccard (J)**	**Dice (D)**
Simple Matching (SM)	******	0.69767	0.72093
Jaccard (J)		******	0.97674
Dice (D)			******

The clustering based on Unweighted Pair Group Method of Arithematic Averages (UPGMA) for Jaccard's coefficient observed to give a highest Cophenetic Correlation value of 0.87407 and lowest cophenetic correlation value 0.82169 was observed for UPGMA clustering using Simple Matching coefficient where as intermediate value 0.83151 was observed for UPGMA clustering using Dice coefficient (Table 4.20).

Table 4.20 Correlation coefficient value (r) obtained for cophenetic values of similarity matrices (Simple Matching, Jaccard's and Dice coefficient) and clusters computed by UPGMA module using MXCOMP (matrix comparison) of NTSYS.

Clustering modules of similarity	**Simple Matching**	**Jaccard**	**Dice**
UPGMA	0.82169	0.87407	0.83151

Therefore, Jaccard's similarity with UPGMA clustering yielded highest correlation coefficient value of 0.87407 followed by Dice and SM similarity respectively. From this results, the three coefficient can be judged in the decreasing order as J> D > SM to be used for interpretation of genetic relationship in *Citrus* spp.

Phenogram of 45 citrus accessions belonging to different species was constructed on the basis of Jaccard's coefficient using SAHN module of NTSYS-PC ver. 2.21i. The genetic relationship among the different *Citrus* spp. was distinct and observed to be clustered separately. The UPGMA clustering revealed the genetic relationship between different *Citrus* spp. Total accessions were clustered into two major (I and II) and two minor (III and IV) clusters. Cluster I consist of grapefruits, sweet oranges, sour oranges, unknown varieties, natural hybrids, mandarins and pummelo. Citron, acid lime and lemon were grouped in cluster II while Citranze and Trifoliate formed cluster III and kumquats formed Cluster IV (Fig 4.9). The genetic variations in cluster I, II and III were not so wide (similarity %, 34, 26 and 36 respectively), where as cluster IV had wider variation (87%). The average genetic similarity among the 45 accessions was 30.0% with a range of 10% to 100%. Cluster group IV had less genetic diversity (i.e. round and oblong kumquats were 87% similar) than I, II and III. The two rootstocks plant i.e. Citranze and Trifoliate were 36% genetically similar. The acid lime (sample 85 and 87) are in same genetic level under the cluster II. The cluster I separated from cluster II at a similarity coefficient of 0.23. Similarly cluster III separated from the cluster I and II at a similarity coefficient of 0.21 and cluster IV separated from rest of the clusters at a similarity coefficient of 0.20. The cluster I showed mandarin is very genetically diverse. The local Hokse (natural hybrid) clustered with sweet oranges. The unknown natural hybrid and unknown varieties were clustered in cluster group I with mandarins (Khoku and Unshiu) (Fig 4.7).

4.7.4. Principal Co-ordinates Analysis (PCO)

Eigen analysis was done in Principal Co-ordinate Analysis (PCO) of Jaccard's matrix by using MVSP 3.21 (Multi-Variate Statistical Package). This shows the distinct clustering of Citrus spp. that is congruent with the phenogram generated by NTSYS. The first principal co-ordinate axis accounted 20.35 % and the second accounted 11.43 % of the total variation. The groups were discriminated with axes 1 and 2 expanding 31.79% of the total variation (Table 4.21 and Fig 4.11).

Table 4.21 Eigenvalues and the percentage for the co-ordinates of PCO.

Eigen Values		
	Axis 1 (Co-ordinate 1)	**Axis 2 (Co-ordinate 2)**
Eigenvalues	80.363	45.141
Percentage	20.358	11.435
Cum. Percentage	20.358	31.793

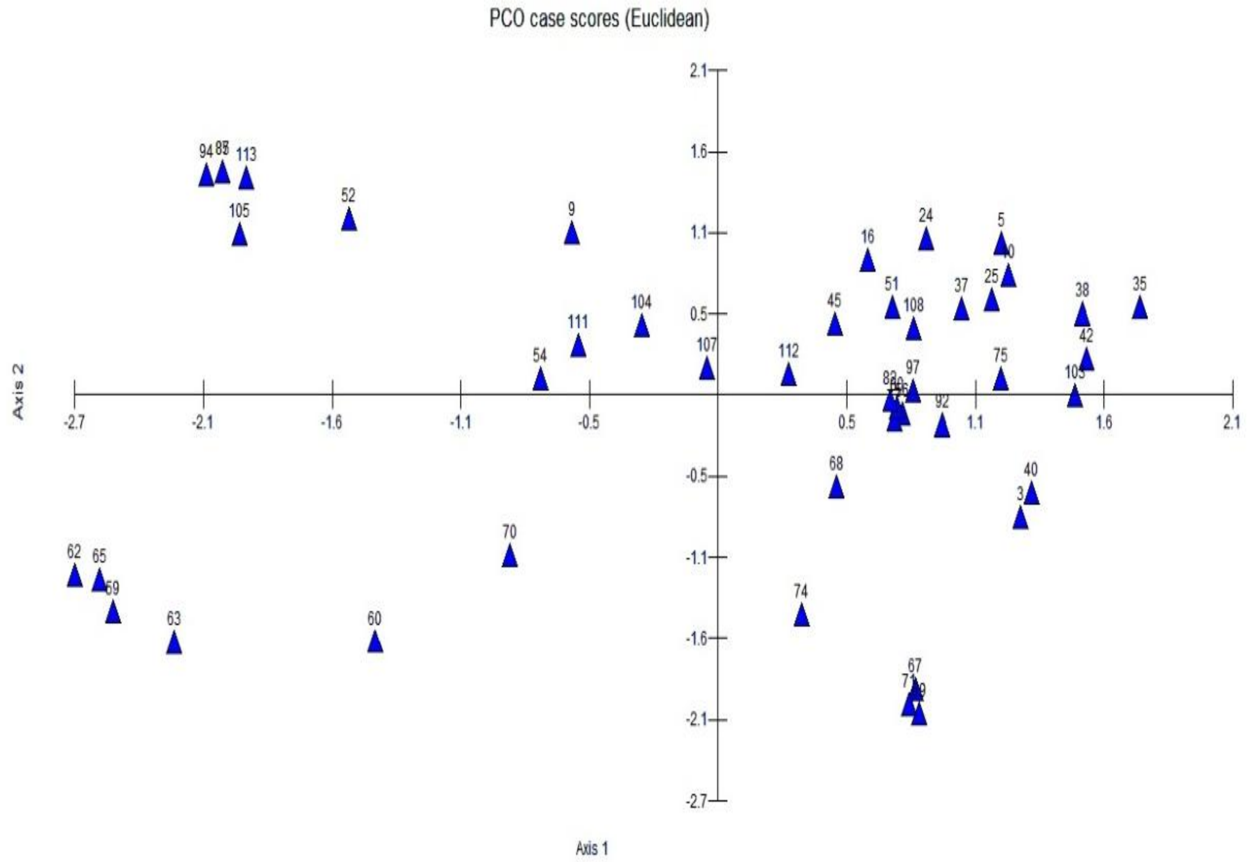


Figure 4.11 Assessment of Principal Co-ordinate Analysis (PCO) of Jaccard's similarity matrix carried out with the MVSP 3.21.

CHAPTER 5: DISCUSSION

Genetic variation exists in most natural populations. Different traits in a population are due to the alleles of gene and/or due to the effects of environment. Hidden genetic variation is even more extensive in population than that observed through the phenotype. It is virtually impossible for two individuals in a population to have the same genotype at all the loci (Karp et al., 1998; Lowe et al., 2004). This genetic variation may be detected through molecular technologies which reveal genetic polymorphisms. The genetic polymorphism can work as genetic markers. A variety of molecular marker techniques have been developed during past few decades, and they allow the analysis and detection of variation at the genomic level, such molecular level information could be used to assess genetic relationships among the major germplasm groups (Biswas et al., 2011). Molecular genetic techniques have a significant impact on plant genetic resources conservation and uses. Initially these techniques were used for the analysis of specific genes, for understanding gene actions, gene mapping and the development of gene transfer technologies. More recently, the techniques have been applied to problems of direct relevance for understanding the distribution and extent of genetic variation within and between species (Ayad et al., 1997). Elucidating relationships, taxonomy, and diversity is important for developing breeding strategies, conserving biodiversity, and improving breeding efficiency. The study of genetic variability in citrus is important for characterizing germplasm, controlling genetic erosion and the registration of new cultivars (Herrero et al., 1996; Barkley et al., 2006). Citrus is a cross-pollinated crops and their populations are highly heterozygous as well as heterogeneous due to free intermating among their related plants and are referred to as random mating populations because each individuals of the population has equal opportunity of mating with any other individual of that population (Singh, 2005). In the present investigation, two main researches have been undertaken. One investigation employed ISSR marker system to assess genetic diversity among different acid lime landraces found in three agro-ecological zone of eastern Nepal while other employed SSR marker system for the assessment of genetic diversity in Nepalese *Citrus* spp.

5.1. Optimization of PCR conditions of ISSR and SSR

Polymerase Chain Reaction (PCR) is a very sensitive assay in which a single DNA molecule can be amplified, and then single-copy genes can be amplified out of complex mixtures of genomic sequences. Since no single protocol will be appropriate for PCR amplification to all situations, so each new PCR application requires optimization. Some often encountered problems include: no detectable product or a low yield of the desired product, the presence of non-specific background bands due to mis-priming or mis-extension of the primers, the formation of "primer-dimers" that compete for

amplification with the desired product, and mutations or heterogeneity due to mis-incorporation (Saiki et al., 1985; Mullis et al., 1986; Mullis and Faloona, 1987; Innis and Gelfand, 1990). In the present investigation, PCR reaction and cycling conditions for acid lime and different *Citrus* ssp have been optimized. These optimized conditions were then subsequently used for the profiling of various samples under study.

5.1.1. Optimization of PCR Reaction conditions for ISSR and SSR

In order to maintain reproducible banding patterns among the laboratories, optimization of various PCR reaction parameters is necessary (Shrestha, 2001; Weising et al., 2005). The following are the sensitive reaction parameters *viz.* template DNA concentration, $MgCl_2$ concentration, Primer concentration, dNTPs concentration, *Taq* polymerase concentration and cycling conditions which were critically optimized for the development of standard PCR protocol for Citrus crops.

The minimum concentration of template DNA that gives the best crispy banding patterns (clear heterozygous bands in case of SSR analysis) was observed at 25 ng DNA, therefore, the DNA concentration was selected and used in further experiments i.e. ISSR and SSR profiling of acid lime and *Citrus* spp. (Plate 4.2 and Plate 4.6). Template may have a considerable influence on the results of a PCR, and their quality and quantity were the main factors affecting reproducibility (Micheli et al., 1997; Weising et al., 2005). For most of the species, good results have been achieved for concentrations ranging from 50 – 100 ng in 25 - 50µl PCR reaction volume (Caetano-Anolles et al., 1991; Shrestha, 2001; Padmalatha and Prasad, 2006). So, in the present investigation DNA concentration ranging from 12.5 to 100 ng in 25 µl reaction volume were tested. Negative control (no DNA) was included and no bands were observed indicating free of false positive results. High amount of DNA usually inhibit PCR-amplification creating hindrance for primer annealing (Micheli et al., 1997). So, minimum concentration was taken.

$MgCl_2$ is a cofactor of *Taq* polymerase enzyme and its concentration affect primer annealing, strand dissociation temperatures of both template and PCR product, product specificity, formation of primer-dimer artifacts, and enzyme activity and fidelity (Innis and Gelfand, 1990). In the present investigation, the range of $MgCl_2$ concentration tested i.e. 1.5 mM to 4.5 mM, the crispy profile with more bands was observed at 3.0 mM $MgCl_2$ concentration for ISSR and crispy clear heterozygous for SSR experiments. Therefore, 3.0 mM concentration of $MgCl_2$ was considered as optimum for the PCR amplification (Plate 4.2 and Plate 4.6). Generally, increasing amounts of Mg^{++} results in the accumulation of non-specific amplification products, and insufficient Mg^{++} will reduce the yield (Wiliams et al., 1991). Though strong and reproducible bands are

obtained over a wide range of Mg^{++} concentration, a change in concentration often result in a qualitative change of fragment patterns (Wiliams et al., 1993). The $MgCl_2$ can be optimized either by determining the optimal primer/template combination (primer screening) in the fixed $MgCl_2$ concentration or varying $MgCl_2$ concentration for fixed primer/template combination (Shrestha, 2001). In the present investigation the later strategy was followed.

Tested ranges of primer concentration were 0.1 µM to 1.6 µM. The banding pattern at 0.4 µM, which is minimum concentration, was crispy and reveled more no. of amplified bands for ISSR and clear heterozygous bands for SSR experiments, so selected as optimum concentration and used for further PCR reactions (Plate 4.3 and Plate 4.7). Primer concentration plays significant role in the outcome of PCR. High concentration of primer promotes mis-priming, accumulation of non-specific products and formation of primer dimer where as low concentration leads to amplification failure. The primer-dimer artifacts are themselves substrates for PCR and compete with the desired product for enzyme, dNTPs, and primers, resulting in a lower yield of the desired product (Innis and Gelfand, 1990; Padmalatha and Prasad, 2006). So, minimum concentration was taken.

In the present investigation, dNTPs were tested in the range of 0.1 to 0.5 mM concentration, among which the concentration at 0.4 mM gave clear and crispy bands for ISSR-PCR and at 0.2 mM gave clear heterozygous bands for SSR-PCR. Hence, 0.4mM and 0.2mM dNTPs concentrations were selected as optimum for further profiling (Plate 4.4 and Plate 4.8). Concentration from 0.1 to 0.2 mM of dNTPs have been quoted as being optimal for the most reactions (Wiliams et al., 1990; Caetano-Anolles et al., 1991). dNTP concentrations were reported to have effect on the pattern of DNA amplification. The four dNTPs should be used at equivalent concentrations to minimize misincorporation errors. Both the specificity and the fidelity of PCR are increased by using lower dNTP concentrations. Low dNTP concentrations minimize mis-priming at non-target sites and reduce the likelihood of extending mis-incorporated nucleotides (Innis et al., 1988). dNTPs are known to chelate Magnesium and high concentration of dNTPs increase the error rate of *Taq* polymerase, interfering its activity (due to limited freedom for Mg^{++}) (Gelfland, 1989; Padmalatha and Prasad, 2006).

In the present investigation, *Taq* polymerase concentrations were tested in the range of 0.5 to 2.5U. The clear, crispy bands were observed in all these concentration for ISSR-PCR. For the present experiment, 1.5 U was selected as optimum for further ISSR experiment. While in case of SSR-PCR, a clear and crispy heterozygous band was

observed at 2U polymerase concentration and therefore selected as optimum (Plate 4.4 and Plate 4.8). A recommended concentration range for *Taq* DNA polymerase is between 1 and 2.5 units per 100 ul reaction when other parameters are optimum (Lawyer et al., 1989). *Taq* DNA polymerase from different suppliers may behave differently because of different formulations, assay conditions, and/or unit definitions. If the enzyme concentration is too high, non-specific background products may accumulate, and if too low, an insufficient amount of desired product is made (Innis and Gelfand, 1990; Weising et al., 2005). However, the choice of enzyme depends on the requirements of the PCR experiment especially with respect to specificity, efficiency or fidelity (Cha and Thilly, 1993).

5.1.2. Optimization of PCR cycling conditions of ISSR and SSR

PCR cycling parameters like temperatures, duration and ramping rate during denaturation, annealing, extension step and number of amplification cycles play significant roles for optimal banding patterns (RycWik et al., 1990; Weising et al., 2005; Shrestha et al., 2010). The most likely cause for failure of a PCR is incomplete denaturation of the target template and/or the PCR product. An applicable annealing temperature is 5°C below the true Tm (melting temperature) of the primers (Innis et al., 1988; Innis and Gelfand, 1990). The optimum number of cycles will depend mainly upon the starting concentration of target DNA when other parameters are optimized. Too many cycles can increase the amount and complexity of non-specific background products (Plateau Effect) and too few cycles give low product yield (Innis and Gelfand, 1990).

Out of four randomly selected ISSR-PCR programs, cycling conditions described by Shahsavar et al. (2007) (Table 4.2) produced the best banding pattern for ISSR profiling of acid lime DNA in the present investigation (Plate 4.1). Though, others programs also gives appreciable banding patterns but best, reproducible, crispy and clear bands were given by the program described by above paper. This may be because number of cycles in this program is high i.e. 40 cycles in comparison to others and low annealing temperature i.e. 50°C. During the present investigation, very few non-reproducibility cases were encountered. This reproducibility was checked by repeating the doubtful experiments.

Out of four randomly selected SSR-PCR programs, cycling conditions described by Snoussi et al. (2012) (Table 4.4) produced best banding patterns for SSR profiling of *Citrus* spp. DNA in the present investigation (Plate 4.5). The causes may be one or more reasons described above.

5.2. Primer Screening for ISSR

For the PCR-based methods, sufficiency of discriminating different genotypes varies considerably among the primers. Usually, it is advisable to screen a considerable number of primers on few samples in order to find an optimal set, which can then be used for the entire materials under study (Weising et al., 2005). This step is important as scorable bands and desirable markers depends on the best primer/template combinations (Shrestha, 2001). The primers having 18 to 28 nucleotides in length and 50 to 60% G+C composition considered as efficient primer (Innis and Gelfand, 1990). The ISSR primers are usually 16-25 bp long, and hence have high reproducibility in comparison to RAPD primers which permits the subsequent use of high annealing temperature (45 – 60 ^{o}C) leading to higher stringency (Reddy et al., 2002). Further, some of the primers used in this investigation are anchored at 5′ or 3′ end which might be the reason for high fidelity and reproducibility of the banding patterns. In past, number of primers used for genetic diversity study in *Citrus* spp. varies according to different investigators. For example, out of screened 105 ISSR markers, 15 primers which gave best amplification results were selected for polymorphism analysis of Indian wild orange and related wild species (Kumar et al., 2010). Similarly, 6 out of 10 ISSR primers amplified reproducible and polymorphic bands in phenetic analysis between *C. indica* and few commercially important *Citrus* spp. (Chibame et al., 2010).

In case of ISSR-PCR profiling, 49 primers were screened using fresh genomic DNA of LS-15 acid lime (C. *aurantifolia*) sample, based on number of fragments (bands) and their visibility using optimized ISSR-PCR reaction and cycling condition. Twenty one were selected on the basis of reproducibility and crispy nature of bands amplified in ISSR-PCR analysis involving all 60 acid lime accession (Plate 4.9). The experiments were repeated twice for the conformation of primers that produce crispy bands. While for the SSR-PCR profiling, all of the 12 screened primer pairs gave reproducible, considerable and scorable crispy bands for the *Citrus* spp. and hence all 12 primers were selected for the profiling of *Citrus* spp. samples.

5.3. ISSR Profiling and Genetic Diversity Assessment of Acid lime

Molecular markers are imperative in the assessment of genetic diversity, ranging from nucleotide level (SNPs) to gene and allele frequencies (genotype information), and devising various germplasm conservation, characterization and utilization (breeding, molecular diagnostics etc.) programs (Sarwat, 2012). The amplified fragments using

different primers ranged from 250 bp to 3200 bp which is comparable as suggested by Reddy et al. (2002) (Table 4.6).

The banding patterns are qualitative and are rarely informative on its own, so they are evaluated statistically (i.e. for translation into biological meaning). The banding patterns originating from different samples have to be compared to each other and for this, individual bands within a lane are assigned to particular positions (molecular weight marker-assisted sizing), and the different lanes are screened for comigrating (i.e. matching) bands (Weising et al., 2005). Scoring of ISSR loci were done visually corresponding to an individual for each ISSR primer assessed. Binary data matrix was generated on the basis of presence (1), absence (0) and missing (9) bands.

The present research is novel as no such work (using ISSR marker for genetic diversity assessment of acid lime accessions) has been carried out before on Nepalese accessions. From the total amplification profiles for every accession with each ISSR primers (selected 21 primers), monomorphic and polymorphic ISSR loci were observed. Among the total 234 band, 204 were observed polymorphic revealing 87.18% polymorphism in 60 acid lime accessions (Table 4.6). In literature, ISSR markers were mostly utilized for revealing polymorphisms among the *Citrus* spp. (not only limited within acid lime). Polymorphism reported in *Citrus* spp. is comparable with the present investigation such as 89.4% (Shahsavar et al., 2007), wild *Citrus* spp., 87% (Kumar et al., 2010), *C. indica* and few commercially important *Citrus* spp., 100% (Chibame et al., 2010) and so on.

