Identifying morphological diversity and superior germplasm of *Citrus reticulata* Blanco (Mandarin) in Sikkim and Darjeeling Himalayas, India: an application of DIVA-GIS and cluster analysis

Subhankar Gurung, Arun Chetri*, Meera Tamang and Mamta Chettri

Taxonomy and Biodiversity Laboratory, Department of Botany, Sikkim University, Gangtok, Sikkim, 737102, India

*Corresponding author: achettri01@cus.ac.in

Abstract

*Citrus reticulata* is an important cash crop for the farmers in the Sikkim and Darjeeling Himalayas, India. The lack of knowledge of its diversity has only resulted in the lack of uniformity in the fruit quality. It has become imperative to identify superior varieties that meet the demands of the market to improve the citrus industry. Hence, a total of 105 accessions of mandarin were collected randomly from different locations to assess the morphological diversity using cluster analysis and DIVA-GIS. The orchards were randomly visited in each district of the state of Sikkim and two hilly districts of the state of West Bengal (WB). A sample tree was selected from each orchard and the quantitative and qualitative characters of its leaves, flowers, and fruits were measured. A significant variation was observed in the quantitative characters with a positive correlation between fruit weight and length, fruit diameter and weight, total soluble solids (TSS)/acidity and fruit diameter. The first 6 components of Principal component analysis (PCA) exhibited 69.34% of the total variation. DIVA-GIS showed the highest diversity index for fruit weight, fruit diameter and TSS/acidity in East district, Sikkim. The highest coefficient variation for fruit diameter was observed in the East district and Darjeeling district, WB and TSS/acidity and fruit weight in East district. The dendrogram generated divided the accessions into two major clusters. The grid maps generated identified diverse accessions in the East district and Darjeeling district, which can be a source of superior germplasm.

**Keywords**: *Citrus reticulata*; morphological; germplasm; quantitative; diversity; DIVA-GIS.

**Abbreviations**: DIVA-GIS_Data-Interpolating Variational Analysis – Geographical Information System.

Introduction

*Citrus* belongs to the family, Rutaceae and is one of the most important cash crops in the world (Swingle and Reece, 1967). The origin of *Citrus* cannot be determined but reports suggest that it originated from the south and southeast tropical regions of Asia (Moore, 2001; Sharma et al., 2004; Ladaniya, 2008; Singh, 2010). The natural and commercially cultivated citrus include sweet orange, mandarin, lemons and lime. The North-Eastern part of India is a treasure house of citrus species and their varieties (Sharma et al., 2004) where 11.7% of its total area is under its cultivation (NHB, 2006). Among all the citrus fruits, mandarin *Citrus reticulata* Blanco, is an important cash crop and constitutes about 41% of the total citrus fruits that are produced in India (Das et al., 2005). The North-Eastern part of India comprises of diverse forms of mandarin and could be considered as its center of origin (Ray and Deka, 2000). Mandarin is the most important cash crop in Sikkim (Kishore et al., 2010) and the adjoining Darjeeling hills. Sikkim alone contributes 2% of the total citrus production and covers an area of 5% (NHB, 2006). Mandarin is mainly grown in the lower hills of Sikkim and Darjeeling Himalayas under sub-tropical humid climatic conditions (Kishore et al., 2010). Anecdote from the farmers of mandarin suggests that mandarin is cultivated for several decades. Unfortunately, due to the lack of data and knowledge about the varieties that their forefathers sowed in the field, very little is known about mandarin varieties and their relatives. This has resulted in Sikkim mandarin being referred to as Darjeeling mandarin and vice versa (Gurung et al., 2017). There are several reports which suggest that *C. reticulata* Blanco comprises of ecotypes which include Darjeeling mandarin and Khasi mandarin (Das et al., 2004), while there are certain reports that suggest *C. reticulata* comprise of Sikkim mandarin as well (Kishore et al., 2010). Therefore, in order to clear this confusion, it is imperative to closely observe the morphological characters of the cultivated mandarin species and derive a relationship between them. Furthermore, natural hybridization and spontaneous mutations are a common phenomenon in *Citrus* spp. (Dorji and Yapwattanaphun, 2011) which has led to the diversification of the species over the years. The evidence of the occurrence of zygotic twins (Das et al., 2007) attributes to greater variation and lack of uniformity in fruit quality in the plant (Dorji and Yapwattanaphun, 2011). The level of diversity that exists is
unclear and a comprehensive study related to it has not been conducted in Sikkim and Darjeeling Himalayas. To achieve this, morphological characters are used to identify accessions that are superior in quality. DIVA-GIS has been used to identify potential areas for collection of mandarin germplasm to resolve the problem of lack of uniformity of fruit quality. This would ensure proper utilization of the available diversity. DIVA-GIS is a Geographic Information System that assists the plant genetic resources and biodiversity communities to map the range of distribution of species in which they are interested (Hijmans et al., 2002). DIVA-GIS has been used to identify areas of high diversity in wild potatoes (Hijmans et al., 2000), Jatropha curcas (Sunil et al., 2009), Piper (Parthasarathy et al., 2006). Mandarin ensures livelihood security to the farmers of the region whose marketability is determined by the increase in fruit size and the ratio of total soluble solids (TSS) to acidity ratio (Dorji and Yapwattanaphun, 2011). The identification of superior accessions could help in breeding programs, ensure the quality of fruits that meet the demand of the market. Hence, this study focuses on studying the diversity of mandarins grown in Sikkim and Darjeeling Himalayas in India using DIVA-GIS for morphological traits and to study the relationship among accession using cluster analysis.

