Genetic and Genomic Foundations For Capsicum Improvement And Development

Nkwiza M. Nankolongo[1*], Orlex Baylen Yllano[2], Leilani D. Arcer[3], Neil John V. Vegafria[4], Ephraim A. Evangelista[5], Ferdinand A. Esplana[5], Lester Harris R. Catolico[6], Merbeth Christine L. Pedro[7] and Edgar E. Tubilag[8]
Adventist University of the Philippines, College of Science and Technology, Biology Department

*2052288@aup.edu.ph**

ABSTRACT

Capsicum is a genetically diverse eudicot, diploid, and self-pollinating plant that grows well in slightly warmer environments. This paper critically reviews *capsicum* biology, horticultural characteristics, genetic diversity, ploidy levels, chromosome structure, genome organisation and phylogenetic relationships. The related literature presented in this paper was obtained from scientific sites and journals available in resources such as NCBI, ScienceDirect, SAGE journals, Google Scholar, ResearchGate and many more. Using Molecular Evolutionary Genetics Analysis (MEGA 11) 22 NADH dehydrogenase sequences were extracted from NCBI and then aligned. The findings discuss that *capsicum* has a chromosome number 2n=2x=24. In contrast, wild species are known to have a ploidy level of 2n=2x=26. The *capsicum* genome is estimated to be between 1498cM and 2268cM. Hot peppers' genomic sequencing data revealed 37.989 scaffolds with an estimated size of 3.48Gb, 34771 in tomatoes, and 39031 genes in potatoes. Based on the *capsicum* phylogenetic relationship, 22 species of *capsicum* were rooted. The information discussed in this study is indispensable in *Capsicum* frontier research, breeding, development, management and utilization of this economically important and highly regarded crop worldwide.

Keywords: *Capsicum*, genetic diversity, genome, chromosomes, ploidy levels.

INTRODUCTION

Over the past centuries, *Capsicum* species have been used for household, agrarian, medical, and industrial issues. Even if this could be the case, there have not been many scientific investigations on this genus. Five domesticated *Capsicum* species have been the subject of the majority of studies even though there are still several others that have not yet been researched. Due to a lack of data such as genetic sequences, alignments are difficult to construct but not impossible. To the field of Bioinformatics: The information discussed in this chapter can also be used in *Capsicum* frontier research, breeding, development, management, and utilization of this economically important and highly regarded crop worldwide. To the field of medicine and pharmaceutical industries: the research carries significant implications for the medicine and pharmaceutical industries, as they shed light on the potential applications of *capsicum* in developing new drugs and treatments for various health conditions. To future researchers: Future research can use this data to properly match the sequences of different *Capsicum* species. To Biology students and teachers: The knowledge presented in this study is useful for the biological discipline as a reference for related research. How can the alignment of domestic and wild species of peppers reveal the genetics and genomics of *capsicum* species? To analyze published literature on *Capsicum's* ploidy levels; To analyze the published literatures on *Capsicum's* chromosome structures; To analyze the published literature on *Capsicum's* genome organization and To determine the phylogenetic relationships of *Capsicum's* NADH dehydrogenase. The related literature presented in this paper where obtained from scientific sites and journals available in resources such as NCBI, ScienceDirect, SAGE journals, Google Scholar, ResearchGate and many more. Using Molecular Evolutionary Genetics Analysis (MEGA 11) 22 NADH dehydrogenase sequences were extracted from NCBI and then aligned. It is estimated that the genomes of domestic and wild varieties of *capsicum* are almost identical. The genome sequencing data of hot peppers is similar to those of other plant species.

LITERATURE REVIEW

Invasive species compendium (2022) states the following, *Capsicum annum* grows to a height of 0.5-1.5 m, is heavily branched, and has extremely strong taproots. The stem is normally green to brown-green in color, with purple patches near the node on occasion. The leaf design is alternate, basic, and highly changeable, with petioles up to 10cm long. The leaf blades can be up to 10^(-16) cm x 5^(-8) cm in size. The apex is acuminate, and the edge is whole, pale dark green, and subglabrous. The flowers are arranged singly, and the pedicel is around 4 cm long when in bloom. The fruits may grow up to 8 cm in length. The calyx is cup-shaped, persistent and enlarging in fruit, usually with 5 conspicuous teeth; corolla campanulate to rotate with five to seven lobes, 8-15 mm in diameter, usually color white. It can develop five to seven stamens with pale blue to purplish anthers. The ovary is 2(-4) locular, style fifliform, and has a white or purplish, stigma capitate. Fruit is a non-pulpy berry, which is variable in its size, shape, color, and degree of pungency. It may be more or less conical, up to 30cm long, green, yellow, cream, or purplish, when it is not yet fully developed. But when it is matured it becomes red, orange, yellow or brow. The seeds are orbicular and flattened, about 3-4.5 mm in diameter, 1 mm thick and the color is pale yellow. The plant is considered an annual, herbaceous, perennial, seed propagated, shrub, and climber.

