UNIVERSIDAD AUTÓNOMA CHAPINGO

INSTITUTO DE HORTICULTURA

NEW APPROACHES FOR *Opuntia* sp. CHARACTERIZATION AND TAXONOMY: SEED MORPHOMETRIC DESCRIPTORS, PROTEINS AND DNA MARKERS

TESIS

QUE PRESENTA:

SAMIR SAMAH

QUE COMO REQUISITO PARCIAL PARA OBTENER EL GRADO DE:

DOCTOR EN CIENCIAS EN HORTICULTURA

Diciembre de 2014

Chapingo, Estado de México

NUEVOS ENFOQUES PARA LA CARACTERIZACIÓN Y TAXONOMÍA DE ACCESIONES DE NOPAL (*Opuntia* sp.): DESCRIPTORES MORFOMÉTRICOS DE LAS SEMILLAS, MARCADORES PROTEICOS Y DE DNA

Tesis realizada por **Samir Samah** bajo la dirección del comité asesor indicado, aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

DOCTOR EN CIENCIAS EN HORTICULTURA

PRESIDENTE

DRA. ERNESTINA VALADEZ MOCTEZUMA

ASESOR

DRA. ELSA VENTURA ZAPATA

ASESOR

DR. MIGUEL ANGEL SERRATO CRUZ

LECTOR EXTERNO

DR. JESÚS LÓPEZ REYNOSO

AGRADECIMIENTOS

Expreso mi más sincero agradecimiento a los contribuyentes de México quienes por conducto de la **Secretaria de Relaciones Exteriores (SRE)** financiaron esta etapa de mi formación profesional de doctorado.

A la Universidad Autónoma Chapingo, en especial al Instituto de Horticultura y a cada uno de los doctores que lo conforman, por todas las enseñanzas y experiencias recibidas que contribuyeron a una formación académica y profesional de excelencia.

A la **Dra. Ernestina Valadez Moctezuma** por su atinada atención y apoyo, asi como por haberme dado todas las facilidades para realizar esta investigación y la confianza mostrad durante la estancia en este doctorado.

A la Dra. Elsa Ventura Zapata por el apoyo invaluable, por la siempre amable disposición para asesorarme en la realización de esta investigación.

Al Dr. Miguel Ángel Serrato Cruz por su apoyo, consejos y disposición.

Al Dr. José de Jesús López Reynoso por su dedicación en la revisión de la tesis y su apoyo.

A mi familia por su apoyo, amor, compresión y la fortaleza para sobrellevar la lejanía durante estos tres años.

A mis compañeros y amigos del **laboratorio de Biología Molecular**, de la **UACH** y de **México** en general quienes fueron siempre atentos, amables y dispuestos a apoyarme en cualquier momento.

DATOS BIOGRÁFICOS

Samir Samah, nacido en el pueblo el Ouassaf, Grarem Gouga, Wilaya de Mila, Argelia. Cursó sus estudios preparatorios en ciencias naturales y de la vida en la preparatoria de Rabah Khadrouch, Grarem, Gouga.

En el año 2002 curso la carrera de ingeniero agrónomo en la Escuela Nacional Superior de Agronomía (anteriormente, Instituto Nacional Agronómico) de Argel, Argelia, donde obtuvo el grado de Ingeniero Agrónomo Especialidad en Fitotecnia en el año 2007. Fue premiado por ocupar el primer lugar de 114 estudiantes para la generación 2002/2007 (5 años de estudios) con una calificación de 14.1 sobre un máximo de 20.

Se trasladó a España al Instituto Agronómico Mediterráneo de Zaragoza (España) del Centro Internacional de Altos Estudios Agronómicos Mediterráneos (CIHEAM). Universidad de Lleida, España, en el periodo del 2008-2010 donde obtuvo el grado de Master en Mejoramiento Vegetal.

En el 2012 ingresó al Doctorado en Ciencias en Horticultura, Universidad Autónoma Chapingo.

NEW APPROACHES FOR *Opuntia* sp. CHARACTERIZATION AND TAXONOMY: SEED MORPHOMETRIC DESCRIPTORS, PROTEINS AND DNA MARKERS

NUEVOS ENFOQUES PARA LA CARACTERIZACIÓN Y TAXONOMÍA DE ACCESIONES DE NOPAL (*Opuntia* sp.): DESCRIPTORES MORFOMÉTRICOS DE LAS SEMILLAS, MARCADORES PROTEICOS Y DE DNA

Samir Samah[1] and Ernestina Valadez-Moctezuma[2]

ABSTRACT

The genus *Opuntia* is cultivated basically for fruit and vegetable production. Different species offer three forms: sweet ("tunas") and acid fruits ("xoconostles") and tender cladodes ("nopalitos"). However, few studies on the genetic structure of Mexican Opuntia are available. Identification of Opuntia genotypes is still mainly based on morphological traits. Nevertheless, molecular studies have suggested the taxonomic revision of this genus. In this study, three approaches were undertaken for genetic characterization and contribution to the taxonomy of the most recognized genotypes in Mexico: morphological seed analysis, total and storage seed protein and DNA markers. It was confirmed that species classification and accession assignation are erroneous. Seed variables have high discriminatory power and it is recommended that they be taken into account as potential parameters for genotype assignation. Similarly, we propose that protein molecular markers be included in genetic diversity studies and for the taxonomic revision of the genus. In the present study the combination of data on both total and reserve seed protein profiles was needed to differentiate all the Opuntia accessions studied. Genetic distances were very narrow and most of the genotypes had the same genetic profile based on the phylogenetic relationship of seven genomic regions and the FRUITFULL gene. The genetic divergence between "tunas" and "xoconostles" is very narrow and a clear separation between the two fruit variants was obtained only with SSR analysis. The genetic structure of Opuntia was found to be complex as both linear and reticulate ties among the Mexican Opuntia germplasm were revealed.

Key words: Nopal, systematic, genetic diversity, phylogenetic relationship.

RESUMEN

El género *Opuntia* es considerado un cultivo frutícola y hortícola ya que produce frutos dulces (tunas), frutos ácidos ("xoconostles") y "nopalitos". Sin embargo, estudios sobre la estructura genética de nopal mexicano son escasos. La identificación de genotipos de Opuntia se ha basado principalmente en caracteres morfológicos. Sin embargo, los estudios moleculares sugieren la revisión taxonómica de este género. En el presente estudio, se llevaron a cabo tres enfoques para la caracterización genética y la contribución a la taxonomía de los genotipos más reconocidos en México: análisis morfológico de las semillas, proteínas de reserva y totales de las semillas y marcadores de DNA. Se confirmó la errónea clasificación de las especies y la equívoca asignación de accesiones a sus respectivas especies. Las variables de semillas tienen un alto poder discriminatorio y se sugiere considerarlos como parámetros para la asignación taxonómica de genotipos. Del mismo modo, se propone incluir los marcadores proteicos para estudios de la diversidad genética y para la revisión taxonómica del género. En este estudio fue necesaria la combinación de datos de los perfiles de proteínas totales y de reserva para diferenciar todas las accesiones estudiadas de Opuntia. La distancia genética fue muy pequeña y la mayoría de los genotipos tuvieron el mismo perfil genético basado en la relación filogenética de las siete regiones genómicas y del gen FRUITFULL. La divergencia genética entre las tunas y xoconostles es muy estrecha y la clara separación entre ambos tipos se obtuvo solamente en el análisis SSR. La estructura genética de Opuntia resultó compleja ya que ambas formaciones lineales y reticulares fueron reveladas

Palabras clave: Nopal, sistemática, diversidad genética, relación filogenética.

[1]Estudiante de doctorado, Doctorado en Ciencias en Horticultura. Departamento de fitotecnia, Universidad Autónoma Chapingo. 38.5 Carretera México-Texcoco. Chapingo. México. E-mail: samirsamah2001@yahoo.fr
[2] Profesor-investigador del Departamento de fitotecnia. Universidad Autónoma Chapingo. 38.5 Carretera México-Texcoco. Chapingo. México. E-mail: evaladez@correo.chapingo.mx

TABLE OF CONTENT

GENERAL INTRODUCTION AND OBJECTIVES

INTRODUCTION

Mexico, as it is known, is one of the centers of origin and domestication of the most important plants in the world. Particularly, the genus *Opuntia* (prickly pear or nopal) is represented by 66-83 species with a wide range of varieties and several degrees of domestication (Reyes-Agüero *et al.,* 2005). The economic and agronomic prominence of the nopal in Mexico and other countries and their ability to grow in semiarid environments has motivated several studies to understand the morphological and physiological variation, crop cultivation and molecular studies. These studies stand out that: (1) the Mexican nopal presents a broad genetic variablility manifested in several phenotypes adapted to different environments, (2) likely over time, some of these features have been set due to human selection and (3) even though the majority of the materials studied to date are located in many species; molecular analyzes suggest otherwise.

The number of species included in the genus is not known (Caruso *et al.,* 2010), but the number of species identified in the genus ranges from 160 to 250. This difference is mainly due to nomenclatural problems occurring not only in Opuntia, but also within other genera of the Opuntioideae subfamily. The main reasons for such taxonomical chaos are the scarcity of morphological characters, the lack of solid descriptors, the high level of phenotypic, the recent diversification, the presence of apomixis and the occurrence of polyploidy. As a result of the incorrect assignment, the same varieties are often classified as belonging to different species, and in other cases are considered hybrids between unknown parentals (Caruso *et al.,* 2010).

In Mexico, attempts to organize and classify the cactus dates back to mesoamerican era where they already established two groups: the "Comitl" group includes plants with spherical and cylindrical stems; and "Napalli" group that includes species with jointed, flattened and discoid stems. The different classes are distinguished by adding the word *nochtli* specifying one or more morphological qualities (Granados and Castañeda, 1991). During the next two centuries, the taxonomy was based on this system of classification and it was until the twentieth century that biochemical, physiological and cytogenetic aspects were considered. Recently, modifications of actual classification have been suggested as DNA markers results were published (Valadez-Moctezuma *et al.*, 2014 a; b).

Molecular markers, which are reproducible and stable, might be useful tools to address doubts in classification that are not addressed by morphological characterization. In the past three decades, few studies have been performed to characterize existing opuntia collections using random molecular markers (Wang *et al.*, 1998; Luna-Paez *et al.*, 2007). Two studies employed different molecular tools to elucidate taxonomical aspects of the genus, particularly the origin of *O. ficus indica* (Labra *et al.*, 2003; Griffith, 2004). Recently, other ambiguities in the taxonomic classification of *Opuntia* species emerged in a study that used microsatellite polymorphisms to try to be discriminate between two morphologically distinct *O. echios* varieties (echios and gigantea) native to the Galapagos Islands (Helsen *et al.*, 2009). Once again, the authors highlighted that the current taxonomic differentiation between these taxa was not supported by molecular data. Caruso *et al.* (2010) conducted a study to determine the effectiveness of microsatellites markers in assessing the genetic diversity of the genotypes of *O. ficus-indica*, and determined, their relationship with wild accessions and related species. Particularly, the erroneous taxonomic

assignment of nopal species in Mexico was reported (Valadez-Moctezuma *et al.*, 2014a; b) applying RAPD and ISSR analysis.

From the above background, we emphasize that 1) the morphological analyses previously published didn't considered the discriminative value of seed characteristics; 2) the potential use of the seed protein profiles for genetic diversity study is unknown; 3) the genetic analsis based on DNA markers is scarce and the number of genotypes studied has been limited; 4) the phylogenetic relationship between the Mexican Opuntias and genetic separation between "tunas" and "xoconostles" is not revealed.

Given the importance and representativeness of the genus *Opuntia* in Mexico, collections with commercial or agronomic importance are safeguarded in germplasm banks; these collections are described from a morphological and anthropocentric point of view. Strengthening the characterization of these plant resources with alternative tools will reveal more efficiently the existing phenetic and genetic diversity, and support their correct taxonomic assignment, and also contributing to breeding programs for the creation of new varieties (Illoldi-Rangel *et al.*, 2012). Today, to market the Opuntia fruits and its derived products, it is necessary to specify the scientific name of the species or the horticultural forms. The lack of this information hinders and even prevents commercial transactions.

OBJECTIVES

Main objective

This research was undertaken to study the genetic relatedness of Mexican Opuntia accessions through morphological, biochemical and DNA markers, and to support their differentiation and taxonomy.

Specific objectives

1. Determine the discriminatory power of the characteristics derived from internal and external morphology of the seeds (Chapter 2).
2. Characterize an extensive collection of Opuntia using seed proteins analysis (Chapter 3).
3. Apply SSR analysis to differentiate Opuntia genotypes, estimate genetic diversity and determine the population structure (Chapter 4, Part 1).
4. Elucidate the phylogenetic relationship among Opuntia accessions by means of the PCR-RFLP technique (Chapter 4, Part 2).
5. Estimate the genetic diversity of the Fruitfull fruit gene through CAPS marker (Chapter 4, Part 3).
6. Separate xoconostles from prickly pear genotypes (Chapter 2; Chapter 4, Part 1 and Part 2)
7. General discussion about the findings of molecular taxonomy in the genus *Opuntia* and propose a systematic of the genotypes studied (Chapter 5).

REFERENCES

Caruso, M., Currò, S., Las Casas, G., La Malfa, S., and Gentile, A., 2010. Microsatellite markers help to assess genetic diversity among *O. ficus-indica* cultivated genotypes and their relation with related species. Plant Syst. Evol. 290: 85-97.

Granados, S.A. and Castañeda, P.A.D. 1991. El Nopal. Historia, Fisiología, Genética e Importancia Frutícola. Editorial Trillas. D.F., México. 227 p.

Griffith, M. P. 2004. The origins of an important cactus crop, *Opuntia ficus-indica* (Cactaceae): new molecular evidence. Am. J. Bot. 91: 1915-1921.

Helsen, P., Verdyck, P., Tye, A., and Van Dongen, S. 2009. Low levels of genetic differentiation between *Opuntia echios* varieties on Santa Cruz (Galapagos). Plant. Syst. Evol. 279: 1-10.

Illoldi-Rangel, P., Ciarleglio, M., Sheinvar, L., Linaje, M., Sánchez-Cordero, V. and Sarkar, S. 2012. *Opuntia* in México: Identifying Priority Areas for Conserving Biodiversity in a Multi-Use Landscape. PLoS ONE 7: e36650. doi:10.1371/journal.pone.0036650.

Labra, M., Grassi, F., Bardini, M., Imazio, S., Guiggi, A., Citterio, S., Banfi, E., Sgorbati, S. 2003. Relationships in *Opuntia* Mill. Genus (Cactaceae) detected by molecular marker. Plant Sci. 165: 1129-1136.

Luna-Paez, A., Valadez-Moctezuma, E., Barrientos-Priego, A. F., and Gallegos-Vázquez, C. 2007. Characterization of *Opuntia* spp. by Means of Seed with RAPD and ISSR Markers and Its Possible Use for Differentiation. JPACD 9: 43-59.

