Establishing Genetic Association Among Selected Members of *Citrus* Species Through Protein Profiling Obtained from Comparative Electrophoresis

J.I.B. Janairo, K. F. Saulog, and N.Lazaro –Llanos*
Department of Chemistry, College of Science, De La Salle University, Manila, Philippines
Center for Natural Sciences and Ecological Research, De La Salle University, Manila, Philippines
*Corresponding author: Tel/Fax: 062-5360230; E-mail: nancy.lazaro-llanos@dlsu.edu.ph

The use of seed proteins in the evaluation of genetic diversity is cost effective and less time consuming. Water and salt-soluble proteins from the seeds of *Citrus* species: *C reticulata, C. aurantifolia, C grandis, C. sinensis, C. limon,* and *C. microcarpa* were extracted and variations in electrophoretic profiles were determined. Finely grounded seeds were homogenized in 50 mM Tris buffer, pH = 7.0, centrifuged at 14,000 rpm for 10 minutes at 4ºC, and the supernate collected as the water-soluble fraction. The pellet was resuspended in 50mM Tris buffer, pH = 7.0 in 150 mM NaCl, homogenized, then centrifuged at 14,000 rpm for 10 minutes at 4ºC, and the supernate collected as the salt-soluble fraction. Samples were analyzed using SDS PAGE. Nine protein bands were found common to all samples and six bands were used as genetic markers in the water-soluble fractions. A dendrogram was constructed for the water-soluble proteins. In the salt-soluble fraction, ten protein bands were found to be common to all samples and five bands were concluded to be genetic markers. *C.reticulata* and *C.sinensis* possessed identical water and salt soluble protein profiles and *C. grandis* were found to be most distinct among the samples.

Keywords: *Citrus* fruits, protein profiles, genetic markers, Discontinuous SDS – PAGE

1. INTRODUCTION

Molecular data, such as DNA base pairing and sequences have long complemented traditional taxonomy in proper classification of different species wherein most cases these data have shown underlying differences that traditional taxonomy fail to show (Redondo, 2008). Hence, the use of DNA as genetic markers have gained wide use and acceptability. However, the use of proteins to serve a similar purpose is also a reliable method and an excellent alternative especially in plant systematics. Protein markers have been regarded as important taxonomic tool since these proteins are highly conserved in evolution (Kouts, 1976) and these proteins are not susceptible to damage or mutation caused by environmental factors (Yelamo, 1992). Furthermore seed proteins are not influenced by external conditions rendering them invariable. Therefore, their electrophoretic protein profile obtained from SDS PAGE can serve as reliable genetic markers. (Grib, 1984; Konarev, 1988; Konarev, 1998; Poperelyn, 2000)

The use of seed proteins as genetic markers have been established in the following studies: 1) Genetic Characterization of Iranian Wild *Prunus* Species (Gomez, 2007); 2) on wheat

varieties (Shuiab, 2007; Metakowsky, 1990; Dvoracek, 2003; Anderson 1993); 3) on pine (*Pinus,* species) (Allona, 1992; Landgridge, 2004; Bahrman, 1995; Gerber, 1993); 4) on millet (*Pennisetum,* species) (Lagudah, 1990); 5) on sunflower (Raymond, 1991; Aksyanov, 2005); 6) on *Capsicum* [Zecevic, 2002; Posch, 2002]; 7) on pea (*Pisum* species) (Gottschalk, 1982); 8) on rice (*Oryza*) (Padmavathi, 2001); 9) on sorghum (Chaunhan, 2002); 10) on legumes (Gog, 1982); 11) on maize (Burstin, 1994).