Currently, the number of *Capsicum* cultivars and taxonomic varieties worldwide is around fifty thousand (Basu & De, 2003; Jarret, 2008). In the genus *Capsicum*, there are five domesticated species and around 25 identified species according to Aguilar-Melendez et al. (2009), while Tripodi & Kumar (2019) mentioned that over 30 species make up the genus *Capsicum.* Morphological characteristics to differentiate the five *Capsicum* taxa (*C. annum, C.baccatum, C. chinense, C frutescens, and C. pubescens*) include bloom and seed color, calyx form, number of flowers per node, and flower orientation (Basu & De, 2003; Hawkes et al., 1979 ; Aguilar-Melendez et al., 2009) . *C. annum* is considered to have been domesticated from *C.annum* populations in the wild, *Annum glabriusculum* in Mexico, perhaps from many geographically distinct wild populations (Aguilar-Melendez et al., 2009).

Fruit form has been examined extensively in the Solanaceae family, including tomato, pepper, and eggplant (Taher et al., 2017 ; Passam & Karapanos, 2016; Doganlar et al., 2002). On chromosomes 2, 3, 7, 8, and 10, (Taher et al., 2017 ; Zygier et al., 2005 ; Hurtado et al., 2013; Brewer et al., 2006 ; Barchi et al., 2021 ; Nunome et al., 2001 ; Knapp, 2002 ; García-González & Silvar, 2020) . Allelic variation in the Sun, Ovate, Fascinated (FAS), and Locule Number (LC) genes determine the form of the tomato fruit (Rodríguez et al., 2011).

Multiple QTLs for fruit length, width, and the fruit shape ratio (length: width) have been discovered on chromosomes 1- 4, 8, 10, and 11, according to (Zygier et al., 2005 ; Knapp, 2002 ; Chaim et al., 2006 ; (Chaim et al., 2006). Chaim linked two key fruit QTLs, fs 3.1 (fruitshape) and fe 10.1 (fruit elongation), to chromosome 3 and 10, respectively, in a BC4F2 population segregating for fruit-shaped (Knapp, 2002). These QTLs accounted 67 and 44% of the variance in fruit form and elongation found in the population, respectively (Knapp, 2002). Fruit trait inheritance in connection to pericarp form, colour thickness, and total soluble solids was investigated by (Vilarinho et al., 2015). The round form characteristic was governed by a single gene based on segregation ratios and found five QTLs contributing to fruit form and one QTL for pericarp thickness on chromosomes 1, 2, 4, 10, and 3, respectively, explaining 4 to 26% of the diversity in the serrano by jalapeno recombinant inbred line F8 population (Chaim et al., 2006). The expression of a gene that resembled the tomato gene Ovate and discovered substantial variations between round and elongated pepper cultivars was compared by (Adhikari et al., 2020), and carried out a study on five domesticated species by (Moulin et al., 2015).

METHODS

Data Collection

The related literature presented in this paper where obtained from scientific sites and journals available in resources such as NCBI, ScienceDirect, SAGE journals, Google Scholar, ResearchGate, and many more.

Phylogenetic Analysis

Using Molecular Evolutionary Genetics Analysis (MEGA 11), 22 NADH dehydrogenase sequences were extracted from NCBI and then aligned. Thereafter, a phylogenetic tree was created, which then produced clusters and branches based on their relatedness. According to Misawa & Tajima, (2000), on the basis of Klastorin's technique, a new straightforward method for categorizing genes is suggested. With the help of evolutionary changes, this technique divides genes into monophyletic groupings that are different from one another. As long as the phylogenetic tree of the genes is produced, the approach is usable. Bootstrap was applied to show the percentage of relatedness among the branches.