Reyes-Agüero, J. A., Aguirre, J. R. and Hernández, H. M. 2005. Systematic notes and a detailed description of *Opuntia ficus-indica* (L.) Mill. (Cactaceae). Agrociencia 39: 395-408.

Scheinvar, L. 1995. Taxonomy of utilized opuntias. In: Barbera, G., Inglese, P., Pimienta-Barrios, E. (Eds.) Agroecology, cultivation and uses of cactus pear. FAO plant production and protection paper 132. FAO, Rome, pp 20-27.

Valadez-Moctezuma, E., Samah, S., Luna-Paez, A. 2014a. Genetic diversity *of Opuntia* spp. varieties assessed by classical marker tools (RAPD and ISSR). Plant Syst. Evol. doi: 10.1007/s00606-014-1112-y.

Valadez-Moctezuma, E., Ortiz-Vásquez, Q., Samah, S., 2014b. Molecular based assessment of genetic diversity of xoconostle accessions (*Opuntia* spp.). Afr. J. Biotechnol. 13: 202-210.

Wang, X., Felker, P., Burow, D. M. and Paterson, H. A. 1998. Comparison of RAPD marker patterns to morphological and physiological data in the classification of *Opuntia* accessions. J. PACD 3: 3-14.

Chapter 1

OPUNTIA CHARACTERIZATION AND TAXONOMY

OPUNTIA CHARACTERIZATION AND TAXONOMY

EVOLUTION AND SYSTEMATICS OF THE CACTACEAE

Cactaceae is one of the most important plant families of the New World's arid regions. Cactaceae is also one of the most popular plant families in horticulture and has been the subject of interest of many botanists since 18th century. The Cactaceae is morphologically distinct and doubtless supported as monophyletic by morphological synapomorphies and molecular data (Gibson and Nobel, 1986; Hernández-Hernández *et al.*, 2011). It belongs to the order Caryophyllales in which it is characterized by succulent plant. It is a morphologically very diverse family. It has evolved a variety of growth-forms ranging from tree-like, large columnar forms to shrubby forms or succulent climbers and to small globular forms, present areoles in the stem, in the pericarp of the flower and in the surface of most of the fruits; with the CAM metabolism; nitrogenous pigments (betalains) (Scheinvar, 1995).

Although the monophyly of the Cactaceae has hardly ever been questioned, the establishment of taxonomic units within the Cactaceae has been always difficult and controversial. Currently, the Cactaceae is subdivided into four subfamilies: Pereskioideae, Maihuenioideae, Opuntioideae and Cactoideae. The Cactoideae is the largest subfamily, representing seven tribes: Cacteae (25 genera), Cereeae (15), Echinocereeae (25), Hylocereeae (six), Notocacteae (seven), Rhipsalideae (four), and Trichocereeae (23). Opuntioideae is the next largest subfamily and includes two tribes, Opuntieae and Cylindropuntieae, comprising ten and seven genera, respectively, and 192 species of which 75 species are placed in the largest genus *Opuntia s.s*. Two other subfamilies, Pereskioideae and Maihuenioideae, each comprise a single genus (Hunt, 2002). The major

Cactaceae lineages have been identified in molecular phylogenetic studies. But at the same time, it was found that most of the tribes and genera as traditionally defined are not monophyletic (Korotkova, 2011).

Defining species limits in Cactaceae is difficult. Individual populations of one species can vary considerably in their morphology due to phenotypic plasticity and responses to the environmental conditions. Consequently, interpretation of morphological characters is often troublesome. Moreover, many species have been described based on one or only few individuals. As a result, Cactaceae had been heavily over described approximately 15000 binomials exist (Anderson, 2001). This is ten times higher than the number of currently accepted species (Hunt, 2006).

EVOLUTION AND SYSTEMATICS OF THE OPUNTIOIDEAE

The subfamily Opuntioideae is widespread throughout the Americas from Canada to southern Patagonia. It has traditionally been recognized as a monophyletic taxonomic entity (Griffith and Porter, 2009). It is characterized by a number of synapomorphies: (1) presence of glochids: small, deciduous barbed spines (2) woody funicular tissue surrounding the seed (funicular envelope) (3) high amounts of calcium oxalate monohydrate druses and monoclinic cluster crystals in the outer hypodermis of stems, and (4) polyporate pollen grains with peculiar exine structures. The seeds of the Opuntoideae are unique, in being entirely encased by a hard aril derived from the funiculus (Stuppy, 2002). Five tribes (Wallace and Dickie, 2002) are recognized within Opuntioideae. Tribe Opuntieae (platyopuntioids) is a well-supported clade within Opuntioideae (Hernández-Hernández *et al.*, 2011) that consists of *Brasiliopuntia* (K. Schumann) A. Berg., *Consolea* Lemaire, *Miqueliopuntia* Frič ex F. Ritter, *Nopalea* Salm-Dyck, *Opuntia* s.s., *Salmiopuntia*

Frič ex Guiggi, *Tacinga* Britton & Rose, and *Tunilla* Hunt and Illiff. The platyopuntioids were so named for the flat, photosynthetic stem segments (i.e., cladodes) characteristic of most members. Species of *Maihueniopsis* s.l. were also recovered in Opuntieae (Griffith and Porter, 2009), but this genus is often placed in the tribe Cumulopuntieae (Hunt, 2002). Thus, *Opuntia* s.s. has been reduced drastically in size with many segregate genera [e.g., *Austrocylindropuntia* Backeb., *Brasiliopuntia* (K. Schum.) A. Berger, *Cylindropuntia* (Engelm.) F. M. Knuth] now recognized (Griffith and Porter, 2009).

Opuntia s.s. (nopales, prickly pears) is the largest genus in Opuntioideae and the most widespread genus in Cactaceae, distributed natively from Canada to Argentina (Anderson, 2001). There are 215 species (Hunt, 2002) within the genus, of which, there are 66-83 reported only in Mexico (Guzmán *et al.*, 2003). This genus is suggested to have originated as recently as 5.6 (± 1.9) Ma (Arakaki *et al.*, 2011). Members of *Opuntia* s.s. are cultivated worldwide as fruit and vegetable crops and are increasingly used as forage and fodder for livestock in arid areas of the world, such as parts of Brazil, Mexico, western Asia, and northern and southern Africa. In Mexico, where species of *Opuntia* have been cultivated for at least the last 14 000 yr (Casas and Barbera, 2002), they represent an iconic national figure, illustrated on the country's flag. The Aztecs and other mesoamerican civilizations used the cactus pads as a vegetable or fodder and the prickly pear as a seasonal fruit; sweet (cactus pear) or acid fruits (xoconostle).

According to some authors, Opuntia comes from Opus, Opunte or Opuncia population of ancient Greece in the Leócrida region, where growing some thorny plants like cacti (Bravo, 1978). According to others, the origin of the name dates back to Pliny (23-79 A.C.), who called thorny plant (artichoke) *opuncia* which growing near Opuns, India.

Matthiolus (1565) is the first to apply the term to a cactus (Bravo, 1978). Indigenous Mexicans had their own nomenclature. The Aztecs distinguished between two distinct groups of cacti; Náhuatl named *nochtli* the cactus, species with green stems, jointed, flattened, and discoid and *comitls* include species with globose and/or cylindrical stems (Scheinvar, 1995). The "tuna" term, originated from Antillean, is still used at the date in the Caribbean islands, Central and South America to designate both plants and fruits of Opuntia, and "xoconostle" from Nahuatl: xoco = acid, noxtle = tuna.

BOTANICAL DESCRIPTION OF OPUNTIA

The plants of the *Opuntia* genus are shrub, tree, sometimes creeping, generally they have four growth habits that can become erect, spreading, decumbent or pendant (Luna-Paez, 2008); height of 0.6 to 5.5 m and with from 0.3 to 5.5 m; branches from the base or sometimes the first branch at a height of 1.5 m, the branches can be upright, diffuse, extended or recumbent; chestnut bark, dark, yellow, green, brown or gray (Reyes-Agüero *et al.,* 2009). The stems are presented as cladodes oblanceolate, obovate, oblong or ovate, sometimes rhombic or circular; rarely undulate margins; length from 17 to 63 cm, width from 12 to 32 cm and thickness ranged between 0.5 and 4 cm. These modified stems to have areolas that are the characteristic structures of the Cactaceae family. Thorns and glochids are present in cladodes, is from these structures were originate new vegetative shoots and usually from the apical areolas arise the floral buds (Bravo, 1978). The buds located on the borders often produce small ephemeral leaves, where new stems, flowers and spines are occurred. Spines up to 10 per areola, sometimes, none, to 153 with thorns´s cladode face; and with radial and diffuse upright orientation.

Flowers has cylindrical pericarp, conical truncated, obovate or oval, sometimes lanceolate or oblong. The floral axis displays, from its base up, the pedicel that joins the flower with the cladode, the pericarp surrounding the lower ovary and the receptacle containing the perianth, androecium and gynoecium (Bravo, 1978). Opuntia flour is generally yellow, but there are also orange, pink, purple, red, white or mottled. Hermaphrodite flowers are the most common (Gibson and Nobel, 1986). Stamens are numerous; 265 to 59; they are generally yellow or green with a circular or spiral arrangement around the style (Reyes-Agüero *et al.,* 2006). The style is simple, hollow, usually green or yellow, although some are pink, red or orange (Bravo, 1978). The stigma appears above the anthers at the center of the flower. Stigma lobes are broad and green, orange or yellow. The unilocular ovary contains numerous ovule or seed primordia (Reyes-Agüero *et al.,* 2006).

The fruit is a unilocular berry, fleshy or juiced (Luna-Paez, 2008), takes several forms: turbinates, obovoid, elliptical, sometimes globose, ovoid or pyriform, sometimes pedunculated. The "tunas" have high respiration rates in response to accelerated metabolic processes and this causes a short postharvest life. Fruits present a wide range of colors: green, yellow, purple, pink, red, etc.; and with hard to soft consistency. Sweetness ranged from 1.7 to 20 °Brix. Xoconostles are fruits produced by a group of plants included in *Opuntia* genus. Xoconostles have a wide inner edible wall, a thin outer wall that is not easily removed, and seeds arranged in the center of the fruit with dry funiculus (Gallegos-Vázquez *et al.,* 2012). Ten species that produce xoconostle fruits had been reported by Scheinvar *et al.* (2009), nine of these belong to the genus *Opuntia* (*O. heliabravoana* Scheinvar, *O. elizondoana* E. Sánchez and Villaseñor, *O. joconostle* F.A.C. Weber, *O. matudae* Scheinvar, *O. spinulifera* Salm-Dyck, *O. leucotricha* DC, *O. zamudioi* Scheinvar,

O. durangensis Britton and Rose, *O. oligacantha* C.F. Förster), and one more specie to the genus *Cylindropuntia* (*Cylindropuntia imbricata* DC).

Opuntia seeds usually are small, except in *O. megasperma* (species not included in the thesis) that has seeds up to 1.3 cm length. Generally take an ovoid or lentiform shape. They are covered on the outside with an indoor hard funicular; the seed coat is thin and not woody. The embryo occupies a huge part of the seed and it is curved, sometimes spiral; the cotyledons are large (Bravo, 1978). The number of seed reached 450 per fruit; normal seed varied from 0 to 306 with a weight fruit reaches 9 grams, length between 0.2 to 0.6 cm, diameter 0.2 to 0.5 cm and hardness of 9-456 kgf. There is considerable confusion in the literature concerning impermeability of seeds with 'hard' seed coats, because the ability to take up (imbibe) water has not been tested in most of them. Seeds of some *Opuntia* species were reported recently to have a water-impermeable seed coat *sensu lato* (i.e. physical dormancy), in combination with physiological dormancy. However, physical dormancy is not known to occur in Cactaceae. Although dormancy in seeds of this species can be broken by scarification, they have physiological dormancy only (Orozco-Segovia *et al.*, 2007).

MOLECULAR TAXONOMY

Taxonomy and systematic

Taxonomy is the science that explores, describes, names, and classifies all organisms (Rouhan and Gaudeul, 2012). However, systematic can be considered to have two major goals: (1) to discover and describe species and (2) to determine the phylogenetic relationships of these species (Hebert and Gregory, 2005). It is not before 1813 that the

Swiss botanist A. P. De Candolle (1778-1841) invented the neologism “taxonomy” from the Greek ταξις (order) and νόμος (law, rule) and published it for the first time in his book “Théorie élémentaire de la Botanique” (“Elementary Theory of Botany”). He defined this scientific discipline as the “theory of the classifications applied to the vegetal kingdom”. Then, the Global Biodiversity Assessment of the United Nations Environment Programme defined taxonomy as “the theory and practice of classifying organisms”, including the classification itself but also the delimitation and description of taxa, their naming, and the rules that govern the scientific nomenclature.

Depending on the authors, taxonomy is viewed either as a synonym for the systematic science also called biosystematic, including the task of classifying species, or only as a component of systematic restricted to the delimitation, description, and identification of species. This latter meaning of taxonomy has emerged lately, with the advent of phylogenetics as another component of systematic that allows classifications based on the evolutionary relationships among taxa (Rouhan and Gaudeul, 2012).

Different methods were developed often based on newly available technologies and based on distinctive philosophies or novel thinking to obtain the same end product (taxa). The first developed tool was simply the visual description, i.e., morphology or habit. With technological advances, i.e., lenses and the microscope, micromorphology and anatomy became added as tools. Over time taxonomists have added many tools to their repertoire in order to achieve better taxonomies (Baum, 2012). Today, most recognized species have been delineated and described based on morphological evidence: in general, they have been delimited based on one or more qualitative or quantitative morphological characters that show no or very little overlap with other species (Wiens, 2007). The initial enthusiasm

for molecular taxonomy most probably came from the additional and complementary information that it provided. It makes use of tools that are not specific to a particular group of plants, and it may appear more prone to scientific publications in peer-reviewed journals than more traditional, taxonomic studies.

Phylogeny inference at species level

The species is often seen as the fundamental unit of evolution but species delimitation depends on what a species is; since species definition has long remained highly debated. The existence of species itself is somewhat controversial, especially in plants where asexuality, hybridization, and polyploidy may render the definition and delimitation of species complex. Some argue that species are "arbitrary constructs of the human mind"; others claim that they are objective, discrete entities (Korotkova, 2011).

The proliferation of species concepts started in the 1970s. It gave rise to several decades of debate and taxonomic instability because many concepts were incompatible in that they lead to the recognition of different species boundaries and diverse number of species. Morphological approaches have dominated species delimitation for centuries, starting with the purely typological (i.e., essentialist) pre-Darwinian view. But most contemporary biologists are familiar with the idea that species are groups of actually or potentially interbreeding natural populations, which are reproductively isolated from other groups (the "Biological Species Concept"), whether or not they differ in phenotypic characters that are readily apparent (Rouhan and Gaudeul, 2012). Other concepts are summarized in Table 1.

Table 1. Some alternative contemporary species concepts/criteria (Rouhan and Gaudeul, 2012).

