Peer-reviewed version available at Plants 2018, 7, 46; doi:10.3390/plants7020046

Title: ISSR-markers Assisted Genetic Diversity Assessment of Acid Lime [*Citrus aurantifolia* (Christm.) Swingle] Germplasm of Eastern Nepal

Nabin Narayan Munankarmi[1,*], Neesha Rana[2], Tribikram Bhattarai[1], Ram Lal Shrestha[3], Bal Krishna Joshi[3], Bikash Baral[4,*], Sangita Shrestha[2]

[1]Central Department of Biotechnology, Tribhuvan University (TU), Kirtipur, Kathmandu, Nepal
[2]Molecular Biotechnology Unit, Faculty of Science, Nepal Academy of Science and Technology (NAST), Khumaltar, Lalitpur, Nepal
[3]Nepal Agriculture Research council (NARC), Lalitpur, Nepal
[4]Department of Biochemistry, University of Turku, FIN- 20014 Turku, Finland

*Corresponding Authors' e-mail:

Nabin Narayan Munankarmi (e-mail: nabin2045@gmail.com) ORCID Number: 0000-0002-1050-0901

Bikash Baral (e-mail: bikash.baral@utu.fi) (ORCID: 0000-0003-2085-9528)

Abstract

Acid lime [*Citrus aurantifolia* (Christm.) Swingle] is a fruit crop, enriched with high commercial value and is cultivated in 60 out of 75 districts representing all geographical landscapes of Nepal. Lack of high yielding cultivars is probably one of the main reason for its extremely reduced productivity which warrants a deep understanding of genetic diversity in existing germplasm. Hereby, we aim to access the genetic diversity of acid lime germplasm cultivated at 3-different ecological gradients of eastern Nepal employing PCR-based Inter-Simple Sequence Repeats markers (ISSR). Altogether, 21 polymorphic ISSR markers were used to assess the genetic diversity in 60 acid lime cultivars sampled from different geographical locations. Analysis of binary data matrix was performed on the basis of bands obtained, scoring of the data was done accordingly, and principal coordinate analysis and phenogram were constructed using different computer algorithms. ISSR profiling yielded 234 amplicons, of which 87.18% were found to be polymorphic. The number of amplified fragments ranged from 7-18 with amplicon size ranging from 250-3200 bp. The NTSYS based Cluster analysis using UPGMA algorithm taking Dice Similarity coefficient separated 60 accessions into 2-major and 3-minor clusters. The genetic diversity analysis revealed the highest for Terai and the lowest for High-hill zone. Cluster I comprised of accessions from High-hill and Mid-hill regions revealing the close genetic relationship, whereas cluster II comprised of accessions from all three agro-ecological zones and the exotic varieties. Furthermore, our results revealed the accessions harvested from different geographical gradients were not genetically distinct, but highest diversity was observed in Terai accessions in comparison to the regions belonging to the High and Mid-hills. Thus, our data indicate that the ISSR provides a better option for evaluating the genetic diversity of Nepalese Acid Lime cultivars and furnished significant information, assisting parental selection in current and future breeding programs and germplasm conservation which ultimately may help to provide a technological breakthrough for the farmers of the developing country like Nepal.

Keywords: Citrus breeding, diversity, genetic similarity, Lime, molecular markers, PCR

1. Introduction

Citrus, an important genus from Rutaceae family [1], is an ancient perennial crop more often cultivated in tropical and sub-tropical parts of the globe [2]. Nepal is one of the countries in Asia where citrus is thought to have been originated [3], and has crucial role in the horticultural industry. The diploid citrus plants (2n=2x=18) are hybridized through cross pollination, in turn generating the hybrids, plus increasing the ploidy levels [4,5].

Acid lime [*Citrus aurantifolia* (Christm.) Swingle] and lemon [*Citrus limon* (L.) Burm. f.] are important fruit crops in Asia, and India is the largest producer of lime and lemon [6]. Acid lime commonly known as "Kagati" in Nepal is enriched in vitamin 'C' with multi uses as for preparing juice, pickles, and salad. Its medicinal properties are attributed by its preventive and curative measures against various diseases of joints and bones, cold, influenza, dysentery, piles, scurvy, cold, and constipation [7]. As Lemons and Limes juice have profound amount of citric acid comprising 1.38 and 1.44 g/oz resp. [8], beverages with citric acid are reported to reduce the content of calcium, and ultimately enhancing urinary citrate excretion. Hence, it could be a good dietary supplement for preventing and managing calcium Urolithiasis (kidney stone) [7,9].

Acid lime proves to be a crucial commercially cultivated fruit crop of Nepal, with a ranking of 3rd after mandarin and sweet orange in terms of area coverage (2,731 ha). The cultivation of Acid lime is done in several districts of Terai (60 out of 75 ; low-lying land on the outer foothills of Himalayas) to the land lying on High-hill landscapes of Nepal, particularly concentrating in Eastern Nepal [10]. Unlike Mandarin and sweet orange, acid lime can be successfully raised from High-hill to Terai regions of Nepalese land [10]. Three different cultivars of lime have been grown in Nepal, viz., Acid Lime, Eureka, and hybrids. Among them Acid lime bear high commercial value in the market due to its size, better aroma and enriched medicinal value [11]. The favorable season for the production of lime in Nepal is from September - November, however, the demand of this fruit is throughout the year [12].

In Nepal, the production of acid lime is 8.3 ton/ha [13], which is itself very low as compared to the productivity of other countries like Argentina (19 ton/ha), and India (12.2 ton/ha) [14]. The reason for this low production may be due to several stressors, like biotic stress (Pests and diseases) and abiotic stress (salinity, drought, and temperature). In addition, prevailing Climate Change is putting more pressure on gross crop productivity [15]. Considering these scenarios, elite Acid Lime varieties with desirable qualities, like nematode resistance, resistance to disease, juice content, higher yield,

Peer-reviewed version available at Plants 2018, 7, 46; doi:10.3390/plants7020046

stress tolerance etc., holds better promises. Development of such cultivars with desirable qualities can be achieved via breeding programs (conventional and non-conventional) *viz.,* molecular marker-assisted breeding, protoplast fusion, mutation breeding and genetic engineering [16-18]. High level of variation in fruit quality, seasonality in flowering, harvesting time, productivity and disease resistance among acid lime accessions of different agro ecological zones have been made [19]. Therefore, the study of genetic diversity at the molecular level and conservation of acid lime germplasm of Nepal are crucial tasks that remain to be performed for its breeding and cultivar development program.

The major motive of plant breeders lies in improving the qualitative and quantitative traits of the existing cultivars. This has been achieved via conventional breeding involving the whole genomes followed by the selection of highest quality recombinants among several segregating individuals. However, this is highly time-consuming and tedious that involves multiple crosses and several generations, vigilant linkage drag and phenotypic selection [20]. Recently developed DNA-based molecular marker technologies have become immensely useful to plant breeders as complementary tools for conventional breeding. They have evolved as efficient tools for genetic diversity assessment, cultivar identification, marker-assisted selection and breeding, and recently to genomics-assisted breeding for crop improvement [21-25]. Genetic diversity estimation employing molecular marker tools is a fundamental task to be performed as it provides baseline information to plant breeders for detecting a unique germplasm that is required for the improvement and selection of horticultural traits and introduction of disease resistant cultivars to improve both the quality and quantity production of the fruit [26].