PIC value is the probability of a given marker being informative in a random mating. It indicates heterozygosity and is associated with degree of polymorphism. The value of PIC between 0 and 1 estimates the degree of polymorphism of the marker (Arya et al., 2011). In the present investigation, highest PIC value of 0.93 for primer UBC 857, lowest of 0.74 for UBC 807 and average PIC value of 0.85 were found. PIC value of less than 0.25 indicates low polymorphism, a value between 0.25 and 0.50 indicate average polymorphism and a value higher than 0.5 indicates a highly polymorphic locus (Botstein et al., 1980). It has been reported that primers with comparably higher PIC value were more useful than others for distinguishing accessions (Teklewood and Becker, 2006). Primers C1, C2, C4, C5, C7, C8, C9, UBC 810, UBC 812, UBC 825, UBC 834, UBC 835, UBC 836, UBC 841, UBC 842, UBC 857, UBC 873, UBC 888, UBC 889 gave PIC value of greater than 0.80 suggesting they are highly useful for analysis of genetic diversity of acid lime accessions (Table 4.7). This high PIC value supports the high genetic variability in acid lime accessions of Nepal. Resolving power (Rp) is an index developed to compare the value of different primers in terms of the informative bands obtained in a given set of

germplasm (Prevost and Wilkinson, 1999). Band informativeness (I_B) and the primer resolving power (R_P) provide quantitative data allowing direct comparisons between primers (Sokal, 1979). In the present investigation, the primer (UBC 857) that has highest PIC value (0.93) also has highest Rp value (21.16) which indicates the same primer highly suitable for diversity assessment of acid lime as it gives informative bands (Table 4.7).

5.3.1. Analysis of Genetic Diversity using NTSYS and MVSP Statistical Packages for Acid lime

From the binary data matrix, pair wise similarity matrices were deduced by means of Simple Matching (SM), Jaccard's (J) and Dice (D) coefficient of similarity using SimQual (Similarity for Qualitative Data) module of NTSYS pc ver. 2.21i. For scoring, only visibly distinguished bands in agarose gel picture were considered. In some profiles, confused ISSR bands were observed and experiments were repeated for verification. In addition, few samples failed to give bands with some ISSR primers. Those experiments were repeated to ensure amplification failure and scoring was done as '9' for still missing bands which was assigned as an indicator of missing data by NTSYS-PC (Transue et al., 1994). Since, estimates of the similarity for different coefficient are different (Jaccard, 1908; Dice, 1945; Sokal and Michener, 1958), ranged of similarity coefficient were obtained as Simple Matching (0.54-0.94), Jaccard's (0.42-0.90) and Dice (0.57-0.95). The SM coefficient is the sum of shared positive bands plus the proportion of shared negative bands whereas, Jaccard's and Dice coefficient considered only the proportion of positive bands shared by samples and neglecting the proportion of shared negative bands. Jaccard's coefficient takes into consideration only those matches between the bands that are present (Jaccard, 1908) while Dice coefficient measures the proportion of bands shared as the result of being inherited from a common ancestor and represents the proportion of bands present and shared in both samples divided by the average of the proportion of bands present in each sample (Dice, 1945; Nei and Li, 1979).

In the present investigation, ISSR dominant markers were utilized for the genetic diversity analysis of acid lime landraces of eastern Nepal. For the dominant markers, it is generally assumed that each bands represents a different bi-allelic locus (Wiliams et al., 1990). The three most commonly used similarity coefficients are SM, J and D which differ in the amount of bias produced by the level of artifactual bands (Lamboy, 1994). SM coefficient seems appropriate when two alleles exist at a locus and one produces a band while the other doesn't and both alleles are present in the materials being compared (Dudley, 1994). Genetic variation analysis of accessions using ISSR marker based on SM coefficient is not well documented. Dice similarity coefficient had been

used for the phylogenetic analysis among selected Citrus germplasm accessions (Fang et al., 1998), phenetic relationship between *Citrus indica* and few commercially important *Citrus* spp. (Chibame et al., 2010) using ISSR markers. Jaccard's coefficient was used for polymorphism analysis in Indian wild orange and related wild species using ISSR markers (Kumar et al., 2010). Therefore, all these similarity coefficients used in present investigation has its own kind of significance according to the goodness of fit that they provide. So, comparison of all these similarity coefficient matrices with their phenotypic illustrations is essential for selection of competent one with proper goodness of fit which finally leads to unbiased assessment.

Original matrices were compared by applying Mantel test (Mantel, 1967) in the option of MXCOMP in NTSYS-PC ver. 2.21i program by implementing two similarity matrixes at a time. The obtained correlation value is then tabulated (Table 4.8). The correlation between J and SM as well as SM and Dice (0.98143 and 0.98318 respectively) were observed significantly low in comparison to the correlation between J and Dice i.e. 0.99710. The consensus fork index (CI_C), is a numerical value that indicates the degree to which the consensus tree is resolved, used to compare two rooted labeled trees at a time using CONSEN option in NTSYS-PC (Rohlf, 2009). Significantly high consensus fork index of CI_C = 1.00000 was observed for Jaccard's and Dice coefficient (Table 4.9). Again, for the conformation of competent one, cophenetic correlation matrix and similarity matrix on which it was based was compared using the same MXCOMP option. Dice similarity with UPGMA clustering method revealed the highest cophenetic correlation value of 0.90356 followed by Jaccard with 0.89800 and the smallest for Simple Matching with 0.88396 (Table 4.10). The cophenetic correlation can be used as a measure of goodness of fit for a cluster analysis (Rohlf, 2009). The degree of fit can be interpreted subjectively as follows:

Level	Interpretation
r ≥≥ 0.9	Very good fit
0.9 > r ≥≥ 0.8	Good fit
0.8 > r ≥≥ 0.7	Poor fit
r < 0.7	Very poor fit

Where, 'r' is the cophenetic correlation coefficient between clusters analyzed in accordance with the similarity matrix.

Therefore, from all the comparison carried out, Simple Matching coefficient has suitable value of 'r' (0.88396) for good fit but still low in comparison to Jaccard's and Dice coefficient. On the other hand, Dice coefficient has significantly high 'r' value (0.90356), which is considered to be very good fit of the UPGMA cluster pattern to the data. Hence,

Dice similarity was coefficient selected as the most efficient coefficient for the discrimination of genetic diversity study of acid lime accessions.

In the present investigation, each similarity coefficient matrices (SM, J and D) were employed for the construction of phenograms to represent the genetic relationship among the acid lime accessions under study using Unweighted Pair Group Method of Arithmetical Averages (UPGMA) with SAHN (sequential agglomerative hierarchical nested clustering) clustering module. SAHN clustering module is used for the constructing phenogram for similarity and dissimilarity matrices (Day and Edelsbrunner, 1984). Clustering was done to group individuals with similar characteristics in one group and those with diverse characteristics in the different group. The most commonly employed distance matrix algorithm is UPGMA for cluster analysis (Moodie et al., 1997; Weising et al., 2005). The overall topologies of three phenograms are more or less similar except for the SM tree in which position of accession LD-31 is altered. This may be due to different approaches for estimation of similarity coefficients (Fig 4.1, 4.2 and 4.3). As Dice similarity coefficient found to be most efficient coefficient, discussion of the present investigation will be based on this coefficient. The phenogram constructed using Dice coefficient shows clustering of accessions from different agro-ecological zones i.e. no separate cluster for different agro-ecological zones were observed. Sixty acid lime accessions were separated into two major and three minor clusters. This types of clustering patterns were observed among the acid lime accessions of different agro-ecological zone of eastern Nepal using SSR markers (Shrestha et al., 2012a). Cluster I comprises of accessions from High-hill and Mid-hill showing close genetic relationship, cluster II comprises of accessions from all the agro-ecological zone showing close genetic relationship in which the exotic varieties *viz*. Vanarasi, Madrasi and Rampur (LKv-60, LKm-61 and LKr-62 respectively) are clustered together. These varieties also clustered together on assessing genetic diversity of acid lime using SSR markers (Shrestha et al., 2012a). The cluster III and IV consist of single accessions so may be considered as minor clusters and cluster V consists of accessions (LS-56, LD-58 and LS-57) from Terai ecological zone. The LS-56, LD-58 and LS-57 also clustered together on assessing genetic diversity using SSR markers (Shrestha et al., 2012a) which indicated that they are closely related. The intermixing of accessions in the phenogram may be attributed to the distribution of diversity of acid lime genotypes in different agro-ecological domains. The genetic diversity among different genotypes may be attributed to gene flow as a result of cross-pollinating nature of acid lime. The cluster groups III, IV and V have high genetic distance because these accessions are morphologically different from other groups and were collected from the Terai (except LT-9 i.e. from high hill) zone (Shrestha et al., 2012a). Greater genetic diversity leads to greater productivity, greater nutrient

retention in ecosystem as well as ecosystem stability (Tilman, 2000). In order to improve varieties, the ideal parent for hybridization should have high level of diversity (Jeya Prakash et al., 2006). Acid lime is a cross-pollinated crop and large genetic variation exist among the accessions and these variations will be potential sources for the selection of elite genotypes for the breeding and variety development programs. The considerable high genetic diversity observed might be synergistic consequences of cross-pollination between acid lime landraces (Slatkin, 1987; McDermott and Mcdonald, 1993). The similarity percent ranged from 57% to 95% with an average of 81% were observed and acid lime accessions from Terai were found to have high genetic diversity in comparison with that of high-hill and mid-hill agro-ecological zone (Appendix 9). This result is comparable with the result obtained using SSR markers where average genetic similarity among these acid lime were found to be 77% ranging from 54% to 100% (Shrestha et al., 2012a). In addition, percentage similarity for acid lime landraces of different agro-ecological zones based on ISSR similarity matrix (Dice) was computed. Within and between agro-ecological zones showed wide range of similarity value in Terai agro-ecological zone (57-94%) in comparison with Mid-hill (70-94%) and High-hill (75-95%). The acid lime landrace from Terai and High-hill were found to be having an average of 82.68% similarity. Likewise, from High-hill and Mid-hill have 83.6%, and Terai and Mid-hill have 79.98% similarity. Among three agro-ecological zones, the accessions from High and Mid-hills were found to be closely related with a high average genetic similarity (86.02% and 82.94%) as compared to Terai (79.33%) (Table 4.11). The result is comparable with the result obtained using SSR markers where accessions from high and mid hills have high average genetic similarity (73% and 81%) in comparison to Terai (69%) (Shrestha et al., 2012a). Our results showed that, the accessions collected from different altitudinal range were not genetically distinct but high variation was observed in Terai accessions than High and Mid-hills. This may be due to the planting materials carried by the farmers from different hill districts with migration and introducing from neighboring country in Terai agro-ecological zone. While low level of genetic variability observed in mid hill and high hill is due to acid lime trees were established in natural conditions in these zones (Shrestha et al., 2012a).

3D-Plot of PCO was constructed with Eigen vector using NTSYS-PC ver. 2.21i for the 60 acid lime accessions (Fig 4.4) and Eigen analysis was done with Principal Co-ordinate Analysis (PCO) using MVSP ver. 3.22 (Fig 4.5). In both cases, Dice (or Nei and Li) similarity coefficient were employed. The pattern of genetic relationships of the acid lime accessions shown by UPGMA based SAHN clustering phenogram was similar with the PCO. Accessions from the different agro-ecological zones showed wide range of genetic diversity. The first two principal co-ordinates with eigen value was greater than 22.85%

of the total variance among the accessions. Therefore, PCO using first two principal co-ordinates provide a good grouping of the accessions in the co-ordain system confirming the relationship with the phenogram.

Further, direct assessment based on monomorphic and polymorphic bands amplified by 21 ISSR primers were also evaluated. This revealed high average percent polymorphism in accession of Terai (80.69) followed by Mid-hill (73.23) and High-hill (65.15). This shows more genetically diverse acid lime in Terai agro-ecological zone and less genetically diverse acid lime found in High-hill agro-ecological zone (Table 4.13). The result is in agreement with the one that is shown by the phenogram.

5.3.1.1. Genetic Differentiation Using Analysis of Molecular Variance (AMOVA) for Acid lime of agro-ecological zones

At present, the genetic structure of agro-ecological zones is investigated by analysis of variance instead of using the F_{ST} and G_{ST} estimators (Weir and Cockerham, 1984; Weising et al., 2005). The AMOVA (Analysis of Molecular Variance) was carried out using GenAlEx 6.5. It allows analysis for many types of genetic marker and offers statistical testing by random permutation, and often used with dominant marker data (Peakall and Smouse, 2012). It works with distances between individuals and calculates the variance between and within pre-defined groups, and allows one to test various particular genetic structures (Weising et al., 2005). Although the principles behind F_{ST} and G_{ST} on the one hand and the AMOVA on the other hand are rather different, they usually produce very similar estimates when applied to the same set of marker data (Nybom, 1991). It will describe how the marker variance is partitioned within and among populations (Stewart and Excoffier, 1996). ϕ_{PT}, an indicator of genetic differentiation among the three agro-ecological zones was found to be 0.139 and this result suggested genetic differentiation of 0.86 within acid lime agro-ecological zones. Breeding system, genetic drift or genetic isolation of agro-ecological zones can cause high level of genetic differentiation among plant species population (Hogbin and Peakall, 1999), and low genetic differentiation among agro-ecological zone suggests limited effects of such consequences. There was significant partitioning of the genetic variation ($p < 0.001$), with very high difference between the genetic variation occurring among agro-ecological zone (14%) and within agro-ecological zone (86%) (Table 4.14).

5.4. SSR Profiling and Genetic Diversity Assessment of *Citrus* spp.

Molecular markers are useful tools for assaying genetic variation and provide an efficient means to link phenotypic and genotypic variation (Varshney et al., 2005). The

SSR-PCR profiles of 45 citrus accessions belonging to different species from Dhankuta and Kathmandu, Nepal by the selected 12 SSR primers were observed to be reproducible and scorable for diversity analysis. The pairs of the markers used in this study amplified successfully all *Citrus* spp. which is in accordance with the earlier reported high level of conservation of SSR primers across *Citrus* spp. and related genera (Jarrell et al., 1992; Kijas et al., 1997; Corazza-Nunes et al., 2002; Ahamad et al., 2003). The amplified fragments using different SSR primers ranged from 80 bp to 270 bp (Table 4.15). The bands (alleles) amplified by these primer pairs showed 100% polymorphism. This indicates that microsatellite sequences on which the primers were designed are ubiquitous in the genome of the *Citrus* spp. of present study, and also that the sequences are highly variable among these species.

Replication slippage occurs during PCR amplification of simple sequence repeats in vitro leading to formation of stutter bands i.e. minor products that differ in size from the main product by multiples of the length of the repeat unit (Hauge and Litt, 1993; Murray et al., 1993). Quantitative experiments show that the *Taq* polymerase slippage rate increases with the number of repeat units and is inversely correlated with repeat unit length (Shinde et al., 2003). PCR-induced stutter bands have been observed by many microsatellite users; tetra-nucleotide repeat markers typically give fewer stutter bands than di-nucleotide and, in particular, mononucleotide repeats and they generally appear as products that are shorter than the size of allele being amplified (Ellegren, 2004). Therefore, in present investigation more tri-nucleotides repeat markers were used in comparison to di-nucleotide repeat markers. In our investigation, stutter bands were observed including few mis-amplification (product size more than even 1 kb) but the bands of size generally in range of 80-270 bp were considered as generally length of simple sequence repeats in *Citrus* ranges 100-300 bp (Golein et al., 2005; Barkley et al., 2006). SSR marker can be used to differentiate homozygous and heterozygous alleles between lines from the same origin (Shehata et al., 2009). This suggest the usefulness of SSR markers for molecular differentiation of plant species at the taxonomic level (Kenis and Keulemans, 2000).

5.4.1. Alleles per locus, Polymorphic Information Content and Probability of Identity

The average number of alleles per locus provides complementary information of polymorphism (IPGRI, 1991). In present investigation, 60 putative alleles were detected at 12 SSR loci in 45 samples varying from 2 (TAA 27) to 8 (TAA 52) with an average of 5 per locus. Therefore, the most polymorphic markers are TAA 52, CAT 01, TAA 3 and TAA 41 with 8, 6, 6 and 6 alleles respectively and least polymorphic marker is TAA 27 with 2

alleles. In literature, a total of 275 alleles were detected at 24 SSR loci with an average of 11.5 alleles per locus in the assessment of genetic diversity and population structure in a citrus germplasm collection using 24 SSR markers (Barkley et al., 2006). Likewise, a total of 52 putative alleles were detected at 7 SSR loci with an average of 7.42 alleles per locus in the assessment of genetic variability in some Iranian sweet oranges and mandarins (Golein et al., 2005). Maximum of 5 alleles amplified by CAT01, and TAA14 primer pairs while a minimum of two alleles amplified by TC26 primers in Satsuma orange (Ghanbari et al., 2009). Average of 6 alleles were observed in Mandarin and lemon accessions (Li et al., 2006), whereas 4.4 alleles per locus were observed in the diversity of *Citrus* spp. by SSR markers (Santos et al., 2003). Likewise 32 alleles were amplified by 9 SSR markers on Mandarin germplasm in Brazil (Cao et al., 2007).

The Probability of Identity (PI) provides an estimate of the average probability that two unrelated individuals, drawn from the same randomly mating population, will by chance have the same multilocus genotype. PI is widely used in DNA forensics as an indication of the statistical power of a specific set of marker loci (Peakall and Sydes, 1996; Peakall and Smouse, 2012). Comparison of SSR markers with regard to their information content (number of alleles and PI values) showed that most informative loci for the present investigated set of species were loci TAA 41 (6 alleles and PI value 0.075) and TAA 52 (8 alleles and PI value 0.089), where as less informative one was TAA 27 (2 alleles and PI value 0.385).

The data on microsatellite loci were used for PIC value calculation for the examination of SSR primers and extent of information on diversity. PIC provides an estimation of discriminatory power of loci by taking into account the number of alleles and their relative allele frequencies (Ogunkanmi et al., 2008). The high discrimination power of SSR marker is essential for variations analysis in gene pool of crops and ability to distinguish between closely related individuals (Ogunkanmi et al., 2008). In the present investigation, the PIC value ranged from 0.497 (Primer TAA 27) to 0.802 (Primer TAA 41) with an average of 0.662. The PIC values of the primers TAA 41 and TAA 52 are 0.802 and 0.785 respectively, which was considered as highly polymorphic loci according to the classification proposed by Botstein et al. (1980). So, above mentioned primers are more useful than other remaining primers for genetic variability study of Nepalese *Citrus* spp. and their cultivars. These results support the inference obtained from the PI (Probability of Identity) values i.e. those primers that has highest PIC values, has lowest PI value (Table 4.15). These PIC value can be compared with those reported in earlier studies. In the study carried out by Golein et al. (2005), PIC value ranged from 0.505 to 0.950 with an average of 0.819 while in the research carried out by Barkley et al. (2006),

average PIC value was found to be 0.625 which is comparable with our results since the former research focus on sweet orange and mandarin only while the later one encompasses maximum *Citrus* spp. as in our experiments. Likewise, PIC value ranging from 0 to 0.59 (Mean = 0.19) in citrus and its related genera (Weising and Gardner, 1999). Similarly, PIC value of SSRs primer ranged from 0.49 to 0.73 (Mean = 0.61) in *Citrus* spp. (Bryan et al., 1999).

5.4.2. Allele frequency, Observed and Expected Heterozygosity, Shannon's Information Index and Fixation Index

These parameters were calculated using GenAlEx ver. 6.5, and results were tallied results obtained from PowerMarker ver. 3.25. The same result obtained indicates the reliability and reproducibility of the results.

Allele frequency is the frequency at which alleles of interest are found at any locus, and is used to estimate frequency of given genetic profile and also used to characterize the genetic diversity or richness of the gene pool in a population (NFSTC, 2010). In the present investigation, variable allele frequencies at each locus were observed (Table 4.17). This variation in allele frequency can be attributed to genetic drift and reproductive isolation (Bokoume et al., 2007). The informativeness of a genetic marker is also measured by allele frequencies. The evenness of allele or genotype frequencies is accounted for by the measures of average observed heterozygosity, expected heterozygosity, and effective number of alleles (Mohammadi and Prasanna, 2003). Effective numbers of alleles are the number of alleles that can be present in a population and tells about the number of alleles that would be expected in a locus in each population (IPGRI, 2003). The effective number of alleles ranged from 1.850 to 4.844 with an average of 3.009.