RESULTS AND DISCUSSION

Capsicum is a well-diverse shrub commonly grown annual in different parts of the world. It is genetically diverse, with numerous cultivars and taxonomic derivatives. *Capsicum* species are diploids with most having 24 chromosomes (n=x=12), but several wild species have 26 chromosomes (n=x=13). The majority of *Capsicum* species have 24 chromosomes (n=x=12), however, many wild species have 26 chromosomes (n=x=13). Variances in the form, size, and number of chromosomes are prevalent in populations of the same species or interspecific taxa, and these differences are categorized into cytotypes or chromosomal races. Researchers confirmed that such variations are common in the *Capsicum* genus, whose cytotypes differ mostly in karyotypic formula and chromosomal size. Secondary constrictions were found in the homologous pairs 1 and 12 and 6 and 11 of the frutescens accessions respectively. Genome sequence information of hot pepper revealed 37,989 scaffolds with an estimated size of 3.48 Gb. These protein-coding genes of *Capsicum* were relatively the same as other Solanaceae species – tomato (34,771 genes) and potato (39,031 genes). The genetic maps of tomato and pepper are nearly comparable in length, with 1275 cM in tomato and 1246 cM in pepper (Livingstone et al., 1999). The total length of the chili pepper genome has been estimated to be between 1498 cM and 2268 cM, which is approximately two to three times larger than the tomato genome. Phylogenetic analysis using Unweighted Pair Group Method with Arithmetic Mean (UPGMA) computed the Maximum Composite Likelihood with 1000 bootstrap replications revealed that the *C. ciliatum, C. lanceolatum, C. lycianthoides,* and *C. geminifolium* grouped together. More so, *C. minutiflorum* and *C. ceratocalyx* were strongly clustered together, the same with *C. chinensis* and *C. frutescens. C. pubescens, C. galapagoence, C. chacoense,* and *C. cardenasi* formed a cluster adjacent to *C. annuum* on top and *C. baccatum* species below. Other *Capsicum* species like *C. eximium, C. coccineum, C. flexuosum*, and *C. hunzikerianum* had distinct branches and separated from other *Capsicum* taxa.

Genetic diversity is crucial for *Capsicum* development and management. This will enable the *Capsicum* populations to adapt to ever-changing environments, a resource of valuable alleles and genes in the population, and contribute to ecosystem diversity (Reed & Frankham, 2003; Leimu et al., 2006). Studies have shown that genetic diversity was correlated with average fitness across populations (Reed & Frankham, 2003; Leimu et al., 2006).

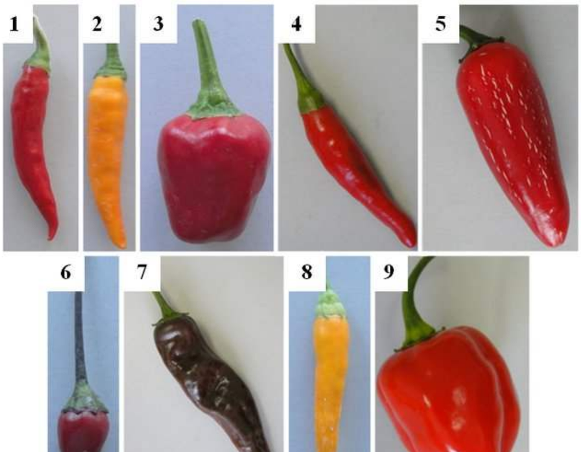

Figure 1**. Cultivars of *C.annum L.* (Adopted Slovakia et al., 2021)** *Cultivars of Capsicum annuum L.: 1 – Black Cobra, 2 –Kilian, 3 – Pepperoncini Greek, 4 – Tabasco, 5 – Jalapeno, 6 –Black Prince, 7 – Chocho, 8 – Medusa, 9 – Habanero Red* (Slovakia et al., 2021)