The taxonomy of C*itrus* fruits has drawn the attention of botanists and horticulturists because no formal agreement has been reached when it comes to the proper speciation of this genus. The confusion arises from a genetic trait that is when present in a biotype will cause deviation in normal reproduction (Barret, 1976, Ramon-Laca, 2005). This trait was observed to be that the *Citrus* fruits sometimes produce multiple adventitious nucellar embryos. To make matters even more complicated, *Citrus* fruits can easily cross with one another to produce hybrids, resulting to the deviation of taxonomic limits once set by geographic locations (Moore, 2001, Araujo, 2000). The use of seed proteins of *Citrus* species as genetic markers will help identify the differences among these fruits on the molecular level.

2. MATERIALS AND METHODS

2.1. Samples

Seeds from the ripe fruits of the following *Citrus* species were collected: *C. aurantifolia* (lime), *C. grandis* (pomelo), *C. limon* (lemon), *C. microcarpa (*calamondin), *C. reticulata* (mandarin), and *C. sinensis* (Valencia orange)

2.2. Extraction of Water- and Salt-Soluble Proteins

The seeds were collected and dried using a filter paper, and were segregated to their respective mortars; subsequently their seed coats were removed and quickly frozen with liquid nitrogen. The frozen seeds were finely powdered, and a known volume of Tris-HCl buffer at pH 7 was added to each vials were each seed types were placed. The mixtures were homogenized and centrifuged at 14,000 rpm for 10 minutes at 4 ^{0}C using a high speed refrigerated centrifuge. The supernate was collected and labeled as Supernate 1. The pellets were transferred to their respective mortar and pestle and a known volume of Tris-HCl buffer at pH 7 in saline solution was added to each vial were the samples were placed. The mixtures were homogenized and centrifuged at 14,000 rpm for 10 minutes at 4 ^{0}C using a high speed refrigerated centrifuge. The supernate was collected and labeled as Supernate 2.

2.3. Protein Profiling

The samples were analyzed using discontinuous SDS-PAGE. Three trials were done per protein type. The resulting bands from SDS-PAGE were documented. The relative mobility (Rm) values were determined from which a dendogram and a similarity index table were constructed. The presence of a band at a particular molecular weight will be indicated by a (+) value while the absence of the band will be indicated by a (-) value.

3. RESULTS AND DISCUSSION

Comparative electrophoresis is an essential technique in plant biosystematics because the obtained electrophoretic homology exhibits theoretical significance that is used to evaluate phylogenetic and taxonomic relationships and affinity between taxa [8]. This biochemical technique has been applied to various species which aims to verify or revise known classifications (Shechter, 1975; Caraso, 1977; Ayala, 1974; Hill, 1977; Crawford, 1976; Gomez, 2007). The use of protein banding analysis to identify diversity among taxa is considered accurate and reliable especially when analyzing affinity among very close species such as *Citrus* (Gomez, 2007; Cherry, 1970; Sanahai, 1977).

In this study, the water – soluble and salt – soluble proteins of the sample *Citrus* species were subjected to discontinuous SDS-PAGE

and the resulting banding patterns were used to evaluate their affinity with one another. The actual gel matrix is shown in Figures 1 and 2. A similarity index table was constructed as well in order to facilitate quick identification of protein bands that can be used as a genetic marker. A protein genetic marker is a protein band that is found exclusive to one species or to a number of species but it should not be common to all samples. It is also possible that such genetic markers can also be in the form of an exclusive absence of a specific protein band that is common to all samples except for one species. A (+) sign indicates that the particular protein is present in that species of interest. The given molecular weight on the first column of the table corresponds to the mean molecular weight of the identified similar protein among the six samples including the second trial. The 205, 136, 116, 94, 82, 67, 49, 36 and 29 kDa protein were present in all samples indicating that such proteins are characteristic of the *Citrus* family. The 174 and 155 kDa proteins were exclusively found in *suha*. This translates to how distinct *suha* is compared to the other members of the genus and the morphology of *suha* can attest to these findings. Orange and ponkan share identical protein profiles, with the 147 kDa protein exclusive to them. This means that they resemble each very closely not only in their physical characteristics but in their protein make up as well, but can be differentiated from other members of the genus.