Heterozygosity refers to proportion of individuals in a population that carry two different alleles at a locus and is a measure of genetic variability (Ellegren, 2004). It is important to both natural and cultured populations as it provides a large spectrum of genotypes for adaptive response to changing conditions, and heterozygous individuals are usually superior to less heterozygous individuals in many economically important characteristics like growth, fertility and disease resistance (Jannati et al., 2009). The measure of the amount of heterozygosity across loci can be used as a general indicator of the amount of genetic variability (NFSTC, 2010). In present investigation, observer heterozygosity (Ho) and expected heterozygosity (He) were computed for each individual marker as a measure of marker diversity (Table 4.16). Heterozygosity (H) is a

widely used measure of the allelic diversity or informativeness of a genetic marker. The informativeness of a genetic marker increases as H increases. The present investigation also substantiated these facts i.e. those primers that has highest PIC values and lowest PI values also has highest heterozygosity (He) (Table 4.16). A locus is considered a polymorphic marker if its heterozygosity, H = 0.1 and a locus is considered highly polymorphic if H = 0.7 (Lui, 1998). The information content of SSRs, as measured by expected heterozygosity, is a function of the number of alleles and the frequency of each allele in the population under study (Taramino and Tingey, 1996). The average He over all loci is an estimate of the extent of genetic variability in the population which is 0.638 in present investigation (Table 4.16). The difference in He and Ho may be due to sampling drift, out-crossing (cross-pollination) mode of reproduction in citrus and their maintenance by vegetative propagation (Sefc et al., 1999). The hypervariability observed at SSR loci was expected because of the unique mechanism by which this variation is generated: replication slippage is thought to occur more frequently than single nucleotide mutations and insertion/deletion events, which generate the polymorphisms detectable by AFLP and RAPD analyses (Powell et al., 1996; Milbourne et al., 1997). The codominant nature of these markers permits the detection of a high number of alleles per locus and contributes to higher levels of expected heterozygosity being reached (Belaj et al., 2003).

Shannon's Information Indices have been widely employed in ecology but largely overlooked in genetics. The performance, power and theoretical expectation of Shannon's indices for estimating genetic diversity were assessed and concluded that the Shannon's information framework offers an alternative method of quantifying biological diversity across multiple scales (genes to landscapes) (Sherwin, 2006; Peakall and Smouse, 2012). The Shannon's Information Index (I) value ranged from 0.677 to 1.670 with an average of 1.226 in present investigation, suggesting high level of diversity in the samples analyzed. The advantages of this index are that it takes account of the number of alleles and evenness of the alleles, and is increased either by having additional unique alleles, or by having greater allele's evenness. The obtained values of Shannon's index in present investigation supports the considerable variation of the studied accessions at 12 SSR loci analyzed.

Fixation index (F) describes the reduction in heterozygosity within populations relative to the total population due to selection or drift (Lowe et al., 2004). The mean value of F found to be 0.446, which supports the heterozygostiy values obtained (Table 4.6). Heterozygosity of genotype is a key component for genetic mapping on F1 progenies

and based on a unique F1 progeny obtained from a cross between two heterozygous genotypes, it is possible to develop a genetic map for each parent (Luro et al., 2008).

5.4.3. Analysis of Genetic Diversity using NTSYS and MVSP Statistical Packages for *Citrus* spp.

Phenetic and cladistic are the two main approaches used to infer taxonomic relationships in plants (Sneath and Sokal, 1973; FAO, 1998). The numerical taxonomic system is the phenetic approach to infer taxonomic relationship. A phenetic approach using UPGMA distance method was used to generate a tree from a distance matrix. The UPGMA algorithm computes the average similarity or dissimilarity of a candidate operational taxonomic unit (OTU) to an extant cluster giving equal weight to each OUT in that cluster, disregarding its structural sub-division (Sneath and Sokal, 1973). Using SSR-PCR data, similarity matrices has been generated by pair-wise comparisons of each species (Appendix 10, 11 and 12) and cluster analysis performed to generate phenograms (Fig 4.6, 4.7 and 4.8) to show the genetic relationships between various *Citrus* spp. under study. Among the three similarity matrices i.e. SM, J and Dice matrix and three phenograms generated using these matrices, the best one was selected using the same criteria as discussed in the analysis of acid lime diversity using ISSR markers. In this SSR-*Citrus* spp. investigation, Jaccard's similarity with UPGMA clustering method revealed the highest cophenetic correlation value of 0.87407 followed by Dice with 0.83151 and the lowest for Simple Matching with 0.82900. Therefore, Jaccard's coefficient has significantly high 'r' value (0.87407) which is considered to be good fit of the UPGMA cluster pattern to the data (Rohlf, 2009). Hence, Jaccard's coefficient was selected and the phenogram developed from this coefficient was utilized for further diversity analysis and used as most efficient coefficient for the discrimination of genetic diversity study of *Citrus* spp.

The present study is the first attempt to evaluate genetic variation patterns in Nepalese Citrus spp. using SSR as molecular marker techniques. In any germplasm centre like NCRP (National Citrus Research Program), Dhankuta, the knowledge of degree of similarity or difference among individual genotypes in an accession or among accessions is essential for relatedness analysis. Similarly, structure analysis needs peer knowledge of amount of genetic variation present and its partitioning among accessions and individual genotypes is required (Ayad et al., 1997). Cluster analysis was performed by the distance-based method—SAHN (sequential agglomerative hierarchical nested clustering), where series of successive mergers were used to group individuals. Assessment of genetic diversity was done using Jaccard's coefficient and phenogram.

Because of the nature of cluster formation, a phenogram does not show the affinity interrelationship of its component OTU's (Barett and Rhodes, 1976). The similarity percentages varied from 10% to 100% which is an indication of high level of genetic variation among the *Citrus* spp*.* and their cultivars (Appendix 11). Citrus taxonomy and phylogeny are very complicated, controversial and confusing, mainly due to sexual compatibility between Citrus and related genera, the high frequency of bud mutations and the long history of cultivation and wide dispersion (Nicolosi et al., 2000). The present investigation provides few supports for the taxonomy and phylogeny.

The phenogram grouped the 45 citrus accessions belonging to different species into four distinct clusters (two major and two minor). All the mandarins, sweet oranges, sour oranges, grapefruits, pummelos, unknown varieties and natural hybrids are grouped in cluster I. The Lemons, acid limes and citrons are grouped in cluster II where as root stock *Citrus* (citranze and trifoliate) and kumquats form cluster III and cluster IV respectively. The grouping of the accessions in cluster I and cluster II supports the results obtained in the numerical taxonomy study of affinity relationship in cultivated citrus and its clones where *C. aurantium*, *C. grandis*, *C. paradisi*, *C. reticulata*, and *C. sinensis* forms a larger group and *C. aurantifolia*, *C. limon*, and *C. medica* forms a smaller group (Barett and Rhodes, 1976). These two clusters shared 23% similarity. Mandarin was reported as most phenotypically and genotypically heterogenous (Moore, 2001) and was proposed as one of the 'three ancestral species' in cultivated citrus (Barett and Rhodes, 1976). Swingle and Reece (1967) classified all the mandarins as one species, where as Tanaka (1977) classified them into 36 species. Barkley et al. (2006) reported that mandarins were most polymorphic among the ancestral species. The AFLP, SAMPL, SSAP and SSR data showed that mandarin accession were scattered in the phenogram (Biswas et al., 2011). In present investigation, SSR data showed mandarins were scattered within the cluster group I in the phenogram, indicating they might be more divergent at the molecular level. The divergence within the mandarin group has been reported by many researchers (Koehler-Santos et al., 2003; Campos et al., 2005). Sweet orange originated as a natural hybrid of mandarin and pummelo and most of their varieties are obtained by somatic mutation from one ancestor tree. So, despite of differences in morphological characters, genetic variation of sweet orange was low (Barett and Rhodes, 1976; Fang et al., 1997; Novelli et al., 2006; Uzun, 2009). The present investigation also supports this facts and showed high level of genetic similarity in sweet oranges. In other study, it was found identical microsatellite profiles at 9 out of 10 SSR loci among analyzed orange cultivars and clones (Hvarleva et al., 2008). Sweet orange has a majority of its genetic makeup from mandarin and only a small proportion from pummelo (Barkley et al., 2006). In our investigation, majority of sweet oranges accessions clustered with

mandarin than pummelo which supports the results. Tangelo is a hybrid derived from mandarin X pummelo and Tagore from mandarin X sweet orange (Uzun and Yesiloglu, 2012), which gets clustered as expected in present investigation. Citron was one of the progenitor of lemons (Nicolosi et al., 2000; Gulsen and Roose, 2001). It was reported that lemons were thought to be natural hybrids of a citron and a lime (Scora, 1975; Barett and Rhodes, 1976) or a hybrid of citron and sour orange (Nicolosi et al., 2000; Gulsen and Roose, 2001). This result is supported by our findings as lemons get clustered with citrons and limes. Limes are apparently hybrids of citrons and papedas (Scora, 1975) or a tri-hybrid cross of citron, pummelo, and Microcitrus, and had the highest observed heterozygosity of all the taxonomic groups (Barett and Rhodes, 1976; Nicolosi et al., 2000; Barkley et al., 2006). The current investigation substantiated above facts as shown in phenogram. Kumquat (*Fortunella margarita* and *F. hindsii*) and rootstock citrus (*Poncirus trifoliata* and Old Citranze) formed separate clusters from the rest of the species supporting they are from different genus. The natural hybrid of unknown variety and khoku mandarin is 72.2% genetically similar which suggest khoku mandarin as it one of progenitor. The natural hybrid (variety local hokse) gets clustered with sweet oranges (delicious seedless and Valencia late varieties) suggesting these sweet oranges as its possible parent. Clustering of unknown *Citrus* spp. with the mandarin (Khoku and Unshiu) in cluster I suggest these mandarins (Unshiu has more contribution) as their probable parents (Fig 4.7).

Eigen analysis was done with Principal Co-ordinate Analysis (PCO) using MVSP 3.21 for all the 45 citrus accessions belonging to different species (Fig 4.11). The pattern of genetic relationships among *Citrus* spp*.* shown by PCO substantiated similar pattern as shown by the UPGMA based SAHN clustering phenogram. The first two principal co-ordinates with eigen value was greater than 31.79% of the total variance among the accessions. Therefore, PCO using first two principal co-ordinates provide a good grouping of the accessions in the co-ordain system confirming the relationship with the phenogram.

CHAPTER 6: CONCLUSION

Nepal offers a diversified geo-climatic conditions and the natural resource base for low cost production of a wide variety of horticultural commodities of which citrus is one of the major fruit crops in terms of nutritional and economic value. However, Significant national requirement of citrus fruits are fulfilled from India even in seasonal period because of low production, narrow production period and poor infrastructure facilities for storage. Besides these, production and quality of citrus fruits are decreasing every generation due to high level of genetic erosion which may ultimately result in the loss of the original genetic characters. Another problem in *Citrus* spp. is apomixis that is responsible for maintenance of same genetic constitution in hybrid plants as that of their parental plants. In this context, information about the genetic diversity has promising application to tackle these problems. Genetic diversity assessment based on molecular marker tools finds its application in plant breeding for the production of elite genotypes in terms of disease resistance, drought tolerance, off season production as well as for enhancing other qualitative and quantitative traits. However, for latter purposes, molecular marker have to be integrated with MAS and Marker Assisted Introgression, gene pyramiding and linkage analysis. The generated information will be helpful for plant breeders for selection of the parental genotypes for hybridization followed by subsequent breeding and variety development programs.

In the first part of the present investigation, ISSR-PCR reactions and cycling conditions for acid lime landraces at different agro-ecological zone of Nepal were optimized and utilized for ISSR profiling of 60 acid lime landraces under study. Genetic diversity analysis were then performed on ISSR data using NTSYS-PC, GenAlEx and MVSP statistical packages, and found high level of diversity in acid lime landraces of Terai agro-ecological zone in comparison to other zones. Among the tested ISSR primers, Primers C1, C2, C4, C5, C7, C8, C9, UBC 810, UBC 812, UBC 825, UBC 834, UBC 835, UBC 836, UBC 841, UBC 842, UBC 857, UBC 873, UBC 888, UBC 889 gave PIC value greater than 0.80 suggesting that they are highly useful for analysis of genetic diversity of acid lime accessions. Prior to this molecular study, genetic diversity assessment of acid lime landraces of different agro-ecological zones of eastern Nepal based on SSR markers has been carried out. However, use of genome based ISSR marker for the assessment of genetic diversity of acid lime landraces is the first initiative in the Horticulture field of Nepal. The result of this investigation will provide genetic information about intraspecific diversity in acid lime and help in identification of parental genotypes for crossing and development of new elite cultivars through future breeding programs.

Although conventional approaches of morphological, fruit characterization, consumer's preference and biochemical analysis aided in the documentation of genetic diversity of

Nepalese *Citrus* spp. and their cultivars, DNA based molecular markers adds a new and reliable dimension to cultivar identification, germplasm characterization and genetic diversity assessment within and between species. In the second part of the present investigation, SSR-PCR reactions and cycling conditions for *Citrus* spp. of Nepal were optimized and utilized for SSR-PCR amplification and profiling of 45 *Citrus* accessions belonging to different species. Genetic diversity analyses on SSR data were thus performed using NTSYS-PC, GenAlEx, Powermarker and MVSP statistical packages. Twelve's SSR loci were employed for determination of the level and organization of genetic diversity and elucidating relationships among *Citrus* spp. and found an average of 30% similarity among them. All the 12 SSR primers amplified the tested samples of Citrus spp. and this result showed conservation of tested microsatellite loci among the *Citrus* spp. These SSR markers showed moderate level of heterozygosity (average He, 0.638) in Citrus spp. under study. As *Citrus* spp. is known to hybridize among themselves without much difficulty, hybridization seems to be one of the major sources of variation in these *Citrus* spp. However, other sources of variations (such as mutations, polyploidy, apomixis etc.) should also be considered while dealing with genetic diversity. Among the tested SSR primer pairs, TAA 45, CAT 01, TAA 52 and TAA 41 had higher PIC value and were found to be more informative for diversity study in *Citrus* spp. The present investigation on the genetic diversity assessment of *Citrus* spp. of Nepal based on SSR marker is the first initiative in the field of Horticulture of Nepal. Our investigation showed that SSR markers are powerful tools that can differentiate closely related *Citrus* spp., provide information about interspecific, intraspecific and even intergeneric diversity estimates of *Citrus spp.* and also can pave the way to tackle the long standing problem of *Citrus* classification and identification in Nepal. Besides these, this investigation will also help in future breeding programs for the development of hybrids with quality traits such as disease resistance, high shelf-life and enhancing other quantitative traits such as higher yield, higher vitamin C content etc. for the development of improved varieties of various *Citrus* spp.

Hence, SSR and ISSR markers which are recent advances in molecular biology were found to be the powerful tools for evaluation of genetic variation in *Citrus* germplasm of Nepal including acid lime landraces. The findings of this investigation have opened up avenues for the academic research in the fields of germplasm characterization, classification, breeding, sustainable utilization, documentation and conservation in Nepal.

Recommendations

1. Though the ISSR markers proved to be useful for genetic diversity assessment of acid lime landraces, use of other molecular marker techniques such as RAPD, CAPS, AFLP, SCAR and SNPs should also be considered for finer and comparative molecular analysis and to solve the discrepancies left unsolved by ISSRs.
2. ISSR markers can be correlated with different qualitative and quantitative traits of acid lime and can be used to develop genetic maps.
3. The present investigation has provided overall information on genetic diversity and relationships of citrus germplasms of Nepal. Based on this relationship, future plant breeding experiments aiming at production of elite cultivar of various citrus crops could be executed.
4. The result can be further used to manipulate genetic determinants of horticulturally important traits and to characterize basis of productivity of citrus.
5. The results of this investigation may be useful to tackle the long standing problems of citrus classification and identification in Nepal.
6. Genetic variability either occurs naturally in various populations or can be induced by scientific interventions. Genetic variabilities in germplasm collections can be introduced by various means *viz*. conventional breeding, mutation breeding, molecular breeding and biotechnology (somaclonal variation). Present investigations attempted to study naturally occurring genetic variation in acid lime accessions and various *Citrus* spp. using two main marker systems. The information generated from these investigations would be useful to commercial plant breeders for the development of elite cultivars of various *Citrus* spp. as well as for the *ex situ* and *in situ* gene bank managers. Therefore, it is recommended that the findings of this research to be cited by concerned scientific personnel.

CHAPTER 7: REFERENCES

Adamblondon AF, Sevignac M, Dron M and Bannerot H (1994) A genetic-map of common bean to localize specific resistance genes against anthracnose. *Genome*. **37**: 915–924

Ahamad R, Struss D and Southwick SM (2003) Development and characterization of microsatellite markers in citrus. *j Am Soc Hort Sci*. **128**: 584–590

Ajibade SR, Weeden NF and Chite SM (2000) Inter-simple sequence repeat analysis of genetic relationships in the genus Vigna. *Euphytica*. **111**: 47–55

Akagi HY, A. Y, Inagaki A, Nakamura and Fujimura T (1996) A co-dominant DNA marker closely linked to the rice nuclear restorer gene, Rf-1, identified with inter-SSR fingerprinting. *Genome*. **39**: 1205–1209

Akkaya MS, Bhagwat AA and Cregan PB (1992) Length polymorphisms of Simple Sequence Repeat DNA in soybean. *Genetics*. **132**: 1131–1139

Alvarez I and Wendel JF (2003) Ribosomal ITS sequences and plant phylogenetic inference. *Molecular phylogentic Evolution*. **29**: 417-434

Angela K, Peter G, Isaac. and David SI (1998) Molecular Tools for Screening and Biodiversity of Plants and Animals. (Chapman and Hall. pp. Pages

Anonymous (2008) Food Security in Mountains: High Time for Action. Convention of Biological Diversity Montreal, *UNEP*.

Anonymous. (2008a) Monoculture and the Irish Potato famine: Understanding Evolution [Online] <http://en.wikipedia.org/wiki/Genetic_diversity>

APP (1995) *Agriculture perspective plan (APP)*, Agricultural Project Service Center and John Mellor Associates) pp.

Arcade A, Anselin F, Rampant PF, Lesage MC, Paques LE and Prat D (2000) Application of AFLP, RAPD and ISSR markers to genetic mapping of European and Japanese larch. *Theor Appl Genet*. **100**: 299–307

Armour JAL, Alegre SA, Miles S, Williams LJ and Badge RM (1999) *Minisatellites and mutation processes in tandemly repetitive DNA In: Goldstein DB, Schlotterer C (eds) Microsatellites: evolution and applications.*, Oxford University Press) pp.

Arya V, Yadav S and Yadav JP (2011) Intra-specific Genetic Diversity of different Accessions if cassia occidentalis by RAPD Markers. *Genet Eng Biotechnol J*. **22**: 1-8

Aubert B, Tontyaparn S and Suwan DB (1990) *"Rehabilitation of citrus industry" in the Asian Pacific Regions*, Proceedings of 4th International Asia Pacific Conference on Citrus Rehabilation) pp.

Ayad WG, Hodgkin T, Jaradat A and Rao VR. (1997) Molecular genetic techniques for plant genetic resources, Report of an IPGRI Workshop, 9-11 October, 1995. Rome, Italy. pp 3-79

Ayala EJ and Kiger JA (1984) *Modern Genetics*, Renjamin-Cumings) pp.