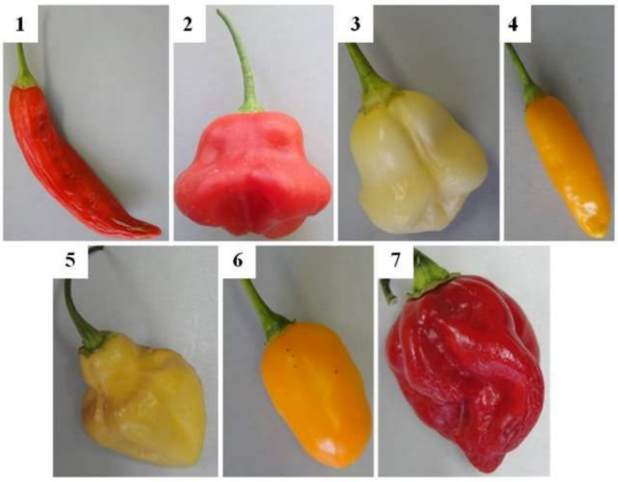

Figure 2. **The cultivars of *Capsicum baccatum L.* (Adopted Slovakia et al., 2021)** *Cultivars of Capsicum baccatum L.: 1 – Escabeche, 2 – Bishops Crown Red, 3 – Aji Fantasy Sparkly White, 4 – Aji Amarillo, 5 – Puerto Rican, 6 – El Oro de Ecuador, 7 – Habanero Red Savina* (Slovakia et al., 2021)

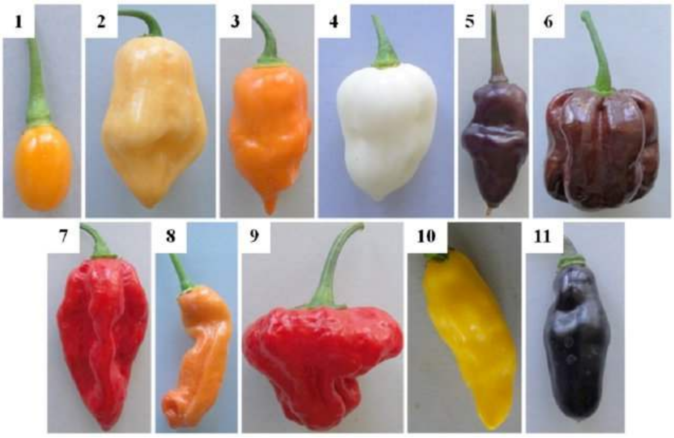

Figure 3. **The cultivars of *Capsicum chinense Jacq*. (Adopted Slovakia et al., 2021** *Cultivars of Capsicum chinense Jacq.: 1 – Aji Charapita, 2 – Habanero Peach, 3 – Trinidad Scorpion Peach, 4 – Jolokia White, 5 – Fidalgo Roxa, 6 – Habanero Chocolate, 7 – Fatalii Red, 8 – Peter Orange, 9 – Red Mushroom, 10 – Citron, 11 – Pimenta de Neyd* (Slovakia et al., 2021)

For phylogenetic relationships of *Capsicums,* 22 Nicotinamide Adenine Dinucleotide Dehydrogenase (NADH dehydrogenase) DNA based sequences from 19 different species were obtained from NCBI. NADH dehydrogenase is a flavoprotein-containing oxidoreductase that catalyzes the conversion of NADH to NAD. The enzyme may be found in eukaryotes as a part of the mitochondrial electron transport complex I and in transferring electrons from photoproduced stromal reductants like NADPH and ferredoxin to the intersystem plastoquinone pool. NADH dehydrogenase is the primary enzyme complex in the electron transport chain in mitochondria. Nicotinamide adenine dinucleotide (NAD) is transformed from its reduced form, NADH, to its oxidized form, NAD+ (Ifuku et al., 2011). The phylogenetic study on the waxy gene (Rodríguez et al., 2011) indicated that *C. chinense and C. frutescens* were grouped together. More so, a study on the trnC-rpoB intron, trnH-psbA intron, and waxy gene sequence data from seven *Capsicum spp.* also revealed that *C. chinense, C. annuum*, and *C. frutescens* grouped in the same cluster (Carrizo García et al., 2016). The above grouping supported our *Capsicum's* phylogenetic analysis based on NADH Dehydrogenase (Figure 4). *Capsicum hunzikerianum* was the outgroup in this phylogenetic tree. These grouping where considered based on the bootstrap values above 50 percent. Before analysis, portions of the alignment with gaps in at least 50% of the taxa were eliminated from the data set (these regions were almost universally phylogenetically uninformative).