Dayap, is somewhere in between lemon and *calamansi* with respect to the similarity index and their protein profiles. All three of them have almost identical protein profiles except that *dayap* shares the distinct 44 kDa protein of *calamansi* and the distinct 58 kDa protein of lemon. It can be observed from the structural morphology of *dayap* that indeed it has close resemblance both to *calamansi* and lemon. *Dayap* has the size and shape of lemon; on the other hand, it has the color of *calamansi* and it grows optimally under conditions similar to that of *calamansi*. Other protein markers are the 63 kDa protein found in orange, ponkan and *suha*; the 58 kDa protein found in *dayap*, lemon, orange, ponkan and *suha*; and the 44 kDa protein found in *calamansi*, *dayap*, orange, ponkan and *suha.* In general, the water soluble seed proteins of the samples exhibit high-heterogeneity.

The water-soluble proteins can be assumed to be proteins that are localized within the cell and/or part of the transmembrane protein that is attached to the polar part of the membrane bilayer. Solubility in an aqueous environment is essential for such proteins because this property facilitates the movement of such proteins within the cell as well as the retention of its functional structure. The structure of proteins is very important since function is dependent on structure and vice versa. Any change in protein structure, or even an amino acid substitution can alter the function of the protein. Since these proteins are water-soluble, they are composed of predominantly polar amino acid residues such as threonine, serine and glycine. Solubility in an aqueous media is caused by the noncovalent interaction, specifically, H-bonding existing between the polar amino acid and the solvent

As stated earlier, the extracted water-soluble proteins are said to be localized within the cell. Therefore, most of these proteins are metabolic proteins or proteins that play a role in the metabolic activities of the cell. Since the samples were centrifuged at 14000 rpm, it is certain that the cell along with its organelles are completely ruptured [16] making the possible sources of the proteins enzymes, membrane proteins, proteins from organelles or synthesized proteins that will be exported out of the cell or will be used locally. It is evident that the water-soluble seed proteins are more diverse compared to the salt-soluble seed proteins with respect to the number of protein bands that are not common to all samples which implies that each sample species has a metabolic activity that is unique.

In order to study the relatedness of the sample species, a dendrogram was constructed by inputting the values of the distance matrix to MATHLAB, a statistical software (see figure 3). The distance matrix was based on the values

obtained in the coefficient of similarity which were calculated using the formula:

$$Cs = \frac{p}{(p+d)}$$

where p is the number of common bands in both samples and d represents the number of bands present in only one sample.

Similar to the electrophoretic profiles of the water-soluble proteins, a similarity index table for the salt-soluble proteins was also constructed in order to easily compare the protein bands of one sample to another (refer to Table 2.). The first step was to select the bands which were well separated and most likely to be reproducible. Very faint bands which were not present on both trials were disregarded. Then, the molecular weights of the bands which were considered to be the same polypeptide in trials 1 and 2 were averaged. Finally, molecular weight of protein bands in all the samples that correspond to the same polypeptide were grouped and averaged in order to come up with the values seen in the similarity index table. This was carefully done by analyzing the intensity of the bands, relative mobility and intensity. When the positions of the bands were compared to the molecular standard using DigiGenius SynGene Tools Software, it was found that the salt-soluble proteins exhibited less band variation compared to water-soluble proteins.

The 186, 156, 140, 124, 83, 69, 63, 56, 46, 40, 33 and 27 kDa proteins were common to all *Citrus* species. Highlighted rows only show that the band is present in all six samples and cannot be used as genetic markers. Furthermore, only three protein bands can be utilized to differentiate the samples as compared to six bands in the water-soluble proteins. Apparently, the 92 and 60 kDa proteins were exclusive to *suha* while the 103 kDa protein band was common to 5 other species. This further shows how different *suha* is compared to other samples. Consequently, only *suha* can be distinguished from other *Citrus* species using this electrophoretic profile since the remaining 5 samples have the same sets of protein bands. Furthermore, minor variations in band pattern in salt-soluble proteins entails that it will not readily provide enough data that will facilitate in differentiating one Citrus species from another.