Azizi A, Wagner C, Honermeier B and Friedt W (2009) intraspecific diversity and relationship between subspecies of *Origanum vulgare* revealed by comparative AFLP and SAMPL marker analysis. *Plant Syst Evol.*: 10-19

Baldwin BG, Sanderson MJ, Porter JM, Wojciechowski MF, Campbell CS and Donoghue MJ (1995) The ITS region of nuclear ribosomal DNA: A valuable evidence on angiosperm phylogeny. *Annals of the Missouri Botanical Garden*. **82**(2): 247-277

Barcaccia G, Alberini E and Rosellini D (2000) Inheritance and mapping of 2n-egg production in diploid alfalfa. *Genome*. **43**: 528-537

Barett HC and Rhodes AM (1976) A numerical taxonomic study of affinity relationship in cultivated Citrus and its clone relatives. *Syst Bot*. **1**: 105-136

Barkley NA, Dean RE, Pittman RN, Wang ML, Holbrook CC and Pederson GA (2007) Genetic diversity of cultivated and wild-type peanuts evaluated with M13-tailed SSR markers and sequencing. *Genet Res*. **89**: 93-106

Barkley NA, Newman ML, Wang ML, Hotchkiss MW and Pederson GA (2005) Assessment of the genetic diversity and phylogenetic relationships of a temperate bamboo collection by using transferred EST-SSR markers. *Genome*. **48**: 731-737

Barkley NA, Roose ML, Krueger RR and Federici CT (2006) Assessing genetic diversity and population structure in a Citrus germplasm collection utilizing simple sequence repeat markers (SSRs). *Theor Appl Genet*. **112**: 1519-1531

Barreneche T, Casasoli M, Russell K, Akkak A, Meddour H, Plomion C, Villani F and Kremer A (2004) Comparative mapping between Quercus and Castanea using simple-sequence repeats (SSRs). *Theor Appl Genet*. **108**: 558–566

Becker J and Heun M (1995) Mapping of digested and undigested random amplified microsatellite polymorphisms in barley. *Genome*. **38**: 991–998

Beckmann JS and Soller M (1990) Toward a unified approach to genetic mapping of eukaryotes based on sequence tagged microsatellite sites. *Nat Biotechnol*. **8**: 930-932

Belaj A, Satovic Z, Cipriani G, Baldoni L, Testolin R, Rallo L and Trujillo I (2003) Comparative study of the discriminating capacity of RAPD, AFLP and SSR markers and of their effectiveness in establishing genetic relationships in olive. *Theor Appl Genet*. **107**: 763-744

Bernet GP, Mestre PF, Pina JA and Asins MJ (2004) Molecular Discrimination of Lemon Cultivars. *Hortscience*. **39**(1): 165-169

Bhattacharya SC and Dutta S (1996) Classification of Citrus fruits of Assam. *Indian J Genet*. **11**: 57-62

Biswas MK, Chai L, Amar MH, Zhang X and Deng X (2011) Comparative analysis of genetic diversity in Citrus germplasm collection using AFLP, SSAP, SAMPL and SSR markers. *Sci Hortic*. **129**: 798-803

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 sativaL). *Theor Appl Genet*. **98**: 780–792

Bokoume C, Wickneswari R, Rajanaidu N, Khushairi A, Amblard PH and Billotte N (2007) Allelic diversity of natural oil palm (*Elaeis guineensis* Jacq.) populations detected by microsatellite markers: Implication for conservation. *Plant Genet Resour*. **5**(2): 104-107

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

Bretó MP, Ruiz C, Pina JA and Asíns MJ (2001) The diversification of Citrus clementina Hort. ex Tan., a vegetatively propagated crop species. *Mol Phyl and Evol*. **21**: 285-293

Bryan GJ, McNicoll J, Ramsay G, Meyar RC and Jong WSD (1999) Polymorphic Simple Sequence Repeatmarker in Chloroplast Genome of Solanaceious Plant. *Theor Appl Genet*. **99**(5): 859-867

Caetano-Anolles G, Bassam BJ and Gresshoff PM (1991) DNA amplification fingerprinting using very short arbitrary primers. *Biotechnol*. **9**: 553-557

Cameron SW and Frost HB (1968) *Genetic, breeding and nucellar embryony*, 2) pp. 325–89

Campos ET, Espinosa MAG, Warburton ML, Varela AS and Monter AV (2005) Characterization of Mandarin (*Citrus* spp.) using morphological and AFLP markers. *Interciencia*. **30**(11): 687-693

Cao QQ, Meng HJ, Wen XP, Yi HL and Deng XX (2007) Genetic Diversity of Male Sterile and Low Fertility Germplasm of Citrus Revealed Using SSR Markers. *Chinese J Agric Biotechnol*. **4**(2): 99-104

Capparelli R, Viscardi M, Amoroso MG, Blaiotta G and Bianco M (2004) Inter- simple sequence repeat markers and flow cytometry for the characterization of closely related Citrus limon germplasms. *Biotechnol Lett*. **26**: 1295–1299

Cha RS and Thilly WG (1993) Specificity, efficiency and fidelity of PCR. *PCR Meth Appl*. **3**: 18-29

Chambers GK and MacAvoy ES (2000) Microsatellites: consensus and controversy. *Comparative Biochem Physiol*. **126**(B): 455-476

Charters YM and Wilkinson MJ (2000) The use of self-pollinated progenies as 'in-groups' for the genetic characterization of cocoa germplasm. *Theor Appl Genet*. **100**: 160–166

Chawla HS (2002) *Introduciton to Plant Biotechnology*, Oxford and IBH Publishing Co. Pvt. Ltd.) pp. 329-358

Chen Y, Hausner G, Kenaschuk E, Procunier D, Dribnenki P and Penner G (1998) Identification of microspore-derived plants in anther culture of flax (Linum usitatissimum L.) using molecular markers. *Plant Cell Reports*. **18**: 44–48

Chibame K, Marak. and Laskar MA (2010) Analysis of phenetic relationship between Citrus indica Tanaka and a few commercially important citrus species by ISSR markers. *Sci Hortic*. **124**: 345–348

Condit R and Hubbell SP (1991) Abundance and DNA sequence of two-base repeat regions in tropical tree genomes. *Genome*. **34**: 66-71

Cooke RJ and Reeves JC (1998) New approach to potato variety identification. *Potato Res*. **42**: 529-539

Corazza-Nunes MJ, Machado MA, Nunes WMC, Cristofani M and Targon MLPN (2002) Assessment of genetic variability in grapefruits (C. paradise Macf.) and pummelos (C. maxima (Burm.) Merr.) using RAPD and SSRs markers. *Euphytica*. **126**: 169-76

Cox JE (1974) Citrus. *Proc Fla Sta Hort Soc*. **74**: 95-104

Cuc LM, Mace ES, Crouch JH, Quang VD, Long TD and Varshney RK (2008) Isolation and characterization of novel microsatellite markers and their application for diversity assessment in cultivated groundnut (Arachis hypogaea). *BMC Plant Biol*. **8**: 55

Day WH and Edelsbrunner H (1984) Efficient algorithm for agglomerative hierarchical clustering methods. *J Classif*. **1**(1): 7-24

de Rosas ARP, Segura EL, Fichera L and Garcia BA (2008) Macrogeographic and microgeographic genetic structure of the Chagas' disease vector Triatoma

infestans (Hemiptera: Reduviidae) from Catamarca, Argentina. *Genetica*. **133**: 247-260

Dhakal DD and Bhattarai S (2002) Production system of lime and lemon in Nepal: A survey report. Proceedings of Workshop on 'Production and Marketing of Lime and Lemon in Nepal' organized by IAAS. Rampur at Lamjung campus 26-28

Dice LR (1945) Measures of the amount of ecologuc association between species. *Ecology*. **26**: 297-302

Dorji K and Yapwattanapun C (2011) Assessment of morphological diversity for local mandarin (Citrus reticulata Blanco.) accessions in Bhutan. *Journal of Agriculture Technology*. **7**(2): 485-495

Dudley JW (1994) *Comparison of genetic distance estimators using molecular marker data* American Scoiety of Horticultural Science) pp. 3-7

Dugo G and Di Giacomo A (2002) *Citrus: The Genus Citrus, Medicinal and Aromatic Plants-Industrial Profiles*, Taylor & Francis group) pp.

Edwards A, Civitello A, Hammond HA and Caskey CT (1991) DNA typing and genetic-mapping with trimeric and tetrameric tandem repeats. *Am J Hum Genet*. **49**: 746–756

Ellegren H (2004) Microsatellite: Simple Sequences with Complex Evolution. *Nat Rev Genet*. **5**: 435-445

Excoffier L, Smouse PE and Quattro JM (1992) Analysis of molecular variance inferred from metric distance among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*. **131**: 479-491

Exeler N, Kratochwil A and Hochkirch A (2008) trong genetic exchange among populations of a specialist bee, Andrena vaga (Hymenoptera: Andrenidae). *Conserv Genet*. **9**: 1233-1241

Fang DQ, Krueger RR and Roose ML (1998) Phylogenetic relationships among selected Citrus germplasm accessions revealed by inter-simple sequence repeat (ISSR) markers. *J Am Soc Hort*. **123**(4): 612–617

Fang DQ and Roose ML (1997) Identification of closely related citrus cultivars with inter-simple sequence repeat markers. *Theor Appl Genet*. **95**: 408–417

Fang DQ, Roose ML, Krueger RR and Federici CT (1997) Fingerprinting trifoliate orange germplasm accessions with isozymes, RFLPs and inter-simple sequence repeat markers. *Theor Appl Genet*. **95**: 211–219

FAO. (1998) Plant Protection Bulletin, Annual Report of FAO. Malaysia. pp

FAO. (2002) Plant Protection Bulletin, Annual Report of FAO. Malaysia. pp

Federici CT, Fang DQ, Scora RW and Roose ML (1998) Phylogenetic relationships within the genus Citrus (Rutaceae) and related genera as revealed by RFLP and RAPD analysis. *Theor Appl Genet*. **96**(6): 812-822

Ferdinandez YSN, Somers DJ and Coulman BE (2001) Estimation of the genetic relationship of hybrid bromegrass to smooth bromegrass and meadow bromegrass using RAPD markers. *Plant Breeding*. **120**: 149-153

Fraser LG, Tsang GK, Datson PM, De Silva HN, Harvey CF, Gill GP, Crowhurst RN and McNeilage MA (2009) A gene rich linkage map in the dioecious species Actinidia chinensis (kiwifruit) reveals putative X/Y sex-determining chromosomes. *BMC Genomics*. **10**: 102

Gelfland DH (1989) *Taq DNA polymerase*, Pub Stockton Press) pp. 17-22

Ghanbari A, Jelodar NB and Rahiman H (2009) Studying of Genetic Diversity in Satsuma (*Citrus unshiu*) Mandarin utilizing Microsatellite Markers. *Int J Agric Res*. **4**: 88-96

Gilbert JE, Lewis RV, Wilkinson MJ and Caligari PDS (1999) Developing an appropriate strategy to assess genetic variability in plant germplasm collections. *Theor Appl Genet.* **98**: 1125–1131

Gmitter F and Hu X (1990) The possible role of Yunnan, China, in the origin of contemporary *Citrus* species (Rutaceae). *Econ Bot*. **44**(2): 267-277

Gmitter FG, Grosser JW and Moore AG (1992) *Citrus Biotechnology in Agriculture and Forestry*, 8 (CAB Int) pp. 335–369

Goldstein DB and Clark AG (1995) Microsatellite variation in North American populations of Drosophila melanogaster. *Nucleic Acids Res*. **23**: 3882-3886

Goldstein DB and Pollock DD (1997) Launching microsatelites: a review of mutation processes and methods of phylogenetic inference. *J Hered*. **88**: 335-342

Golein B, Talaie A, Zamani Z, Ebadi A and Behjatnia A (2005) Assessment of genetic variability in some Iranian sweet oranges (*Citrus sinensis* [L.] Osbeck) and Mandarins (*Citrus reticulata* Blanco) using SSR markers. *Int J Agric Biol*. **2**: 167–170

Green RM, Vardi A and Galun E (1986) The plastome of Citrus cultivars and species and comparison with related genera. *Theor Appl Genet*. **72**: 170–177

Gulsen O and Roose ML (2001) Lemons: diversity and relationship with selected Citrus genotypes as measured with nuclear genome markers. *J Amer Soc Hort Sci*. **126**: 309–327

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. *Theor Appl Genet*. **89**: 998–1006

Gur-Arie R, Cohen CJ, Eitan Y, Shelef L, Hallerman EM and Kashi Y (2000) Simple sequence repeats in Escherichia coli: abundance, distribution, composition, and polymorphism. *Genome Res*. **10**: 62-71

Hair JHJ, Anderson RE, Tatham RL and Black WC (1995) *Multivariate Data Analysis*, Prentice Hall) pp.

Hamrick JL and Godt MJW (1996) *Conservation Genetics of Endangered Plant species*, Pub Chapman & Hall) pp. 281-304

Hancock JM (1999) *Microsatellites and other simple sequences: genomic context and mutational mechanisms*, Oxford University Press) pp.

Hantula J, Dusabenyagasani M and Hamelin RC (1996) Random amplified microsatellites (RAMS) a novel method for characterizing genetic variation within fungi. *Eur J Path*. **26**: 159–166

Haque I, Bandopadhyay R and Mukhopadhyay Y (2009) Intraspecific variation in Commiphora wightii population based on internal transcribed spacer (ITS1-5.8s-ITS2) sequences of rDNA. *Diversity*. **1**: 89-101

Hauge XY and Litt M (1993) A study of the origin of 'shadow bands' seen when typing dinucleotide repeat polymorphisms by the PCR. *Hum Mol Genet*. **2**: 411-415

He C, Poysa V and Yu K (2003) Development and characterization of simple sequence repeat (SSR) markers and their use in determining relationships among Lycopersicon esculentum cultivars. *Theor Appl Genet*. **106**: 363-373

He XY, Zhang YL, He ZH, Wu YP, Xiao YG, Ma CX and Xia XC (2008) Characterization of phytoene synthase 1 gene (Psy1) located on common wheat chromosome 7A and development of a functional marker. *Theor Appl Genet*. **116**: 213-221

Helgi Ö, Stephen AR, Arthur JW and Herbert ES (2005) *The Physiology of Flowering Plants*, Cambridge University Press) pp. 309

Herrero R, Asins MJ, Carbonell EA, Navarro L and (1996) Genetic diversity in the orange subfamily Aurantioideae. I. Intraspecies and intragenus genetic variability. *Theor Appl Genet*. **92**: 599-609

Hiebert CW, Thomas JB, Somers DJ, McCallum BD and Fox SL (2007) Microsatellite mapping of adult-plant leaf rust resistance gene Lr22a in wheat. *Theor Appl Genet*. **115**: 877-884

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

Hooker JD (1872) *The Flora of British India*, 1 (Rev and Co.) pp. 484-517

Huang J and Sun SM (2000) Genetic diversity and relationships of sweet potato and its wild relatives in Ipomoea series Batatas (Convolvulaceae) as revealed by inter-simple sequence repeat (ISSR) and restriction analysis of chloroplast DNA. *Theor Appl Genet.* **100**: 1050–1060

Hughes CR and Queller DC (1993) Detection of highly polymorphic microsatellite loci in a species with little allozyme polymorphism. *Mol Ecol*. **2**: 131-137

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

Hvarleva TT, Kapari-Isaia L, Papayiannis A, Atanassov A, Hadjinicoli A and Kyriakou A (2008) Characterization of Citrus cultivars and clones in Cyprus through microsatellite and RAPD analysis. *Biotechnol Biotec Eq*. **22**: 787-794

ICIMOD. (2003) Cash Crop Farming in Nepal: The Importance of Pollinators Diversity and Managed Pollination in Citrus. Jawalakhel, Kathmandu. pp

Innis MA and Gelfand DH (1990) PCR Protocols: A Guide to Methods and Applications. Innis MA. San Diego, CA, Academic Press**:** 3-12

Innis MA, Myambo KB, Gelfand DH and Brow MAD (1988) DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA. *Proc Natl Acad Sci USA*. **85**: 9436-9440

IPGRI (1991) Descriptors for Cirrus. Revised ed. IPGRI, Rome. Isolated from a Size-Fractionated Genomic Library of *Brassica napus* L. (rapeseed). *Theor Appl Genet*. **91**: 206-211

IPGRI. (2003) Genetic diversity analysis with molecular marker data: Learning module [Online] <irc.igd.cornell.edu/MolecularMarkers/04_Measures.pdf>

Isabel N, Beaulieu J, Theriault P and Bousquet J (1999) Direct evidence for biased gene diversity estimates from dominant random amplified polymorphic DNA (RAPD) fingerprints. *Mol Ecol*. **8**: 477-483

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

Jakse J, Stajner N, Kozjak P, Cerenak A and Javornik B (2008) Trinucleotide microsatellite repeat is tightly linked to male sex in hop (Humulus lupulus L.). *Mol Breed*. **21**: 139–148

Jannati M, Fotouhi R, Abad AP and Salehi Z (2009) Genetic diversity analysis of Iranian citrus varieties using micro satellite (SSR) based markers. *J Hortic Forest*. **1**(7): 120-125

Jarrell DC, Roosse ML, Traugh SN and Kupper RS (1992) A genetic map of citrus based on the segregation of isozymes and RFLP in an intergeneric cross. *Theor Appl Genet*. **84**: 49-56

Jeffreys AJ, Wilson DJ and Thein SL (1985) Hypervariable 'minisatellite' regions in human DNA. *Nature*. **314**: 67-73

Jeya Prakash Sp, Biji KR, Mimchael Gomez S, Ganesha Murthym K and Chandra Babu R (2006) Genetic diversity analysis of sorghum (*Sorghum bicolor* L. Moench) accessions using RAPD markers. *Indian J Crop Sci*. **1**(1-2): 109-112

Joshi K (2011) Molecular Differentiation and Phylogeny of Swertia (Gentiannaceae) of the Himalayan Region, Nepal. *Int J Biotechnol Biochem*. **7**(2): 265-277

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. *Theor Appl Genet*. **100**: 1311–1320

Joshi SP, Ranjekar PK and Gupta VS (1999) Molecular markers in plant genome analysis Pune, India, National Chemical Laboratoy

Kalia RK, Rai MK, Kalia S, Singh R and Dhawan AK (2011) Microsatellite markers: an overview of the recent progress in plants. *Euphytica*. **177**: 309-334

Karp A, Edwards KJ, Bruford M, Funk S, Vosman B, Morgante M, Seberg O, Kremer A, Boursot P, Arctander P, Tautz D and Hewitt GM (1997a) Molecular technique for biodiversity evaluation: oppertunities and challenges. *Nat Biotechnol*. **15**: 625-628

Karp A, Isaac PG and Ingram DS (1998) *Molecular Tools for Screening Bidiversity of Plants and Animals*, Chapman and Hall) pp. 301-318

Karp A, kresovich S, Bhat KV, Ayad WG and Hodgkin T (1997b) *Molecular tools in plant genetic resources conservation: a guide to a technologies: IPGRI Technical Bulletin No. 2*, 2 (International Plant Genetic Resources Institute) pp.

Karp A, Seberg O and Buiatti M (1996) Molecular technique in the assessment of botanical evaluation. *Ann Bot*. **78**: 143-149

Kenis K and Keulemans J (2000) The use of microsatellites to investigate the homozygous status of apple plants obtained by anther culture and parthenogensis *in situ*. *Acta Hortic*. **538**: 581-585

Khan I (2007) *Citrus Genetics, Breeding and Biotechnology*, CAB International) pp. 19-44

Kijas JMH, Fowler JCS and Thomas MR (1995) An evaluation of sequence tagged microsatellite site markers for genetic analysis within Citrus and related species. *Genome*. **38**(2): 349-355

Kijas JMH, Thomas MR, Fowler JCS and Roose ML (1997) Integration of trinucleotide microsatellites into a linkage map of Citrus. *Theor Appl Genet*. **94**: 701–706

Kim KS, Ratcliffe ST, French W, Liu L and Sappington TW (2008) Utility of EST-derived SSRs as population genetic markers in a beetle. *J Hered*. **99**: 112-124

Kochhar SL (1998) *Economic Botany in Tropics*, Macmillan India Ltd) pp.