Name of the species concept/criterion	**Definition of the species**	**Major contributor(s)**
Interbreeding species concept [forms the basis for the General (metapopulation) Lineage Concept]	A group of potentially interbreeding populations	Wright (1940); Dobzhansky (1950)
Isolation species concept a [often called the biological species concept]	A group of potentially interbreeding populations that is reproductively isolated from other such groups	Poulton (1904); Mayr (1942); Dobzhansky (1970)
Phenetic species concept	A group that forms a phenetic cluster (quantitative difference)	Sokal and Crovello (1970)
Ecological species concept	A group that shares the same niche or adaptive zone	Van Vaalen (1976)
Evolutionary species concept a [corresponds closely to the General (metapopulation) Lineage Concept]	A lineage (i.e., an ancestral-descendant sequence of populations) evolving separately from others and with its own evolutionary role and tendencies	Simpson (1951); Wiley (1978)
Phylogenetic species Concept-character diagnosability version	An irreducible (basal) cluster of organisms, diagnosably distinct from other such clusters, and within which there is a parental pattern of ancestry and descent (fixed qualitative character) The diagnostic character can be from any trait (e.g., morphological or molecular) and of any significance (e.g., a single base pair)	Cracraft (1989)
Phylogenetic species Concept-reciprocal monophyly version	A group that shows monophyly (consisting of an ancestor and all of its descendants and commonly inferred from the possession of shared derived character states)	Rosen (1979); Donoghue (1985); Mishler (1985)
Genealogical species concept	A group that shows monophyly for all (or at a consensus of) gene genealogies in the genome	Baum and Shaw (1995)
Genotypic species concept	A group recognizable on the basis of multiple, unlinked, inherited genetic markers A pair of such genotypic clusters is recognizable if the frequency distribution of genotypes is bimodal or multimodal, and strong heterozygote deficits and linkage disequilibria are evident between the clusters	Mallet (1995)
Cohesion species Concept[a]	A group that is characterized by cohesion mechanisms, including reproductive isolation, recognition mechanisms, ecological selection, as well as genealogical distinctiveness	Templeton (1998)

The commonly observed incompatibility between different criteria stems from the fact that various properties actually arise at different stages in the process of speciation: as lineages diverge, they become distinguishable in terms of quantitative traits, diagnosable in terms of fixed character states, reproductively incompatible, they evolve distinct ecologies, and they pass through polyphyletic, paraphyletic, and monophyletic stages. These changes

commonly do not occur at the same time, and they are not even necessarily expected to occur in a specific order. Nevertheless lineages diverge; the number of species criteria satisfied will increase and allow a highly validated hypothesis of lineage separation and species delimitation. The biological and phylogenetic species concepts are most favored as species definitions (Korotkova, 2011).

Marker choice for species-level studies

Species-level phylogenies are among the most interesting but probably also are the most difficult-to-address issues in systematic (Helsen *et al.*, 2009). The evolutionary considerations are of primary importance when one wants to use DNA markers to infer phylogenetic relationships between a set of accessions.

The rule to keep in mind is that the further we need to go in evolutionary times, the slower molecule must evolve. Going too far with too much diverging sequences will lead to homoplasy (characters identical by state, not by descent) through convergence or reversion. On the opposite, slow evolving sequences will not be enough in discriminating for groups that have evolved recently. Figure 1 illustrates this rule: if a very slow evolving sequence is used, it might be unable to differentiate the two hypothetical species under study (Fig. 1a). A sequence with an intermediate rate of evolution and concerted evolution would allow the identification of each species, but would be unable to reveal any intraspecific variability (Fig. 1b). To reach such a level of information, we would need to use a single-copy gene (Fig. 1c) or a microsatellite marker (Fig. 1d), but the latter, due to high evolutionary rate, may generate homoplasy (*) which could lead to erroneous interpretations if comparing species A and B, as individual B4 would appear more related to species A than to individuals from species B. Such rapidly evolving sequences are

therefore not appropriate for studying relationships at too high taxonomic levels (Besse, 2014).

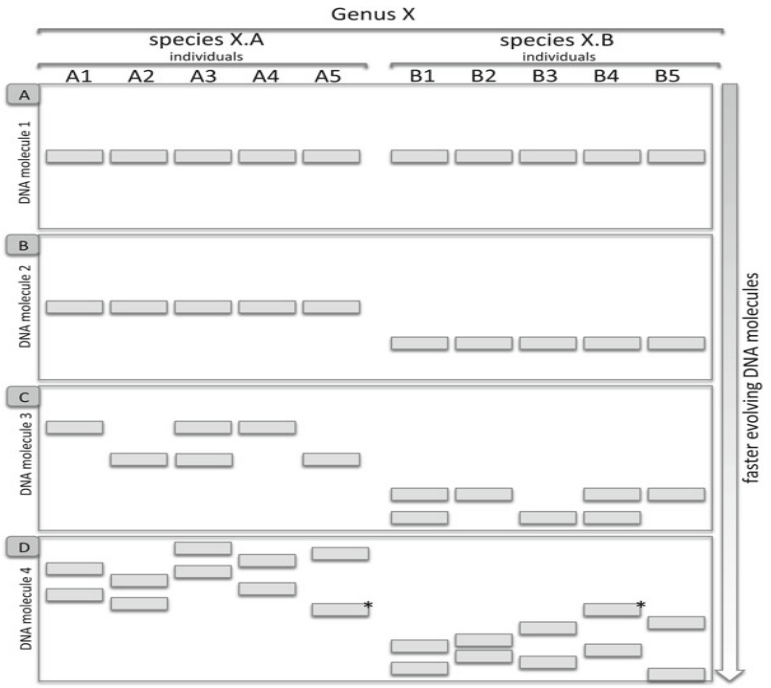


Fig. 1. General guidelines for the choice of markers to be used for plant taxonomy (Besse, 2014).

Guidelines for the choice of sequences to be used depending on the level of taxonomic divergence are illustrated in Figure 2. Since the level of taxonomic differentiation can vary considerably depending on the taxa; therefore one always needs to perform preliminary tests of various sequences on a representative subset of accessions to assess their power in differentiating our own individuals, species, or genera of interest.

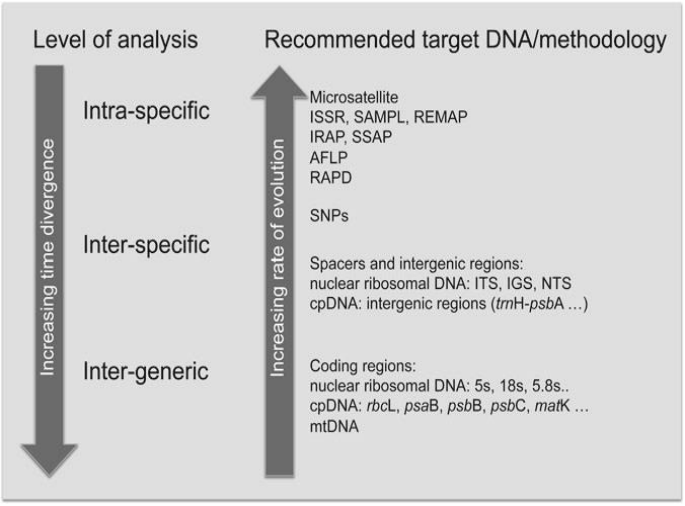


Fig. 2. General guidelines for the choice of markers to be used for plant taxonomy (Besse, 2014).

Knowledge of the mode of inheritance of the molecules under study is also important. Nuclear sequences are inherited in a Mendelian inherence, with the contribution from both parents. Organelle (chloroplast and mitochondrial) sequences are almost always uniparentally inherited. This can have important consequences when building a molecular phylogeny, as individuals or species of interspecific origin will appear inconsistently on the trees generated by/for each type of markers (Fig. 3): a species B of hybrid origin will be grouped with its mother species A using cytoplasmic sequences, although it will appear different from it on the nuclear tree. AFLP, RAPD, ISSR, and other multi-locus profiling methods generate dominant markers. The polymorphism revealed is mainly due to mutations in the hybridization region of one of the primers, leading to either amplification of the locus (presence) or null allele (absence of amplification), i.e., a dominant system. Consequently, such methods provide only biallelic markers. On the other hand, microsatellites are very powerful monolocus markers as they are multiallelic and codominants (Besse 2014).

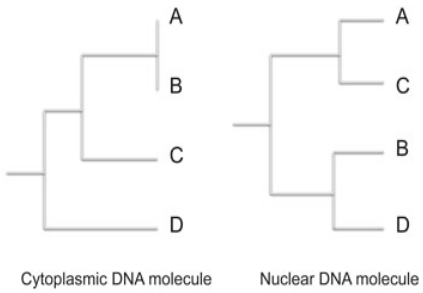


Fig. 3. A hypothetic phylogeny involving a hybrid species B whose maternal parent is species A.

Rouhan and Gaudeul (2012) informed that a solid taxonomic understanding of the study group is also desirable to guide the taxon sampling. Current sequencing techniques allow the inclusion of more taxa and sequences in a given study. Even large sequence data sets with thousands of nucleotides can be generated in short time and with reasonable effort. Nevertheless, the outcome of the phylogenetic study will depend on the markers used, not just on the pure amount of data generated (Besse 2014)..

Another important aspect in molecular taxonomy is the statistical analysis of data (Laurentin, 2009). Fragment length patterns (different band sizes visualized and coded after electrophoretic separation) will only be analyzed using distance-based methods (e.g., UPGMA or neighbor joining), whereas sequence data will be analyzed either using distance-based methods or more powerfully using character-based methods (e.g., maximum parsimony or maximum likelihood), allowing true phylogenetic trees to be constructed rather than phenetic trees (Laurentin, 2009).

Strengths of Molecular Taxonomy (reviewed in Rouhan and Gaudeul, 2012)

1. It is based on neutral markers that are independent of environmental conditions. Furthermore, it is not limited in number, and avoids bias due to the subjectivity of a given taxonomist.
2. Molecular tools sometimes allow the detection of additional, so-called cryptic, species that could not be distinguished on morphological analyses only.
3. It allows for better understanding the evolutionary process at work within taxonomically complex groups.
4. It can be performed on any life stage.

Limitations of Molecular Taxonomy (reviewed in Rouhan and Gaudeul, 2012)

1. It is difficult to isolate only DNA from the target material and exclude DNAs of any other animal, vegetal, or fungal organisms living around or in the plant under study.
2. The cost, as molecular lab facilities and often rather expensive consumables are needed.
3. The acquisition of a molecular data set also implies some more or less subjective choices, e.g., on the distinction of orthologs vs. paralogs, sequences´s lengths alignment or on the statistical analysis to carry out.
4. The possible lack of genetic divergence when sister species have very recent origins complicate the genetic analyses because they will share alleles, if reproductive isolation is not complete.
5. Molecular markers can also suffer from homoplasy, i.e., markers can show similar character states that, however, do not derive from a common ancestor.

CHARACTERIZATION METHODS AND TAXONOMY IN OPUNTIA

Morphological and morphometric markers

The morphologic characteristic on Cactaceae reflects the environmental conditions where these plants are growing. These characteristics are acquired through inheritance of old selections. However, some characters can be derived from recent evolutionary changes (Bravo, 1978). The morphological description often relies on the phenotypic characteristics of a group of plants "type" to differentiate them from other groups. This characterization method is widely applicable when carried out with care and dedication at all phases of evaluation; training and experience are required. This is why sometimes the morphological characterization led to underestimate or overestimate of some characteristics leading to erroneous conclusions (Luna-Paez, 2008). In the case of the genus *Opuntia*, cladodes are the traditional vegetative organ used to describe and characterize accessions; also, flower and fruit are used (Reyes-Agüero *et al.*, 2005); flout and fruits are also included (Valdéz-Cepeda *et al.,* 2003).

Morphological characterization has been used to a) know the performance of Opuntia for crop cultivation and their potential uses, b) to make clear support or enhance their taxonomy, c) to quantify the genetic variability, and d) to characterize the processes of intra-specific hybridization (Mondragón and Pérez, 2002). However, this description is time consumption due to the large number of characters to survey and the need to have controlled conditions to minimize the effect of environment. The most extensive published paper on Opuntia morphology is the reported by Reyes-Agüero *et al.* (2005). These authors analyzed 42 morphological attributes of cladodes, "nopalitos" and fruits of 243 Opuntia

variants. Nevertheless, the discriminative potential values of seed internal and external characteristics are received little attention.

Molecular markers

Biochemical molecular markers

Seed protein markers are polymorphisms present in proteins detected by biochemical techniques. The development of biochemical markers produced a revolution in genetic studies in plants, which so far had based exclusively on a limited number of morphological markers (Ladizinsky and Hymowitz, 1979). Isozymes and storage proteins are the most often used markers. **Isozymes** are defined as different molecular forms of an enzyme which have a common catalytic activity. DNA mutations that encoding these enzymes can result in changes in the amino acid composition, producing proteins with the same biological activity but different net charge and therefore with distinct migration in an electrophoretic field (Gepts, 1990). These markers have the disadvantages such as they are not being very polymorphic, limited in number, and as morphological characters, isozymes are affected by environmental conditions and plant tissue of origin. **Storage proteins** are a heterogeneous group of proteins present in the plant seed with function is to provide energy to the embryo in the early stages of growth. These proteins have been extensively studied in cereals and oilseed (Shewry *et al.,* 1995). No enzyme activity is required to detect polymorphisms in polyacrylamide gels. The seed storage proteins are classified based on their solubility in albumins (water-soluble), globulins (saline solution-soluble), prolamin (alcohol-soluble) and glutelins (acid or alkaline solution-soluble) (Osborne, 1924).

Several electrophoretic studies revealed a high degree of polymorphism in the number and the electrophoretic mobility of these proteins (Kumar *et al.,* 2012). Its composition is highly conserved within a taxonomic group of plants, so that the determination of the electrophoretic patterns of these proteins are useful as genetic markers to identify hybrids, study population structures, gene flow and ploidy (Ladizinsky and Hymowitz, 1979), as well as taxonomic markers in systematic studies and to separate plant taxa (Kumar et al., 2012).

Protein and enzyme markers have been successfully applied in the taxonomical classification of diverse plant families such as Poaceae (Duvall and Biesbor, 1989), cucurbitaceae (Pasha and Sen, 1991) and Fabaceae (Misset and Fontenelle, 1992), among others. Research reported by Chessa *et al.* (1997), using the analysis of five isoenzyme systems was led to group Opuntia genotypes in several groups; however, they could not clearly identify the cultivated varieties (*O. ficus-indica*) from wild (*O. amyclaea* and *O. littoralis*). Uzan (1997) studied enzyme systems to differentiate cultivars and ecotypes of *O. ficus-indica* in Turkey without comparing these with Silvestre species.