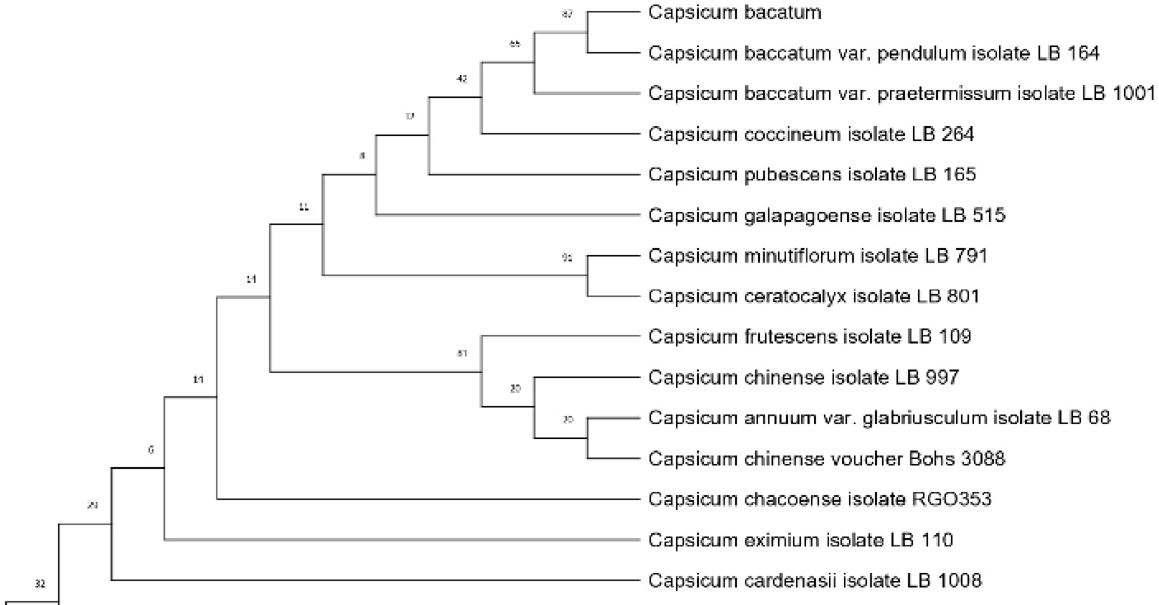

Figure 4. **Phylogenetic relationships (UPGMA) of 22 sequences of Capsicum based on NADH dehydrogenase.**

CONCLUSION, IMPLICATION, SUGGESTION AND LIMITATIONS

Capsicum is a well-diverse shrub commonly grown annual in different parts of the world. It is enetically diverse, with numerous cultivars and taxonomic derivatives. It has a widely described chromosome number of 2n = 24; however, some wild *Capsicum* species, such as *C.buforum, C. capylopodiume*, and *C. capylopodiume* have a ploidy level of 2n = 26. *Capsicum* had a relatively similar genomic content with tomato, as evidenced by the presence of pepper sequences complementary to many tomato cDNAs. However, the pepper genome had been significantly rearranged compared with the tomato genome. To date, important *Capsicum* gene sequences are available in the GenBank for future research on improving the yield, biotic, and abiotic traits of *Capsicum*. This comprehensive review of *Capsicum* genetics and genomics is essential in understanding the biology of *Capsicum* towards improving its horticultural traits and nutritional and medicinal values.

Based on this study, more research on this genus is vital for further studies to take place. Given the widespread usage of these plants, it would be of utmost significance for scientists to thoroughly examine the genetics of as many species as possible, and cultivars. Further studies about the genetics and genomics of *Capsicum* are highly recommended as they the findings would be essential to the improvement of this genus and its applications. No map of *Capsicum* has yet been developed thus thoroughly defining and saturating the pepper chromosomes is highly needed.

ACKNOWLEDGEMENT

I would like to express our sincere gratitude to our adviser, Dr. Orlex Yllano for the constant support he rendered throughout this thesis. I would like to appreciate and recognize his patience and enthusiastic encouragement as well as his critiques of this work.

I would also like to extend our thanks to the following Leilani D. Arce, Neil John V. Vegafria, Ephraim A. Evangelista, Ferdinand A. Esplana, Lester Harris R. Catolico, Merbeth Christine L. Pedro and Edgar E. Tubilag, the faculty of the Biology department of the Adventist University of the Philippines, who have always encouraged me to study hard and remain focused.

Lastly, we are more grateful to my parents for their unceasing support throughout my studies, both financially and emotionally, and also for setting a good example of handwork and perseverance.