Three different classes of molecular markers are currently available for plant breeders to expedite crop improvement. These include: *(i)* hybridization-based markers such as Restriction Fragment Length Polymorphisms (RFLPs); *(ii)* PCR-based like Random Amplified Polymorphic DNAs (RAPDs), Amplification Fragment Length Polymorphisms (AFLPs), Microsatellite or Simple Sequence Repeats (SSRs); and *(iii)* Sequencing-based such as Single Nucleotide Polymorphisms (SNPs) [20]. These molecular markers prove useful for various purposes in crop improvement programs, such as *(i)* for constructing saturated molecular genetic linkage maps (physical and genetic) in various species [27,28]; *(ii)* for identification of markers associated with genes/ Quantitative Trait Loci (QTL) controlling traits of economic importance for indirect Marker Assisted Selection (MAS); *(iii)* gene introgression through backcrossing; *(iv)* germplasm characterization,

genetic diversity assessment and cultivar identification; *(v)* genome organization and phylogenetics, etc. [29-31].

Of the various molecular-marker systems, PCR-based ISSR marker system have wide usage in studying genetics [32]. These markers are polymorphic nature [33], abundance in the genome [34], and have the advantages of SSR markers, circumventing the major obstacle of the development of SSR markers, i.e. the requirement of flanking sequences for primer design and enjoying the advantages of random markers [35]. ISSR technique combines the benefits of AFLP and SSR markers with RAPD's universality [22]. The ISSR markers are informative for species where genome sequences are unavailable [36]. The scoring is done as dominant markers, and inherited in Mendelian fashion [37]. ISSR, a PCR-based marker has a capacity to rapidly screen and differentiate between closely related individuals [32].

Several molecular marker-based studies like RAPD and ISSR have been conducted in different Citrus germplasm by [24,38-41]. Although the codominant SSR marker system based genetic study was carried out prior to this study using the similar samples [39], however, it was not based on advanced Capillary Electrophoresis (CE) and was carried out using conventional Agarose Electrophoretic system. Therefore, based on this as well as ISSR being more robust dominant marker system than RAPDs and condition where no prior genome sequence was known, our aim was to evaluate the genetic diversity and relationship among the acid lime germplasm of different localities of Eastern Nepal employing the ISSR markers and comparative analysis performed based on the results obtained from 3-different marker systems. An overall objective of this research was to set-up a baseline data to assist future breeding and conservation programs of Acid Lime in Nepal. In addition, our main research revolves around different questions, such as *(i)* the extent of genetic diversity in the Acid Lime cultivars of Eastern Nepal, and *(ii)* the agro-ecological zone which harbored most genetically diverse Acid Lime cultivars based on our present ISSR study.

2. Material and Methods

2.1. Plant Materials

Altogether 60 young and healthy leaf samples (6 to 8 weeks old) of acid lime were harvested from farmer's plantation areas of Eastern-zone Nepal and were stored in an airtight zip-lock bags with silica gel. Random sampling was done from the selected trees of all agro-ecological zones, viz., Terai, Mid-hills, and High-hills (Suppl. Table 1).

148 *2.2. DNA Extraction and PCR amplification*

149 Dried leaf samples (100 mg) were ground to a fine powder using liquid nitrogen, and the genomic
150 DNA was extracted according to the protocol of DNeasy plant DNA extraction mini-kit (QIAGEN).
151 ISSR-PCR amplification was performed in 25µL total reaction volume having 25ng of genomic
152 DNA, 3.0mM $MgCl_2$, 2.5µL (10mM) of 10× PCR reaction buffer (Fermentas), 0.4µM primer,
153 0.4mM dNTPs and 1.5U *Taq* polymerase (Fermentas, Life Sciences). The PCR cycling conditions
154 consisted of initial denaturation of 94°C (2 m) followed by 40 cycles of denaturation at 94°C (30 s),
155 annealing at 50°C (45 s), elongation at 72°C (2 m) and a final elongation at 72°C (7 m), followed by
156 a hold at 4°C (for infinity) [24].
157
158 The PCR products obtained were analyzed in 2% (w/v) agarose gel comprising Ethidium Bromide
159 (0.5µg/mL, Promega Co.) [42] after running in 1× TAE Buffer (50V; 2 h). Gel-doc system (Ingenius,
160 Syngene Bioimaging, UK) was employed for gel visualization and documentation of ISSR bands.
161 The size of the obtained PCR products was analyzed by using Gene ruler^{TM} 100 bp plus DNA ladder
162 (Fermentas Life Sciences).
163
164 *2.3. ISSR Profiling and Scoring of the Data*

165 Using optimized ISSR-PCR reactions and cycling conditions, 49-different ISSR oligos were
166 screened using fresh genomic DNA samples of acid lime. All experiments comprised 3-biological
167 replicates and 3-technical replicates. Out of 49 oligos, 21 oligos that provide crispy, multiple,
168 scorable, and reproducible bands were selected for further ISSR profiling. The ISSR profiles
169 generated by each of the 21 oligos were used to score the bands, and the creation of binary data
170 matrix. Scoring of all polymorphic and monomorphic bands was performed. Scoring of the markers
171 as '0', '1' and '9' was performed for absence, presence and the failure of PCR amplification
172 respectively [43-45].
173
174 *2.4. Data Analysis*

175 The binary data matrix was analyzed using Microsoft-Excel 2007. It estimates the banding
176 characteristics; such as *(i)* Total number of bands obtained (TNB); *(ii)* Number of polymorphic bands
177 (NPB); *(iii)* Percentage Polymorphism (PP); *(iv)* Polymorphic Information Content (PIC); *(v)* Band
178 Informativeness (I_B); and *(vi)* Resolving Power (R_P) for each primer used {PP = NPB/TNB generated
179 by each primer; PIC = $1-\sum (P_{ij})^2$, where P_{ij} is the frequency of the i^{th} pattern revealed by j^{th} primer
180 summed across all patterns revealed by the primers [46]; $I_B = 1 - [2 \times (0.5 - P)]$, where P represents
181 the proportion of accessions comprising the band; $R_p = \sum I_B$ [47]}.

Genetic diversity in acid lime was computed using Numerical Taxonomy and Multivariate System (NTSYS, version 2.21i, New York, USA). Similarity indices were calculated applying a similarity to qualitative data. From these similarity indices, sequential, agglomerative, hierarchical and nested (SAHN) clustering was performed using the unweighted pair group method of arithmetic averages (UPGMA) algorithm [48]. Similarity coefficients were computed-based on 3-different measures: Simple Matching Coefficient (SM) [49], Dice's Coefficient of Similarity (D) [43,50] and Jaccard's Coefficient (J) [44]. The matrices of SM, J and D coefficients were compared by Mantel test [51] using MXCOMP option in NTSYS program. The cophenetic correlation test was applied for estimating the correlation between each of the similarity matrix and its corresponding phenogram. The estimated correlation coefficient values show the goodness of fit of cluster analysis performed on the basis of each of SM, J and D. In order to evaluate trees constructed from UPGMA clustering by genetic similarity coefficients, Consensus Fork indices (CI_C) were calculated using Strict Consensus method of NTSYS program for each combination of similarity coefficient and UPGMA clustering. CI_c measures how resolved the tree is [47]. The best-fitted similarity matrix coefficient was then employed for the genetic diversity assessment.

The genetic diversity and relationship among the acid lime cultivars of different agro-ecological zones of Eastern Nepal were also studied through a Principal Coordinate Analysis (PCoA) using Multivariate Statistical Package (MVSP, version 3.21) and in terms of Percentage of Polymorphic Bands (PPB), Nei's Genetic Diversity (H) and Shannon's Information Index (I) using Popgene (Version 1.32).