4. CONCLUSION

In conclusion, seed proteins can be utilized as a tool to identify affinity and diversity among species within taxa. In this study however, the water – soluble proteins were the ones that can be used since it is the one that exhibited high heterogeneity and polymorphism. The salt – soluble proteins, on the other hand, showed otherwise. This study was successful in illustrating relatedness and differences among the *Citrus* species which will take *Citrus* taxonomy one step further.

REFERENCES

Aksyanov, I.V. (2005). Protein markers specificity of sunflower inbred lines, *HELIA*, *28*, 49-54.

Anderson, N.G., et. al. (1993). Two-dimentional Electrophoresis of wheat seed proteins, *Crop Science. 25*, 667-6674.

Allona, I., et al. (1992). Seed storage protein from *Pinus pinaster*: homology of major components with11s proteins from angiospersm *Plant Science, 87***,** 9-18

Araujo, E., et al. (2000). What is Citrus? Taxonomic implications from a study of cp-DNA evolution in the tribe Citreae (Rutaceae subfamily Aurantioideae). *Org. Divers. Evol. 3*, 55 – 62.

Ayala, R. (1974). Single seed disc electrophoresis of *maize,* wheat and sorghum**.** *Bulletin of the Torrey Botanical Club. 101*, (4) 207-210.

Bahrman, N., & Petit R. (1995). Genetic polymorphism in maritime pine (Pinus pinaster) assessed by 2-D gel electrophoresis of needle, bud, and pollen proteins, J. *Mol Evolution*,

Barett, H., & Rhodes, A. (1976). A numerical

taxonomic study of affinity relationships in cultivated *Citrus* and its close relatives. *Systematic Botany*. *1* (2) 105-136.

Burstin, J., et.al. (1994). Molecular markers and protein quantities as genetic description in maizeII. Diversity among 21 inbred lines. *Theor. Appl. Genetics*, pp. 943-950

Carasco, J., et al. (1977). Seed protein profiles of African representatives of *Vigna* spp. *Taxon.* Vol. 26, No. 5/6, pp. 513-516.

Chaunhan, P., et.al. (2002). Molecular weight analysis of seed proteins of forage sorghum (Sorghum bicolor L.) *Seed Science and Technology*, pp. 11-14

Cherry, J., et al. (1970). Comparative studies of seed proteins of species *Gossypium* by gel electrophoresis. *Evolution.* Vol. 24, No. 2, pp. 431-447.

Crawford, D., & Julian, E. (1976). Seed protein profiles in the narrow – leaved species of *Chenopodium* of the Western United States: Taxonomic value and comparison with the distribution of flavanoid compounds. *Amer. J. Bot.* Vol. *63* (3), 302-308.

Dvoracek, V., & Curn. V. (2003). Evaluation of protein fractions as biochemical markers for the identification of spelt wheat cultivars (*Triticum spelta* L.), *Plant Soil Environment*, *49*, 99-105.

Gerber S., et.al. (1993). Seed-protein variation in maritime pine (*Pinus pinaster* Ait.) revealed by two-dimensional electrophoresis : genetic determinism and construction of a linkage map. *Theor Appl Genet*, *85*, 521-528.

Gomez, P., et al. (2007). Comparison of the use of morphological, protein and DNA markers in the genetic characterization of Iranian wild Prunus species.. *J.Scientia*. (XXX) pp. XXX – XXX.

Gorg, A., et. al. (1982). SDS electrophoresis of legume seed proteins in horizontal ultrathin-layer pore gradient gels, *Zeitschrift fur Levensmitteluntersuchungun-Forschung A*, *17***,** 282-285.