Koehler-Santos P, Dornelles ALC and de Freitas LB (2003) Characterization of mandarin citrus germplasm from southern Brazil by morphological and molecular analysis. *Pesq Agropec Bras*. **38**: 797-806

Koirala SD (2010) *Genetic Diversity of Potato (Solanum tuberosum L.) assessed by Microsatellite (SSR) markers*. PhD Thesis submitted in Tribhuvan University, Kirtipur, Kathamandu: 1-19

Kojima T, Nagaoka T, Noda K and Ogihara Y (1998) Genetic linkage map of ISSR and RAPD markers in Einkorn wheat in relation to that of RFLP markers. *Theor Appl Genet*. **96**: 37–45

Kovach WL (2007) MVSP - A Miltivariate Statistical Packages for Windows ver 3.21. Pentraeth, Wales, U.K, Kovach Computing Services

Krezdorn AH Classification of Citrus. University of Florida, Department of Fruit Crops

Krueger RR and Roose ML (2003) Use of molecular markers in the management of Citrus germplasm resources. *J Am Soc Hort Sci*. **128**: 827–837

Kumar S, Jena SN and Nair NK (2010) ISSR polymorphism in Indian wild orange (Citrus indica Tanaka, Rutaceae) and related wild species in North-east India. *Sci Hortic*. **123**: 350–359

Kurane J, Shinde V and Harsulkar A (2009) Application of ISSR marker in pharmacognosy: Current update. *Pharmacogn Rev*. **3**: 216-228

Ladaniya MS (2008) *Fruit biochemistry*, San Diego Academic Press) pp. 125-190

Lalitha SK (1999) DNA markers in plant improvement: An overview. *Biotechnol Adv*. **17**: 143–182

Lallan R and Singh S (2006) MEDICINAL IMPORTANCE OF CITRUS PRODUCTS AND BY-PRODUCTS - A REVIEW. *Agr Rev*. **27**(3): 170-180

Lama TK and kayastha RS (1999) Diversity of Citrus species in Pokhara Valley and its surrounding Areas. Proceeding of III National Conference of Science and Technology, *Royal Nepal Academy of Science and Technology*. 2: 1138-1142

Lamboy WF (1994) Computing genetic similarity coefficients from RAPD data: the effects of PCR artifacts. *Genome Res*. **4**: 31-37

Lanham PG and Brennan RM (1998) Characterization of the genetic resources of redcurrant (Ribes rubrum: subg. Ribesia) using anchored microsatellite markers. *Theor Appl Genet*. **96**: 917–921

Lanoue KZ, Wolf PG, Browning S and Hood EE (1996) Phylogenetic analysis of restriction site variation in wild and cultivated Amatanthus species (Amaranthaceae). *Theor Appl Genet*. **93**: 722-732

Larry KJ (1999) Citrus cultivation. Citrus Health management. Timmer LW and Duncan LW. Lake Alfred APS Press

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**: 6427-6437

Levin IN, Gilboa E, Yeselson S, Shen. and Schaffer AA (2000) Fgr, a major locus that modulates the fructose to glucose ratio in mature tomato fruits. *Theor Appl Genet*. **100**: 256–262

Levinson G and Gutman GA (1987) Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Mol Biol Evol*. **4**: 203–221

Lewontin RC (1972) The apportionment of human diversity. *Evol Biol*. **6**: 381-398

Li P, Lin S, yang X, Hu G and Jiang Y (2009) Molecular phylogeny of Eriobotrya lindl. (loquat) inferred from internal transcribed spaver sequences of nuclear ribosome/ Pak. *J. Bot*. **41**(1): 185-193

Li YZ, Cheng YJ, Yi HL and Deng XX (2006) Genetic Diversity in Mandarin Landraces and Wild Mandarins from China Based on Nuclear and Chloroplast Simple Sequence Repeat Markers. *J Hort Sci Biotechnol*. **81**(3): 371-378

Litt M and Luty JA (1989) A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac-muscle actin gene. *Am J Hum Genet*. **44**: 397–401

Love JM, Knight AM, McAleer MA and Todd JA (1990) Towards construction of a high resolution map of mouse genome using PCR-analyzed microsatellites. *Nucleic Acids Res*. **18**: 4123-4130

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

Luce C, Noyer JL, Tharreu D, Ahmadi N and Feyt H (2001) The use of microstellite markers to examine the genetic resources of rice (*Oryza sativa* L.) adapted to European conditions. Proceeding of international symposium on molecular markers. *Acta Hortic*. **546**: 221-234

Lui BH (1998) *Statistical genomics: linkage, mapping, and QTL analysis*, CRC Press LLC) pp. 611

Lui XL, Quian ZG, Liu FH, Yang YM and Pu CX (2011) Genetic diversity within and among populations of *Neopicrorhiza scrophulariiflora* (Scrophulariaceae) in China, an endangered medicinal plant. *Biochem Syst Ecol*. **39**: 297-301

Luro FL, Costantino G, Terol J, Argout X, Allario T, Wincker P, Talon M, Ollitrault P and Morillon R (2008) Transferability of the EST-SSRs developed on Nules clementine (*Citrus clementina* Hort ex Tan) to other Citrus species and their effectiveness for genetic mapping. *BMC Genomics*. **9**: 287-1287

Lynch M and Milligan BG (1994) Analysis of population genetic structure with RAPD markers. *Mol Ecol*. **3**: 91-99

Lynch M and Walsh B (1998) *Genetics and analysis of quantitative traits*, Sinauer Associates) pp.

MacPherson JM, Eckstein PE, Scoles GJ and Gajadhar AA (1993) Variability of the random amplified polymorphic DNA assay concentration. *Mol Cell Probes*. **7**: 229–293

Maeda M, Uryu N, Murayama N, Ishii H, Ota M, Tsuji K and Inoko H (1990) A simple and rapid method for HLA-DP genotying by digestion of PCR-amplified DNA with allele specific restriction endonucleases. *Hum. Immunol*. **27**: 111–121

Mantel NA (1967) The detection of disease clustering and a generalized regression approach. *Cancer Res.* **27**(2): 209-220

Martin JP and Sanchez-Yelamo MD (2000) Genetic relationships among species of the genus Diplotaxis (Brassicaceae) using inter-simple sequence repeat markers. *Theor Appl Genet*. **101**: 1234–1241

Matsumoto A and Tsumura Y (2004) Evaluation of cleaved amplified polymorphic sequence markers. *Theor Appl Genet*. **110**: 80–91

Maxam AM and Gilbert W (1997) A new method for sequencing DNA. *Proc Natl Acad Sci U.S.A.* **74**: 560-564

McDermott JM, Brandle U, Dutly F, Haemmerli UA, Keller S, Muller KE and Wolf MS (1994) Genetic variation in powdery mildew of barley: Development of RAPD, SCAR and VNTR markers. *Phytopathology*. **84**: 1316-1321

McDermott JM and Mcdonald BA (1993) Gene flow in plant photosystems. *Annu Rev Phytopathol*. **31**: 353-373

Metais I, Aubry C, Hamon B and Jalouzot R (2000) Description and Analysis of Genetic Diversity between Commercial Bean Lines (*Phaseolus vulgaris* L.). *Theor Appl Genet*. **101**(8): 1207-1214

Meyer W, Mitchell TG, Freedman EZ and Vilgays 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**: 2274–2280

Micheli MR, Bora R and D'Ambrosio E (1997) *Random Amplified Polymorphic DNA assay*, Springer - Verlag) pp. 53-63

Milbourne D, Meyer R, Bradshaw JE, Baird E, Bonar N, Provan J, Powell W and Waugh R (1997) Comparison of PCR-based markersystems for the analysis of genetic relationships in cultivated potato. *Mol Breed*. **3**: 127-136

Miller K and Tangley L (1991) *Trees of Life: Saving tropical forests and their biologucal wealth*, 6 (New York Avenue) pp.

Misra A, Shasany AK, Shukla AK, Darokar MP, Singh SC, Sundaresan V, Singh J, Bagchi GD, Jain SP, Saikia D and Khanuja SPS (2010) AFLP markers for identification of Swertia species (Gentianaceae). *Genet Mol Res*. **9**(3): 1535-1544

MoAC. (2011) Statistical information of Nepalese agriculture. Agri. Business Promotion and Statistical Division, Ministry of Agriculture and Cooperatives. Singh Durbar, Kathmandu, Nepal. pp

Mohammadi SA and Prasanna BM (2003) Review & Interpretation: Analysis of Genetic Diversity in Crop Plants—Salient Statistical Tools and Considerations. *Crop Sci*. **43**: 1235-1248

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

Moodie M, Finch RP and Marshall G (1997) Annalysis of genetic variation in wild mustard (*Sinapis arvensis*) using molecular markers. *Weed Sci*. **45**: 102-107

Moore GA (2001) Oranges and lemons: clues to the taxonomy of Citrus from molecular markers. *Trends in Genetics* **17**(9): 536-540

Moore SS, Sargeant LL, King TJ, Mattick JS, Georges M and Hetzel DJ (1991) The conservation of dinucleotide microsatellites among mammalian genomes allows the use of heterologous PCR primer pairs in closely related species. *Genomics*. **10**: 654-660

Moreno S, Martin JP and Ortiz JM (1998) Inter-simple sequence repeats PCR for characterization of closely related grapevine germplasm. *Euphytica*. **101**: 117–125

Morton J (1987) *Mexican Lime*, Hort. Purdue) pp. 168-172

Mullis KB, Faloona E, Scharf S, Saiki R, Horn G and Erlich HA (1986) pecific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harbor Symp Quant Biol*. **51**: 263-273

Mullis KB and Faloona FA (1987) Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol*. **155**: 335-350

Mullis KB, Faloona FA, Scharf SJ, Saiki SK, Horn GT and Erlich HA (1986) Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harbor Symp Quant Biol*. **51**: 263-273

Murray V, Monchawin C and England PR (1993) The determination of the sequences present in the shadow bands of a dinucleotide repeat PCR. *Nucleic Acids Res*. **21**: 2395–2398

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. *Theor Appl Genet*. **94**: 597–602

NCRP. (2012) Annual Report 2067/68 (2010/11), National Citrus Research Program, NARC. Paripatle, Dhankuta. pp

Neale DB and Williams CG (1991) Restriction fragment length polymorphism mapping in conifers and applications to forest genetics and tree improvement. *Canadian J For Res*. **21**: 545-554

Neeraja CN, Maghirang-Rodriguez R, Pamplona A, Heuer S, Collard BC, Septiningsih EM, Vergara G, Sanchez D, Xu K, Ismail AM and Mackill DJ (2007) A marker-assisted backcross approach for developing submergence-tolerant rice cultivars. *Theor Appl Genet*. **115**: 767-776

Nei M (1978) Estimation of average heterozygosity and genetic distance fraom a small number of individuals. *Genetics*. **89**: 583-5890

Nei M (1987) *Molecular Evolutionary Genetics*, columbia University Press) pp.

Nei M and Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc Natl Acad Sci USA*. **76**: 685-686

NFSTC. (2010) Heterozygosity [Online] <http://www.nfstc.org/pdi/Subject07/pdi_s07_m01_01_a.htm>

Ni J, Colowit PM and Mackill DJ (2002) Evaluation of genetic diversity in rice sub species using microsatellite markers. *Crop Sci*. **42**: 601-607

Nicolosi E, Deng ZN, Gentile A, La Malfa S, Continella G and Tribulato E (2000) Citrus phylogeny and genetic origin of important species as investigated by molecular markers. *Theor Appl Genet*. **100**(8): 1155–1166

Novelli VM, Cristofani M, Souza AA and Machado MA (2006) Development and characterization of polymorphic microsatellite markers for the sweet orange (*Citrus sinensis* L. Osbeck). *Gene Mol Biol*. **29**: 90-96

Nybom H (1991) *Applications of DNA fingerprinting in plant breeding*, pp. 294–311

Ogunkanmi LA, Ogundipe OT and Fatokun CA (2008) Genetic diversity in wild relatives of Cowpea (Vigna unguiculata) as revealed by simple sequence repeats markers. *J Food Agric Environ*. **6**(3-4): 264-268

Ollitrault P (1990) Isozyme and DNA Restriction Fragment Length Polymorphisms (RFLP) as genetic markers in Citrus selection. Proceedings of Asia Pacific International conference on Citriculture. 57-69

Olmstead RG, Sweere JA, Spangler RE, Bohs L and Palmer J (1999) *Phylogeny and provisional classifications of the Solanaceae based on chloroplast DNA*, Royal Botanic Gardens) pp. 111-117

Oluwaponmile OO (2009) *Genetic diversity and charaacterization of some selected Vigna species using molecular and morphological markers*. PhD Thesis submitted in Unversity of Nottingham: 1-20

Padmalatha K and Prasad MNV (2006) Optimization of DNA isolation and PCR protocol for RAPD analysis of selected medicinal and aromatic plants of conservation concern from Peninsular India. *Afr J Biotechnol*. **5**(3): 230-234

Pahari S and Bimb HP (2004) Assessment of genetic diversity of Citrus reticulate Blanco and Citrus limon (L.) Burm.f. of Nepal through Isozyme analysis. IV National Conference on Science and Technology, Khumaltar, Nepal, *NAST*. 2: 785-792

Paran I, Kesseli R and Michelmorem R (1991) Identification of Restriction Fragment Length Polymorphism and Random Amplified Polymorphic DNA markers linked to downy mildew resistance gene in lettuce, using near-isogenic lines. *Genome*. **34**: 1021-1027

Parasnis AS, Ramakrishna W, Chowdari KV, Gupta VS and Ranjekar PK (1999) Microsatellite (GATA)n reveals sex-specific differences in papaya. *Theor Appl Genet*. **99**: 1047–1052

Parida SK, Kalia SK, Sunita K, Dalal V, Hemaprabha G, Selvi A, Pandit A, Singh A, Gaikwad K, Sharma TR, Srivastava PS, Singh NK and Mohapatra T (2009) Informative genomic microsatellite markers for efficient genotyping applications in sugarcane. *Theor Appl Genet*. **118**: 327-338

Pasakinskiene I, Griffiths CM, Bettany AJE, Paplauskiene V and Humphreys MW (2000) Anchored simple-sequence repeats as primers to generate species-specific DNA markers in Lolium and Festuca grasses. *Theor Appl Genet*. **100**: 384–390

Pasquale FD, Siragusa M, Abbate L, Tusa N, Pasquale CD and Alonzo G (2006) Characterization of five sour orange clones throughmolecular markers and leaf essential oils analysis. *Sci Hort*. **109**: 54–59

Paudyal KP (2000) Farmer's participatory selection of superior pummelo (Citrus grandis L. Osbeck) trees in Nepal. Proceeding of the third National Horticulture Research Workshop, Khumaltar, Kathmandu, *NARC*.

Paudyal KP and Chalise B (2007) Evaluation of Satsuma mandarin (*Citrus unshiu*) varieties for early season production in Nepal. Proceeding of Fourth National Horticulture Seminar on Horticulture for Food Security, Khumaltar, Lalitpur, Nepal, *NARC*.

Paudyal KP and Shrestha RL (2004) Diversity Study and selection of lime (Cirtus aurantifolia) genotypes for expansion of production period in Nepal. 4th National Horticulture Workshop Kathmandu, Nepal, *Nepal Agriculture Research Council*.

Paudyal KP and Subedi H (2008) Selection of sweet orange (*Citrus sinensis* Osbeck) varieties for production period expansion. *Nepal Hort Soc*. **6**

Paul S, Wachira FN, Powell W and Waugh R (1997) Diversity and genetic differentiation among populations on Indain and Kenyan tea (Camellia sinensis (L.) O. Kuntze) revealed by AFLP markers. *Theor Appl Genet*. **94**: 255-263

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

Peakall R and Sydes MA (1996) *Defining priorities for achieving practical outcomes from the genetic studies of rare plants, in Back from the Brink: refining the threatened species recovery process*, Surrey Beatty and Sons) pp.

Pol RG, de Casenave JL, Feldhaar H, Milesi FA and Gadau J (2008) Polyandry in two South American harvester ants. *Insect Soc*. **55**: 91-97

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S and Rafalski A (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Mol Breed*. **2**: 225-238

Prevost A and Wilkinson MJ (1999) A new system of comparing PCR primers applied to ISSR fingerprinting of potato cultivars. *Theor Appl Genet*. **98**: 107–112

Protopapadakis E and Papanikolaus Y (1999) Use of four isozymatic systems in lemon and lemon-like citrus cultivars to detect their genetic diversity. *J Hort Sci Biotechnol*. **74**: 26–29

Provan J, Russell JR, Booth A and Powell W (1999) Polymorphic chloroplast simple sequence repeat primers for systematic and population studies in the genus Hordeum. *Mol Ecol*. **8**: 505-511

Qian W, Ge S and Hong DY (2001) Genetic variation within and among populations of a wild rice Oryza granulata from China17 detected by RAPD and ISSR markers. *Theor Appl Genet*. **102**: 440–449

Queller DC, Strassman JE and Hughes CR (1993) Microsatellites and Kinship. *Trends Ecol Evol*. **8**: 285-288

Radha TH and Mathew L (2007) *Fruits crops, Horticulture Science Series-3*, pp.

Rahman MM and Nito N (1994) Phylogenetic relationships among the ‘‘true citrus fruit trees’’ by glutamate oxaloacetate transaminase isozymes analysis. *J Jpn Soc Hort Sci*. **62**: 755–760

Rahman MM, Nito N and Isshiki S (2001) Cultivar identification of ‘Yuzu’ (Citrus junos Sieb. Ex Tanaka) and related acid citrus by leaf isozymes. *Sci Hort*. **87**: 191–198

Ranjit M and Gokarna GC (1997) *Citrus Research and Development Action Plan*, Chemonics International Inc) pp. 1-30

Rao SP, Bounphanousay C, Schiller JM and Jackson MT (2002) Collection, classification of cultivated and wild rices of the Lao PDR. *Genet Resour Crop Ev*. **49**: 75-81

Ratnaparkhe MB, Tekeoglu M and Muehlbauer FJ (1998) Intersimple- sequence-repeat (ISSR) polymorphisms are useful for finding markers associated with disease resistance gene clusters. *Theor Appl Genet*. **97**: 515–519

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

Reif JC, Melchinger AE and frisch M (2005) Genetical and Mathematical properties of similarity and dissimilarity coefficient applied in plant breeding and seed bank management. *Crop Sci*. **41**: 1-7

Reuther W, Batchelor LD and Webber HJ (1967) *The Citrus Industry*, Univ. Calif. Div. Agril. Sci.) pp. 534-537

Reuther W, Calavan EC and Carman GE (1967-1989) *The Citrus Industry*, 1-5 (Calif. Div. Agril. Sci.) pp.

Richard I and Beckman JS (1995) How neutral are synonymous codon mutations? *Nat Genet*. **10**: 259

Rijal SP (2002) *Phenotypic variation in pummelo (Citrus grandis L.) at Sankhuwasabha and Bhojpur district of Nepal*. PhD Thesis submitted in Tribhuvan University, Chitwan, Nepal:

Rode J, In-Chol K, Saal B, Flachowsky H, Kriese U and Weber WE (2005) Sex-linked SSR markers in hemp. *Plant Breed*. **124**: 167–170

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

Roistacher CN (1991) *Graft-transmissible disease of citrus, Handbook for detection and diagnosis*, Publication division food and agriculture organization of the United Nations) pp.

Roose ML, Soost RK and Cameron JW (1995) *Citrus*, 2 (Longman) pp. 443–449

RycWik W, Spenser WJ and Rhoads RE (1990) Optimization of annealing temperature for DNA amplification in vitro. *Nucleic Acids Res*. **18**: 6409-6412

Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, Erlich HA and Arnheim N (1985) Enzymatic amplification of globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*. **230**: 1350- 1354

Salimath SS, de Oliveira AC, Godwin ID and Bennetzen JL (1995) Assessment of genome origins and genetic diversity in the genus Eleusine with DNA markers. *Genome*. **38**: 757–763

Sambrook J and Russell DW (2001) *Molecular cloning: A Laboratory Manual*, III (Pub Cold Spring Harbour Laboratory Press) pp.