3. Results

3.1.ISSR Polymorphism in Acid Lime

Altoether, 234 loci were amplified by using 21 primers across the 60 acid lime accessions with an average amplification of 9.72 bands per primer (Fig. 1). Among total amplified bands, 204 (87.18%) were found to be polymorphic and 30 (12.82%) were monomorphic. The polymorphic bands produced by different oligos ranged from 55.56 -100% (8 oligos revealed 100% polymorphisms). The number of scorable bands produced per primer ranged from 7-18 with variation in amplicon size ranging from 250-3,200 bp. The highest number of ISSR loci (18) was produced by primer UBC 857, whereas the lowest number of ISSR loci (7) was produced by primer C1 and UBC 834 in the total accessions (Suppl. Table 2).

The Polymorphic Information Content (PIC) value ranged from 0.74 (UBC 807) to 0.93 (UBC 857) with an average of 0.85. The Band informativeness (I_B) of the 21 ISSR primers ranged from 0.42 (C 4) to 1.77 (C 1) with an average of 1.12 and the Resolving Power (R_P) ranged from 4.63 (UBC 807) to 23.16 (UBC 857) with an average of 12.03. Here we found, the primers that have the highest PIC value also gave the highest R_P score (Suppl. Table 2).

3.2.Genetic Diversity in Acid Lime Cultivars

The varied range of similarity indices 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 D (0.57 - 0.95) with an average similarity coefficient value of 0.79, 0.69 and 0.81 respectively. The Mantel test (Matrix comparison) result of original matrices showed a correlation value between J and D to be the highest and significant (0.99710) in comparison to SM and J (0.98143) and SM and D (0.98318) (Suppl. Table 3).

The highest CI_C value (CI_C = 1.00000) was observed for J and D coefficients (Suppl. Table 4). Cophenetic correlation coefficient value (r) between the genetic similarity matrices and cophenetic matrices are presented in Suppl. Table 5. Unweighted Pair Group Method of Arithmetic Averages (UPGMA) distance for D coefficient gave the highest Cophenetic Correlation value (r = 0.90356) (Suppl. Table 5). Compared to J and SM coefficients, D coefficient was evaluated to be the best coefficient for deducing the genetic diversity and relationship among various acid lime cultivars as shown by the highest cophenetic correlation coefficient value (r = 0.90356) indicating very good fit for cluster analysis (Suppl. Table 5).

On comparative analysis made for the similarity coefficients, Dice coefficient was revealed to be the best, which was subjected further for interpreting genetic diversity, and the relationships among various accessions of acid lime representing different geographical gradients. Based on the Dice similarity coefficient, genetic similarity within 60 acid lime accessions ranged from 57% to 95% with an average of 81% (Suppl. Table 6). The individual genetic similarity/distance among various *C. aurantifolia* accessions have been assessed from the pair-wise comparison of Dice similarity matrix, which revealed that High-hill accessions LT-21 and LT-20 to be the most genetically similar (0.95) and Terai accessions LS-56 and LS-35 to be the most genetically distant (0.571) genotypes. Considering these similarity indices, Terai accessions were shown to have wider genetic base followed by Mid-hill and High-hill.

Genetic diversity of acid lime germplasm from different zones assessed based on Percentage of Polymorphic Band (PPB), Nei's Gene Diversity (H) and Shannon's Information Index (I) using Popgene ver. 1.32 revealed highest diversity indices t in Terai accessions (PPB, 69.660%; H, 0.215; I, 0.325) followed by Mid-hill (PPB, 66.670%; H, 0.202; I, 0.308) and High-hill (PPB, 55.130%; H, 0.173; I, 0.262) (Suppl. Table 7).

3.3.The genetic relationship based on UPGMA Cluster Analysis and PCoA

The 60 genotypes were separated into two major clusters (I and II) and three minor clusters (III, IV, and V) in the phenogram (Fig. 2). The accessions from High-hill, Mid-hill, and Terai zones were intermingled in different clusters. The cluster I comprised of 30 accessions from High-hill and Mid-hill agro-ecological zones. In this cluster, the highest genetic similarity coefficient was observed for the accession LT-20 and LT-21 (0.95), and the lowest similarity coefficient (i.e., highest genetic distance) was observed between LS-56 and LS-35 (0.571). Cluster II comprised of 25 accessions from High-hill, Mid-hill and Terai agro-ecological zones along with the exotic varieties of Vanarasi, Madrasi and Rampur (LKv-60, LKm-61, and LKr-62 respectively). In this cluster, accessions LS-37 and LS-39 had the highest similarity value of 0.993 followed by 0.940 (between LS-42 and LD-45), 0.927 (between LD-48 and LD-50) and so on. Cluster III and cluster IV consisted of single accession, LT-9, and LD-59 from High-hill and Mid-hill respectively. Cluster V comprised of three accessions (LS-56, LD-58, and LS-57) that belongs to Terai agro-ecological zone (Suppl. Table 8). The cluster II has been separated from cluster I at a similarity coefficient of 0.803 and cluster IV separated from rest of the group at a similarity coefficient of 0.66. There was only a small genetic variation between cluster groups I and II (similarity %, 81.4 and 81.8 respectively) and clusters IV and V (similarity %, 73.7 and 73.9 respectively), whereas wider variation was observed between cluster IV and II (Fig. 2).

Two-dimensional plots of the principal coordinate analysis (PCoA) classified the 60 acid lime accessions based on ISSR allelic variation (Fig. 3). The first principal co-ordinate axis accounted for 14.51%) (Eigen value = 323.27; percentage of variance = 14.51%) and second accounted for 8.34% (Eigen value = 185.77; percentage of variance = 8.34%) of the total genetic variation with a cumulative variation of 22.85%. Therefore, groups were discriminated, with axes 1 and 2 expanding 22.85% of the total variation.

4. Discussion

4.1.ISSR polymorphism and genetic diversity estimation in Nepalese Acid Lime

Polymorphism reported in *Citrus* spp. is comparable with our present investigation (PP=87.18%) such as 89.4% in wild *Citrus* spp. [24], 87% in *C. indica* [52], and 100% of few commercially important *Citrus* spp. [53]. The total amplification profiles generated by the 21 ISSR primers yielded 234 bands of which 204 were polymorphic and 30 were monomorphic, which gives us a clue about the existence of high level of genetic diversity among selected acid lime germplasm from 3-different ecological zones of Eastern Nepal. RAPD screens whole genome as revealed by 94.94% polymorphism in corresponding samples, much higher compared to present findings [41]. However, ISSR amplicons correspond to specific inter SSR loci and ISSR- PCR is more stringent than RAPD because of the use of longer oligos (16-25 bp), that allows the use of high annealing temperatures [22]. PIC value provides information about heterozygosity and is associated with the degree of polymorphism. Primers with comparably higher PIC values are useful in discriminating accessions [54]. In our study, the PIC value ranged from 0.74 to 0.93 with an average of 0.89. The highest PIC value of 0.93 was observed for primer UBC 857 and lowest of 0.74 for UBC 807 with an average value of 0.85. PIC value of >0.80 indicates their usefulness for the assessment of genetic diversity of acid lime accessions (Suppl. Table 2). Interestingly, PIC value of the SSR based study of the similar germplasm revealed comparatively low values [39], which might be due to SSR markers being codominant and specific PCR-based marker system. Also, as the SSR-based study was conducted in conventional Agarose Gel Electrophoretic system (in contrast to Polyacrylamide or Capillary Electrophoresis), small allele size differences (varying in few bases) might not have been properly resolved. Capillary Electrophoresis has been shown to be the superior technique for SSR-based genetic analysis [55]. However, the PIC value of the present study is comparable to that reported in our previous RAPD-based study (ranged from 0.78 to 0.88 with an average of 0.80) [41]. Our current investigation revealed the primer (UBC 857) that had highest PIC value (0.93) to have highest Resolving power value (23.16) (Suppl. Table 2), which provides us a glimpse of quantitative data that allows us to make direct comparisons between the primers [47].