Seed protein patterns of nine species, representatives of the genera *Stetsonia, Cereus, Harrisia, Opuntia* and *Tephrocactus* were analyzed by electrophoresis. The results showed that the protein pattern for each species is stable, regardless of the place and date of sample collection (Carreras *et al.,* 1997). In the latter work, only four species of *Opuntia* from Argentina were included and any Mexican genotype was studied. The first work focused on the use of storage proteins of the stems to differentiate *Opuntia* species in Mexico was made by Estrada-Galván *et al.* (2000) where the use of total seed protein was discarded. Galvez *et al.* (2009) differentiated six *Opuntia* species from Argentina using total seed

protein. The cultivated species *O. ficus indica* was clustered with wild species (*O. salagria, O. quimilo* and *O. sulphurea*) at a similarity coefficient of 94 %. Three years later, Galvez *et al.* (2012), using the same marker, found an association between *O. ficus indica* and *O. quimilo* and *O. megapotamica* species at a similarity level of 82 %.

DNA molecular markers

Currently, molecular markers are applied to many fields of biology as evolution, ecology, biomedicine, forensics and diversity studies. An ideal DNA marker must be multiallelic, polymorphic, codominant and no epistatic (Labra *et al.*, 2003).

In recent years, molecular markers based on polymerase chain reaction (PCR) amplification of specific genomic sequences have been proposed as a direct and effective tool to estimate intergeneric and interspecific relationship between different taxa of Opuntia. Wang *et al.* (1998) showed that RAPD analyses can be useful in cultivar differentiation and identification of Opuntia duplicate accessions in collections. Labra *et al.* (2003), combining AFLP and cpSSR markers, found a high similarity between *O. ficus indica*, *O. megacantha*, *O. amyclaea*, *O. undulata*, *O. spinulifera* species. These authors suggested that *O. ficus indica* will be regarded the domesticated form of *O. megacantha*, based on the fact that individuals from different populations of the two species showed a similar genetic constitution. Griffith (2004) investigated the origin of *O. ficus-indica* by Bayesian phylogenetic of nrITS sequences. The seven specimens of *O. ficus-indica* were located in a well-defined group (100 % probability); together with a group of arborescent plants with fleshy fruits *(O. hypiacantha, O. leucotricha, O. megacantha, O. streptacantha* and *O. tomentosa*). These inferences led to assume the taxonomic concept of *O. ficus indica* may include multiple lineages derived clones selected for their reduction in the

number of spines and desirable characteristics of the fruit. This hypothesis was supported by the molecular analysis reported by Caruso *et al.* (2010) using SSR markers. These authors found that the accessions of *O. ficus-indica* were not separated from other species of *Opuntia* tree as *O. amyclaea, O. megacantha, O. streptacantha, O. fusicaulis* and *O. albicarpa*. Similarly, the RAPD technique was used to assess the genetic diversity of *O ficus-indica* genotypes in Tunisia (Zhoghlami *et al.,* 2007).

In Mexico, region of origin and diversification of many *Opuntia* species with agronomic and economic interests, few molecular based studies were carried out. Molecular characterization of some genotypes from central Mexico was reported by Mondragon-Jacobo *et al.* (2003). Luna-Paez *el al.* (2007) characterized 22 Opuntia cultivars with RAPD and ISSR markers. Recently, Valadez-Moctezuma *et al.* (2014 a; b) applied RAPD and ISSR markers to differentiate 52 cultivated accessions of Opuntia form Mexico revealing a great genetic diversity and pointed on the taxonomic misclassification. Thus, relationships and population structure of Opuntia germplasm are still little studied.

From the above, the molecular tools are definitely the most suitable instruments for assessing the level of genetic diversity in Opuntia germplasm. The molecular analysis should be a prerequisite step for planning breeding programs. The use of these markers is highly recommended to reclassify the cultivated accessions of cactus, which exhibit a great degree of variation, regardless of the current taxonomic classification.

Taxonomy is an open-ended science of classification and identification; it is constantly renovating itself. Taxonomy entails analysis and synthesis. When a new technology becomes available, taxonomy is quick to adopt it for its never-ending quest of producing

more refined and more accurate classification and identification. The more recent tools including genomics enable more precise authentication based on plurality of data taken from the variation inherent in the taxa (Baum, 2012).

STUDY OUTLINE

The work presented here represents an initiative to integrate three types of markers; morphological, biochemical and DNA to characterize an extensive Opuntia germplasm from Mexico.

Chapter 2 offerings a morphologic survey of *Opuntia* s.s. based on external and internal seed´s characteristics. A detailed study of the use of characters derived from analysis of seeds images, entire seeds and longitudinal sections, is presented with an attempt to differentiate accessions and to determine the most discriminating characters. Two derived works are accessible in annexes section of the thesis; the first describe the use of seeds features for its possible integration as characters stats for Opuntia varieties registration; and the second was addressed to separate the “xoconostles” from the “tunas” accessions based on the seed morphometric characteristics.

The discriminative power of molecular biochemical markers, viz. seed proteins (total and storage seed protein) is illustrated in Chapter 3. These markers are used to differentiate accessions of Opuntia. Similarly, preferential accumulation of the four fractions of storage proteins in the seeds is reported.

Chapter 4 is devoted to presenting the results of DNA molecular markers, viz. SSRs, PCR-RFLP and CAPS markers. In a separated three parts, the genetic diversity,

differentiation of genotypes and phylogenetic relationships in an extensive collection of accessions of Mexican nopal are determined.

Chapter 5 is reserved to integrate the results of the researchers presented here for a general discussion of the thesis.

REFERENCES

Anderson, E. F. 2001. The Cactus Familly. Oregon, USA 776 p.

Arakaki, M., Christin, P. A., Nyffeler, R., Lendel, A., Eggli, U., Ogburn, R. M., Spriggs, E., Moore, M. J., and Edwards, E. J. 2011. Contemporaneous and recent radiations of the world's major succulent plant lineages. PNAS 108: 8379-8384.

Baum, B. R. 2012. A Taxonomist's View on Genomic Authentication. In. Nikolaus, J. Sucher, *et al*. (eds.), Plant DNA Fingerprinting and Barcoding: Methods and Protocols, Methods in Molecular Biology pp: 1-12.

Besse, P. 2014. Guidelines for the Choice of Sequences for Molecular Plant Taxonomy. P. Besse (Ed). Molecular Plant Taxonomy. Methods and Protocols, pp 39-52.

Bravo, H. H. 1978. Las Cactáceas de México. 2ª ed. UNAM. México. 743 p.

Britton, N. L. and Rose, J. N. 1963. Cactaceae: Descriptions and Illustrations of Plants of the Cactus Family. Courier Dover Publications, New York. 268 p.

Carreras, M. E., Fuentes, E., and Merino, E. F. 1997. Seed protein patterns of nine species of cactaceae. Bioch. Syst. Ecol. 25:43-49.

Caruso, M., Currò, S., Las Casas, G., La Malfa, S., Gentile, A., 2010. Microsatellite markers help to assess genetic diversity among *O. ficus-indica* cultivated genotypes and their relation with related species. Plant Syst. Evol. 290: 85-97.

Casas, A., and Barbera, G. 2002. Mesoamerican domestication and diffusion. In: Nobel, P. S. (ed) Cacti: Biology and uses. University of California Press, USA, pp: 143-162.

Chessa, I., Nieddu, G., Serra, P., and Ingles, P. 1997. Isozyme characterization of *Opuntia* species and varieties from Italian germplasm. Acta Hort. 438: 45-52.

Duvall, M. R. and Biesboer, D. D. 1989. Comparisons of electrophoretic seed protein profiles among North American populations of Zizania. *Biochem.* Syst. Ecol. 17: 39-43.

Estrada, G., E. B., Scheinvar, L., Alagón, A., Olalde, P. G. 2000. Resultados preliminares de electroforesis en proteínas como método auxiliar en la taxonomía de *Opuntia* spp. Acta Científica Potosina 15: 7-24.

Gallegos-Vázquez, C., Scheinvar, L.**,** Núñez-Colín, C. A., Mondragón-Jacobo, C. 2012. Morphological diversity of xoconostles (*Opuntia* spp.) or acidic cactus pears: a Mexican contribution to functional foods. Fruits 67: 109-120.

Galvez, M. J., Castro, H. A., Villamil, C. B. 2012. Immunochemical properties of seed proteins as systematic markers in Cactaceae. Biochem. Syst. Ecol. 44: 20-26.

Galvez, M. J., Prat, M. I. Villamil, C. B. 2009. Antigenic patterns of seed proteins in Opuntioideae (Cactaceae). Biochem. Syst. Ecol. 37: 91-97.

Gepts, P. 1990. Genetic diversity of seed storage proteins in plants. In Brown, A.H.D., Clegg, M.T., Kahler, A.L., Weir, B.S. (Eds.) Plant population genetics, breeding, and genetic resources 64-82.

Griffith, M. P. 2004. The origins of an important cactus crop, *Opuntia ficus-indica* (Cactaceae): new molecular evidence. Am. J. Bot. 91: 1915-1921.

Griffith, M. P., and Porter, M. 2009. Phylogeny of Opuntioideae (Cactaceae). International Journal of Plant Science 170: 107-116.

Guzmán, U., Arias, S., and Dávila, P. 2003. Catálogo de Cactáceas Mexicanas. UNAM y CNCUB. D.F., México.315 p.

Hebert, P. D. N., and Gregory, T.R. 2005. The promise of DNA barcoding for taxonomy. Syst. Biol. 54: 852-859.

Helsen, P., Browne, R., Anderson, D., Verdyck, P., and Van Dongen, S. 2009. Galápagos' *Opuntia* (prickly pear) cacti: extensive morphological diversity, low genetic variability. Biol. J. Linn. Soc. 96: 451-461.

Helsen, P., Verdyck, P., Tye, A. and Van Dongen, S. 2009. Low levels of genetic differentiation between *Opuntia echios* varieties on Santa Cruz (Galapagos). Plant. Syst. Evol. 279:1-10.

Hernández-Hernández, T., Hernández, H., De Nova, J. A., Puente, R., Eguiarte, L. E., Magallón, S. 2011. Phylogenetic relationships and evolution of growth form in Cactaceae (Caryophyllales, Eudicotyledoneae). Am. J. Bot. 98: 44-61.

Hunt, D. R. 2002. That's *Opuntia*, that was! Succ Plant Res 6: 245-248.

Hunt, D. R. 2006. The New Cactus Lexicon. The Manse, Chapel Lane, Milbourne Port, UK. 526 p.

Korotkova, N. 2011. Phylogeny and evolution of the epiphytic Rhipsalideae (Cactaceae). PhD thesis, Erscheinungsjahr, 177 p.

Kumar, A., Sharma, N., Panwar, P., and Gupta, A. K. 2012. Use of SSR, RAPD markers and protein profiles based analysis to differentiate *Eleusine coracana* genotypes differing in their protein content. Mol. Biol. Rep. 39: 4949-4960.

Labra, M., Grassi, F., Bardini, M., Imazio, S., Guiggi, A., Citterio, S., Banfi, E., Sgorbati, S. 2003. Relationships in *Opuntia* Mill. Genus (Cactaceae) detected by molecular marker. Plant Sci. 165: 1129-1136.

Ladizinsky, G. and Hymowitz, T. 1979. Seed protein electrophoresis in taxonomic and evolutionary studies. Theor. Appl. Genet. 54:145-151.

Luna-Paez, A. 2008. Diferenciación genómica de semillas y cladodios de *Opuntia* ssp. y su posible uso para diferenciación. Tesis doctoral, UACH, 128 p.

Luna-Paez, A., Valadez-Moctezuma, E., Barrientos-Priego, A. F. and Gallegos-Vázquez, C. 2007. Characterization of *Opuntia* spp. by Means of Seed with RAPD and ISSR Markers and Its Possible Use for Differentiation. JPACD 9: 43-59.

Misset, M.T., and Fontenelle, C. 1992. Protein relationships between natural populations of *Ulex europaeus* and *U. galli* (Faboideae, Genisteae) and their hybrids. Plant Syst. Evol. 179: 19-25.

Mondragón, JC. Pérez, GS. 2002. Genetic resources and breeding cactus pear (*Opuntia* spp.) for fodder production. Acta Hortic. 581: 87-93.

Mondragón-Jacobo, C. 2003. Molecular characterization using RAPDs of a cactus (*Opuntia* spp. Cactaceae) collection from central México as a basis for plant breeding. Rev.Chap. Ser. Hortic. 9: 97-114.

Orozco-Segovia, A., Márquez-Guzmán, J., Sánchez-Coronado, M. E., Gamboa de Buen, A., Baskin, J. M., Baskin, C. C. 2007. Seed anatomy and water uptake in relation to

seed dormancy in *Opuntia tomentosa* (Cactaceae, Opuntioideae). Ann. Bot. 99: 581-592.

Pasha, M. K. and Sen, S. P. 1991. Seed protein pattern of Cucurbitaceae and their taxonomic implications. Biochem. Syst. Ecol. 19: 569-576.

Reyes-Agüero, J. A., Aguirre, R. J. R., Carlín, C. F., and González, D. A. 2013. Diversity of wild and cultivated *Opuntia* variants in the meridional highlands plateau of Mexico. Acta Hortic. 995: 69-74.

Reyes-Agüero, J. A., Aguirre-Rivera, J. R., Carlín, F. C., González, D. 2009. Catálogo de las principales variantes silvestres y cultivadas de *Opuntia* en la Altiplanicie Meridional de México. UASLP, SAGARPA y CONACYT. San Luis Potosí, México. 350 p.

Reyes-Agüero, J. A., Aguiere-Rivera, J. R. and Valiente-Banuet, R. A. 2006. Reproductive biology of Opuntia: A review. J. Arid Envir. 64: 549-585.

Reyes-Agüero, J.A., Aguirre, J. R. and Hernández, H. M. 2005. Systematic notes and a detailed description of *Opuntia ficus-indica* (L.) Mill. (Cactaceae). Agrociencia 39: 395-408.

Rouhan, G. and Gaudeul, M. 2014. Plant Taxonomy: A Historical Perspective, Current Challenges, and Perspectives. Besse P. (Ed). Molecular Plant Taxonomy. Methods and Protocols, pp 1-38.

Scheinvar, L., Filardo-Kerstupp, S., Olalde-Parra, G., and Zavaleta-Beckler, P. 2009. Diez especies mexicanas productoras de xoconostles: *Opuntia* spp. y *Cylindropuntia imbricada* (Cacteceae), UNAM, UAEH, UAM, Mexico, 280 p.

Scheinvar, L. 1995. Taxonomy of utilized opuntias. In: Barbera, G., Inglese, P,. Pimienta-Barrios, E. (eds) Agroecology, cultivation and uses of cactus pear. FAO plant production and protection paper 132. FAO, Rome, pp: 20-27.

Shewry, P. R., Napier, J. A., and Tatham, A. S. 1995. Seed storage proteins: structures and biosynthesis. Plant Cell 7: 945-956.