Gottschalk, W., & Muller, H.P. (1982). Seed proteins of Pisum mutants and recombinants, *Plant Foods for Human Nutrition 31*, 297-306.

Grib, O.M. (1984). Inheritance of borderline components in barley hybrids under various types of crossing, *Agricultural Biology*, *12,* 29-32.

Hill, R. (1977). Variablity of soluble seed proteins in populations of *Mentzelia L.* from Wyoming and adjacent states. *Bulletin of the Torrey Botanical Club.* Vol. 104, No. 2, pp. 93-101.

Huber, L., et al. (2003). Implications of Subcellular fractionation in proteomics. *Circulation Research*, *AHA*. *92*:962.

Konarev, A.V. (1998). The use of molecular markers in the investigation of plant genetic resources, *Agricultural Biology, 5*, 3-25.

Konarev, A.V., et.al. (2000). Seed proteins as markers in problem decision of genetic resources of breeding and seed growing. *Cytology and Genetic*, *34*, 91-104.

Kouts, H., & Hillebrand, G. (1976). An electrophoretic and serological investigation of seed proteins in Galeopsis tetrahit L.(Labiatae) and its putative parental species. *Amer. J. Bot. 63* (2) 156-165.

Langridge, P., et. al. (2004). Biochemical genetics of some seed proteins of *Pinus* radiata, *Biochemical Genetics*, *19*, 585-597.

Lagudah, E.S., & Hanna, W.W. (1990). Patters of variation for seed proteins in the *Pennisetum* gene pool*, Journal of Heredity,* 81, pp. 25-29.

Metakowsky, E.V., et.al. (1990). Gluten polypeptides as useful genetic markers of dough quality in Australian wheats, *Australian J. of Agricultural Research*, *41*, 280-306.

Moore, G. (2001). Orange and lemons: clues to the taxonomy of *Citrus* from molecular markers. *Trends in Genetics. 17* (9).

Padmavathi, G., Siddiq, E., & Kole, C. (2001). Inheritance of protein markers detecting polymorphism among rice genotypes with

contrasting host response to green leafhopper, *Current Science*, *80*, 1111-1112.

Poperelyn, A. (2000). The electrophore-grammes genetic interpretation of seed F1 helianthin of sunflower, *Cytology and Genetics*, *34***,** 84-90.

Posch, A., et. al. (2002). Genetic variability of pepper (*Capsicum annumm* L.) seed proteins studied by 2-D electrophoresis with immobilized pH gradients, *Electrophoresis*, *13*, 774-777.

Ramon – Laca, L., & Arias, B. (2005). Pharmacological properties of citrus and their ancient and medieval uses in the Mediterranean region. *J.Jep.* (97) pp.89-95.

Raymond, J., et al. (1991). The seed proteins of sunflower: comparative studies of cultivars, *Phytochemistry*, *30*, 2849-2856.

Sanahai, S., & Rana, R. (1977). Seed protein homology and elucidation of species relationships in *Phaseolus* and *Vigna* species. *New Phytol.* (79) 527-534.

Shuaib, M., et al. 2007. Characterization of wheat varieties by seed storage protein electrophoresis. *African J. of Biotechnology, 6*, 497-500.

Shechter Y., & De Wet, M. (1975). Comparative electrophoresis and isozyme analysis of seed proteins from cultivated races of sorghum**.** *Amer. J. Bot. 62* (3) 254-261.

Redondo, R., et al. (2008). Molecular systematics of the genus *Artebius* (*Chiroptera: Phyllostomidae*). *J.ympev*. Vol. XXX.

Yelamo, M., et al. (1992), Comparative electrophoretic studies of seed proteins in some species of of the genera *Diplotaxis*, *Erucastrum* and *Brassica*. *Taxon*. *41* (3) 477-483.

Zecevic, B. (2002). Seed proteins as genetic markers in the study of phylogenetic relations in pepper, *Acta Horticulturae*, (579), 113-116.