Results

Quantitative traits

The statistical analysis of quantitative characters showed high significant variations among the leaves, flowers, and fruits of different accessions except for the number of segments in a fruit and leaf length/width ratio (Table 2). S-02-08 (South district, Sikkim) had the longest leaf lamina length while D-02-05 (Kalimpong district, WB) had the lowest (Supplementary - Appendix, Table S2). The broadest leaf was recorded in S-03-08 (East district, Sikkim) while the narrowest leaf was recorded in D-01-18 (Darjeeling District, WB). Similarly, the heaviest fruit was recorded in S-03-32 (East district) while S-01-02 (North district, Sikkim) had the lowest fruit weight (Supplementary - Appendix, Table S1). The fruits in D-01-07 (Darjeeling district) were found to have the largest diameter while the fruits in S-01-01 (North district) had the smallest fruit diameter. The highest %TSS was recorded in S-01-11 (North district) while the highest acidity was recorded in S-02-11 (South district). The %TSS to acidity ratio was found to be highest in D-01-09 (Darjeeling District) while the lowest was recorded in S-02-11 (South district). Estimated correlation coefficient showed that TSS/acidity is positively correlated with the diameter of the fruit (0.473*), height is positively correlated with weight (0.772*), diameter is positively correlated with weight (0.549*), and the number of seeds showed a positive correlation with the diameter (0.423*) (Fig. 1).

Principal component analysis

PCA revealed that the six-axis contributed to 69.34% to the total variations. Flower traits such as petal length (0.144), number of stamens (0.098), pedicel length (0.068) and petal width (0.05) were the most important traits that contributed to the variation of 27.15% in PC1. Leaf length (3.379), leaf width (1.001), weight (0.947) and height (0.556) were the most important traits accounting for 16.88% variation in PC2. PC3 was influenced by fruit and flower traits, contributing to 11.54% of the variation. PC4, PC5 and PC6 accounted for 8.35%, 7.28% and 6.50% of the variation respectively (Table 2, Fig. 2).

Cluster analysis

A dendrogram generated from unrooted neighbor-joining method divided 10S accessions into two major clusters (cluster I and cluster II; Figure 4). Further, the accessions from Darjeeling, east and west district formed a distinct sub-cluster while the other accessions clustered with each other irrespective of their geographical association (Fig. 3).

Grid-Mapping

The highest diversity index for TSS/acidity, fruit weight and diameter were observed in the East district, Sikkim (Fig. 4, 5, 6). The highest coefficient of variation in fruit weight was observed in Darjeeling district and East district, Sikkim (Fig. 6) and TSS/acidity and fruit diameter in Darjeeling district (Fig. 4, 5).