Stuppy, W. 2002. Seed characters and the generic classification on the Opuntioideae (Cactaceae). In Hunt, D. and Taylo,r N. (Eds.). Studies in the Opuntioideae. David Hunt, The manse and Chapel Lave, Sherborne, England. pp. 25-58.

Uzun, I. 1997. Fruit and cladode izozumes in cactus pear. Acta Hortic. 438:53-56.

Valadez-Moctezuma, E., Ortiz-Vásquez, Q., Samah, S., 2014b. Molecular based assessment of genetic diversity of xoconostle accessions (*Opuntia* spp.). Afr. J. Biotechnol. 13: 202-210.

Valadez-Moctezuma, E., Samah, S., Luna-Paez, A., 2014a. Genetic diversity *of Opuntia* spp. varieties assessed by classical marker tools (RAPD and ISSR). Plant Syst. Evol. DOI: 10.1007/s00606-014-1112-y.

Valdéz-Cepeda, R. D., and Blanco-Macías, F. 2002. Fractal geometry of daughter young cladodes of *Opuntia ficus-indica* (Cactaceae). Acta Hortic. 581: 169-175.

Wallace, R. S. and Dickie, S. L. 2002. Systematic implications of chloroplast DNA sequence variation in subfam. Opuntioideae (Cactaceceae). In Hunt, D. and Taylor, N. (Eds.). Studies in the Opuntioideae. David Hunt, The manse and Chapel Lave, Sherborne, England. pp: 25-58.

Wang, X., Felker, P. Burow, D. M. and Paterson, H. A. 1998. Comparison of RAPD marker patterns to morphological and physiological data in the classification of *Opuntia* accessions. J. PACD 3: 3-14.

Wiens, J. J. 2007. Species delimitation: new approaches for discovering diversity. Syst. Biol. 56:875-878.

Zoghlami, N., Chrita, I., Bouamama, B., Gargouri, M., Zemni, H., Ghorbel, A., and Mliki, A. 2007. Molecular based assessment of genetic diversity within Barbary fig (*Opuntia ficus indica* (L.) Mill. in Tunisia. Sci. Hortic. 113: 134-141.

Chapter 2

Morphological marker

MORPHOLOGICAL SEEDS DESCRIPTORS FOR CHARACTERIZE AND DIFFERENTIATE GENOTYPES OF *OPUNTIA* (CACTACEAE, OPUNTIOIDEAE)

Samir Samah and Ernestina Valadez-Moctezuma

Article published: Annual Research & Review in Biology, ISSN: 2347-565X ,Vol.: 4, Issue.: 24 (16-31 December) Page 3791-3809. DOI : 10.9734/ARRB/2014/11606

MORPHOLOGICAL SEEDS DESCRIPTORS FOR CHARACTERIZE AND DIFFERENTIATE GENOTYPES OF *OPUNTIA* (CACTACEAE, OPUNTIOIDEAE)

ABSTRACT

Aims: In this paper, a morphometric study was carried out to analyze the variation of seeds of Opuntia accessions using several statistical approaches. The main objective of this work was to determine if the morphological seeds variables are suitable for characterization and differentiation of Opuntia genotypes, and determine if they are useful to clarify the taxonomic status of this group of plants, which is confusing and contradictory.

Methodology: A total of 110 Opuntia accessions (Cactaceae, Opuntioideae), one accession of *Cylindropuntia* sp. (Cactaceae, Opuntioideae) and two other outgroups (Cactaceae, Pachycereae), some of Opuntia accessions are classified in delimited species but other ones have no specific assignation, were used. Nineteen internal and external seeds variables were obtained using image analysis. Basic statistical analysis, analysis of variance, principal component analysis, cluster and discriminant analysis were performed.

Conclusion: The most important findings were: the selected variables were of interest for the characterization and identification of the different Opuntia genotypes. Nevertheless, the grouping of the accessions did not consistent with the current taxonomy. The variables responsible for the separation between genotypes were seed area, major axis length, minor axis length, and Feret diameter and seed weight. These variables have a high discriminatory power and can be taken into account as potential parameters for genotypes assignation within the *Opuntia* genus.

Keywords: *Opuntia*; seed morphology; longitudinal section; embryo; statistics analysis.

INTRODUCTION

The *Opuntia* genus (*sensu stricto*; Cactaceae, Opuntioideae) refers to cacti with flat pseudostems or cladodes, cyathiform tubular perianths with shorter stamens than the tepals (Stuppy, 2002). This genus includes 191-215 species (Anderson, 2001; Hunt, 2002), originating in the north and south of the American continent; some of them were relatively new distributed worldwide. The difference in the number of species is mainly due to the nomenclature problems occurred not only in Opuntia but also within the other genera of Opuntioideae subfamily (Caruso *et al.*, 2010). In Mexico, about 83 species are recognized which renamed "nopal" (Guzmán *et al.*, 2003). Opuntia plants are closely associated with the Mexican culture development; since they were used for human food, such as vegetables and fruit, in semiarid regions of the southwestern areas of Tamaulipas and Tehuacan Valley from 9.000 to 11.000 years ago (Kiesling, 1999). The tender cladodes are also used to prepare juice, jelly, honey, jam and pasta, and the oil is extracted from its seeds. Opuntia plants are also used as fodder and for the restoration and vegetation in arid and semi-arid environments. The cultivated *Opuntia* species include: *O. megacantha*, *O. streptacantha*, *O. albicarpa*, *O. amyclaea*, *O. robusta*, *O. hyptiacantha*, *O. cochinillifera*, *O. joconostle*, and *O. matudae*, among others (Scheinvar, 1995; Caruso *et al.*, 2010).

Today, commercial varieties are generally octaploid but the ploidy level is varied from 2X to 8X (Felker *et al.*, 2006), although their ancestry is unknown. Moreover, many authors report the difficulty of the correct assignment of cultivated genotypes in a defined taxon (Felker *et al.*, 2006; Caruso *et al.*, 2010). The continuous morphological variation, the lack of clear descriptors for each specie, the high phenotypic plasticity and the ploidy variations have led to problems in species delimitation and genotypes assignation (Caruso *et al.*, 2010). As a result of incorrect assignments, the same varieties are often classified as

belonging to different species, and in other cases they are considered to be hybrids among unknown parentals.

The classification of Opuntia genotypes has been based only on morphological characteristics, especially fruits and cladodes variation; and the specie determination is based on taxonomic keys by comparing few wild individuals (Scheinvar *et al.,* 2009). However, the differences that may exist at the time of the identification can be inconsistent and resulting form environmental variation. To overcome this, alternatives approaches are suggested, one of them is based on quantitative approaches to grouping genotypes by similarities between traits measured in cladodes, fruits and flowers (Reyes-Agüero *et al.,* 2005a; Gallegos-Vázquez *et al.,* 2011). Valdez-Cepeda *et al.* (2003) reported that the presence/absence of spines and their lengths are usefull traits for morphological characteristics. However, Felker (2006) suggested that the absence of spines should not be considered as basis for taxonomic classification, because this character has simple inheritance. In this regard, several features of the spine such as length, thickness, inclination, color and layout, as well as their number by areola are partially dependent on the environment conditions, such as availability of nutrients and moisture (Rebman and Pinkava, 2002). For these reasons, spineless genotypes have been classified as *O. ficus-indica* and genotypes with spines as *O. megacantha*, *O. streptacantha* and *O. amyclaea* (Reyes-Agüero *et al.,* 2005b). Unlike at other times, genotypes with spines have been classified as *O. ficus-indica*. Kiesling (1999) considered to *O. amyclaea, O. megacantha* and *O. streptacantha* as synonyms of *O. ficus-indica,* and he divided this latter species into two botanical forms: a) *O. ficus-indica* f. *amyclaea,* with presence of spines; b) *O. ficus-indica* f. *ficus-indica,* spineless. Actually, the presence of spines in the cladodes is an inadequate feature to classify *Opuntia* species (Felker *et al.*, 2005). Caruso *et al.* (2010)

reported that the character of spinescence might have been developed multiple times during the evolution of the genus, and might have been selected from different populations. In other researchs, Opuntia varieties have been differentiated and described using molecular markers such as RAPD (Bendhifi *et al.,* 2013), ISSR (Luna-Paez *et al.,* 2007; Valadez-Moctezuma *et al.,* 2014) and SSR (Helsen *et al.,* 2009; Caruso *et al.,* 2010). Based on molecular data, morphological and biogeographic distribution, Labra *et al.* (2003) suggest that *O. ficus-indica* should be regarded as the domesticated form of *O. megacantha*. Furthermore, based on Bayesian phylogenetic analyzes of nrITS sequences, Griffith (2004) affirmed the hypothesis that *O. ficus-indica* is a close relative of an arborescent group with fleshy fruits of central and southern Mexico, and the taxonomic concept of *O. ficus-indica* may include clones derived from multiple lineages. However, using SSR, Caruso *et al.* (2010) can not separated *O. ficus-indica* from other arborescent species. Moreover, Helsen *et al.* (2009) attempted to distinguish two varieties of *O. echios* (echios and gigantea) using SSR markers, but the results again emphasized that the current taxonomic differentiation was not supported by molecular data.

Despite the rapid advances in molecular techniques and the interest for the characterization of plant genetic resources with these tools, the morphological characterization should always be considered as useful for the use in collections and description studies (Khoury *et al.,* 2010). Morphological characterization is necessary because it provides to users valuable information about individual accessions, the relationship between the characters, and the structure of the collections (Chessa, 2010). Meanwhile, statistical methods, including principal component analysis and cluster, can be used as effective tools to assess variability among genotypes. The lack of a general consensus on the taxonomy of *Opuntia* genus makes difficult the correct assignation of

gentoypes in the collections. Furthermore, the identification of highly discriminating descriptors is important to obtain an efficient and reproducible classification of the species and varieties and to adapt the list of descriptors for specific purposes.

In none of the characterization studies in Opuntia it has been taken into account the differences that may exist between Opuntia seed and its possible discriminatory potential. The potential taxonomic significance of seed morphology has been recognized in several groups of plants (Davitashvili *et al.*, 2010; de Queiroz *et al.*, 2013; Liu *et al.*, 2013), and the delimitation of the genera based on these characters was in agreement with the results of molecular studies. Meanwhile, the seed image analysis has gained great importance for the species identification of wild plants and as well as seeds of species and varieties of agronomic importance (Bacchetta *et al.*, 2011), proving, thus, be a useful tool for taxonomic studies.

Therefore, the objectives of the present research were to: (1) investigate the discriminatory potential of variables of Opuntia seeds, accurately measured using reliable and repeatable method such as image analysis, and (2) determine the potential use of these variables for classification and taxonomic position in this genus.

MATERIAL AND METHODS

Plant materials

Fruit samples of 110 Opuntia accessions were collected at two locations; CRUCEN-UACh, Zacatecas and "Nopalera" UACh, Texcoco germplasm banks (Table 1). Ten fruits from at least three individuals plants of each accession were harvested at commercial maturity, from which all mature seeds were removed manually, then dried in the open air,

cleaned off any remaining pulp and only viable seeds were stored in paper-bags at room temperature until use. One sample of *Cylindropuntia* sp. one other of pitahaya (*Hylocereus undatus)* and one pitaya (*Stenocereus thurberi*) are included as outgroups. Some of Opuntia accessions are classified in delimited species but other ones have no specific assignation (Table 1).

Seed measurements

A total of 19 characteristics of seeds and 113 accessions were used to build the data set and statistical analysis to characterize Opuntia accessions and to determine the potential use of these characteristics for taxonomy. For external morphology; 24 seeds/repetition (three repetitions) of each sample were randomly chosen to take pictures of them with a digital camera. For internal morphology, the technique developed by Guerrero-Muñoz *et al.* (2006) was applied. Five clean and viable seeds/ repetition (three repetitions) were adhered to the surface of a glass slide and oriented parallel to the median section. These seeds were polished symmetrically and parallel to median section (longitudinal section) with fine sandpaper until the mid-section and they were viewed and photographed individually under a Leica EZ4 stereoscope (Leica Microsystems, Switzerland) with an integrated camera.

All obtained images were processed using Photoshop CS5 12.0 program to define the area of seed, embryo, perisperm and funicular seedcoat (testa). The seed variables were then obtained by UTHSCSA ImageTool version 3.00 program (Dove, 2002; http://ddsdx.uthscsa.edu/dig/itdesc.html). The methodology described by Mebatsion *et al.*, (2012) was adopted to improve the contrast. To determine the weight of seeds, 100 fully

developed seeds (three replicates) were counted and weighed with analytic balance (220g. 0.1^{-1} mg) (ABS 220-4; Karn and GmbH).

Table 1 List of prickly pear accessions from Mexico evaluated to study seed morphometric diversity of *Opuntia* sp.