The clustering based on UPGMA analysis revealed the genetic diversity and relationship among acid lime accessions of three geographically diverse agro-ecological zones. No specific cluster was formed for accessions from different agro-ecological zones under study. Our result is congruent with others finding of being separated into two major and three minor clusters, apart from the distribution of different accessions in different clusters [39]. Cluster I of present study comprised of accessions from High-hill and Mid-hill zones showing the close genetic relationship, whereas cluster II comprised of accessions from all three agro-ecological zones and the exotic varieties *viz*., Vanarasi, Madrasi and Rampur (LKv-60, LKm-61, and LKr-62 respectively). These exotic varieties were also

clustered together on phenogram generated from SSR markers [39] and RAPD markers [41]. Accessions (LS-56, LS-57and LD-58) from Terai ecological zone are clustered together, similar to the results obtained in SSR-based phenogram by [39], indicating their genetic closeness. The intermixing of accessions that are grown in different agro-ecological domains in different clusters in the phenogram may be attributed to the genetic similarities among different accessions in various qualitative and quantitative traits. In order to improve varieties, the ideal parent for hybridization should be selected based on the level of genetic diversity estimated using molecular markers [56]. The high usage of morphological traits for the determination of genetic relationship among plants and its varieties exists [57]. However, morphological markers do not often reflect genetic relationships because of their interaction with the environment and epistasis [58]. On the basis of previous study on fruit diversity and vitamin C content, four elite accessions [two from high-hills, and one each from Terai and Mid-hills were confirmed to be of superior quality and recommend for conservation, breeding and various developmental purposes [59]. In our present investigation, the first two accessions (LT-17 and LT-23) are clustered together in I group and remaining two (LD-49 and LM-44) are clustered in II group.

Dice similarity matrix based genetic diversity estimates within each of the three agro-ecological zones revealed wide genetic base in accessions from Terai agro-ecological zone (0.57-0.94) in comparison with Mid-hill (0.70-0.94%) and High-hill (0.75-0.95%). However, regarding between agro-ecological zones genetic base, the highest value was observed between High-hill and Terai accessions (57-95%) followed by Mid-hills and Terai accessions (57-94%). Terai accessions being most genetically diverse had a comparable genetic base to that of Mid-hill vs. Terai and High-hill vs. Terai. The result is comparable with the result obtained using SSR markers where accessions from high and mid-hills have the high average genetic similarity (73% and 81%) in comparison to Terai (69%) [39]. Our results showed that the collected accessions from different agro-ecological zones were not genetically distinct but highest diversity was observed in Terai accessions compared to in High and Mid-hills. In our investigation, the diversity indices like Shannon's information index (I) and Nei's gene diversity (H) were found to be 0.325 and 0.215 respectively in Terai agro-ecological zone which was highest among the three zones studied. This indicates Terai to have diverse gene pool, compared to Mid and High-hills. This discrepancy in diverse gene pool could be due to higher accessibility for the movement of germplasm in Terai within the country, and also from neighboring country India. In contrast, lower level of genetic variability as was observed in Mid-hills and High-hills might be due to acid lime trees being established in natural conditions in these zones [39]. Our

Peer-reviewed version available at Plants 2018, 7, 46; doi:10.3390/plants7020046

351 previous study using similar germplasm based on RAPD markers shows almost similar diversity
352 indices values [41].
353
354 *4.2.Application of ISSR-based Genetic Diversity Estimates in Acid Lime Breeding Program*
355 In the present investigation, ISSR dominant markers were utilized for the genetic diversity analysis
356 of acid lime cultivars of eastern Nepal. The advantage in the use of ISSR markers in plant breeding
357 lies in their linkage to SSR loci. Although microsatellites themselves are probably non-functional
358 and selectively neutral, they are linked to the coding regions, so that ISSRs mark the gene-rich
359 regions [60]. Also, because of quicker ISSR analysis procedure that detects over hundreds of bands
360 per primer, this has gained increased attention to both the plant genetics and the breeders' alike [61].
361 To date, more polymorphism has been detected with the use of ISSRs than any other assay procedure
362 [62-64]. However, there are possibilities that the fragments with the same mobility in electrophoresis
363 originate from non-homologous regions, which can contribute to some distortion in the estimates of
364 genetic similarities [61,65].
365
366 Many qualitative and quantitative agronomic traits such as high juice content, fruit size, disease and
367 insect resistance have the genetic basis of inheritance and can be enhanced by the use of molecular
368 markers and marker-assisted selection (MAS) technique. Selection and Improvement of good
369 qualitative and quantitative traits are important steps in the variety of developmental programs.
370 Moreover, breeding of good quality traits requires selection of parents with a wider genetic diversity
371 [66]. For this, sufficient knowledge about genetic diversity in the gene pool is required to adopt the
372 efficient and valuable breeding approach.
373
374 Even though tremendous demand for Acid lime exists in Nepal, the country is not self-reliant in its
375 sole production, thus import of bulk quantities is required to quench the thirst of growing population.
376 Nepal is enriched with a favorable geo-climatic condition for acid lime cultivation, however, its
377 production per hectare is comparatively very low because of the lack of high yielding varieties, rapid
378 diseases and pests' infestations, poor agronomic practices, not an introduction of disease-resistant
379 cultivars, etc. Our current study assesses the genetic diversity of acid lime germplasm of eastern
380 Nepal using ISSR marker technique and underscores the need for the conservation of these resources
381 by the development of elite cultivars.
382
383 Using representative accessions of acid lime from three different agro-ecological zones, we
384 investigated the genetic diversity of Nepalese Acid Lime using ISSR marker. Moderately overall

high genetic polymorphism (87.18%) was detected using 21-primers. The cluster analysis revealed heterogeneous grouping into 2-major and 3-minor clusters and no clustering according to geographical locations was evident. Acid lime elite cultivars selected in the previous study using the same 60 accessions can now be used as a breeding material or genetic stocks for the future breeding program and genetic distance between these and others can be used for parental selection for breeding. In a nutshell, Genomics-Assisted breeding and Metabolomics-assisted breeding strategies may provide tremendous breakthrough and have higher scope in plant breeding.

Acknowledgments

We owe our thanks to Ministry of Science and Technology (MOST), Government of Nepal for providing financial aid for this project. Authors also highly acknowledge staffs of Nepal Academy of Science and Technology (NAST) and to all the members of Molecular Biotechnology Laboratory of NAST for their kind cooperation and support until the accomplishment of the project.

Competing interest

The authors declare to have no competing interest.

Author's contributions

NNM designed the experiment. NNM, NR, TB, and SS performed the experiment. NNM, RLS, BB, BKJ, and SS performed the data analysis. NNM, BB, and SS drafted the manuscript.

References

1. Velasco, R.; Concetta, L. A genealogy of the citrus family. *Nat Biotechnol* **2014**, *32*, 640-642.
2. Hvarleva, T.T.; Kapari-Isaia, L.; Papayiannis, A.; Atanassov, A.; Hadjinicoli, A.; Kyriakou, A. Characterization of citrus cultivars and clones in cyprus through microsatellite and rapd analysis. *Biotechnol. Biotec. Eq* **2008**, *22*, 787-794.
3. Ranjit, M.; Gokarna, G.C. *Citrus research and development action plan*. Chemonics International Inc: USA, 1997.
4. Nicolosi, E.; Deng, Z.N.; Gentile, A.; La Malfa, S.; Continella, G.; Tribulato, E. Citrus phylogeny and genetic origin of important species as investigated by molecular markers. *Theor. Appl. Genet* **2000**, *100*, 1155–1166.
5. Barett, H.C.; Rhodes, A.M. A numerical taxonomic study of affinity relationship in cultivated citrus and its clone relatives. *Syst Bot* **1976**, *1*, 105-136.