Discussion

The study of phenotypic diversity on the basis of morphological characters is less favored due to the morphological plasticity that it exhibits (Dorji and Yapwattanaphun, 2011). However, several studies that include both morphological and molecular markers have suggested them to be independent of each other (Koehler et al., 2003; Campos et al., 2005). Furthermore, studies have reported that there are multiple genes that control the agronomic traits and the ideal way to assess it is through morphological means (Liu and Deng, 2007). GIS mapping helps in predicting distribution, documentation, analyzing diversity and variation, developing new strategies for conservation (Abrahama et al., 2010). It has also become an important tool to visually elucidate the diverse nature of a species and identify its location (Flemons et al., 2007). The grid maps generated through DIVA-GIS showed wide variations and diversity within and between the districts in our present study which suggests that environmental factors alone do not trigger morphological variation. Cross-pollination, propagation through seed without nucellar selection and mutation has produced a large variation in plant type and fruit quality (Dorji and Yapwattanaphun, 2011) and is apparent among the plants in the same orchard. Similar findings have been reported in Bhutan (Dorji and Yapwattanaphun, 2011) wherein wide variation in quantitative characters was observed in Bhutanese mandarin. The highest diversity index and coefficient of variation for characters that are pertinent for marketability were observed for the East and Darjeeling district, indicating the presence of diverse accessions. The TSS content for the analyzed accessions varied significantly among all the accessions. The level of TSS/acidity indicates the sweetness of an orange (Kishore et al., 2010) which is directly proportional to each other. We found that the accessions from D-01-09 (Darjeeling district) had the highest TSS/acidity as compared to the other accessions. The correlation coefficient between TSS/acidity, weight and diameter showed positive correlation (0.610*) and (0.534*) respectively, which is similar to the findings by Gill et al. (2005). The application of nitrogen fertilizers (N) has been shown to contribute positively towards%
Table 1. Result of ANOVA testing for differences in morphological traits of fruit, flower, and leaves collected from different accessions.

<table>
<thead>
<tr>
<th>Traits</th>
<th>DF</th>
<th>P-Value</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of segment</td>
<td>34</td>
<td>0.009114</td>
<td>NS</td>
</tr>
<tr>
<td>Fruit Weight</td>
<td>34</td>
<td>3.73E-20</td>
<td>***</td>
</tr>
<tr>
<td>Fruit Diameter</td>
<td>34</td>
<td>3.69E-14</td>
<td>***</td>
</tr>
<tr>
<td>Fruit Length</td>
<td>34</td>
<td>3.81E-15</td>
<td>***</td>
</tr>
<tr>
<td>Width of Epicarp</td>
<td>34</td>
<td>7.73E-16</td>
<td>***</td>
</tr>
<tr>
<td>No. of seeds</td>
<td>34</td>
<td>7.62E-06</td>
<td>***</td>
</tr>
<tr>
<td>TSS</td>
<td>34</td>
<td>2.59E-15</td>
<td>***</td>
</tr>
<tr>
<td>Acidity</td>
<td>34</td>
<td>2.23E-09</td>
<td>***</td>
</tr>
<tr>
<td>Leaf Length</td>
<td>34</td>
<td>5.39E-05</td>
<td>***</td>
</tr>
<tr>
<td>Leaf Width</td>
<td>34</td>
<td>8.79E-08</td>
<td>***</td>
</tr>
<tr>
<td>Leaf Length/Width</td>
<td>34</td>
<td>0.510286</td>
<td>NS</td>
</tr>
<tr>
<td>Petal Length</td>
<td>34</td>
<td>3.35E-12</td>
<td>***</td>
</tr>
<tr>
<td>Petal Width</td>
<td>34</td>
<td>2.14E-08</td>
<td>***</td>
</tr>
<tr>
<td>Pedicle Length</td>
<td>34</td>
<td>2.89E-14</td>
<td>***</td>
</tr>
<tr>
<td>No. of stamens</td>
<td>34</td>
<td>1.96E-24</td>
<td>***</td>
</tr>
</tbody>
</table>

The P-value of significant traits are in boldface (at P <0.001); NS, Non-significant.