Sanger F, Nicklen S and Coulson AR (1977) DNA sequencing with chain terminating inhibitors. *Proc Natl Acad Sci U.S.A*. **74**: 5463-5467

Sanker AA and Moore GA (2001) Evaluation of inter–simple sequence repeat analysis for mapping in citrus and extension of the genetic linkage map. *Theor Appl Genet*. **102**(2): 206–214

Santos PK, Dornelles ALC and Freitas LB (2003) Characterization of Mandarin Citrus Germplasm from Southern Brazil by Morphological and Molecular Analyses. *Pesqui Agropecu Brasileira*. **38**(7): 797-806

Sapkota DP (2006) *Characterization and Evaluation of acid lime landraces at Rampur, Chitwan condition*. PhD Thesis submitted in Tribhuvan University, Chitwan, Nepal:

Sardesai N, Nemacheck JA, Subramanyam S and Williams CE (2005) Identification and mapping of H32, a new heat gene conferring resistance to hessian fly. *Theor Appl Genet*. **111**: 1167-1173

Sarwat M (2012) ISSR: A Reliable and Cost-Effective Technique for Detection of DNA Polymorphism. Plant DNA Fingerprinting and Barcoding: Methods and Protocols. Sucher NJ, Hennell JR and Carles MC. LLC, Springer Science Business Media. 862

Schlotterer C (1998) *Microsatellites*, IRL Press) pp. 237–261

Schlotterer C (2000) Evolutionary dynamics of microsatellite DNA. *Chromosoma*. **109**: 365-371

Schranz ME, Song BH, Windsor AJ and Mitchell-Olds T (2007) Comparative genomics in the Brassicaceae: a family-wide perspective. *Curr Opin Plant Biol*. **10**: 168–175

Scora RW (1975) On the history and origin of Citrus. *Bull Torrey Bot Club*. **102**: 369-375

Scora RW (1988) *Biochemistry, taxonomy and evolution of modern cultivated Citrus*, 1 (Oxford & IBH Publishing Company) pp. 277–289

Sefc KM, Regner F, Turetschek E, Glössl J and Steinkellner H (1999) Identification of microsatellite sequences in *Vitis riparia* and their applicability for genotyping of different *Vitis* species. *Genome*. **42**(3): 367-373

Semagn K, Bjornstad A and Ndjiomdjop MN (2006) An overview of molecular marker methods for plants. African Journal of Biotechnology. *Afr Biotechnol*. **5**(25): 2540-2568

Sensi E, Vignani R, Scali M, Masi E and Cresti M (2003) DNA fingerprinting and genetic relatedness among cultivated varities of Olea europaea L. estimated by AFLP analysis. *Sci Hortic*. **97**: 379-388

Shah RB (1992) *Citrus Fruits (Training Mannual)*, DOA (Department of Agriculture)) pp. 1-16

Shahsavar AR, Izadpanah K, Tafazoli E and Tabatabaei BE (2007) Characterization of citrus germplasm including unknown variants by Inter-simple sequence repeat (ISSR) markers. *Sci Hortic*. **112**: 310-314

Shand H. (1997) Human Nature: Agriculture Biodiversity and Farm based food security, An Independent Study of Rafi. Ottawa, Ontario, Canada. pp

Shankar AA and Moore GA (2001) Evaluation of ISSR analysis for mapping in Citrus and extension of the genetic linkage map. *Theor Appl Genet*. **102**: 206–214

Shannon CE (1948) A Mathematical Theory of Communication. *Bell Systems Technical Journal*. **27**: 379-523, 623-656

Shehata AI, Al-Ghethar HA and Al-Homaidan AA (2009) Application of simple sequence repeat (SSR) markers for molecular diversity and heterozygosity analysis in maize inbreed lines. *Saudi J Biol Sci*. **16**: 57-62

Sherwin W (2006) Measurement of biological information with applications from genes to landscapes *Mol Ecol*. **15**: 2857-2869

Shinde D, Lai Y, Sun F and Arnheim N (2003) Taq DNA polymerase slippage mutation rates measured by PCR and quasi-likelihood analysis: (CA/GT)n and (A/T)n microsatellites. *Nucleic Acids Res*. **31**: 974-980

Shrestha JKC (2012) *Genetic Diversity Analysis of Swertia chirayita (Rixb. Ex Flem) Karst Population od Nepal Using Polymerase Chain Reaction Based Random Amplified Polymorphic DNA (RAPD-PCR) Technique*. PhD Thesis submitted in Tribhuvan University, Kirtipur, Kathmandu: 1-50

Shrestha RL, Dhakal D, Gautam D, Paudyal KP and Shrestha S (2012a) Genetic Diversity Assessment of Acid Lime (Citrus aurantifolia) Landraces in Nepal, Using SSR Markers. *Am J Plant Sci*. **3**: 1674-1681

Shrestha RL, Dhakal D, Gautam D, Paudyal KP and Shrestha S (2012b) Study of Fruits Diversity and Selection of Elite Acid lime (*Citrus aurantifolia* Swingle) Genotypes in Nepal. *Am J Plant Sci*. **3**: 1098-1104

Shrestha S (2001) *Molecular Systematic of Weedy Sporobolus species of Australia*. PhD Thesis submitted in The University of Queensland, Australia:

Shrestha S, Adkins SW, Graham GC and Loch DC (2003) Phylogeny of the Sporolobous indicus complex based on Internal Transcribed Spacers (ITS) sequences. *Aust Syst Bot*. **16**: 165-176

Shrestha S, Graham GC, Loch DS and Adkuns SW (2010) Molecular identification of weedy *Sporobolus* species by PCR-RFLP. *Weed Res*. **50**: 383-387

Shrestha TH. (1999) Key Achievement in Citrus section: Annual Report, Horticulture Development Project. Kirtipur, Nepal. pp

Shrestha YH, Shrestha SD and Thapa BB. (2011/12) Barshik Pragati Pratibedan. Programme NCD, Kirtipur, kathmandu. pp 14-29

Simon JC, Carre S, Boutin M, Prunier-Leterme N, Sabater-Mun B, Latorre A and Bournoville R (2003) Facultative Symbiont Infections Affect Aphid Reproduction. *P Roy Soc B-Biol Sci*. **270**: 1703-1712

Singh BD (2005) *Plant Breeding: Principles and Methods*, Kalyani Publishers) pp. 1-434

Singh S and Naqui SAMH (2001) *Citrus*, International Book Distributing Company) pp. 69-80

Slatkin M (1987) Gene flow and geographic structure of natural populations. *Science*. **236**: 787-792

Slotis DE and Slotis P (2003) The role of phylogenetics in comparative genetics. *Plant Physiol*. **132**: 1790-1800

Sneath PHA and Sokal RR (1973) *Taxonomy, the principle and practice of numerical classification*, Pub W.H freeman and company) pp.

Snoussi H, Duval MF, Garcia-Lor A, Belfalah B, Froelicher F, Risterucci AM, Perrier X, Jacquemoud-Collet JP, Navarro L, Harrabi M and Ollitrault P (2012) Assessment of the genetic diversity of the Tunisian citrus rootstock germplasm. *BMC Genet*. **13**(16): 1-12

Sokal RR (1979) *Ecological parameter inferred from spatial correlograms*, ICPH) pp. 167-196

Sokal RR and Michener CD (1958) A Statistical Method for Evaluating systematic relationships. *Uni Kansas Sci Bull*. **38**: 1409-1438

Soltis DE, Soltis PS and Doyle JJ (1998) Contribution of PCR-based methods to plant systematics and evolutionary biology. Molecular Systematics of Plants II. Wolfe AD and Liston A. New York, Kluwer Academic Publishers. II**:** 43-86

Somers DJ, Kirkpatrick R, Moniwa M and Walsh A (2003) Mining single nucleotide polymorphism from hevaploid wheat ESTs. *Genome*. **49**: 431-437

Song QJ, Quigley CV, Carter TE, Nelson RL, Boerma HR, Strachan J and Cregan PB (1999) A selected set of trinucleotide simple sequence repeat markers for soybean cultivar identification. *Plant Var Seeds*. **12**: 207-220

Spigler RB, Lewers KS, Main DS and Ashman TL (2008) Genetic mapping of sex determination in a wild strawberry, Fragaria virginiana, reveals earliest form of sex chromosome. *Heredity*. **101**: 507–517

Spooner D, Treuren R and de Vicente MC (2005) Molecular marker for genebank management, IPGRI Technical Bulletin no. **10:** 1-126

Spooner DM, Anderson GJ and Jansen RK (1993) Chloroplast DNA evidence for interrelationships of tomatoes, potatoes and pepinos (Solanaceae). *Am J Bot*. **80**: 676-688

Spooner DM, Nunez J, Trujillo G, Herrera MR, Guzman F and Ghislain M (2007) Extensive simple sequence repeat genotyping of potato landraces supports a major reevaluation of their gene pool structure and classification. *P Natl Acad Sci USA*. **104**: 19398-19403

Stagel A, Portis E, Toppino L, Rotino GL and Lanteri S (2008) Gene-based microsatellite development for mapping and phylogeny studies in eggplant. *BMC Genomics*. **9**: 357

Staub JE, Kuhns LJ, May B and Grun P (1982) Stability of potato tuber isozymes under different storage regimes. *J Am Sci*. **107**: 405–408

Stewart CN and Excoffier L (1996) Assessing population genetic structure and variability with RAPD data: Application to *Vaccinium macrocarpon* (American Cranberry). *J Evol Biol*. **9**: 153-171

Stuber CW, Polacco M and Senior ML (1999) Synergy of Empirical Breeding, Marker Assisted selection and Genomics to Increase Crop Yield Potential. *Crop Sci*. **39**: 1571-158

Subedi H, Paudyal KP, Srivastava SP, Chalise B and Acharya UK (2010) Selection and dissemination of elite genotypes of acid lime (*Citrus aurantifolia* Swingle) for off season production in Mid and East terai region of Nepal. proceedings of nineth National Outreach Research Workshop, Khumaltar, Lalitpur, *Nepal Agriculture Research Council*.

Subedi HP (2010) *Advance Acid lime Farming Technology in Nepal*, National Citrus Research Programme) pp. 1-10

Sunyaev S, Hanke J, Aydin A, Wirkner U, Zastrow I, Reich J and Bork P (1999) Prediction of nonsynonymous single nucleotide polymorphisms in human disease-associated genes. *J Mol Med*. **77**: 754-760

Swingle WT and Reece PC (1967) *The botany of Citrus and its wild relatives*, 1 (University of California Press) pp. 190–430

Taillon-Miller P, Gu Z, Li Q, Hillier L and Kwok PY (1998) Overlapping genomic sequence: a treasure trove of single nucleotide polymorphisms. *Genome Res*. **8**: 748-754

Talon M and Fred GGJ (2008) Citrus Genomics: A review article. *Int J Plant Gen*

Tanaka T (1977) Fundamental discussion of Citrus classiWcation. *Stud Citrologia*. **14**: 1-6

Taramino G and Tingey SV (1996) Simple sequence repeats for germplasm and analysis and mapping in maize. *Genome*. **39**: 277-287

Tatikonda L, Wani SP, Kannan S, Beerelli N, Sreedevi TK, Hoisington DA, Devi P and Varshney RK (2009) AFLP based molecular characterization of an elite germplasm collection of *Jatropha curcas* L.: A biofuel plant. *Plant Sci*. **176**: 505-513

Tautz D (1993) *Notes on the definition and nomenclature of tandemly repetitive DNA sequences*, Birkhaiiser Verlag) pp. 21–28

Tautz D, Trick M and Dover GA (1986) Cryptic simplicity in DNA is a major source of genetic variation. *Nature*. **322**: 652-656

Teklewood A and Becker HC (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

Templeton AR (1994) Bio-diversity at Molecular Genetics Level: Experiences from disparate micro-organisms. *Phil Trans R Soc London*. **345**: 59-64

Tilman D (2000) Causes, consequences and ethics of biodiversity. *Nature*. **405**(6783): 208-211

Tisserland R (1995) *Essential oil safety*, Churchill Livingstone) pp. 64

Torress AM, Soost RK and Diedenhofen U (1978) Leaf isozymes as genetic markers in citrus. *Am J Bot*. **65**(8): 869-881

Transue DK, Fairbanks DJ, Robison LR and Anderson WR (1994) Species Identification by RAPD analysis of grain amaranth genetic resources. *Crop Sci*. **34**: 1385-1389

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 Douglasfir (*Pseudotsuga menziesii*) and sugi (*Cryptomeria japonica*). *Theor Appl Genet*. **92**: 40–45

Tyrka M, Perovic D, Wardynska A and Ordon F (2008) A new diagnostic SSR marker for selection of the Rym4/Rym5 locus in barley breeding. *J Appl Genetics*. **49**: 127-134

Uzun A (2009) *Turunçgillerde genetik çeşitliliğin SRAP markırları ile karakterizasyonu (Characterization of genetic diversity of Citrus by SRAP markers)*. PhD Thesis submitted in University of Cukurova, Institute of Natural and Applied Sciencies:

Uzun A and Yesiloglu T (2012) *Genetic Diversity in Citrus*, InTech) pp. 213-230

Varghese JP, Rudolph B, Uzunova MI and Ecke W (2000) Use of 5'-anchored primers for the enhanced recovery of specific microsatellite markers in Brassica napus L. *Theor Appl Genet.* **101**: 115–119

Varshney RK, Graner A and Sorrells ME (2005) Genic microsatellite markers in plants: features and applications. *Trends Biotechnol*. **23**: 48-55

Verma SK (1994) Role of centre and organization to conserve and use of plant genetic resources (PGR) their structure and functions. Plant Genetic Resources, Nepalese Prespective, NARC and IPGRI. Upadhya MP, Saiju HK, Banya BK and Bista BS. NARC, Kathmandu, NARC**:** 154-164

Vodenicharova M (1989) Use of proteins as molecular-genetic markers in plants. *Genet Sel*. **22**: 269–277

Vos P, Hogers R, Bleeker M, Reijans M, Van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M and Zabeau M (1995) AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Res*. **23**: 4407-4414

Vrigenhoek RC (1994) Genetic diversity and fitness in small population. *Conserv Genet*. **1**: 37-54

Wang G, Mahalingan R and Knap HT (1998) (C-A) and (G-A) anchored simple sequence repeats (ASSRs) generated polymorphism in soybean, *Glycine max* (L.) Merr. *Theor Appl Genet*. **96**: 1086–1096

Wang ML, Barkley NA and Jenkins TM (2009) Microsatellite Markers in Plants and Insects. Part I: Applications of Biotechnology. *Genes, Genomes and Genomics*. **3**(1): 1-10

Wang ML, Huang L, Bongard-Pierce DK, Belmonte S, Zachgo EA, Morris JW, Dolan M and Goodman HM (1997) Construction of an approximately 2 Mb contig in the region around 80 cM of Arabidopsis thaliana chromosome 2. *Plant J*. **12**: 711–730

Weir BS and Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution*. **38**: 1358-1370

Weising K and Gardner RC (1999) A Set of Conserved PCR Primers for the Analysis of Simple Sequence Repeat Polymorphism in Chloroplast Genomes of Dicotyledonous Angiosperm. *Genome*. **42**(1): 9-19

Weising K, Nybom H, Wolff K and Kahl G (2005) DNA fingerprinting in plants Principles, Methods and Applications. New York, Singapore, CRC Press, Taylor & Francis Group

Welsh J and McClelland M (1990) Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res*. **18**: 7213-7218

Whitkus R, Doebley J and Wendel JF (1994) Nuclear DNA markers in systematics and evolution. DNA-Based Markers in Plants. Netherlands, Kluwer Academic Press

Wiliams JGK, Hanafey MK, Rafalski JA and Tingey SV (1993) Genetic analysis using random amplified polymorphic markers. *Method Enzymol*. **218**: 704-740

Wiliams JGK, kubelik AR, Livak KJ, Rafalski JA and Tingey SV (1990) DNA polymorphism amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res*. **18**: 6531-6535

Wiliams JGK, Kunelik AR, Livak KJ, Rafalski JA and Tingey SV (1990) DNA polymorphism amplified by arbitrary primers are useful as genetic-markers. *Nucleic Acids Res*. **18**: 6531-6335

Wiliams NMV, Pande N, Nair S, Mohan M and Bannett J (1991) Restriction fragment length polymorphism analysis of polymerase chain reation products amplified from mapped loci of rice (*Oryza satva* L.) genomic DNA. *Theor Appl Genet*. **82**: 489-498

Wolfe AD, Xiang Q-Y and Kephart SR (1998) Diploid hybrid speciation in Penstemon (Scrophulariaceae). *Proc Natl Acad Sci USA*. **95**: 5112–5115

Wolff K and Morgan-Richards M (1998) PCR markers distinguish Plantago major subspecies. *Theor Appl Genet*. **96**: 282–286

Wolff K, Zietkiewicz E and Hofstra H (1995) Identification of chrysanthemum cultivars and stability of DNA fingerprint patterns. *Theor Appl Genet*. **91**: 439–447

Wu K, Jones R, Dannaeberger L and Scolnik PA (1994) Detection of microsatellite polymorphisms without cloning. *Nucleic Acids Res.* **22**: 3257–3258

Xu H, Wilson DJ, Arulsekar S and Bakalinsky AT (1995) Sequence-specific polymerase chain reaction markers derived from randomly amplified polymorphic DNA markers for fingerprinting grape (Vitis) rootstocks. *J Am Soc Hortic Sci*. **120**: 714-720

Xue DW, Feng SG, Zhao HY, Jiang H, Shen B and Shi NN (2010) The linkage maps of Dendrobium species based on RAPD and SRAP markers. *J Genet Genomics*. **37**

Zane L, Bargelloni L and Patarnello T (2002) Strategies for microsatellite isolation: a review. *Mol Ecol*. **11**: 1-6

Zhang D, Chen S, Hen S, Zhang D and Gao Q (2007) Patterns of Genetic Variation in Swertia przewalskii, an Endangered Endemic Species of the Qinghai-Tibet Plateau. *Biochem Genet*. **45**((1/2)): 33-50

Zietkiewicz E, Rafalski A and Labuda D (1994) Genome fingerprinting by simple sequence repeat (SSR) – anchored polymerase chain reaction amplification. *Genomics*. **20**: 176–183

Zuo Y, Chen Z, Kondo K, Funamoto T, Wen J and Zhou S (2010) DNA barcoding of Panax species. *Planta Med*. **77**: 182-187

APPENDICES

Appendix 1 Comparison of different characteristics of most frequently used molecular marker systems in plants.

S.N.	Properties	RFLP	Microsatellite	RAPD	AFLP	ISSR
1	Genomic abundance	high	medium	very high	very high	medium
2	Part of genome surveyed	low copy coding regions	whole genome	whole genome	whole genome	whole genome
3	Amount of DNA required	high	low	Low	medium	low
4	Type of polymorphism	single base changes, insertion, deletion	changes in length of repeats	single base changes, insertion, deletion	single base changes, insertion, deletion	single base changes, insertion, deletion
5	Level of polymorphism[a]	Medium	high	high	very high	medium
6	Effective multiplex ratio[b]	Low	medium	medium	high	medium
7	Marker index[c]	Low	medium	medium	high	medium
8	Inheritance	Codominan	codominan	dominant	Dominant	dominant
9	Detection of alleles	Yes	Yes	no	No	no
10	Ease of use	labour intensive	easy	easy	difficult initially	Easy
11	Automation	low high				
12	Reproducibility (reliability)	High	high	intermediate	high	medium to high
13	Type of probes/primers	low copy genomic DNA or cDNA clones	specific repeat DNA sequence	usually 10 bp random nucleotides	specific sequence	specific repeat DNA sequence
14	Cloning and/or sequencing	Yes	Yes	no	No	No
15	Radioactive detection	usually yes	No	no	yes/no	No
16	Development/start-up costs	High	high	low	medium	Low
17	Utility for genetic mapping	species specific	species specific	cross specific	cross specific	cross specific
18	Proprietary rights status	No	No (some are licensed)	licensed	licensed	No

[a]Level of polymorphism (average heterozygosity) is an average of the probability that two alleles taken at random can be distinguished

[b]Effective multiplex ratio is the number of polymorphic loci analysed per experiment in the germplasm tested. [c]Marker index is the product of the average expected heterozygosity and the effective multiplex ratio

Source: Semagn et al. (2006)

Appendix 2 List of Citrus spp. collected from Paripatle, Dhankuta and Kathmandu valley, Nepal.