6. Singh, A.K. Probable agricultural biodiversity heritage sites in india: V. The garo, khasi, and jaintia hills region. *Asian Agri-History* **2010**, *14*, 133-156.
7. Dhillon, B.S.; Randhawa, J.S. *Fruit growth and development in citrus: Advance in horticulture*. Malhotra Publishing House: Mew Dehli, 1993; Vol. 3.
8. Penniston, K.L.; Nakada, S.Y.; Holmes, R.P.; Assimos, D.G. Quantitative assessment of citric acid in lemon juice, lime juice, and commercially-available fruit juice products. *J Endourology* **2008**, *22*, 567-570.
9. Aras, B.; Kalfazade, N.; Tugcu, V.; Kemahh, E.; Ozbay, B.; Polat, H.; Tasci, A.I. Can lemon juice be an alternative to potassium citrate in the treatment of urinary calcium stones in patients with hypocitraturia? A prospective randomized study. *Urol Res* **2008**, *36*, 313-317.
10. NCRP. *Annual report 2067/68 (2010/11)*; Paripatle, Dhankuta, 2012.
11. Paudyal, K.P.; Shrestha, R.L. In *Diversity study and selection of lime (cirtus aurantifolia) genotypes for expansion of production period in nepal*, 4th National Horticulture Workshop Kathmandu, Nepal, 2004; Nepal Agriculture Research Council: Kathmandu, Nepal.
12. Dhakal, D.D.; Bhattarai, S. Production system of lime and lemon in nepal: A survey report. In *Proceedings of Workshop on 'Production and Marketing of Lime and Lemon in Nepal' organized by IAAS*, Rampur at Lamjung campus 2002; pp 26-28.
13. MoAC. *Statistical information of nepalese agriculture. Agri. Business promotion and statistical division, ministry of agriculture and cooperatives*; Singh Durbar, Kathmandu, Nepal, 2011.
14. FAO. *Food and agriculture organization of the united nations developments in international citrus trade in 2004-2005*; 2006.
15. Varshney, R.K.; Graner, A.; Sorrells, M.E. Genic microsatellite markers in plants: Features and applications. *Trends Biotechnol* **2005**, *23*, 48-55.
16. Rauf, S.; Iqbal, Z.; Shahzad, M. Genetic improvement of citrus for disease resistance. *Arch. Phytopathol. Plant Protection* **2013**, *46*, 2051-2061.
17. Dominquez, A.; de Mendoza, A.; Guerri, J.; Cambra, M.; Navarro, L.; Moreno, P.; Pena, L. Pathogen-derived resistance to citrus tristeza virus (ctv) in transgenic mexican lime (citrus aurantifolia (christ.) swing.) plants expressing its p25 coat protein gene. *Mol. Breed* **2002**, *10*, 1-10.
18. Viloria, Z.; Grosser, J.W. Acid citrus fruit improvement via interploid hybridization using allotetraploid somatic hybrid and autotetraploid breeding parents. *J. Am. Soc. Hort. Sci* **2005**, *130*, 392-402.

19. Sapkota, D.P. Characterization and evaluation of acid lime landraces at rampur, chitwan condition. Tribhuvan University, Chitwan, Nepal, 2006.
20. Varshney, R.K.; Graner, A.; Sorrells, M.E. Genomics - assisted breeding for crop improvement. *Trends Plant Sci* **2006**, *10*, 621-630.
21. Nei, J.; Colowit, P.M.; Mackill, D.J. Evaluation of genetic diversity in rice sub species using microsatellite markers. *Crop Sci* **2002**, *42*, 601-607.
22. Reddy, M.P.; Sarla, N.; Siddiq, E.A. Inter simple sequence repeat (issr) polymorphism and its application in plant breeding. *Euphytica* **2002**, *128*, 9–17.
23. Shrestha, S. Molecular systematic of weedy *sporobolus* species of australia. The University of Queensland, Australia, 2001.
24. Shahsavar, A.R.; Izadpanah, K.; Tafazoli, E.; Tabatabaei, B.E. Characterization of citrus germplasm including unknown variants by inter-simple sequence repeat (issr) markers. *Sci. Hortic* **2007**, *112*, 310-314.
25. Fang, D.Q.; Roose, M.L. Identification of closely related citrus cultivars with inter-simple sequence repeat markers. *Theor Appl Genet* **1997**, *95*, 408–417.
26. Baig, M.N.R.; Grewal, S.; Dhillon, S. Molecular characterization and genetic diversity analysis of citrus cultivars by rapd markers. *Turk. J. Agric* **2009**, *33*, 375-384.
27. Kijas, J.M.H.; Thomas, M.R.; Fowler, J.C.S.; Roose, M.L. Integration of trinucleotide microsatellites into a linkage map of citrus. *Theor. Appl. Genet* **1997**, *94*, 701–706.
28. Sanker, A.A.; Moore, G.A. Evaluation of inter–simple sequence repeat analysis for mapping in citrus and extension of the genetic linkage map. *Theor. Appl. Genet* **2001**, *102*, 206–214.
29. Fang, D.Q.; Roose, M.L.; Krueger, R.R.; Federici, C.T. Fingerprinting trifoliate orange germplasm accessions with isozymes, rflps and inter-simple sequence repeat markers. *Theor. Appl. Genet* **1997**, *95*, 211–219.
30. Filho, H.D.C.; Machado, M.A.; Targon, M.L.P.N.; Moreira, M.C.P.Q.D.G.; Pompeu, J. Analysis of the genetic diversity among mandarins (*citrus* spp.) using rapd markers. *Euphytica* **1998**, *102*, 133-139.
31. Novelli, V.M.; Machado, M.A.; Lopes, C.R. Iso-enzymatic polymorphism in *citrus* spp and *p. Trifoliata* (l.) raf. (rutaceae). *Gen. Mol. Biol* **2000**, *23*, 163-168.
32. Zietkiewicz, E.; Rafalski, A.; Labuda, D. Genome fingerprinting by simple sequence repeat (ssr) – anchored polymerase chain reaction amplification. *Genomics* **1994**, *20*, 176–183.
33. Bornet, B.; Branchard, M. Non-anchored simple sequence repeat markers: Reproducible and specific tools for genome fingerprinting. *Plant Mol Biol Rep* **2001**, *19*, 209–215.