Fig 1. Correlation coefficients among fourteen quantitative characters of fruits of *C. reticulata*
Abbreviation: No. of Segments (Sgm), Weight (Wt), Diameter (Dtr), Height (Ht), Width of Epicarp (EW), No. of Seeds (NoS), TSS/Acidity (TA), Leaf Length (LL), Leaf Width (LW), Leaf Length/Width (LL.W), Petal Length (PL), Petal Width (PW), No. of Stamens (NOS), Pedicel Length (PeL).

Table 2. Eigen values, proportion of variation and contribution 14 quantitative characters to the total variation in the first six PC axes of the 105 accessions of *C. reticulata*.

<table>
<thead>
<tr>
<th>PC1</th>
<th>PC2</th>
<th>PC3</th>
<th>PC4</th>
<th>PC5</th>
<th>PC6</th>
<th>Total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sgm</td>
<td>-0.058</td>
<td>0.023</td>
<td>-0.019</td>
<td>2.346e-02</td>
<td>-0.066</td>
<td>0.041</td>
</tr>
<tr>
<td>Wt</td>
<td>-12.519</td>
<td>0.947</td>
<td>0.174</td>
<td>1.369e+00</td>
<td>-0.408</td>
<td>0.236</td>
</tr>
<tr>
<td>Dtr</td>
<td>-4.107</td>
<td>-2.033</td>
<td>2.967</td>
<td>-3.522e+00</td>
<td>0.158</td>
<td>-0.031</td>
</tr>
<tr>
<td>Ht</td>
<td>-2.528</td>
<td>0.556</td>
<td>-0.015</td>
<td>3.298e-01</td>
<td>1.301</td>
<td>-1.271</td>
</tr>
<tr>
<td>EW</td>
<td>-0.102</td>
<td>0.059</td>
<td>-0.036</td>
<td>1.409e-03</td>
<td>0.017</td>
<td>-0.069</td>
</tr>
<tr>
<td>NoS</td>
<td>-0.398</td>
<td>-0.371</td>
<td>0.605</td>
<td>-1.217e-07</td>
<td>0.051</td>
<td>-0.148</td>
</tr>
<tr>
<td>TA</td>
<td>-1.595</td>
<td>-7.024</td>
<td>-3.397</td>
<td>1.576e+01</td>
<td>0.035</td>
<td>-0.048</td>
</tr>
<tr>
<td>LL</td>
<td>-1.598</td>
<td>3.379</td>
<td>-4.821</td>
<td>-2.193e+00</td>
<td>-0.485</td>
<td>-0.239</td>
</tr>
<tr>
<td>LW</td>
<td>-0.993</td>
<td>1.001</td>
<td>-1.394</td>
<td>-2.238e+01</td>
<td>1.856</td>
<td>0.928</td>
</tr>
<tr>
<td>LL.W</td>
<td>0.019</td>
<td>-0.103</td>
<td>0.114</td>
<td>-4.000e-01</td>
<td>-0.079</td>
<td>0.830</td>
</tr>
<tr>
<td>PL</td>
<td>0.144</td>
<td>-0.075</td>
<td>0.109</td>
<td>2.085e-05</td>
<td>-0.068</td>
<td>-0.021</td>
</tr>
<tr>
<td>PW</td>
<td>0.056</td>
<td>-0.028</td>
<td>0.060</td>
<td>1.433e-02</td>
<td>-0.063</td>
<td>-0.049</td>
</tr>
<tr>
<td>NOS</td>
<td>0.098</td>
<td>-0.176</td>
<td>0.205</td>
<td>-9.358e-02</td>
<td>-0.113</td>
<td>-0.081</td>
</tr>
<tr>
<td>PeL</td>
<td>0.068</td>
<td>-0.066</td>
<td>0.155</td>
<td>2.616e-02</td>
<td>-0.014</td>
<td>-0.079</td>
</tr>
<tr>
<td>Variance (%)</td>
<td>27.146</td>
<td>16.874</td>
<td>11.544</td>
<td>8.350</td>
<td>7.284</td>
<td>6.496</td>
</tr>
</tbody>
</table>

P.C: Principle Component
Fig 2. Principal Component analysis depicting the contribution of 14 quantitative characters for variation among the 105 accessions of *C. reticulata*. The two components of principal component analysis (Dim1 and Dim2) are plotted on the axes.