N	Accessions	*Opuntia* species	N	Accessions	*Opuntia* species
1	Alfajayucan	*O. albicarpa* Scheinvar	58	Liso Forrajero	*Opuntia* sp.
2	Alteña Blanco	*Opuntia* sp.	59	Mango	*O. albicarpa* Scheinvar
3	Alteña Rojo	*Opuntia* sp.	60	Mansa Amarilla	*Opuntia* sp.
4	Amarilla 2289	*Opuntia* sp.	61	Memelo	*O. affinis hyptiacantha*
5	Amarilla 3389	*Opuntia* sp.	62	Milpa Alta	*O. ficus-indica* (L.) Mill.
6	Amarilla China	*Opuntia* sp.	63	Montesa	*Opuntia* sp.
7	Amarilla Jalpa	*O. ficus-indica* (L.) Mill.	64	Morada	*O. megacantha* Salm-Dyck.
8	Amarilla Jarro	*O. megacantha* Salm-Dyck.	65	Morada T10	*O. megacantha* Salm-Dyck.
9	Amarilla Milpa Alta	*Opuntia* sp.	66	Naranjón Legítimo	*O. albicarpa* Scheinvar
10	Amarilla Miquihuana	*O. lasiacantha* Pfeiffer	67	Naranjona	*Opuntia* sp.
11	Amarilla Montesa	*O. megacantha* Salm-Dyck	68	*O. cochillinifera*	*O. cochillinifera*
12	Amarilla Oro	*O. albicarpa* Scheinvar	69	Oreja de Elefante	*O. undulata* Griffiths
13	Amarillo Plátano	*O. megacantha* Salm-Dyck.	70	Pabellón	*O. ficus-indica* (L.) Mill.
14	Amarilla San Elías	*Opuntia* sp.	71	Pachón	*Opuntia* sp.
15	Amarilla Zacatecas	*O. megacantha* Salm-Dyck.	72	Pelón Rojo	*O. ficus-indica* (L.) Mill.
16	Amarillo	*O. megacantha* Salm-Dyck.	73	Pico Chulo	*O. megacantha* Salm-Dyck.
17	Amarillo Aguado	*Opuntia* sp.	74	Pico de Oro	*Opuntia* sp.
18	Atlixco	*O. ficus-indica* (L.) Mill.	75	Pitahaya	*Hylocereus undatus*
19	Bam	*Opuntia* sp.	76	Pitaya	*Stenocereus thurberi*
20	Blanca de Castilla	*Opuntia* sp.	77	Plátano	*Opuntia* sp.
21	Blanca del cerro	*Opuntia* sp.	78	Princesa	*Opuntia* sp.
22	Blanca San José	*O. albicarpa* Scheinvar	79	Red Villa Puebla	*Opuntia* sp.
23	Blanco Atlacomulco	*Opuntia* sp.	80	Reyna	*O. albicarpa* Scheinvar
24	Blanco Huexotla	*Opuntia* sp.	81	Reyna Crucen	*Opuntia* sp.
25	Bola de Masa	*O. albicarpa* Scheinvar	82	Roja Azteca	*O. megacantha* Salm-Dyck.
26	Burrona	*O. albicarpa* Scheinvar	83	Roja San Martín	*O. megacantha* Salm-Dyck
27	Cacalote	*O. cochinera* Griffits	84	Rojo 3589	*Opuntia* sp.
28	Camuezo	*O. megacantha* Salm-Dyck.	85	Rojo Lirio	*O. megacantha* Salm-Dyck.
29	Cardón	*O. streptacantha* Lem.	86	Rojo Liso	*Opuntia* sp.
30	Cardón Blanco	*O. streptacantha* Lem.	87	Rojo UACh	*Opuntia* sp.
31	Cardona de Castilla	*O. streptacantha* Lem.	88	Rojo Vigor	*O. ficus-indica* (L.) Mill.
32	Cascarón	*O. chaveña*	89	Rosa de Castilla	*O. megacantha* Salm-Dyck.
33	Chapeada	*O. albicarpa* Scheinvar	90	Rubí Reyna	*O. megacantha* Salm-Dyck.
34	Charola Tardía	*O. streptacantha* Lem.	91	San Juan	*Opuntia* sp.
35	Chicle	*O. ficus-indica* (L.) Mill.	92	Sangre de Toro	*Opuntia* sp.
36	Col. Barr. Chica	*Opuntia* sp.	93	Sanjuanera	*O. lasiacantha* Pfeiffer
37	Col. Barr. Grande	*Opuntia* sp.	94	Solferino	*Opuntia* sp.
38	Color de Rosa	*O. albicarpa* Scheinvar	95	Tapón Aguanoso	*O. robusta* H.L. Wendland
39	Colorada	*Opuntia* sp.	96	Tapón rojo	*O. robusta* H.L. Wendland
40	Copena CEII	*O. ficus-indica* (L.) Mill.	97	Tapona de Mayo	*O. robusta* H.L. Wendland
41	Copena F1	*O. ficus-indica* (L.) Mill.	98	Tobarito	*Opuntia* sp.
42	Copena T12	*O. ficus-indica* (L.) Mill.	99	Toluca	*Opuntia* sp.
43	Copena T5	*O. ficus-indica* (L.) Mill.	100	Torreoja	*O. megacantha* Salm-Dyck.
44	Copena V1	*O. ficus-indica* (L.) Mill.	101	Trompa Cochino	*Opuntia* sp.
45	Copena Z1	*O. albicarpa* Scheinvar	102	Tuna Mansa	*O. albicarpa* Scheinvar
46	Cristalina	*O. albicarpa* Scheinvar	103	Tuna Rosa	*O. albicarpa* Scheinvar
47	Cylindropuntia	*Cylindropuntia* sp.	104	Tuna Sandia	*Opuntia* sp.
48	Fafayuca	*O. albicarpa* Scheinvar	105	Var S/I	*Opuntia* sp.
49	Gavia	*O. albicarpa* Scheinvar	106	Verdulero	*Opuntia* sp.
50	Green Guanajuato	*Opuntia* sp.	107	Villanueva	*O. albicarpa* Scheinvar
51	Huatusco	*Opuntia* sp.	108	X_Blanco	*O. joconostle* F.A.C. Weber
52	INIFAP	*Opuntia* sp.	109	X_Chivo	*Opuntia* sp.
53	Jade	*Opuntia* sp.	110	X_Colorado	*O. joconostle* F.A.C. Weber
54	Jarilla Grande	*Opuntia* sp.	111	X_Cuaresmero	*O. matudae* Scheinvar
55	Laltus	*Opuntia* sp.	112	X_Manzano	*O. joconostle* F.A.C. Weber
56	Larreguin	*O. ficus-indica* (L.) Mill.	113	X_Rojo	*Opuntia* sp.
57	Liso Amarillo	*Opuntia* sp.			

The variables obtained from entire seeds were: Area = the area of the object measured as the number of pixels in the polygon; Perimeter = the length of the outside boundary of the object; Major Axis Length = the length of the longest line that can be drawn through the object: Minor Axis Length = the length of the longest line that can be drawn though the object perpendicular to the major axis; Elongation = the ratio of the length of the major axis to the length of the minor axis (if the value is 1, the object is roughly circular or square, whereas it is more elongated when the ratio decreases from 1); Roundness = if the ratio is equal to 1, the object is a perfect circle, when the ratio decreases from 1, the object departs from a circular shape, calculated as $R = [(4\pi * \text{area})/\text{perimeter}^2]$; Feret Diameter = the diameter of a circle having the same area as the object, calculated with the formula: $FD=\sqrt{[(4 * \text{area})/\pi]}$; Compactness = provides a measure of the object's roundness: at 1 the object is roughly circular, when it decreases from 1, the object results less circular, calculated as C = FD/Major Axis Length. The variables obtained from the median section of the seeds (internal morphometric) were: Area and Perimeter of embryo, Area and perimeter of perisperm and funicular seedcoat. Ratios between variables were also calculated (Table 2).

Statistical analysis

A total of 19 quantitative variables were analyzed (Table 2). Both internal and external morphometric seed variables were analyzed together because both types of variables may respond in similar ways to environmental and genetic conditions; therefore the two types of data are similar. Descriptive statistics were performed for all variables, and the following parameters were obtained: mean, minimum, maximum and coefficient of variation. The analysis of variance (ANOVA) was applied to detect discriminant variables among genotypes, and multiple comparisons (Tuckey's test) were computed to identify the

difference between each pair of accessions (P = .05). A variable reduction technique was used to select the most discriminating variables among the 19 measured traits. Stepwise discriminant analysis was used to select traits that were included in the classification model. A significance level of 0.001 of an F test from an analysis of covariance was imposed to choose the most discriminating traits. Wilk's lambda (λ) was used as the criterion to determine the classification efficiency with the entry of each trait. The selected traits were then used in the subsequent analyses. To find out the relevant variables for morphological seed description, a correlation matrix was built using Pearson Correlation Coefficients to aid in interpretation of the analysis, and thereafter a Principal Component Analysis (PCA) was performed. PCA was used on the ranged data as a linear dimensionality reduction technique to identify orthogonal directions of maximum variance in the original data set and to project the data into lower dimensions of the highest variance components, and to examine the percentage contribution of each trait to variation. Then, the cluster analysis using the squared Euclidean distance and Ward's minimal variance method was performed. The relationships among the clusters were elucidated. To facilitate the identification of diagnostic variables, significant differences among means of groups were evaluated by variance analysis under the general linear model because there were unequal numbers of accessions per cluster. Differences between means of groups were compared using Tukey's post hoc test (P =.05). Finally, Stepwise Linear Discriminant Analysis (LDA) algorithm was performed to predict the membership of each accession to the corresponding group resulting from cluster analysis. This approach is commonly used to classify/identify unknown groups characterized by quantitative and qualitative variables. The best features for seed sample identification were detected implementing a stepwise LDA method and a statistical classifier to discriminate and classify the seeds on the basis of the selected characters. This method starts with a model that does not include any of the

variables. At each step, the variable with the largest F to enter value that exceeds the entry criteria chosen ($F \geq 3.84$) is added to the model. The variables left out of the analysis at the last step have F to enter values smaller than 3.84, so no more are added. The process was automatically stopped when no remaining variables increased the discrimination ability Bacchetta *et al.* (2011). A cross-validation procedure was applied to verify the performance of the classifiers. All calculations were done using SAS 9.0 software (SAS Institute, 2002) and/or SPSS 20.0 for Windows (SPSS Inc., USA).

RESULTS

Using the wear technique to display the median plane of seeds, together with the variables derived from the external morphology, we obtained 19 quantitative morphometrics data from internal and external features of seeds of 110 accessions of Opuntia and tree outgroups of the Cactaceae family.

ANOVA and descriptive analysis

Analysis of variance showed highly significant differences (P = .05) among Opuntia accessions for all characters studied, indicating the existence of a high degree of morphological diversity of seeds. In this regard, the weight of seeds ranged between 0.103 and 0.26 g, thereof the seed surface between 7.83 and 20.8 mm^2, the major axis length between 3.57 and 5.78 mm and the area of the embryo varied from 3.30 to 6.52 mm^2. Mean values and the amplitude of the other variables are summarized in Table 2. The coefficient of variation ranged from 0.98 (C) to 19.4 % (PA/SA). However, the most of the variables showed a coefficient of variation less than 10 % (Table 2). Tukey's post hoc test separated the accessions into different groups depending on the variable (data not shown). However, the variety Oreja de Elefante was separated from other accessions; since it had greater SA (20.8 mm^2), SP (18.1 mm), MjA (5.78 mm) and FD (0.52). The Larreguin

(*Opuntia ficus-indica*) accession was characterized by their high PA/SA (0.041). The variables SW, MjA, FD, SA, SP and MnA were the most different among the studied characteristics (Table 2).

Stepwise Discriminant Analysis

The discriminating power of 17 morphological seed variables was sufficient to differentiate the Opuntia accessions (Table 2). The significant results (P = .05) using fewer variables confirmed the usefulness of the STEPDISC procedure in selecting a critical subset of features. Considering the variables selected by this statistical method could reduce the cost and time for investigating Opuntia morphological relationships without compromising information gained. The variables PP and PA/EA did not contribute significantly to discrimination of accessions and were eliminated in the STEPDISC procedure. This method can detect redundant characters as reported by Yada *et al.* (2010).

According to the results of the linear correlations, a high positive correlation was obtained between the variables area and weight of seeds (SA *vs* SW), seed area and major axis length (SA *vs* MjA), seed area and minor axis length (SA *vs* MnA), seed weight and major axis length (SW *vs* MjA), seed weight and minor axis length (SW *vs* MnA); while the area of the embryo and perisperm were not associated with either the weight or area seeds (Figure 1). These results suggest that developmental increases in seed size (weight and area) correspond to increases in the width thereof, as well as in its length.

Table 2. ANOVA, descriptive analysis and variables selected by the STEPDISC procedure among 110 Opuntia accessions (Mean: mean value of the continuous variable, Max: maximum value, Min: minimum value, CV: coefficient of variation, F: critical value from F-test, g: gram, mm: millimeter, mm^2: square millimeter).

Variables	Abreviation	ANOVA and descriptive analysis					STEPDISC Procedure					
		Min	Max	Mean	CV (%)	F value	Step	partial R-square[1]	F Value	Pr > F	Wilks' Lambda[2]	Pr < Lambda
100 Seeds Weight (g)	SW	1.03	2.61	1.66	4.57	52.7***	3	0.932	27.1	<.0001	0.00002996	<.0001
Seed Area (mm^2)	SA	7.83	20.8	13.5	4.21	46.8***	2	0.981	99.5	<.0001	0.00044185	<.0001
Seed Perimeter (mm)	SP	11.2	18.1	14.6	2.42	36.2***	12	0.739	5.37	<.0001	0.00000000	<.0001
Major Axis Length (mm)	MjA	3.57	5.78	4.66	2.23	48.7***	4	0.898	17.2	<.0001	0.00000307	<.0001
Minor Axis Length (mm)	MnA	2.90	4.74	3.78	2.58	35.0***	14	0.641	3.35	<.0001	0.00000000	<.0001
Elongation	Elg	1.12	1.35	1.24	2.33	6.99***	15	0.570	2.47	<.0001	0.00000000	<.0001
Roundness	R	0.70	0.86	0.79	2.50	5.34***	11	0.721	4.91	<.0001	0.00000000	<.0001
Feret Diameter	FD	0.31	0.51	0.41	2.14	47.6***	1	0.977	86.1	<.0001	0.02261176	<.0001
Compactness	C	0.85	0.93	0.88	0.98	8.49***	5	0.980	95.7	<.0001	0.00000006	<.0001
Embryo Area (mm^2)	EA	3.30	6.52	5.18	7.56	7.80***	13	0.695	4.30	<.0001	0.00000000	<.0001
Embryo Perimeter (mm)	EP	8.80	13.9	11.2	5.53	6.80***	10	0.715	4.79	<.0001	0.00000000	<.0001
Perisperm Area (mm^2)	PA	0.08	0.41	0.22	16.1	11.5***	7	0.899	17.3	<.0001	0.00000000	<.0001
Perisperm Perimeter (mm)	PP	1.74	4.14	2.93	10.3	8.20***	Removed (no entered)					
Embryo Area/Seed Area	EA/SA	0.20	0.56	0.39	9.06	8.56***	9	0.782	6.90	<.0001	0.00000000	<.0001
Perispem Area/Seed Area	PA/SA	0.01	0.04	0.02	18.3	10.7***	6	0.863	12.2	<.0001	0.00000001	<.0001
Perisperm Area/Embryo Area	PA/EA	0.02	0.12	0.04	19.4	9.00***	Removed (no entered)					
Embryo Perimeter/Seed Perimeter	EP/SP	0.55	0.94	0.77	6.03	6.10***	8	0.803	7.86	<.0001	0.00000000	<.0001
Perisperm Perimeter/Seed Perimeter	PP/SP	0.12	0.30	0.20	10.7	6.26***	17	0.491	1.78	0.0002	0.00000000	<.0001
Perisperm Perimeter/Embryo Perimeter	PP/EP	0.15	0.37	0.26	11.8	5.20***	16	0.543	2.20	<.0001	0.00000000	<.0001

*** Indicates significant difference at 0.001 levels.
[1] The marginal variability accounted for by a variable when all others are already included in the model.
[2] The likelihood ratio measure of a trait's contribution to the discriminatory power of the model.

Principal component analysis (PCA) was used before cluster analysis to determine the relative importance of the 17 traits. PCA revealed that the first four components explained 90.97 % of the total variability (Figure 2). The first three components accounted for 83.35 % of the variability, of which the first component contributed twice the variability (48.12 %) respect to the second component (23.77 %). The variables that defined, according to their eigenvectors (value in parentheses), the first component in the positive direction were MnA (0.98), FD (0.98), SA (0.95), SP (0.95), MjA (0.94), SW (0.84), EA (0.77), EP (0.73), and in the negative direction EA/SA (0.64). The second component was related to the variables PA/SA (0.95), PA/EA (0.90), PP/SP (0.89), PA (0.83) and PP/EP (0.81) in the positive sense. The third component was determined by the variables C (0.69) in the positive direction and by Elg (0.64) in the negative one. These results revealed that the first component was defined by the variables measured directly on the seeds (weight, length,

area and perimeter), while the remaining components were defined by the rations between different variables.