34. Tautz, D.; Renz, M. Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Res* **1984**, *12*, 4127–4138.
35. Joshi, S.P.; Gupta, V.S.; Aggarwal, R.K.; Ranjekar, P.K.; Brar, D.S. Genetic diversity and phylogenetic relationship as revealed by inter-simple sequence repeat (issr) polymorphism in the genus oryza. *Theor Appl Genet* **2000**, *100*, 1311–1320.
36. Meloni, M.; Perini, D.; Filigheddu, R.; Binelli, G. Genetic variation in five mediterranean populations of juniperus phoenicea as revealed by inter-simple sequence repeat (issr) markers. *Ann Bot* **2006**, *97*, 299–304.
37. Ratnaparkhe, M.B.; Tekeoglu, M.; Muehlbauer, F.J. Intersimple- sequence-repeat (issr) polymorphisms are useful for finding markers associated with disease resistance gene clusters. *Theor Appl Genet* **1998**, *97*, 515–519.
38. Biswas, M.K.; Chai, L.; Amar, M.H.; Zhang, X.; Deng, X. Comparative analysis of genetic diversity in citrus germplasm collection using aflp, ssap, sampl and ssr markers. *Sci Hortic* **2011**, *129*, 798-803.
39. Shrestha, R.L.; Dhakal, D.; Gautam, D.; Paudyal, K.P.; Shrestha, S. Genetic diversity assessment of acid lime (citrus aurantifolia) landraces in nepal, using ssr markers. *Am. J. Plant. Sci* **2012a**, *3*, 1674-1681.
40. Kumar, M.; Parthiban, S.; SaralaDevi, D.; Ponnuswami, V. Genetic diversity analysis of acid lime (citrus aurantifolia swingle) cultivarrs. *The Bioscan* **2013**, *8*, 481-484.
41. Munankarmi, N.M.; Shrestha, R.L.; Rana, N.; Shrestha, J.K.C.; Shrestha, s.; Koirala, R.; Shrestha, S. Genetic diversity assessment of acid lime (citrus aurantifolia, swingle) landraces of nepal using rapd markers. *Int J App Sci Biotech* **2014**, *2*, 315-327.
42. Sambrook, J.; Russell, D.W. *Molecular cloning: A laboratory manual*. Pub Cold Spring Harbour Laboratory Press: New York. USA, 2001; Vol. III.
43. Dice, L.R. Measures of the amount of ecologuc association between species. *Ecology* **1945**, *26*, 297-302.
44. Jaccard, P. Nouvelles recherche sur la distribution florale. *Bulletin. Societe Vaudoise Sciences Naturelles* **1908**, *44*, 223-270.
45. Transue, D.K.; Fairbanks, D.J.; Robison, L.R.; Anderson, W.R. Species identification by rapd analysis of grain amaranth genetic resources. *Crop. Sci* **1994**, *34*, 1385-1389.
46. Botstein, D.; White, R.L.; Skolnick, M.; Davis, R.W. Construction of genetic map in man using restriction fragment length polymorphisms. *Am. J. Human Genetics* **1980**, *32*, 314-331.
47. Sokal, R.R. *Ecological parameter inferred from spatial correlograms*. ICPH: Marryland, 1979.

519 48. Sneath, P.H.A.; Sokal, R.R. *Taxonomy, the principle and practice of numerical classification.*
520 Pub W.H freeman and company: San Francisco, CA, 1973.
521 49. Sokal, R.R.; Michener, C.D. A statistical method for evaluating systematic relationships. *Uni*
522 *Kansas Sci Bull* **1958**, *38*, 1409-1438.
523 50. Nei, M.; Li, W.H. Mathematical model for studying genetic variation in terms of restriction
524 endonucleases. *Proc. Natl. Acad. Sci. USA* **1979**, *76*, 685-686.
525 51. Mantel, N.A. The detection of disease clustering and a generalized regression approach.
526 *Cancer Res.* **1967**, *27*, 209-220.
527 52. Kumar, S.; Jena, S.N.; Nair, N.K. Issr polymorphism in indian wild orange (citrus indica
528 tanaka, rutaceae) and related wild species in north-east india. *Sci Hortic* **2010**, *123*, 350–359.
529 53. Chibame, K.; Marak.; Laskar, M.A. Analysis of phenetic relationship between citrus indica
530 tanaka and a few commercially important citrus species by issr markers. *Sci Hortic* **2010**,
531 *124*, 345–348.
532 54. Teklewood, A.; Becker, H.C. Geographic pattern of genetic diversity among 43 ethiopian
533 mustard (*brassica carinata* a. Braun) accession as revealed by rapd analysis. *Genet. Res.*
534 *Crop. Evol* **2006**, *53*, 1173-1185.
535 55. Vemireddy, L.R.; Archak, S.; Nagaraju, J. Capillary electrophoresis is essential for
536 microsatellite marker based detection and quantification of adulteration of basmati rice (oryza
537 sativa). *J Agric Food Chem* **2007**, *55*, 8112-8117.
538 56. Jeya Prakash, S.p.; Biji, K.R.; Mimchael G, S.; Ganesha M, K.; Chandra B, R. Genetic
539 diversity analysis of sorghum (*sorghum bicolor* l. Moench) accessions using rapd markers.
540 *Indian J Crop Sci* **2006**, *1*, 109-112.
541 57. Ortiz, R.; Madsen, S.; Vuylsteke, D. Classification of african plantain landraces and banana
542 cultivars using a phenotypic distance index of quantitative descriptors. *Theor. Appl. Genet*
543 **1998**, *96*, 904-911.
544 58. Smith, J.S.C.; Smith, O.S. The description and assessment of distance between inbred lines of
545 maize, 11. The utility of morphological, biochemical and genetic descriptots and a scheme for
546 testing of distinctiveness between inbred lines. *Maydica* **1998**, *34*, 151-161.
547 59. Shrestha, R.L.; Dhakal, D.; Gautam, D.; Paudyal, K.P.; Shrestha, S. Study of fruits diversity
548 and selection of elite acid lime (*citrus aurantifolia* swingle) genotypes in nepal. *Am. J. Plant.*
549 *Sci* **2012b**, *3*, 1098-1104.
550 60. Kojima, T.; Nagaoka, T.; Noda, K.; Ogihara, Y. Genetic linkage map of issr and rapd
551 markers in einkorn wheat in relation to that of rflp markers. *Theor Appl Genet* **1998**, *96*, 37–
552 45.

61. Godwin, I.D.; Aitken, A.B.; Lawrence, W.S. Application of inter simple sequence repeat (issr) to plant genomics. *Electrophoresis* **1997**, *18*, 1524-1528.
62. Gupta, M.; Chyi, Y.S.; Romero-Severson, J.; Owen, J.L. Amplification of DNA markers from evolutionarily diverse genomes using single primers of simple-sequence repeats. *Theor Appl Genet* **1994**, *89*, 998–1006.
63. Salimath, S.S.; de Oliveira, A.C.; Godwin, I.D.; Bennetzen, J.L. Assessment of genome origins and genetic diversity in the genus eleusine with DNA markers. *Genome* **1995**, *38*, 757–763.
64. Virk, P.S.; Zhu, H.J.; Newbury, G.J.; Bryan, M.T.; Jackson, M.T.; Ford-Lloyd, B.V. Effectiveness of different classes of molecular marker for classifying and revealing variation in rice (*oryza sativa*) germplasm. *Euphytica* **2000**, *112*, 275-284.
65. Sanchez, M.P.; Davila, J.A.; Loarce, Y.; Ferrer, E. Simple sequence repeat primers used in polymerase chain reaction amplifications to study genetic diversity in barley. *Genome* **1996**, *39*, 112-117.
66. Pangali, P.L.; Hossain, M.; Gerpasio, R.V. *Asian rice blows: The returning crises? International rice research institute (irri)*. Los Banus, 1997.

571 **Legends to Figures**

572 Figure 1. ISSR profile generated by Primer UBC 842. Lanes marked with M is 100bp plus molecular
573 weight markers. A: Lanes 1-15 represents acid lime samples 1-15; B: Lanes 16-30 represents acid
574 lime samples 16-30; C: Lanes 31-45 represents acid lime samples 31-45; D: Lanes 46-62k represents
575 acid lime samples 46-62.

576 Figure 2. Phenogram generated for 60 *Citrus aurantifolia* (Acid lime) accessions (please refer to
577 Supplementary table 1 for sample details) by UPGMA cluster analysis using Dice coefficient of
578 similarity computed from 234 ISSR loci generated by 21 primers. The clusters are labeled as I, II, III,
579 IV, and V.