Abbreviation: Height (Ht), Weight (Wt), Diameter (Dtr), No. of seeds (NoS), Width of Epicarp (EW), TSS/Acidity (TA), No. of Segments (Sgm), Leaf Length (LL), Leaf Width (LW), Leaf Length/Width (LL.W), No. of Stamens (NOS), Pedicel Length (PeL), Petal Width (PW), Petal Length (PL), contrib: Contribution

Table 3. Frequency distribution for the measure of qualitative characters in mandarin (*C. reticulata*) accessions.

<table>
<thead>
<tr>
<th>Character</th>
<th>Frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adherence of segment walls</td>
<td>Strong (8), Medium (87), Weak (5)</td>
</tr>
<tr>
<td>Segment Uniform</td>
<td>Yes (8), No (92)</td>
</tr>
<tr>
<td>Fruit Shape</td>
<td>Oblong (82), Spheroid (18)</td>
</tr>
<tr>
<td>Shape of fruit apex</td>
<td>Depressed (10), Truncate (88), Concave (1), Convex (1)</td>
</tr>
<tr>
<td>Shape of fruit base</td>
<td>Depressed (2), Truncate (84), Concave (6), Convex (7), Spheroid (1)</td>
</tr>
<tr>
<td>Seed Shape</td>
<td>Ovoid (9), Clavate (91)</td>
</tr>
<tr>
<td>Intensity of green color</td>
<td>Medium (47), Dark (53)</td>
</tr>
<tr>
<td>Leaf lamina shape</td>
<td>Elliptic (98), Lanceolate (2)</td>
</tr>
<tr>
<td>Leaf lamina margin</td>
<td>Entire (66), Sinuate (34)</td>
</tr>
<tr>
<td>Petiole width</td>
<td>Narrow (50), Medium (50)</td>
</tr>
<tr>
<td>Relative length of stamen to stigma</td>
<td>Medium (52), Short (48)</td>
</tr>
</tbody>
</table>

Fig 3. Unrooted Neighbor-joining tree depicting relationship among the 105 accessions of *C. reticulata* based on qualitative and quantitative variables of trees, leaves, flowers and fruits. Different color legends in the figure depict collection sites of 105 *C. reticulata* accessions from Sikkim and Darjeeling Himalayas.
Fig 4. Grid map showing diversity index and coefficient of variation for TSS/acidity of *C. reticulata*.

Fig 5. Grid map showing diversity index and coefficient of variation for fruit diameter of *C. reticulata*.
Fig 6. Grid map showing diversity index and coefficient of variation for fruit weight of *C. reticulata*.

Fig 7. Collection sites of *C. reticulata* from Sikkim and Darjeeling Himalayas.
TSS, juice and acid content (Zekri et al., 2009). However, Thompson et al. (2005) and Kusakabe et al. (2006) found that nitrogen rates and fertilizers did not significantly affect the fruit quality including the percentage of TSS and acid content. Since, the use of fertilizers has been banned in the state of Sikkim and the farmers of Darjeeling and Kalimpong have been practicing organic farming, the variation in TSS could be due to genetic variation. The aim of citrus breeders over the past 30 years is to develop seedless, easy-peeling, enriched flavor and aroma and extension of ripening period varieties (Deng, 2005). Various methods are being designed to develop seedless varieties (Bilinques, 2004; Liu and Deng, 2007). The inherited seedless trait exhibited continuous variation over generations (Liu and Deng, 2007). Nevertheless, S-03-18 (East Sikkim) which contained the lowest number of seeds per fruit can be used for the multiplication of seedless varieties. The PCA indicated most of the fruit, flower and leaf traits to be responsible for bringing about variation in accessions. The large variation in fruit and flower characteristics could be a determinant factor in selecting superior varieties that can act as parents for breeding purposes. Cluster analysis based on qualitative and quantitative characters of 105 accessions divided them into two clusters (I and II), which were similar to the findings in Bhutanese Mandarin (Dorji and Yapwattanaphun, 2011). The average similarity coefficient (0.49) indicated a wide variation in the morphological traits of the collected accessions. The accessions from Darjeeling formed a separate sub-cluster than the other sub-clusters which indicates that Sikkim and Darjeeling mandarin are ecotypes as suggested by Kishore et al. (2010). The separate sub-clustering of accessions of east and west district respectively could be due to different environmental conditions and the sub-clusters with mix accession could be due to the heterogeneity of the species (Campos et al., 2005), ability to hybridize (Dorji and Yapwattanaphun, 2011) or genetic variation (Das et al., 2005).