The projection of all 113 studied accessions on the first two components (CP 1 and CP 2) showed high dispersion around the origin of the plot (Figure 3). However, *Larreguin* (55) and *Cylindropuntia* sp. (67) accessions were separated from the remaining ones on the positive sense of the second component; since they have greater perisperm area (0.405 and 0.377 mm^2, respectively). In turn, the pitahaya (75) and pitaya (76) accessions were separated on the negative sense of the first component to having small seeds. It is noteworthy that the genotypes corresponding to xoconostles; acidic prickly pear, (108 to 113) were placed together, since their seed dimensions were lower than the most of the other opuntias. However some prickly pear genotypes such as 20, 38, 32, 71 and 98 were placed together with xoconostles, indicating the need to integrate other data such as fruit characteres and/or molecular markers to separate these two Opuntia groups.

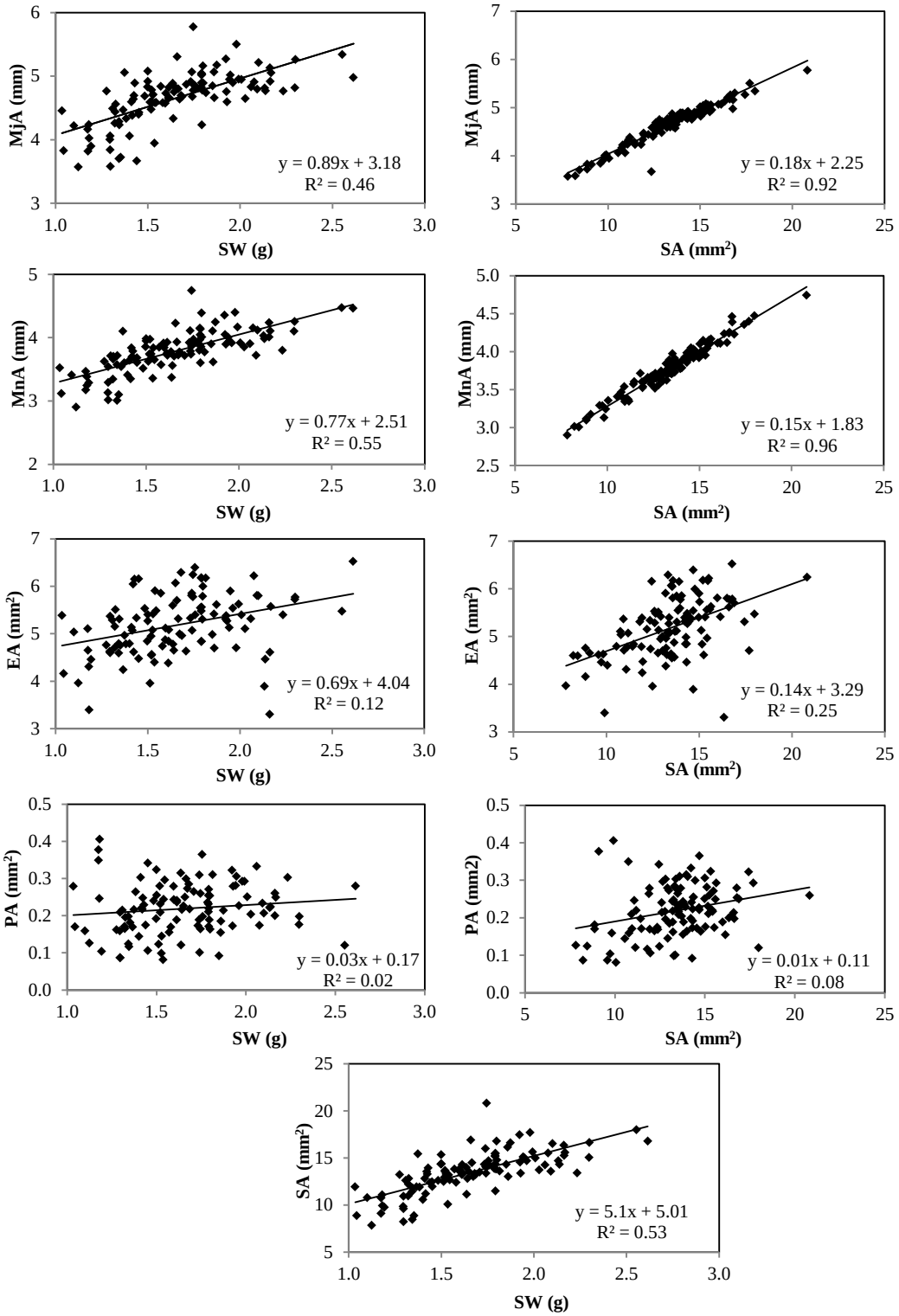


Fig. 1 Linear correlation between seed morphometric variables from Opuntia accessions. SA: seed area, SW: seed weight, MjA: major axis length, MnA: minor axis length, EA: embryo area, PA: perisperm area, g: gram, mm: millimeter, mm^2: square millimeter.

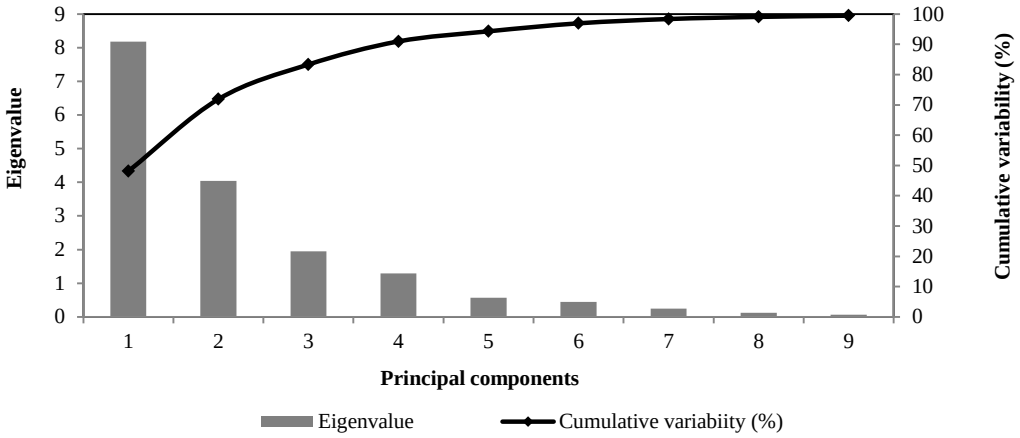


Fig. 2 Representative plot of the cumulative variability and eigenvalues of the first ten PCA components resulting from 17 seed morphometric variables mesured on 110 Opuntia accessions, one sample of *Cylindropuntia* sp. and two samples of pitahya and pitaya.

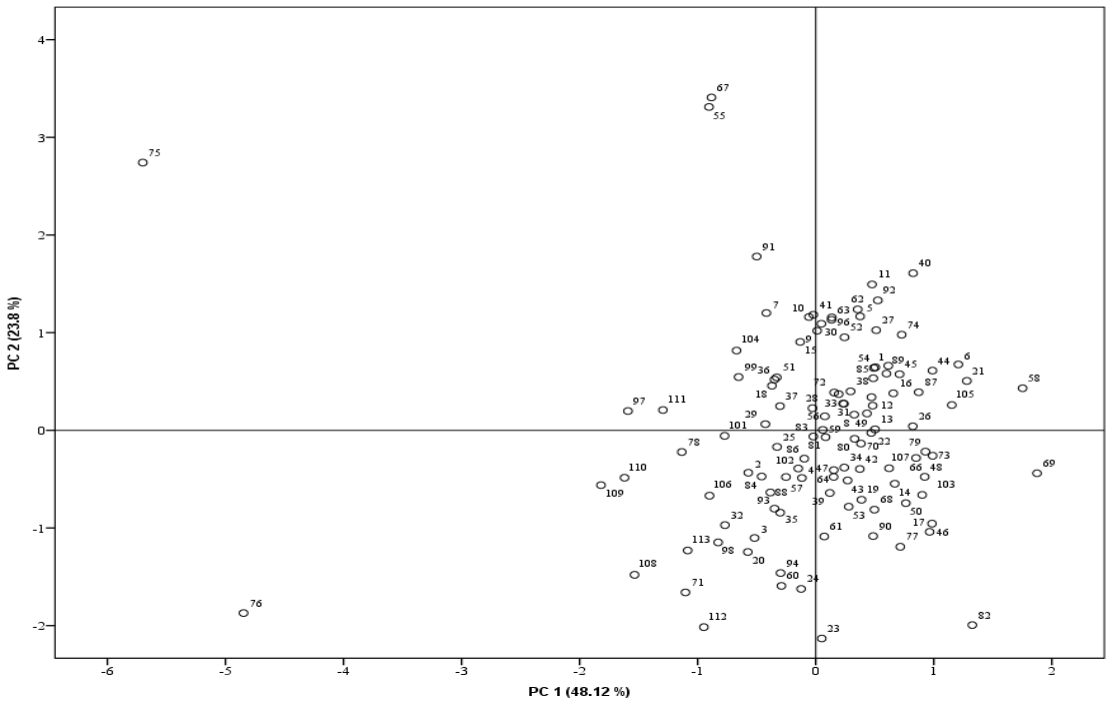


Fig. 3 Plot distribution of the 110 Opuntia accessions, one sample of *Cylindropuntia* sp. and two samples of pitahya and pitaya basad on 17 external and internal seed quantitatives variables.

Cluster analysis separated the 113 accessions studied in seven main groups, of which the group 7 included the two outgroups pitahaya and pitaya. Variance analysis was used to select diagnostic variables between groups, previously defined by cluster analysis. Tukey's test was applied to determine the variables that discriminate between these groups (Table 3). With the exception of the variables PA/SA, PA/EA, and PP/SP, all 14 remaining ones separated the Opuntioideae accesions (groups 1 to 6) from the Pachycereae ones (group 7; *Hylocereus undatus* and *Stenocereus thurberi;* Pitahaya and Pitaya, respectively). Among Opuntioideae accesions, the 6 obtained groups contained different number of accessions (14, 14, 31, 10, 31, 11 in groups 1 to 6, respectively; Table 3). Most of the variables (SW, SA, SP, MjA, MnA, FD, EA, EP, EA/SA) contributed to the separation between the 6 Opuntioideae groups resulting from cluster analysis. Groups 4 (10 genotypes) and 6 (11 genotypes) were characterized by extreme values (highest and lowest, respectively) for the variables SA, SP, MJA, MnA and FD (Table 3). Group 1 (14 genotypes) was characterized by genotypes with high EA and EP. Groups 2 (14 genotypes), 3 (31 genotypes) and 5 (31 genotypes) were characterized by genotypes with intermediate values, in order from lowest to highest, of the variables SA, SP, MjA and MnA.

Table 3. Quantitative variables used to investigate the morphological variation between the seeds groups resulting from the cluster analysis.

Groups (accessions number)	Cluster accessions	SW	SA	SP	MjA	MnA	Elg	R	FD	C	EA	EP	PA	EA/SA	PP/EP
Grp. 1 (14)	1, 5, 11, 27, 31, 38, 52-54, 59, 61, 64, 81, 95	1.67bc	13.8c	14.8bc	4.77bc	3.82c	1.26b	0.79a	0.42c	0.88a	5.79a	12.4a	0.26a	0.42bc	0.26a
Grp. 2 (14)	2, 3, 20, 29, 32, 36, 37, 57, 91, 98, 99, 101, 104, 106	1.38cb	11.3d	13.4d	4.29d	3.49d	1.24b	0.79a	0.38d	0.89a	4.73cd	10.5cd	0.19a	0.42bc	0.25a
Grp. 3 (31)	4, 7, 9, 10, 15, 18, 24, 25, 28, 30, 33-35, 39, 41, 51, 56, 60, 62, 63, 70, 72, 83, 84, 86, 88, 93, 94, 96, 100, 102	1.57bcd	13.0c	14.5c	4.60c	3.71c	1.26b	0.78a	0.41c	0.88a	5.09bc	11.0bcd	0.22a	0.39cd	0.27a
Grp. 4 (10)	6, 21, 46, 48, 58, 69, 73, 79, 82, 105	2.05a	17.4ª	16.7a	5.30a	4.37a	1.22b	0.78a	0.47a	0.89a	5.69ab	12.0ab	0.24a	0.33d	0.27a
Grp. 5 (31)	8, 12, 13, 14, 16, 17, 19, 22, 23, 26, 40, 42, 43, 44, 45, 47, 49, 50, 65, 66, 68, 74, 77, 80, 85, 87, 89, 90, 92, 103, 107	1.87ab	15.0b	15.4b	4.92b	4.01b	1.23b	0.79a	0.44b	0.89a	5.32abc	11.3bc	0.23a	0.36cd	0.27a
Grp. 6 (11)	55, 67, 71, 78, 97, 108-113	1.26d	9.14e	12.2e	3.81e	3.15e	1.22b	0.77a	0.34e	0.89a	4.38d	10.1d	0.17a	0.48ab	0.24a
Grp. 7 (2)	75, 76	0.14e	2.42f	6.69f	2.20f	1.44f	1.55a	0.68b	0.17f	0.79b	1.28e	6.74e	0.05b	0.54a	0.17b

Label in the cluster accessions case refers to the corresponding accession mentioned in the table 1.
Different letters indicate significant differences between groups resulted of Tukey's post hoc comparison, $P = .05$

Data analyzed by Stepwise Linear Discriminant and statistical classifiers were developed in order to distinguish the obtained groups from de the cluster analysis. When several variables are available, the stepwise method can be useful by automatically selecting the best characters on the basis of three statistical variables: Tolerance, F-to-enter and F-to-remove. The Tolerance value indicates the proportion of a variable variance not accounted for by other independent variables in the equation. A variable with very low Tolerance value proves little information to a model. F-to-enter and F-to-remove values define the power of each variable in the model and they are useful to describe what happens if a variable is inserted and removed, respectively, from the current model. The best discriminaning variables selected by the stepwise method among the 17 variables are shown in Table 4. The first two variables selected by the model were the same in cluster analysis. These were Feret diameter (FD) and seed area (SA) that moreover showed values of F to remove clearly higher than other selected features. Using this model, 96.5 % of original grouped cases correctly classified and 92.9 % of the cross-validated samples of the seven clusters were correctly classified (Table 5). Accessions of the group 7 were correctly identified in 100 % of the cases and none of the seeds of other studied accession was mistaken for it. Contrastingly, group 1 showed a lower percentage of correct identification (78.6 %), as accessions were mainly misclassified among those of group 3. The other groups had higher percentages of correctly identification upper of 90 % and only one genotype of each group was wrongly placed (Table 5).

Table 4. Ranking of eight selected variables after stepwise Linear Discriminant Analysis (see Table 2 for the legend of the variables).