580 Figure 3. Principal Co-ordinate Analysis (PCoA) of Dice similarity matrix carried out with MVSP
581 3.21

582

583 **Legends to Supplementary Tables**

584 Supplementary Table 1. The altitudinal range, accession numbers and locality details of acid lime
585 samples.

586 Supplementary Table 2. ISSR primer details, Total Number of amplified Bands (TNB), Number of
587 Polymorphic Bands (NPB), Percent Polymorphism (PP), amplicon size range, Polymorphic
588 Information Content (PIC), Band Informativeness (I_B) and Resolving power (R_P) computed for
589 different ISSR primers used to generate ISSR profiles in 60 *Citrus aurantifolia* (Acid lime)
590 accessions

591 Supplementary Table 3. Correlation Coefficient values generated from Mantel test of original
592 similarity matrices

593 Supplementary Table 4. Consensus Fork Index (CI_C) values generated for the UPGMA based
594 phenograms using different similarity coefficients

595 Supplementary Table 5. Cophenetic correlation coefficients value (r) obtained for different similarity
596 matrices *viz*. Simple Matching, Jaccard's and Dice

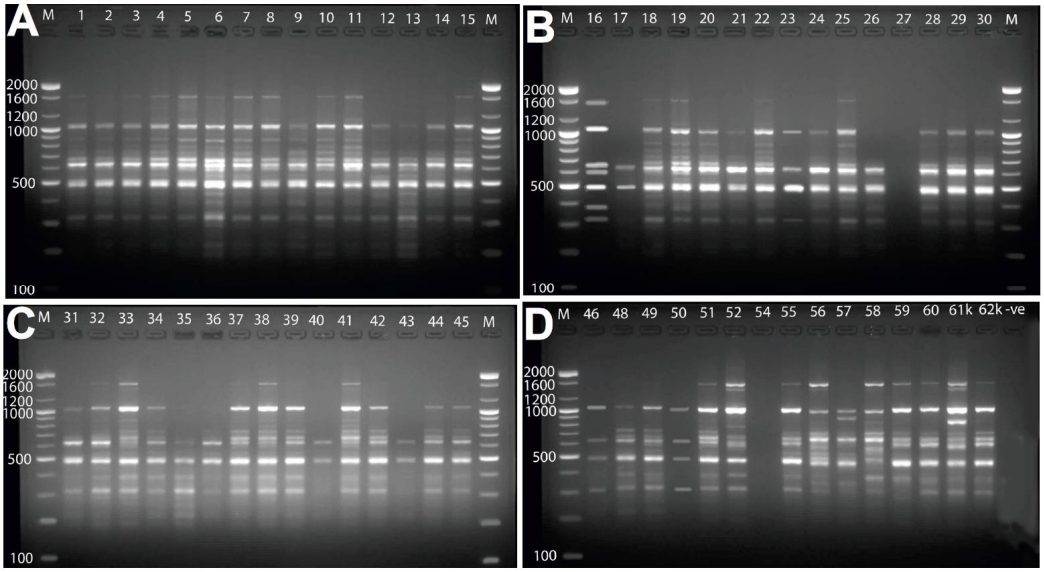
597 Supplementary Table 6. Percentage genetic similarities observed within and between three agro-
598 ecological zones for various acid lime samples based on Dice similarity matrix generated from ISSR
599 profile

600 Supplementary Table 7. Genetic variation observed for *C. aurantifolia* samples representing different
601 agro-ecological zones as revealed by POPGENE ver. 1.32

602 Supplementary Table 8. Acid lime accessions belonging to 2-major and 3-minor clusters of UPGMA
603 phenogram

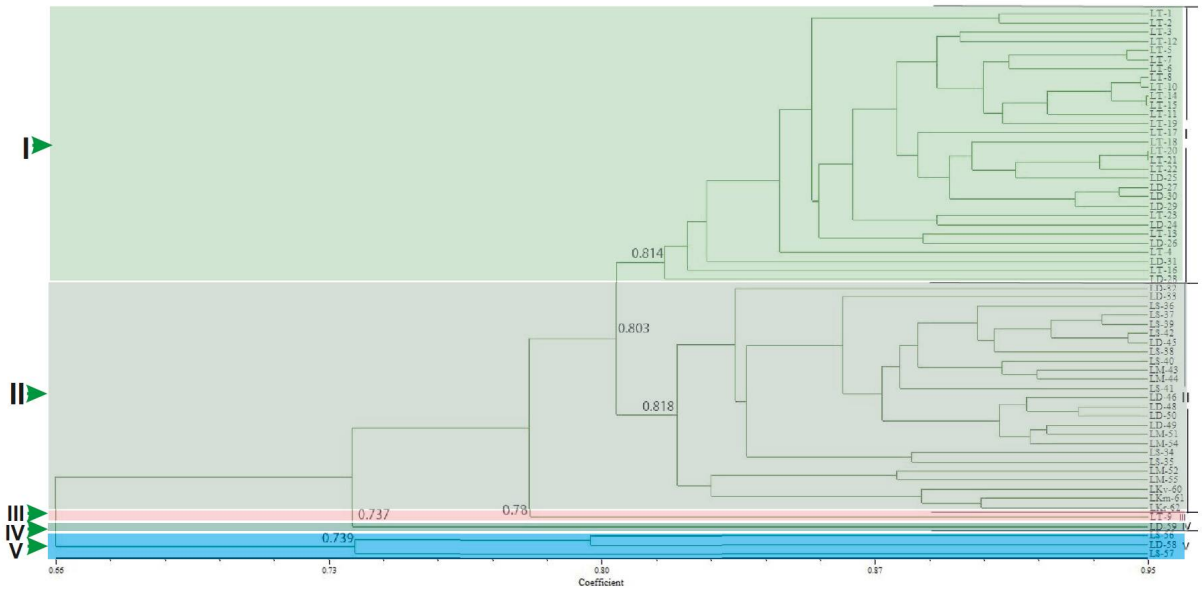
604

605 **Figure 1.**



606
607
608

609 **Figure 2**



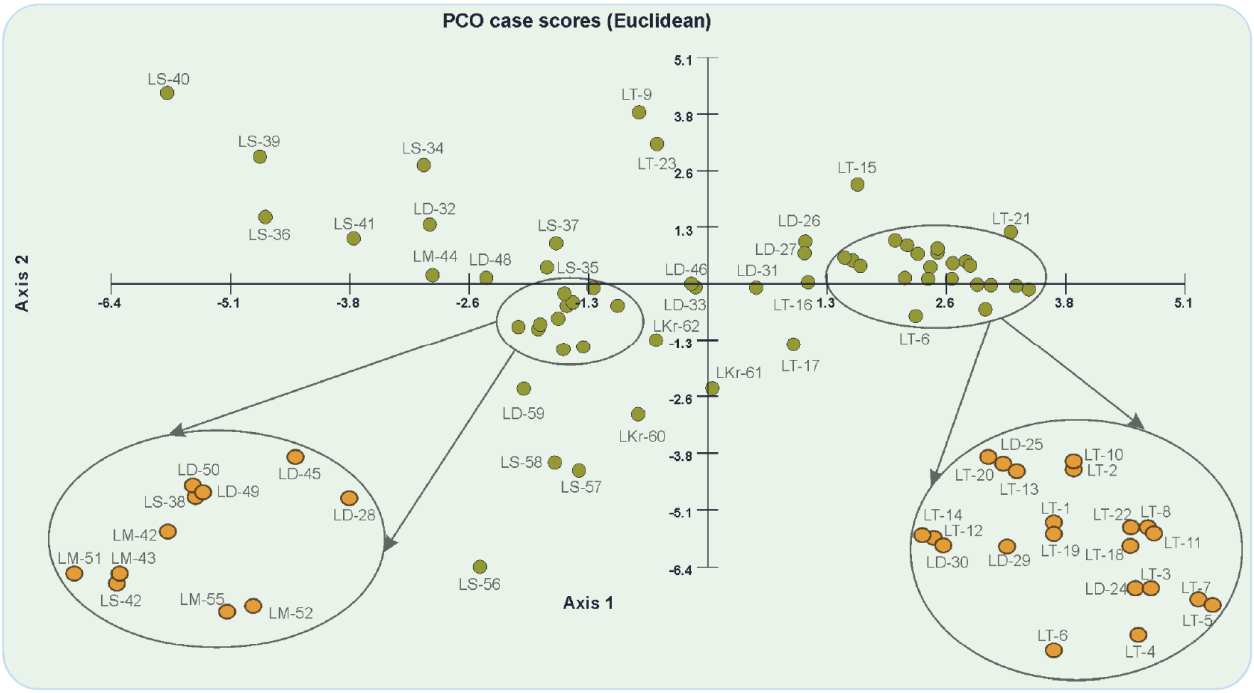
610

611

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 May 2018 doi:10.20944/preprints201805.0144.v1