Materials and methods

Collection of plant Samples

The sampling of fruits was executed in the month of December 2017 to January 2018 and that of leaves and flowers in March of the same year. The survey was conducted from four districts – North, South, East and West districts of Sikkim and Darjeeling, Kalimpong districts of the state of West Bengal (WB), India (Figure 7). Random visits were made to 105 orchards of mandarin growing regions in all of the six districts. A single representative tree propagated through seedling was selected from each orchard and sampling of 5 flowers, 15 leaves and 10 matured fruits were carried out randomly from each of the selected trees. The coordinates and altitude of the orchard were recorded using Global Positioning System (Garmin GPS-12).

Data collection

Data for each accession was logged according to the International Plant Genetic Resource Institute description for citrus (IPGRI, 1999). Data were recorded on fourteen quantitative and ten qualitative characters. Qualitative characters were measured based on coding (Table 3) while the fruit weight was measured using electronic balance with 0.01 g precision. TSS was measured by a refractometer (Erma, Japan) and citric acid content was analyzed by acid-base titration method. To evaluate the relationship between different quantitative characters, ANOVA was performed using Microsoft Excel 2007. Correlation coefficient (r) among the quantitative traits was evaluated using R package “corrplot” version 0.84 (Wei and Simko, 2017).

Cluster and principal component analysis

The phenetic relationship between the accessions was evaluated by Principal Component Analysis (PCA) using the R package “BiodiversityR” (Kindt and Coe, 2005). The relationship between accessions was interpreted through the construction of an unrooted Neighbour Joining (NJ) dendrogram based on simple matching coefficient using Darwin version 6.0 (Perrier and Jacquemound-Collet, 2006). DIVA-GIS for diversity analysis

DIVA-GIS software supports the analysis of gene bank and herbarium databases to elucidate genetic, ecological and geographic patterns in the distribution of crops and wild species (Hijmans et al., 2001). DIVA-GIS software version 7.2.1 (www.DIVAGIS.org) was used to generate grid maps for the analysis of the diversity of three important quantitative traits for marketability viz., weight, diameter and TSS/acidity recorded from various orchards coordinated with geographical coordinates. The point-to-grid option using a circular neighborhood method on the “Analysis Menu” was used to get grids over the points from where fruit samples were collected. Shannon diversity index was used to obtain the diversity patterns of the quantitative traits chosen.

Conclusion

Our study on morphological characters showed a huge variation in the quantitative characters, indicating the diversity of the analyzed accessions. The grid maps generated for selected traits by DIVA-GIS indicated the highest diversity index in East district, Sikkim. The highest coefficient variation was observed in East and Darjeeling district, indicating that these districts could be identified as a source for superior germplasm. Cluster analysis divided the accessions into two major clusters wherein the accession from Darjeeling formed a separate sub-cluster while the other sub-clusters comprised of accessions from Sikkim and Kalimpong. The wide variation in morphological traits as suggested by cluster analysis and PCA suggests that morphological diversity exists in mandarin oranges grown in Sikkim and Darjeeling Himalayas.

Acknowledgments

The authors are immensely grateful to the Department of Botany, Sikkim University for providing all the necessary laboratory facilities. We would also like to thank Mr. Aditya Pradhan and Mr. Patrush Lepcha for their help and suggestions during the study. Lastly, we thank and dedicate this study to the farmers for their cooperation.
References


Wei T, Simko V (2017). R package "corrplot": Visualization of a Correlation Matrix (Version 0.84).