Acc no.	Name of *Citrus* spp.	Variety	Acc no.	Name of *Citrus* spp.	Variety
1	*Citrus limon*	Local	58	*P. pomery*	Unknown
2	*C. aurantifolia*	Unknown	59	*C. aurantifolia*	Unknown
3	*C. paradisi*	Marsheedless	60	*C. limon*	Unknown
4	*C. reticulata*	Local	61	*C. aurantifolia*	Unknown
5	*C. sinensis*	Local	62	*C. aurantifolia*	Unknown
6	*C. reticulata*	Avana	63	*C. aurantifolia*	Unknown
7	*C. grandis*	Star Ruby Red	64	*C. aurantifolia*	Unknown
8	*C. reticulata*	Fruitel Early	65	*C. aurantifolia*	Unknown
9	*C. medica*	Local	66	*P. trifoliata*	Flying Dragon
10	*C. sinensis*	Washington Navel	67	*C. sinensis*	Dhankuta Local
11	*C. sinensis*	Navel Orange	68	*C. maxima*	Phultrac
12	*C. sinensis*	Lanlet Navel	69	*C. sinensis*	Delicious Seedless
13	*Pincirus trifoliata*	Unknown	70	*C. medica*	Dhankuta local
14	*C. noblis*	Kinnow mandarin	71	*C. sinensis*	Valencia late
15	*Fortunella japonicum*	Unknown	72	*C. sinensis*	Tamango
16	*C. aurantium*	Unknown	73	*C. maxima*	Reed
17	*C. aurantifolia*	Unknown	74	Natural hybrid	Local hokse
18	*C. aurantifolia*	Unknown	75	Tangelo	Seminole
19	*C. aurantifolia*	Unknown	76	*C. sinensis*	Sevelle common
20	*C. maxima*	Unknown	77	Mandarin	Kinnow
21	*C. aurantifolia*	Unknown	78	*C. sinensis*	Cara Cara Navel
22	*C. aurantifolia*	Unknown	79	*C. sinensis*	White Rakor
23	*C. maxima*	Unknown	80	*C. sinensis*	Malta Blood Red
24	Unknown	Unknown	81	*C. sinensis*	Salastiana
25	Unknown	Unknown	82	*C. sinensis*	Mosambi
26	*C. jambhera* (Kathe jyamir)	Unknown	83	Tangelo	Minneola
27	*C. unshiu*	Unknown	84	*C. sinensis*	Succari
28	*C. reticulata*	Fortune	85	*C. aurantifolia*	Unknown
29	*C. sinensis*	Vanelle	86	*S. maxima*	Star Ruby
30	*C. unshiu*	Unknown	87	*C. aurantifolia*	Unknown
31	*C. reticulata*	Bansha Kharka Local	88	Tagore	Elendale
32	*C. reticulata*	Bud Muted Khoku	89	Tangor	Ortanique
33	*C. unshiu*	Satsuma wase	90	*C. sinensis*	Lue Gim Gong
34	*C. reticulata*	Nova	91	Tangor	Murkotte (old)
35	*C. reticulata*	Khoku	92	Tangor	Murkotte
36	*C. clementina*	Hernandina	93	*C. aurantifolia*	Unknown
37	Natural hybrid	Unknown	94	*C. aurantifolia*	Unknown
38	*C. unshiu*	Okitsuwase	95	*C. sinensis*	Hamlin
39	*C. reticulata*	Sikkime	96	Tangelo	Orlando
40	*C. clementina*	Nules	97	*C. maxima*	Henderson
41	*C. reticulata*	Pongan	98	*C. maxima*	Nam Roi
42	*C. tangerina*	Dancy	99	*C. sinensis*	New Hall Navel
43	Rangapur lime	Unknown	100	*C. sinensis*	Samauti

44	*C. aurantifolia*	Unknown	101	*C. sinensis*	Washington Navel
45	Old citrange	Unknown	102	Tangor	Ellendale
46	Citrange	Unknown	103	*C. reticulata*	Kamal
47	*C. maxima*	Phodiem	104	*F. margarita*	Unknown
48	*C. aurantifolia*	Unknown	105	*C. aurantifolia*	Unknown
49	*C. unshiu*	Satsuma URSS	106	*C. maxima*	Pink Ruby
50	Volkamerina	Unknown	107	*F. hindsii*	Unknown
51	*C. sinensis*	Valencia Late (local)	108	*C. mitis* (mandarin)	Kalamondin
52	*C. aurantifolia*	Unknown	109	*C. sinensis*	Lane Let Navel
53	Carrizo	Unknown	110	*C. sinensis*	Meisheu-9
54	*P. trifoliate*	Unknown	111	*C. limon*	Unknown
55	Citrumelo	Unknown	112	*C. maxima*	Unknown
56	*C. paradise*	Unknown	113	*C. aurantifolia*	Unknown
57	*C. aurantifolia*	Unknown			

Appendix 3 Quantification of DNA extracted from Acid lime landraces of different agro-ecological region.

Acc No	Farmers Name	Dilution Factor	Absorbance nm				Ratio		DNA concen. µg/ml)
			230	**260**	**280**	**320**	**260/280**	**260/230**	
LT-1	Prem B. Limbu	10:90	0.03	0.013	0.009	0.001	1.44	0.47	6.4
LT-2	Dadhiram Dhakal	10:90	0.02	0.016	0.012	0.006	1.29	0.68	7.9
LT-3	Temple Rai	10:90	0.1	0.065	0.045	0.02	1.43	0.64	32.3
LT-4	Tek B. Katuwal	10:90	0.16	0.097	0.085	0.051	1.14	0.61	48.5
LT-5	Tek B. Katuwal	10:90	0.62	0.035	0.025	0.008	1.38	0.57	17.5
LT-6	Nagandra Ktwal	10:90	0.02	0.013	0.008	0	1.63	0.73	6.4
LT-7	Tun B Katuwal	10:90	0	0.024	0.038	0.001	0.65	0	12.2
LT-8	Hiralal Dhakal	10:90	0.01	0.008	0.005	0.001	1.63	0.68	5.3
LT-9	Su resh Dhakal	10:90	0.5	0.3	0.316	0.23	0.96	0.61	151.7
LT-10	Purna B Tamang	10:90	0.04	0.023	0.018	0.011	1.3	0.58	11.5
LT-11	Nara B Niraula	10:90	0.09	0.056	0.045	0.024	1.23	0.61	27.8
LT-12	Toya Dahal	10:90	0.06	0.043	0.033	0.022	1.28	0.7	21.4
LT-13	Krishna Dahal	10:90	0.06	0.049	0.039	0.027	1.25	0.76	24.3
LT-14	Suresh Dahal	10:90	0.02	0.014	0.009	0.004	1.53	0.64	6.8
LT-15	Guman Sin Karki	10:90	0.06	0.042	0.033	0.023	1.26	0.73	20.8
LT-16	Sunita Khadka	10:90	0.01	0.105	0.073	0.038	1.43	0.89	52.6
LT-17	Gopal Bdr Karki	10:90	0.06	0.035	0.026	0.015	1.34	0.63	17.3
LT-18	Dil Bdr Khadka	10:90	0.08	0.057	0.045	0.029	1.26	0.68	28.3
LKT-19	Min Raj Dahal	10:90	0.02	0.026	0.017	0.001	1.52	1.12	13.1
LT-20	Toran Niraula	10:90	0.02	0.016	0.012	0.001	1.38	0.86	8
LT-21	Mekh Bdr Nirula	10:90	0.01	0.013	0.011	0.004	1.17	1.03	6.3
LT-22	Ganendra Bhujel	10:90	0.01	0.012	0.006	0	2.13	1.07	6.2
LT-23	Dilliram Bhujal	10:90	0.03	0.032	0.022	0.008	1.43	1.26	15.8
LD-24	Eir Bdr Tamang	10:90	0.02	0.02	0.014	0.001	1.48	0.98	10.2

LD-25	NCRP local	10:90	0.02	0.016	0.012	0.003	1.38	1.01	8.1
LD-26	NCRP-52	10:90	0.07	0.032	0.022	0.009	1.44	0.48	15.18
LD-27	NCRP-59	10:90	0.09	0.058	0.044	0.027	1.34	0.62	29.1
LD-28	NCRP-60	10:90	0.02	0.015	0.013	0.004	1.11	0.9	7.4
LD-29	NCRP-49	10:90	0.12	0.083	0.058	0.026	1.43	0.69	41.5
LD-30	NCRP-46	10:90	0.14	0.083	0.067	0.042	1.24	0.6	41.6
LD-31	Prem B Parajuli	10:90	0.01	0.011	0.011	0.004	0.97	0.79	5.4
LD-32	Indra B karki	10:90	0.19	0.11	0.099	0.061	1.11	0.59	54.9
LD-33	Rudra Bdr Karki	10:90	0.02	0.02	0.014	0.001	1.39	1.01	9.9
LS-34	ArbMet- Local-1	10:90	0.02	0.012	0.006	0	0.62	2.02	5.9
LS-35	ArbMet- Local-2	10:90	0.02	0.013	0.004	0	0.56	3.36	6.4
LS-36	ArbMet- 59	10:90	0.02	0.014	0.007	0	0.66	2.12	7.2
LS-37	ArbMet- ncrp53	10:90	0.02	0.011	0.006	0	0.74	1.91	5.6
LS-38	ArbMet- NCRP49	10:90	0.01	0.009	0.004	0	1.12	2.32	4.5
LS-39	ArbMet- NCRP48	10:90	0.02	0.012	0.008	0	0.74	1.57	6.2
LS-40	ArbMet- NCRP50	10:90	0.02	0.012	0.006	0	0.73	1.99	5.8
LS-41	ArbMet- 59	10:90	0.01	0.008	0.002	0	0.69	4.22	4
LS-42	ArbMet- 46	10:90	0.02	0.011	0.003	0	0.63	3.99	5.7
LM-43	Dilip P Parsain	10:90	0.02	0.009	0.004	0	0.51	2.32	4.5
LM-44	Dilip P Parsain	10:90	0.01	0.008	0.003	0	0.58	2.76	4
LM-45	N Situla	10:90	0.02	0.019	0.011	0	0.76	1.75	9.3
LD-46	Ratna B Tamang	10:90	0.02	0.019	0.012	0.002	0.77	1.55	9.4
LD-48	Surendra Rai	10:90	0.02	0.011	0.004	0	0.7	2.87	5.5
LD-49	Ram Bdr Rai	10:90	0.02	0.015	0.011	0.002	0.8	1.42	7.5
LD-50	Rudra Limbu	10:90	0.02	0.016	0.01	0.001	0.75	1.54	7.9
LM-51	KN Timsina55	10:90	0.03	0.025	0.018	0.004	1.38	0.29	12.7
LM-52	KN Timsina59	10:90	0.55	0.284	0.287	0.267	0.99	0.51	142
LM-54	KN Timsina48	10:90	0.02	0.014	0.013	0.006	1.06	0.86	7.1
LM-55	KN Timsina49	10:90	0.33	0.254	0.208	0.112	1.22	0.76	126.8
LM-56	A Mehata (Ur)	10:90	0.04	0.069	0.035	0.001	1.68	1.97	34.5
LS-57	A Mehata (Ro)	10:90	0.02	0.012	0.006	0.001	0.77	1.99	5.8
LM-58	D Parsain (Bkse)	10:90	0.05	0.034	0.025	0.016	1.35	17.7	17.7
LD-59	Ncrp (Ob)	10:90	0.02	0.014	0.009	0.002	0.75	1.63	7.2
LKv60	CRP-1	10:90	0.02	0.021	0.014	0.001	1.46	0.99	10.5
LKm61	CRP-2	10:90	0.03	0.023	0.016	0.003	1.42	0.71	11.6
LKr62	CRP-3	10:90	0.02	0.014	0.011	0.001	1.26	0.71	7.1

Appendix 4 Quantification of DNA extracted from *Citrus* spp. collected from Paripatle, Dhankuta and Kathmandu valley, Nepal.

Sample ID	**Name of *Citrus spp***	**Dilution Factor**	**Absorbance**				**Ratio 260/280**	**Ratio 260/230**	**DNA Concen.**
			230	**260**	**280**	**320**			**µg/ml**
3	*C. paradisi*	10:90	0.298	0.354	0171	0.001	2.07	1.19	176.8

5	*C. sinensis*	10:90	0.211	0.124	0.065	0.004	1.89	0.58	61.8
9	*C. medica*	10:90	0.342	0.290	0.157	0.000	1.85	0.85	145.1
10	*C. sinensis*	10:90	0.185	0.164	0.084	0.000	1.96	0.88	81.8
16	*C. aurantium*	10:90	0.251	0.152	0.083	0.009	1.84	0.60	75.8
24	Unknown	10:90	0.413	0.472	0.274	0.066	1.72	1.14	235.9
25	Unknown	10:90	0.226	0.197	0.106	0.012	1.86	0.87	98.7
35	*C. reticulate*	10:90	0.519	0.773	0.403	0.045	1.92	1.49	386.3
36	*C. Clementina*	10:90	0.564	0.680	0.387	0.091	1.76	1.21	340.2
37	Natural hybrid	10:90	0.429	0.318	0.226	0.080	1.41	0.72	158.9
38	*C. unshiu*	10:90	0.922	0.377	0.750	0.150	1.84	1.49	688.6
40	*C. Clementina*	10:90	0.863	1.380	0.712	0.081	1.94	1.60	689.8
42	*C. tangerine*	10:90	0.832	1.357	0.694	0.069	1.96	1.63	678.7
45	Old Citrange	10:90	0.546	0.668	0.362	0.047	1.84	1.22	333.8
51	*C. sinensis*	10:90	0.696	0.955	0.539	0.096	1.77	1.37	477.3
52	*C. aurantifolia*	10:90	1.044	1.659	0.880	0.125	1.88	1.59	829.3
54	*Poncirus trifoliata*	10:90	0.557	0.349	0.297	0.178	1.17	0.63	174.6
56	*C. paradise*	10:90	0.567	0.834	0.437	0.045	1.91	1.47	416.9
59	*C. aurantifolia*	10:90	0.993	1.592	0.712	0.072	1.96	1.60	796.1
60	*C. limon*	10:90	0.948	1.551	0.812	0.072	1.96	1.64	775.7
62	*C. aurantifolia*	10:90	0.723	1.114	0.572	0.056	1.95	1.54	557.2
63	*C. aurantifolia*	10:90	0.714	1.124	0.569	0.047	1.97	1.57	562.2
65	*C. aurantifolia*	10:90	0.790	1.235	0.644	0.065	1.92	1.56	617.4
66	*P. trifoliate*	10:90	0.393	0.362	0.203	0.029	1.79	0.92	181.2
67	*C. sinensis*	10:90	0.768	1.183	0.625	0.094	1.89	1.54	591.4
68	*C. maxima*	10:90	0.341	0.427	0.220	0.013	1.94	1.25	213.3
69	*C. sinensis*	10:90	0.689	1.032	0.542	0.056	1.90	1.50	516.1
70	*C. medica*	10:90	0.911	1.608	0.823	0.064	1.95	1.77	804.0
71	*C. sinensis*	10:90	0.947	1.641	0.865	0.079	1.90	1.73	820.6
74	Natural hybrid (Hokse)	10:90	0.578	0.850	0.440	0.028	1.93	1.47	424.9
75	Tangelo	10:90	0.944	1.390	0.755	0.117	1.84	1.47	695.2
77	Mandarin	10:90	0.547	0.714	0.396	0.048	1.80	1.30	357.0
80	*C. sinensis*	10:90	0.816	1.37	0.708	0.045	1.94	1.68	687.2
82	*C. sinensis*	10:90	0.781	1.291	0.674	0.046	1.92	1.65	645.3
85	*C. aurantifolia*	10:90	0.674	0.979	0.511	0.070	1.91	1.45	489.3
87	*C. aurantifolia*	10:90	0.623	0.931	0.480	0.030	1.94	1.49	465.3
92	Tangor	10:90	0.439	0.584	0.311	0.031	1.88	1.33	292.2
94	*C. aurantifolia*	10:90	0.599	0.853	0.445	0.044	1.92	1.42	426.5
97	*C. maxima*	10:90	0.744	1.062	0.568	0.067	1.87	1.43	530.8
103	*C. reticulate*	10:90	0.468	0.525	0.287	0.056	1.83	1.12	262.3
104	*Fortunella margarita*	10:90	0.449	0.457	0.256	0.072	1.79	1.02	228.3
105	*C. aurantifolia*	10:90	0.579	0.704	0.376	0.056	1.87	1.22	351.9

107	*Fortunella hindsii*	10:90	0.497	0.578	0.312	0.059	1.85	1.16	289.1
108	*C. mitis*	10:90	0.529	0.635	0.342	0.066	1.86	1.20	317.7
111	*C. limon*	10:90	0532	0.604	0.312	0.036	1.93	1.13	301.9
112	*C. maxima*	10:90	0.383	0.386	0.209	0.025	1.85	1.01	193.2
113	*C. limon*	10:90	0.989	1.618	0.827	0.066	1.96	1.64	808.9

Appendix 13 Examples of ISSR profiles generated by ISSR primers.

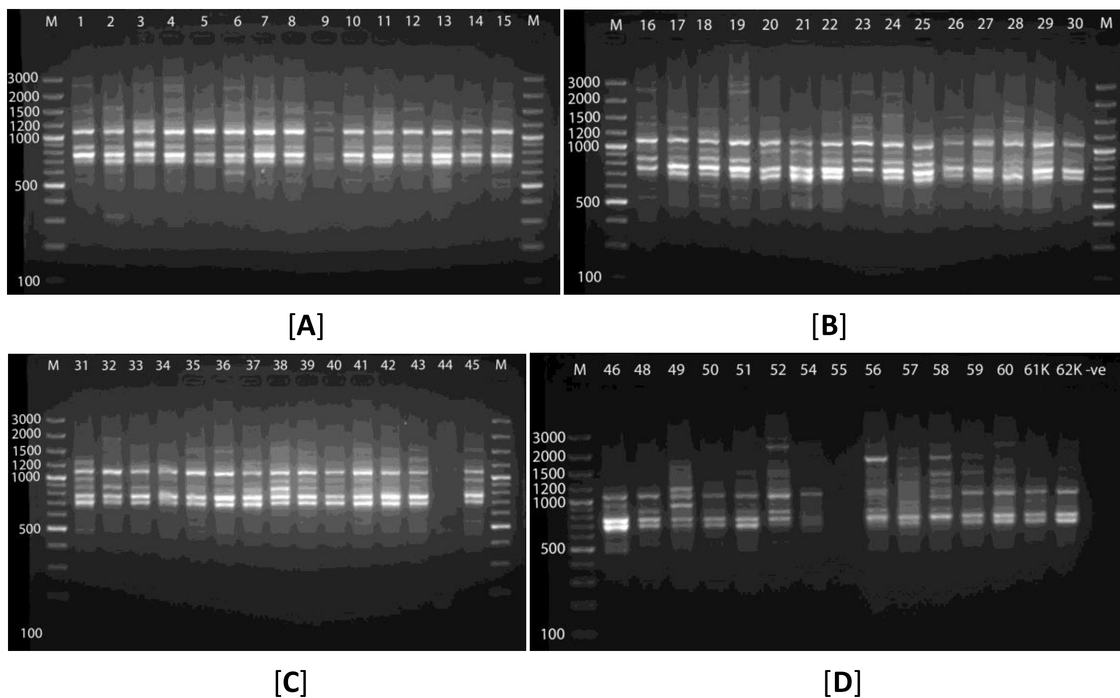


Plate 1 ISSR profiles amplified with **Primer UBC 873.** Lanes marked with M are 100bp plus molecular weight markers. [A] Lanes 1-15 represents acid lime samples 1-15; [B] Lanes 16-30 represents acid lime samples 16-30; [C] Lanes 31-45 represents acid lime samples 31-45; [D] Lanes 46-62k represents acid lime samples 46-62k.