Peer-reviewed version available at Plants 2018, 7, 46; doi:10.3390/plants7020046

612 **Figure 3**



613
614
615

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 May 2018 doi:10.20944/preprints201805.0144.v1

Peer-reviewed version available at Plants 2018, 7, 46; doi:10.3390/plants7020046

Supplementary Table 1. Altitudinal range, accession numbers and locality details of acid lime samples

Above 1200 m asl			**600-1200 m asl**			**Less than 600 m asl**		
Acc. No.	**Altitude (m asl)**	**VDC-Ward No.**	**Acc. No.**	**Altitude, (m asl)**	**VDC-Ward No.**	**Acc. No.**	**Altitude, (m asl)**	**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			

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 May 2018 doi:10.20944/preprints201805.0144.v1

Peer-reviewed version available at Plants 2018, 7, 46; doi:10.3390/plants7020046

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 district; LD, Lime Dhankuta district; LM, Lime Morang district; LS, Lime Sunsari district; LKm, Lime madras; LKr, Lime Rampur; LKv, Lime Varanasi; VDC = Village Development Committee, m = meter, asl = above sea level

616

617

618

Supplementary Table 2. ISSR primer details, Total Number of amplified Bands (TNB), Number of Polymorphic Bands (NPB), Percent Polymorphism (PP), amplicon size range, Polymorphic Information Content (PIC), Band Informativeness (I_B) and Resolving power (R_P) computed for different ISSR primers used to generate ISSR profiles in 60 *Citrus aurantifolia* (Acid lime) accessions

S.N.	Primer Code	Primer Sequence (5' - 3')	Primer Length (bp)	TNB	NPB	PP	Amplicon size range (bp)	PIC	I_B	R_P
1	C1	TCTCTCTCTCTCTCTCTCCC	20	7	5	71.43	550-1800	0.85	1.77	12.40
2	C2	AGCAGCAGCAGCGT	14	10	10	100	500-3200	0.83	0.83	8.33
3	C4	CTCCTCCTCGC	11	14	14	100	300-2200	0.82	0.42	5.87
4	C5	CACCACCACGC	11	12	11	91.67	600-2500	0.83	0.60	7.23
5	C7	HVHGAGAGAGAGAGAGAT	18	17	15	88.23	250-2000	0.89	0.87	14.87
6	C8	TCCTCCTCCTCCTCCRY	17	9	7	77.78	520-2800	0.87	1.47	13.23
7	C9	BDBTCCTCCTCCTCCTCC	18	9	6	66.67	520-2000	0.86	1.50	13.56

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 May 2018 doi:10.20944/preprints201805.0144.v1

Peer-reviewed version available at Plants 2018, 7, 46; doi:10.3390/plants7020046

8	C10	HVHTCCTCCTCCTCCTCC	18	11	8	72.73	500-2000	0.79	0.78	8.63
9	UBC 807	AGAGAGAGAGAGAGAGT	17	9	9	100	450-1300	0.74	0.51	4.63
10	UBC 810	GAGAGAGAGAGAGAGAT	17	16	16	100	390-1980	0.85	1.00	12.00
11	UBC 812	GAGAGAGAGAGAGAGAA	17	12	9	75.00	450-1500	0.91	1.03	16.60
12	UBC 825	ACACACACACACACACT	17	9	5	55.56	500-2000	0.87	1.51	13.63
13	UBC 834	AGAGAGAGAGAGAGAGYT	18	7	5	71.43	310-1550	0.81	1.28	8.96
14	UBC 835	AGAGAGAGAGAGAGAGYC	18	10	10	100	400-2900	0.86	1.29	12.96
15	UBC 836	AGAGAGAGAGAGAGAGYA	18	10	10	100	320-1450	0.87	1.13	11.30
16	UBC 841	GAGAGAGAGAGAGAGAYC	18	10	10	100	320-1700	0.87	1.08	10.83
17	UBC 842	GAGAGAGAGAGAGAGAYG	18	12	12	100	320-1650	0.89	1.04	12.53
18	UBC 857	ACACACACACACACACYG	18	18	17	94.44	300-3000	0.93	1.28	23.16

19	UBC 873	GACAGACAGACAGACA	16	14	13	92.86	470-3000	0.89	0.96	13.43
20	UBC 888	BDBCACACACACAGACA	17	9	7	77.78	480-1450	0.87	1.47	13.23
21	UBC 889	DBDACACACACACACACA	18	9	5	55.56	480-1400	0.87	1.70	15.33
	Total			234	204		Average	0.85	1.12	12.03
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

621

622

623

Supplementary Table 3. Correlation Coefficient values generated 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)			******

624

625

626

Supplementary Table 4. Consensus Fork Index (CI_C) values generated for the UPGMA based phenograms using different similarity coefficients

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

627

628

629

Supplementary Table 5. Cophenetic correlation coefficients value (r) obtained for different similarity matrices *viz*. Simple Matching, Jaccard's and Dice

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

630

631
632

633

Supplementary Table 6. Percentage genetic similarities observed within and between three agro-ecological zones for various acid lime samples based on Dice similarity matrix generated from ISSR profile

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

634 Zone = Agro-ecological zone

635

636

637 **Supplementary Table 7. Genetic variation observed for *C. aurantifolia* samples representing different agro-**
638 **ecological zones as revealed by POPGENE ver. 1.32**

Population of *C. aurantifolia*	**Sample Size**	**Number of polymorphic bands**	**PPB (%)**	**H**	**I**
High-hill	20	129.00	55.130	0.173	0.262
Mid-hill	21	156.00	66.670	0.202	0.308
Terai	19	163.00	69.660	0.215	0.325
Average		149.33	63.820	0.197	0.300
Species level (Multipopulation)	60	204	87.180	0.223	0.348

Where, PPB = Percent Polymorphic Bands/loci, H = Nei's gene diversity, I = Shannon's information index

639

640

641 **Supplementary Table 8**. Acid lime accessions belonging to two major and three minor clusters of UPGMA phenogram

S.N.	Clusters	Accessions
1	I	LT-1, LT-2, LT-3, LT-12, LT-8, LT-10, LT-14, LT-15, LT-11, LT-19, LT-5, LT-7, LT-6 , LT-17, LT-18, LT-20, LT-21, LT-22, LD-25, LD-27, LD-30, LD-29, LT-23, LD-24, LT-13, LD-26, LT-4, LD-31, LT-16, LD-28
2	II	LD-32, LD-33, LS-36, LS-37, LS-39, LS-42, LD-45, LS-38, LS-40, LM-43, LM-44, LS-41, LD-45, LD-48, LD-50, LD-49, LM-51, LM-54, LS-34, LS-35, LM-52, LM-55, LKv-60, LKm-61, LKr-62
3	III	LT-9
4	IV	LD-59
5	V	LS-56, LD-58, LS-57

642

643
644

645