REFERENCES

Adhikari, P., McNellie, J., & Panthee, D. R. (2020). Detection of Quantitative Trait Loci (QTL) Associated with the Fruit Morphology of Tomato. Genes, 11(10), 1117. https://doi.org/10.3390/genes11101117

Aguilar-Melendez, A., Morrell, P. L., Roose, M. L., & Kim, S.-C. (2009). Genetic diversity and structure in semiwild and domesticated chiles (*Capsicum annuum; Solanaceae*) from Mexico. American Journal of Botany, 96(6), 1190–1202. https://doi.org/10.3732/ajb.0800155

Barchi, L., Rabanus-Wallace, M. T., Prohens, J., Toppino, L., Padmarasu, S., Portis, E., Rotino, G. L., Stein, N., Lanteri, S., & Giuliano, G. (2021). Improved genome assembly and pan-genome provide key insights into eggplant domestication and breeding. The Plant Journal, 107(2), 579–596. https://doi.org/10.1111/tpj.15313

Basu, S.K., & De, A.K. (2003). *Capsicum*: historical and botanical perspectives. https://www.semanticscholar.org/paper/Capsicum%3A-historical-and-botanical-perspectives-Basu-De/b1d8db5673dca8e98bed971f4a8e7ec0254fea9f

Botanical Garden, Slovak University of Agriculture in Nitra, Slovakia, M. M. Gryshko National Botanical Garden of Ukraine of National Academy of Sciences, Kyiv, Ukraine, Institute of Biological Conservation and Biosafety, Slovak University of Agriculture in Nitra, Slovakia, Department of Technology and Quality of Plant Products, Slovak University of Agriculture in Nitra, Slovakia, Department of Human Nutrition, Slovak University of Agriculture in Nitra, Slovakia, Mňahončáková, E., Vergun, O., Grygorieva, O., Sedláčková, V. H., Ivanišová, E., Šramková, K. F., Hrúzová, M., & Brindza, J. (2021). Evaluation of the antioxidant potential of *Capsicum annuum L., C. baccatum L. and C. chinense Jacq*. Cultivars [pdf]. Acta Scientiarum Polonorum Technologia Alimentaria, 20(2), 223–236. https://doi.org/10.17306/J.AFS.2021.0941

Brewer, M. T., Lang, L., Fujimura, K., Dujmovic, N., Gray, S., & van der Knaap, E. (2006). Development of a Controlled Vocabulary and Software Application to Analyze Fruit Shape Variation in Tomato and Other Plant Species. Plant Physiology, 141(1), 15–25. https://doi.org/10.1104/pp.106.077867

Carrizo García, C., Barfuss, M. H. J., Sehr, E. M., Barboza, G. E., Samuel, R., Moscone, E. A., & Ehrendorfer, F. (2016). Phylogenetic relationships, diversification and expansion of chili peppers (*Capsicum, Solanaceae*). Annals of Botany, 118(1), 35–51. https://doi.org/10.1093/aob/mcw079

Chaim, A., Borovsky, Y., Rao, G., Gur, A., Zamir, D., & Paran, I. (2006). Comparative QTL mapping of fruit size and shape in tomato and pepper. Israel Journal of Plant Sciences, 54(3), 191–203. https://doi.org/10.1560/IJPS_54_3_191

Doganlar, S., Frary, A., Daunay, M.-C., Lester, R. N., & Tanksley, S. D. (2002). Conservation of Gene Function in the Solanaceae as Revealed by Comparative Mapping of Domestication Traits in Eggplant. Genetics, 161(4), 1713–1726. https://doi.org/10.1093/genetics/161.4.1713

García-González, C. A., & Silvar, C. (2020). Phytochemical Assessment of Native Ecuadorian Peppers (*Capsicum spp.*) and Correlation Analysis to Fruit Phenomics. Plants, 9(8), 986. https://doi.org/10.3390/plants9080986

Hawkes, J. G., Lester, R. N., & Skelding, A. D. (Eds.). (1979). The Biology and taxonomy of the Solanaceae. Published for the Linnean Society of London by Academic Press. Howard LR, Wildman RE. (2007) Antioxidant vitamin and phytochemical content of fresh and processed pepper fruit (*Capsicum annuum*), In: Wildman REC (Ed), Nutraceuticals and Functional Foods, 2nd Ed. CRC Press, Boca Raton, Florida, 165-191.