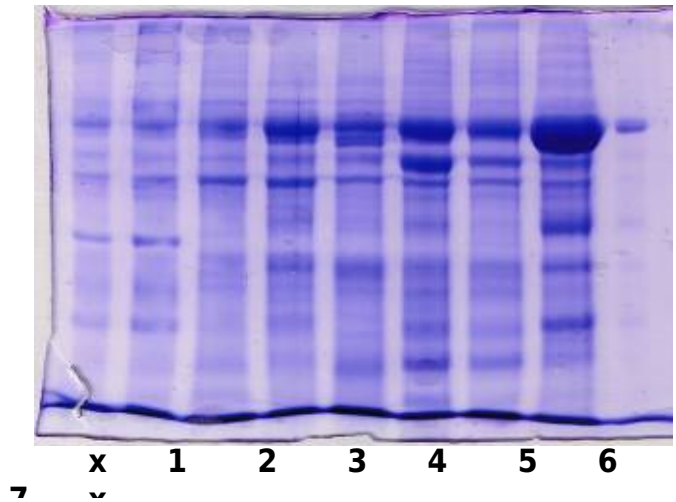

Figure 1. SDS-PAGE of Water-Soluble Proteins from Seeds of Citrus Species. Track 1: Molecular marker (Sigma wide-range marker), 2: C. microcarpa (calamansi), 3: C. aurantifolia (lime), 4: C.limon (lemon), 5: C. sinensis (Valencia orange), 6: C. reticulata (Mandarin Orange/Ponkan), 7: C. grandis (pomelo); x: blank wells.

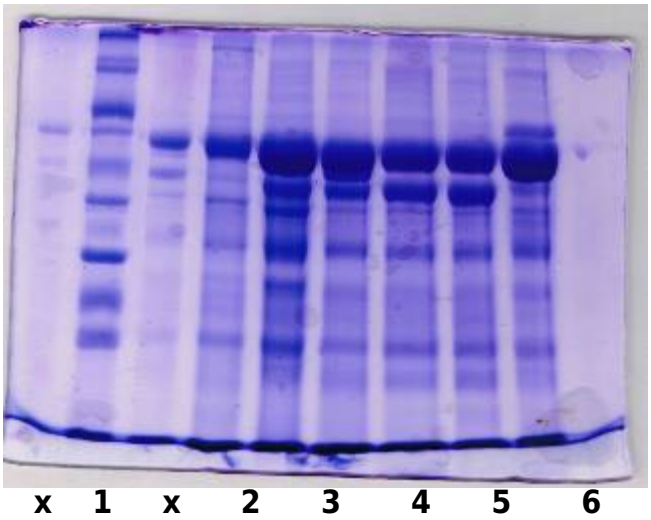

Figure 2. SDS-PAGE of Salt-Soluble Proteins from Seeds of Citrus Species. Track 1: Molecular marker (Sigma wide-range marker), 2: C. microcarpa (calamansi), 3: C. aurantifolia (lime), 4: C.limon (lemon), 5: C. sinensis (Valencia orange), 6: C. reticulata (Mandarin Orange/Ponkan), 7: C. grandis (pomelo); x: blank wells.

Table 1. Similarity Index Table of Water-Soluble Proteins
(blue rows: protein bands common to all; green rows: distinct protein bands)

Water-soluble Proteins						
Molecular Weight (kilodaltons)	*Citrus microcarpa* (calamansi)	*Citrus aurantifolia* (lime/dayap)	*Citrus limon* (lemon)	*Citrus. sinensis* (Valencia orange)	*Citrus reticulata* (Mandarin orange/Ponkan)	*Citrus grandis* (pomelo/suha)
205	+	+	+	+	+	+
174	-	-	-	-	-	+
155	-	-	-	-	-	+
147	-	-	-	+	+	-
136	+	+	+	+	+	+
116	+	+	+	+	+	+
94	+	+	+	+	+	+
82	+	+	+	+	+	+
67	+	+	+	+	+	+
63	-	-	-	+	+	+
58	-	+	+	+	+	+
49	+	+	+	+	+	+
44	+	+	-	+	+	+
36	+	+	+	+	+	+
29	+	+	+	+	+	+