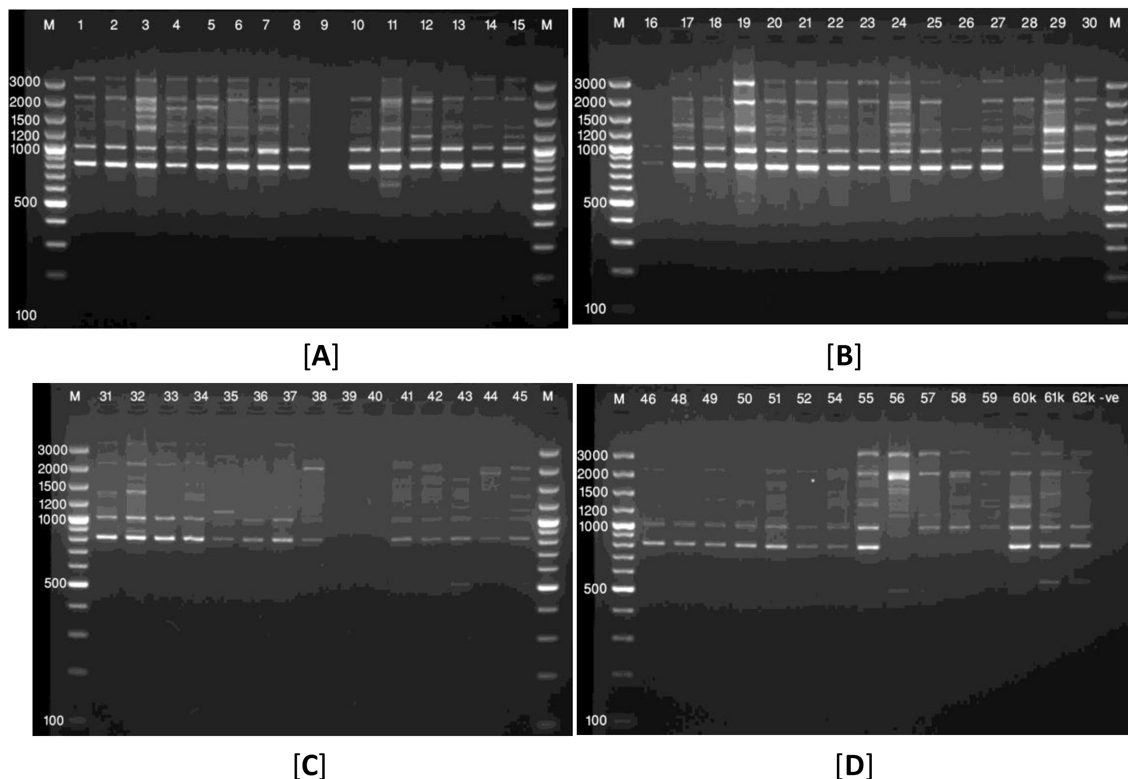
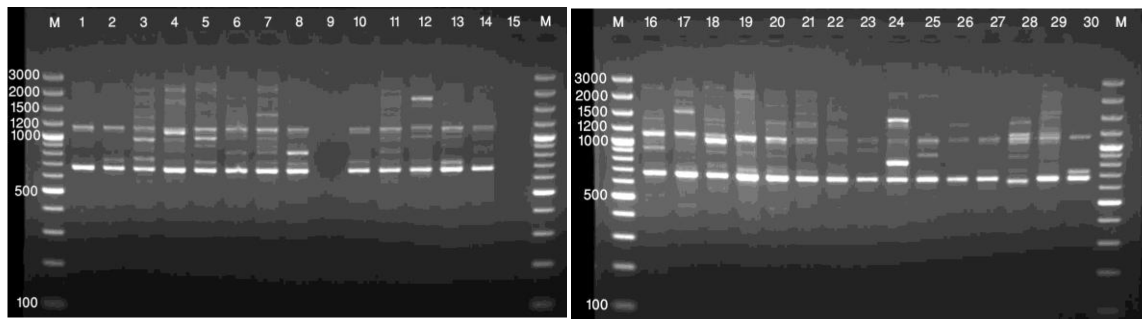
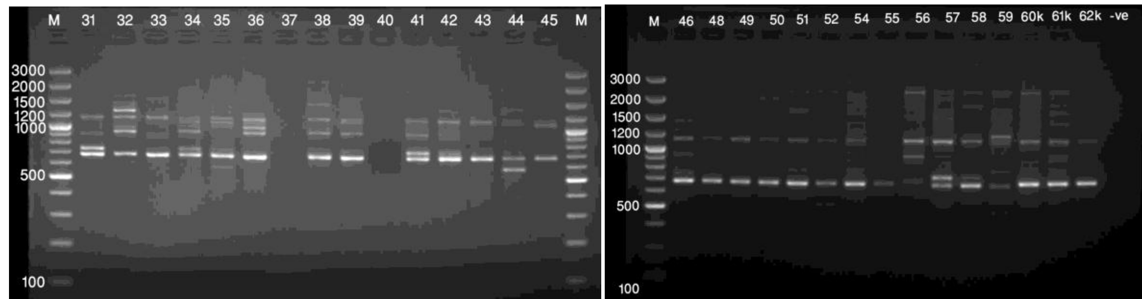


Plate 2 ISSR profiles amplified with **Primer C 2.** Refer legend at plate 1.

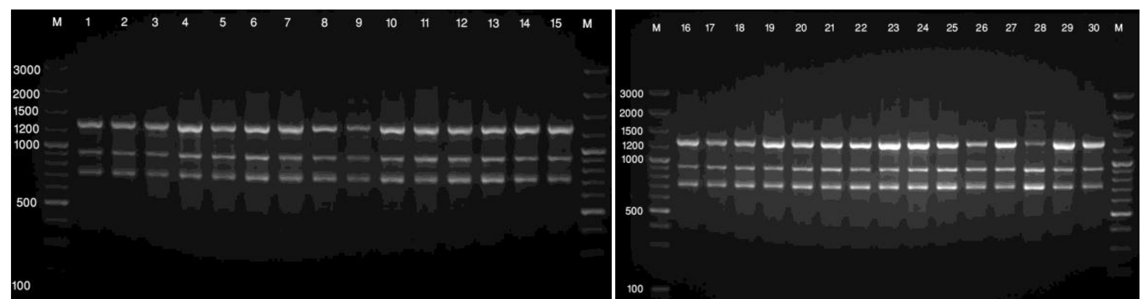


[**A**] [**B**]

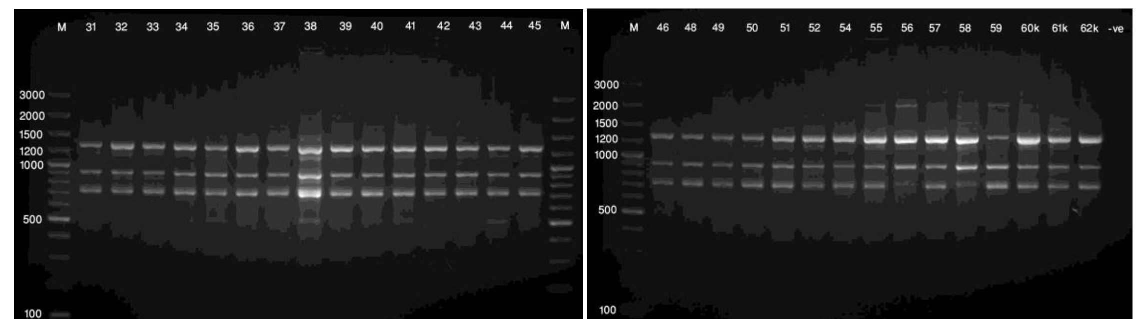


[**C**] [**D**]

Plate 3 ISSR profiles amplified with **Primer C 5.** Refer legend at plate 1.

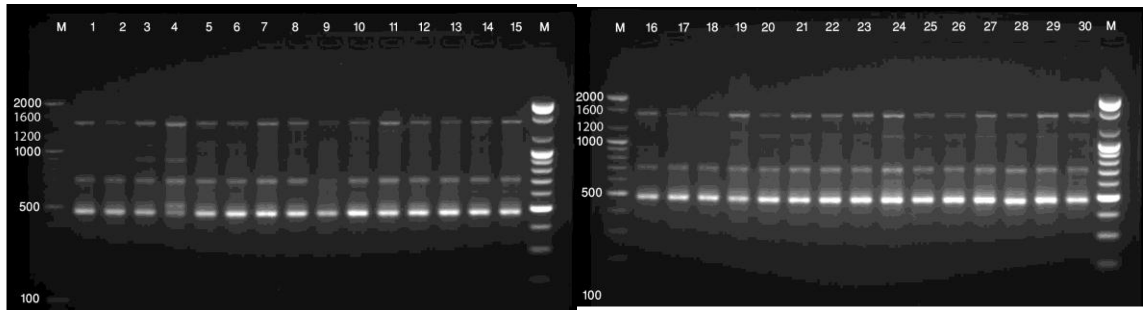


[**A**] [**B**]

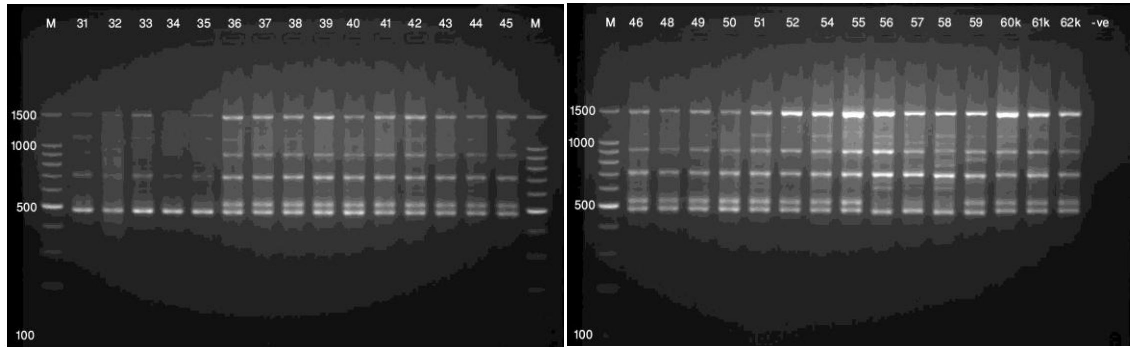


[**C**] [**D**]

Plate 4 ISSR profiles amplified with **Primer C 10.** Refer legend at plate 1.

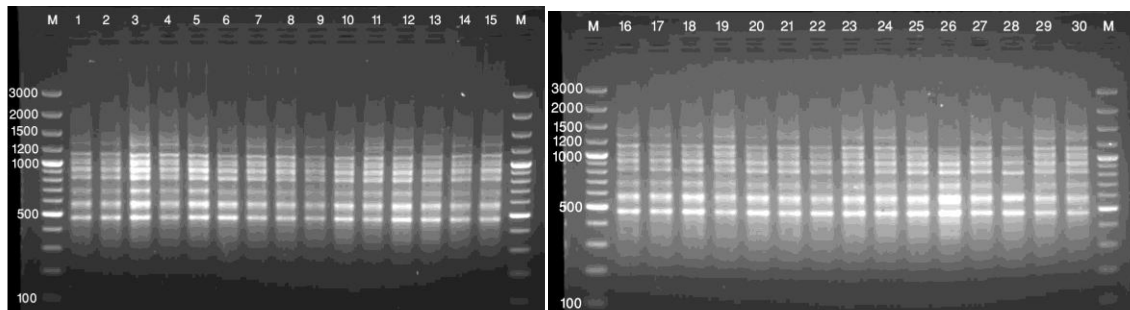


[**A**] [**B**]

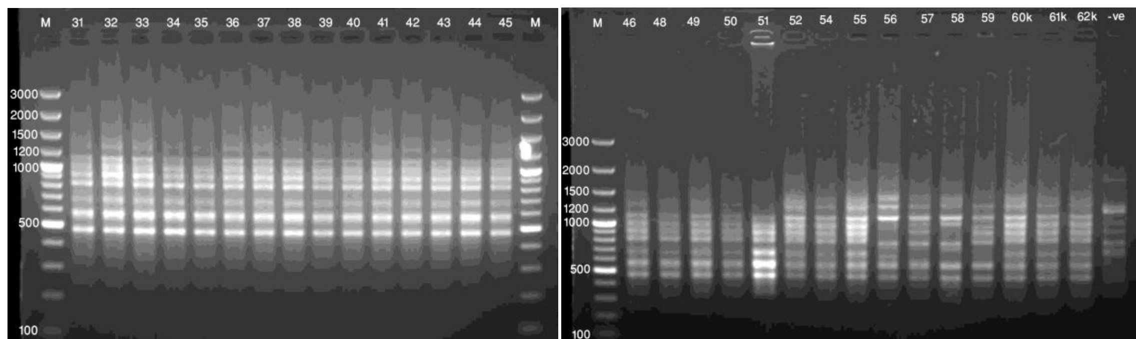


[**C**] [**D**]

Plate 5 ISSR profiles amplified with **Primer UBC 812.** Refer legend at plate 1.



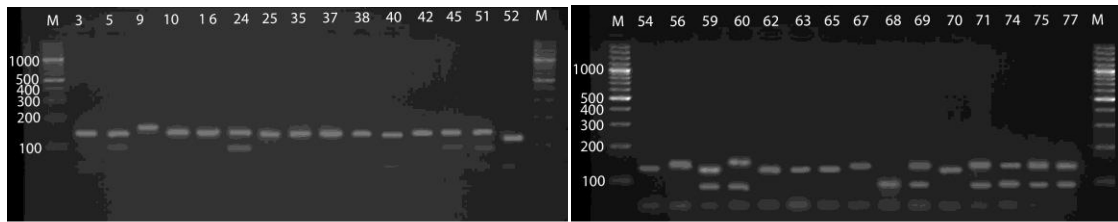
[**A**] [**B**]



[**C**] [**D**]

Plate 6 ISSR profiles amplified with **Primer UBC 889.** Refer legend at plate 1.

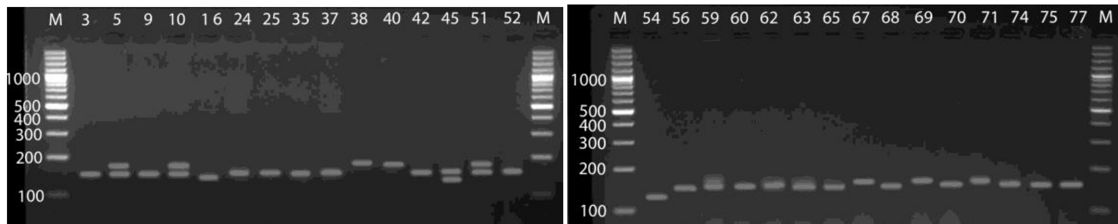
Appendix 14 Examples of SSR profiles generated by SSR primers.



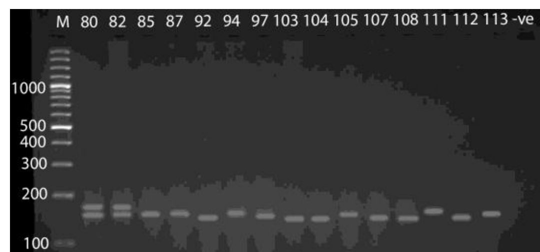
[**A**] [**B**]

[**C**]

Plate 7 SSR profiles amplified with **Primer TAA 45.** Lanes marked with M are 100bp plus molecular weight markers. [A] Lanes 3-52 represents *Citrus* spp. samples 3-52; [B] Lanes 54-77 represents acid lime samples 54-77; [C] Lanes 80-113 represents acid lime samples 80-113.

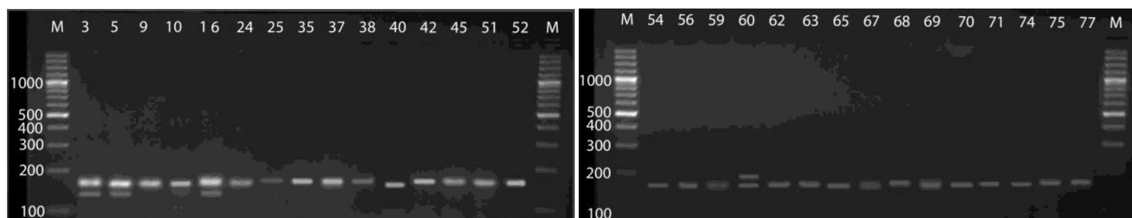


[**A**] [**B**]

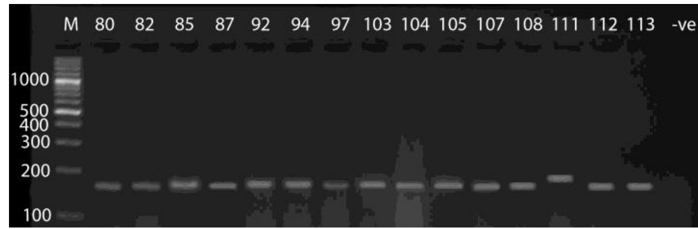


[**C**]

Plate 8 SSR profiles amplified with **Primer CT 19.** Refer legend at plate 7.

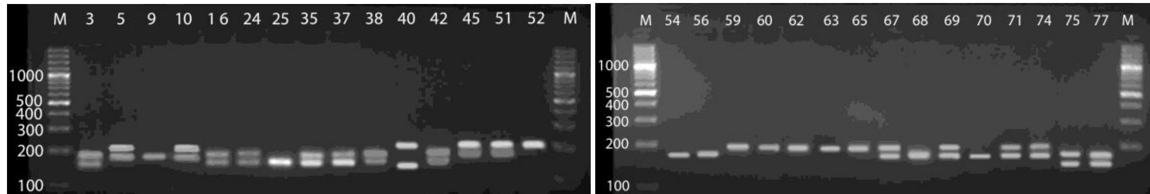


[**A**] [**B**]

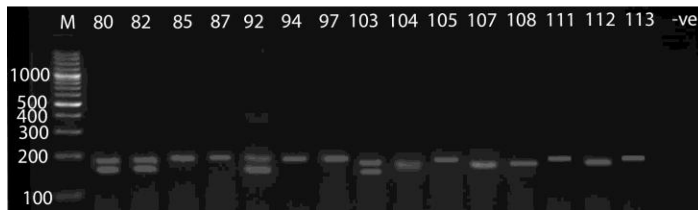


[**C**]

Plate 9 SSR profiles amplified with **Primer CAC 15.** Refer legend at plate 7.

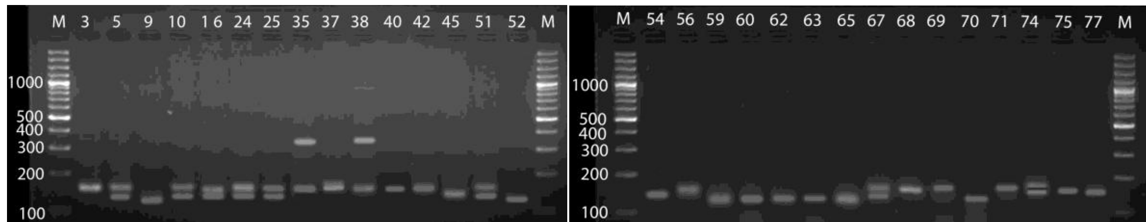


[**A**] [**B**]

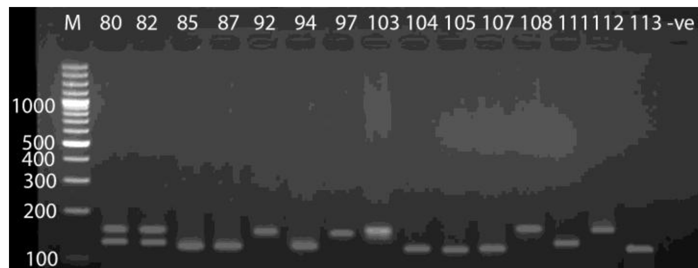


[**C**]

Plate 10 SSR profiles amplified with **Primer GT 03.** Refer legend at plate 7.

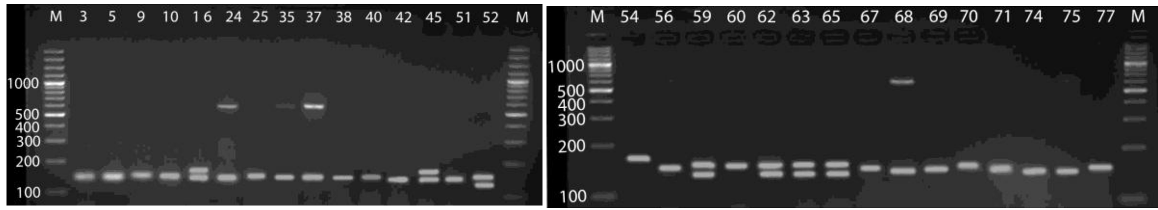


[**A**] [**B**]

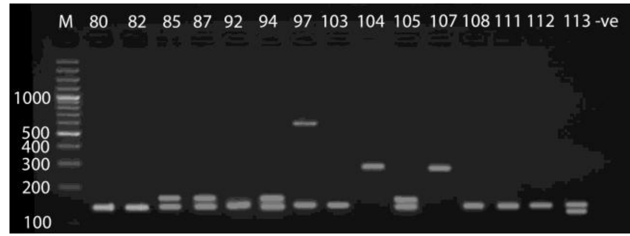


[**C**]

Plate 11 SSR profiles amplified with **Primer TC 26.** Refer legend at plate 7.



[**A**] [**B**]



[**C**]

Plate 12 SSR profiles amplified with **Primer TAA 3.** Refer legend at plate 7.

Appendix 15 Photos Display.

[A] *Citrus reticulata* (Dhankuta, Nepal)

[B] *Citrus aurantifolia* (acid lime) (Kiritpur, Nepal)

[C] Field sample collection (Kirtipur, Nepal)

[D] Field sample collection (Dhankuta, Nepal)

[E] Extraction of DNA

[F] Loading DNA and Gel electrophoresis [G] Marker profile analysis in gel documentation system

[H] Research Team in Molecular Biotechnology Unit, NAST.