Hurtado, M., Vilanova, S., Plazas, M., Gramazio, P., Herraiz, F. J., Andújar, I., & Prohens, J. (2013). Phenomics of fruit shape in eggplant (Solanum melongena L.) using Tomato Analyzer software. Scientia Horticulturae, 164, 625–632. https://doi.org/10.1016/j.scienta.2013.10.028

Ifuku, K., Endo, T., Shikanai, T., & Aro, E.-M. (2011). Structure of the Chloroplast NADH Dehydrogenase-Like Complex: Nomenclature for Nuclear-Encoded Subunits. Plant and Cell Physiology, 52(9), 1560–1568. https://doi.org/10.1093/pcp/pcr098

Jarret, R. L. (2008). DNA barcoding in a crop genebank: the *Capsicum annuum* species complex. The Open Biology Journal, 1(1), 35-42.

Knapp, S. (2002). Tobacco to tomatoes: A phylogenetic perspective on fruit diversity in the Solanaceae. Journal of Experimental Botany, 53(377), 2001–2022. https://doi.org/10.1093/jxb/erf068

Leimu, R., Mutikainen, P. I. A., Koricheva, J., & Fischer, M. (2006). How general are positive relationships between plant population size, fitness and genetic variation?. Journal of Ecology, 94(5), 942-952.

Misawa, K., & Tajima, F. (2000). A Simple Method for Classifying Genes and a Bootstrap Test for Classifications. Molecular Biology and Evolution, 17(12), 1879–1884. https://doi.org/10.1093/oxfordjournals.molbev.a026289

Moulin, M. M., Rodrigues, R., Ramos, H. C. C., Bento, C. S., Sudré, C. P., Gonçalves, L. S. A., & Viana, A. P. (2015). Construction of an integrated genetic map for *Capsicum baccatum* L. Genetics and Molecular Research, 14(2), 6683–6694. https://doi.org/10.4238/2015.June.18.12

Nunome, T., Ishiguro, K., Yoshida, T., & Hirai, M. (2001). Mapping of Fruit Shape and Color Development Traits in Eggplant(Solanum melongena L.) Based on RAPD and AFLP Markers. Breeding Science, 51(1), 19–26. https://doi.org/10.1270/jsbbs.51.19

Passam, Harold & Karapanos, Ioannis. (2016). Eggplants, peppers and tomatoes: Factors affecting the quality and storage life of fresh and fresh-cut (minimally processed) produce. The European Journal of Plant Science and Biotechnology. 2. 156-170.

Reed, D. H., & Frankham, R. (2003). Correlation between Fitness and Genetic Diversity. Conservation Biology, 17(1), 230–237. https://doi.org/10.1046/j.1523-1739.2003.01236.x

Rodríguez, G. R., Muños, S., Anderson, C., Sim, S.-C., Michel, A., Causse, M., Gardener, B. B. M., Francis, D., & van der Knaap, E. (2011). Distribution of SUN, OVATE, LC , and FAS in the Tomato Germplasm and the Relationship to Fruit Shape Diversity. Plant Physiology, 156(1), 275–285. https://doi.org/10.1104/pp.110.167577

Taher, D., Solberg, S. Ø., Prohens, J., Chou, Y., Rakha, M., & Wu, T. (2017). World Vegetable Center Eggplant Collection: Origin, Composition, Seed Dissemination and Utilization in Breeding. Frontiers in Plant Science, 8, 1484. https://doi.org/10.3389/fpls.2017.01484

Tripodi, P., & Kumar, S. (2019). The *Capsicum* Crop: An Introduction. In N. Ramchiary & C. Kole (Eds.), The *Capsicum* Genome (pp. 1–8). Springer International Publishing. https://doi.org/10.1007/978-3-319-97217-6_1

Vilarinho, L. B. O., Henriques da Silva, D. J., Greene, A., Salazar, K. D., Alves, C., Eveleth, M., Nichols, B., Tehseen, S., Khoury, J. K., Johnson, J. V., Sargent, S. A., & Rathinasabapathi, B. (2015). Inheritance of Fruit Traits in *Capsicum annuum*: Heirloom Cultivars as Sources of Quality Parameters Relating to Pericarp Shape, Color, Thickness, and Total Soluble Solids. Journal of the American Society for Horticultural Science, 140(6), 597–604. https://doi.org/10.21273/JASHS.140.6.597

Zygier, S., Chaim, A. B., Efrati, A., Kaluzky, G., Borovsky, Y., & Paran, I. (2005). QTLs mapping for fruit size and shape in chromosomes 2 and 4 in pepper and a comparison of the pepper QTL map with that of tomato. Theoretical and Applied Genetics, 111(3), 437–445. https://doi.org/10.1007/s00122-005-2015-7