Table 2. Similarity Index Table of Salt-Soluble Proteins
(blue rows: protein bands common to all; green rows: distinct protein bands)

Salt-soluble proteins						
Molecular Weight (kilodaltons)	*Citrus microcarpa* (calamansi)	*Citrus aurantifolia* (lime/dayap)	*Citrus limon* (lemon)	*Citrus. sinensis* (Valencia orange)	*Citrus reticulata* (Mandarin orange)	*Citrus grandis* (pomelo/suha)
186	+	+	+	+	+	+
156	+	+	+	+	+	+
140	+	+	+	+	+	+
124	+	+	+	+	+	+
103	+	+	+	+	+	-
92	-	-	-	-	-	+
83	+	+	+	+	+	+
69	+	+	+	+	+	+
63	+	+	+	+	+	+
60	-	-	-	-	-	+
56	+	+	+	+	+	+
46	+	+	+	+	+	+
40	+	+	+	+	+	+
33	+	+	+	+	+	+
27	+	+	+	+	+	+

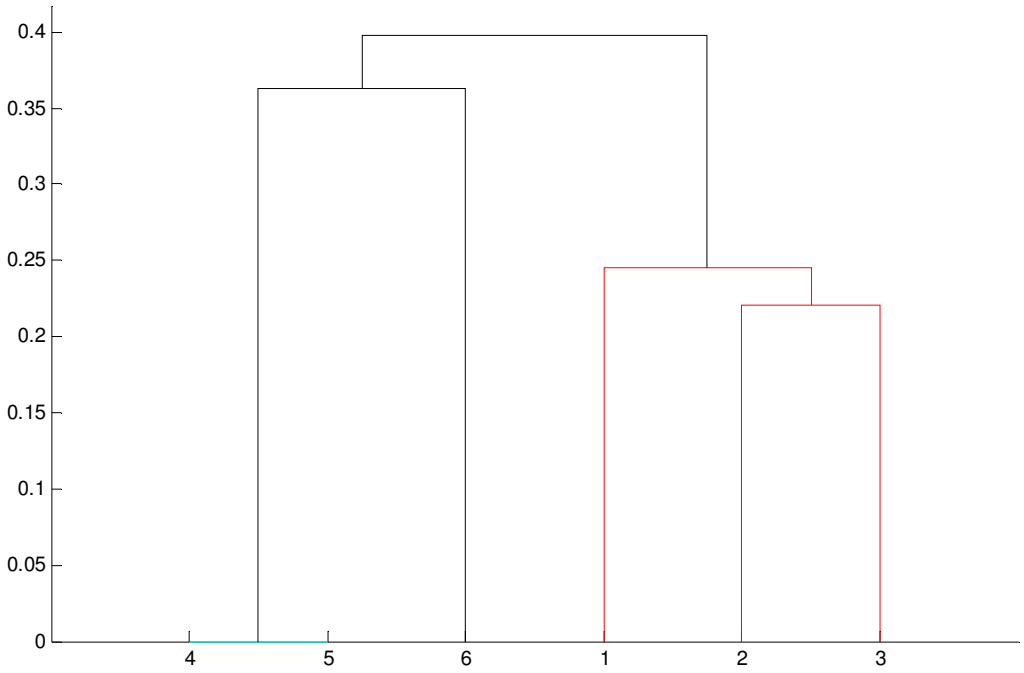

Figure 3. Dendrogram Constructed from the Coefficient of Similarity
1= *C. microcarpa*, 2 = *C. aurantifolia*, 3 = *C. limon*,
4= *C. sinensis*, 5 = *C. reticulata*, 6 = *C. grandis*