Variables	Tolerance	F to Remove	Wilks' Lambda
FD	0.0049	234.9	0.0040
SA	0.0107	176.9	0.0031
EP	0.5214	3.410	0.0003
EA/SA	0.0347	40.42	0.0009
EA	0.0565	26.03	0.0007
Elg	0.0126	11.61	0.0004
MnA	0.0191	12.14	0.0005
C	0.0350	5.900	0.0004

Table 5. Predicted groups membership and cross-validated of correct classification of the Opuntia accessions resulting from the cluster analysis. The number of accessions is indicated in brackets.

Classification Results	Groups	Predicted Group Membership							Total
		1	2	3	4	5	6	7	
Original	1	85.7 % (12)		14.3 % (2)					14
	2		92.9 % (13)	7.1 % (1)					14
	3			100 % (31)					31
	4				90 % (9)	10 % (1)			10
	5					100 % (31)			31
	6						100 % (11)		11
	7							100 % (2)	2
Cross-validated	1	78.6 % (11)		21.4 % (3)					14
	2		92.9 % (13)	7.1 % (1)					14
	3	3.2 % (1)		96.8 % (30)					31
	4				90 % (9)	10 % (1)			10
	5	3.2 % (1)				96.8 % (30)			31
	6		9.1 % (1)				90.9 % (10)		11
	7							100 % (2)	2

DISCUSSION

One of the distinguish characters subfamily Opuntioideae from other subfamilies in Cactaceae is the seed structure. The Opuntioideae seeds are unique, not just in the Cactaceae or even the Caryophyllales but in the whole of the Angiospermae, in being entirely encased by a hard aril derived from the funiculus (Stuppy, 2002). Surprisingly, in view of their uniqueness, the Opuntioideae seed have received little attention. The most remarkable character of Opuntioideae seeds is that they are completely covered by a tissue

derived from the funiculus. Seeds have a thick white funiculus surrounding them, well-developed perisperms and curved embryos. The curvature of the embryo is the result of the campylotropous curvature of the ovule (Stuppy, 2002). Seeds of *Opuntia* species have hard (to-the-touch) seed covers (Orozco-Segovia *et al.*, 2007), and pressures of 440 daN may be required to break them.

In this study, seeds of Opuntia accessions are studied to abtain quantitative variables related to external and internal morphology. These variables were obtained by image analysis and investigated using several statistical analyzes. Andrés-Agustín *et al.* (2006) reported the importance of uni and multivariate analyzes to abtain new association between accessions and species, and supported the importance of these tests to evaluate the taxonomic entities.

The obtained results showed that the seeds of Opuntia have a high range of variation in size (major and minor length) in weight and also in the area. All studies variables were able to discriminate accessions, since the analysis of variance showed highly significant variation. The most discriminating variables were seed weight, major axis length, Feret Diameter, area and perimeter of the seed, and minor axis length. In addition, low coefficient of variation values suggests discriminatory stability of these variables, as well as reported by Guerrero-Muñoz *et al.* (2006). According to Sassone *et al.* (2013), coefficients of variation of 12 %, or less, are acceptable in characterizing plant organs in horticultural species and would be desirable to increase the sample size if this ratio is higher. In our case, out of the 19 studied variables, only three showed high coefficients of variation (PA/SA (19.4 %), PA/SA (18.3 %), PA (16.1 %)); while the remaining ones had

values lower than 10 %. This indicates that the number of used seeds here was appropriate to obtain stable and useful variables for characterization and differentiation purposes.

Estimation of the measured parts of the seed (embryo, testa, perisperm and total area) revealed that the embryo and perisperm area represent 38.4 % and 1.63 % of the total seed area, respectively. Similar values were reported by Stuppy (2002) and by Guerrero-Muñoz *et al.* (2006). A large embryo (whose function is to storage the reserves) produces a seedling with higher photosynthetic productivity and being able to grow faster and compete more successfully (Linkies *et al.*, 2010). Stuppy (2002) reported that the Opuntia seed has small sized, oval, and the embryo has a spiral shape around a folded perisperm strongly reduced, since embryo length increases the storage capacity is increased too.

Out of 19 variables, 17 had a high discriminative power as stepwise discriminant analysis showed, with the exception of two variables (perisperm perimeter and perisperm area/embryo area ratio). Yada *et al.* (2010) reported the usefulness of this statistical technique to reduce the number of characters to be measured; which implies savings in time, effort and expense, without compromising tesults gain; besides detecting redundancy in the variables.

The PCA, based on 17 variables, was performed to study the combination of traits that best explain the variability. The usual procedure to identify the components is to detect the first components that explain the largest proportion of the total variance (Wu *et al.*, 2003). In our case, the components considered with eigenvalues above than 1 (8.18, 4.04, 1.95 and 1.30 for the components 1, 2, 3 and 4, respectively (Figure 1). The CPA results showed the usefulness of the variables minor axis length (MnA), Feret diameter (FD), seed

area (SA), seed perimeter (SP), major axis length (MjA), seed weight (SW) for their ability to differentiate between accessions. The projection of all studied accessions on the first two components (CP 1 and CP 2) showed high dispersion around the origin of the plot, indicating a continuity of variables among accessions without clear boundaries between them. This is due to all the variables used are quantitative.

Cluster analysis separated the 113 accessions studied in seven main groups. Group 7 was composed of two genotypes of Pachycereae included as outgroups. Group 6 was composed mainly of genotypes belonging to xoconostles. Most variables (SW, SA, SP, MjA, MnA, FD, EA, EP, EA/SA) contributed to the separation of the 6 Opuntia groups, resulting from the cluster analysis. However, SA, SP, MjA, MnA, AD and PD and FD variables had the greater power to define this grouping.

The pattern grouping of genotypes did not fit the actual species assignment, nor in PCA neither in cluster analysis. Similar results were found by Reyes-Agüero *et al.* (2005b) and Gallegos-Vázquez *et al.* (2012) using morphological markers as variables derived from cladodes and fruits. This is probably related to the high level of phenotypic plasticity and polyploidy, and also due to the morphological diversity of these accessions. These genotypes had several end use; as fruits, vegetables and/or as forage. For these reasons, many studies have suggested the revision of the classification of the *Opuntia* genus (Labra *et al.,* 2003; Caruso *et al.,* 2010; Valadez-Moctezuma *et al.,* 2014). Moreover, the geographical accessions origin affects their morphological variation, and this has led to very narrow use of the concept of species. Often the location of an accession in a species is arbitrary and lack of solid descriptors; many of the accessions considered in our study have not yet been taxonomically asigned (Table 1). However, accessions representatives of

xoconostles were grouped together (in both analyses), thus showing its distinction from other accessions because them having smaller size for seeds. Studies based on fruit morphology (Gallegos-Vázquez *et al.,* 2011) and molecular markers (Luna-Paez *et al.,* 2007), placed to the xoconostles as sister groups of prickly pears. According to Gallegos-Vázquez *et al.* (2012), the absence of the pulp and the presence of an edible pericarp are the most significant differences between prickly pears and xoconostles. However, the presence of some prickly pears genotypes grouped together with xoconostles suggests the need to use other plant organs and/or molecular markers to differentiate these two Opuntia plants.

The classification test of genotypes to clusters by linear discriminant analysis showed a cross-validation of 92.9 %. Similar results were found by Bacchetta *et al.* (2011), where cross-validation of 92.7 % was fined in samples from five taxa of *Lavatera*. This statistical technique approved the discriminating power of the image analysis derived variables fom Opuntia seed obtained.

For the *Opuntia* genus, the use of plant height, cladodes, fruit and the flower is the traditional way for classifying the genotypes and assign them in their respective species (Scheinvar *et al.,* 2009). These descriptors are considerably affected by the environment geographical conditions and show a low discriminating power. Similarly and although flower attributes are considered stable, Fuentes-Pérez *et al.* (2009) reported that the floral anatomical characteristics of five species of the *Opuntia* genus was not decisive in the taxonomic separation between species. In the present study, we demostrated that many of seeds variables analyzed with images are of potential candidates for use in this complex taxonomic genus. These results can be transferred to state characters useful for cladistic

analysis and can be used as guide selection of taxonomic characters. Seed Opuntia variables are little influenced by environmental pressure and are more affected by the genetic control, which is likely due (i) to the hardness of the seed; (ii) the protective effect offered by the pulp and seeds testa and (iii) the short period of exposure the fruits to environmental factors.

Morphometric characterization of seeds is rapid, reproducible and reliable that accurately identifies the seeds of species from wild plants. Their usefulness in taxonomic studies is promising, due to its efficiency in discriminating between accessions at the level of interpopulation (Bacchetta *et al.,* 2011). This provides new insights into plant taxonomy, and also offers the opportunity to the germplasm banks to identify their accessions through standardized and quickly methods. Our results demonstrated that the image analysis allows estimating the principal dimensions of the seeds (length, width and elongation) with high accuracy. Since the manual measurements are difficult due to the small size of these seeds. Another advantage of this type of analysis is to provide additional features, to be determined objectively and with good discriminating power, such as FD and Elg. Moreover, they are continuous variables, which allow the use of ANOVA statistics (Lootens *et al.,* 2013).

Despite the lower costs associated with the analysis of morphological variables of seeds, molecular analysis remains an essential tool for the investigation of the variability within and between genotypes, and for estimating genetic relationships and assigning genotypes to a defined species.

CONCLUSION

The results presented here proved the utility of the seed variables as grouping characteristics such as weight, size and dimensions of seeds, as well as the several relationships between variables can be generated. We demonstrated the potential discriminatory of seed variables derived from image analysis, in order to their consideration in characterization studies and to assign new identified genotypes in their respective taxa. Not to mention the need for engagement with other morphological and molecular analyzes.

ACKNOWLEDGEMENTS

The first author received a scholarship for doctoral studies from the Ministry of Foreign Affairs, Mexico. We are grateful to Mr. Cruz-Miranda FM (UACh-bank Germplasm "Facundo Barrientos Perez") and to Dr. Gallegos-Vázquez C (CRUCEN-UACh Bank Germplasm, Zacatecas) for providing part of the plant material, and to Dr. J L Rodríguez de La O for the stereomicroscope facilities.

REFERENCES

Anderson, E. F. 2001. The cactus family. Timber. Portland, Oregon, USA.

Andrés-Agustín, J., González-Andrés, F., Nieto-Ángel, R., and Barrientos-Priego, A. F. 2006. Morphometry of the organs of cherimoya (*Annona cherimola* Mill.) and analysis of fruit parameters for the characterization of cultivars, and Mexican germplasm selections. Sci. Hortic.107: 337-346.

Bacchetta, G., García, P. E., Grillo, O, Mascia, F., and Venora G. 2011. Seed image analysis provides evidence of taxonomical differentiation within the *Lavatera triloba* aggregate (Malvaceae). Flora 206: 468-472.

Bendhifi, M., Baraket, G., Zourgui, L., Souid, S., and Salhi-Hannachi, A. 2013. Assessment of genetic diversity of Tunisian Barbary fig (*Opuntia ficus indica*) cultivars by RAPD markers and morphological traits. Sci. Hortic. 158: 1-7.

Caruso, M., Currò, S., LasCasas, G., La Malfa, S., and Gentile, A. (2010) Microsatellite markers help to assess genetic diversity among *O. ficus indica* cultivated genotypes and their relation with related species. Plant Syst. Evol. 290: 85-97.

Chessa I. (2010) Cactus pear genetic resources conservation, evaluation and uses. CACTUSNET Newsletter, Special Issue. 2: 43-53.

Davitashvili, N., and Karrer, G. (2010) Taxonomic importance of seed morphology in Gentiana (Gentianaceae). Bot. J. Linn Soc. 162: 101-115.

de Queiroz, R. T., de Azevedo-Tozzi, A. M. G., and Lewis, G. P. 2013. Seed morphology: an addition to the taxonomy of *Tephrosia* (Leguminosae, apilionoideae, Millettieae) from South America. Plant Syst. Evol. 299: 459-470.

Felker, P., Paterson, A., and Jenderek, M. M. 2006. Forage potential of *Opuntia* clones maintained by the USDA National Plant Germplasm System (NPGS) collection. Crop Sci. 46: 2161-2168.

Felker, P., Rodriguez, S.C., Casoliba, R.M., Filippini, R., Medina, D., and Zapata, R. 2005. Comparison of *Opuntia ficus indica* varieties of Mexican and Argentine origin for fruit yield and quality in Argentina. J. Arid Environ. 60: 405-422.

Fuentes-Pérez, M., Terrazas, T., and Arias, S. 2009. Anatomía floral de cinco especies de *opuntia* (opuntioideae, cactaceae) de México. Polibotánica 27: 89-102.

Gallegos-Vázquez, C., Barrientos, P. A. F., Reyes-Agüero, J. A., and Núñez, C. C. A., 2011. Mondragón JC. Clusters of commercial cultivars of cactus pear and xoconostle using UPOV traits. J PACD 13: 10-23.

Gallegos-Vázquez, C., Scheinvar, L., Núñez, C. C. A., and Mondragón-Jacobo, C. 2012. Morphological diversity of xoconostles (*Opuntia* spp.) or acidic cactus pears: a Mexican contribution to functional foods. Fruits 67: 109-120.

Griffith, M. P. 2004. The origins of an important cactus crop, *Opuntia ficus-indica* (Cactaceae): new molecular evidence. Am. J. Bot. 91: 1915-1921.

Guerrero-Muñoz, P., Zavaleta-Mancera, H. A., Barrientos-Priego, A.F., Gallegos-Vázquez, C., Núñez, C. C. A., Valadez-Moctezuma, E., and Cuevas-Sánchez, J. A. 2006. Técnica para el estudio de la micromorfología interna de semillas duras en *Opuntia*. Rev. Fitotec. Mex. 29: 37-43.

Guzmán, U., Arias, S., and Dávila, P. 2003. Catálogo de cactáceas mexicanas. UNAM, CNCUB, México.

Helsen, P., Verdyck, P., Tye, A., and Van Dongen, S. 2009. Low levels of genetic differentiation between *Opuntia echios* varieties on Santa Cruz (Galapagos). Plant Syst Evol. 279: 1-10.

Hunt, D. 2002. Alphabetical list of currently accepted species. In: Hunt D and Taylor N, editors. Studies in the Opuntioideae. Sherborne, England.

Khoury, C., Laliberté, B., and Guarino, L. 2010. Trends in ex situ conservation of plant genetic resources: a review of global crop and regional conservation strategies. Gen Res. Crop Evol. 57: 625-639.

Kiesling, R. 1999. Domesticación y distribución de *Opuntia ficus-indica*. J PACD 3: 50-59.

Labra, M., Grassi, F., Bardini, M., Imazio, S., Guiggi, A., Citterio, S., Banfi, E., and Sgorbati, S. 2003. Relationships in *Opuntia* Mill. Genus (Cactaceae) detected by molecular marker. Plant Sci. 165: 1129-1136.

Linkies, A., Graeber, K., Knight, C., and Leubner-Metzger, G. 2010. Review. The evolution of seeds. New Phytologist. 86: 817-831.

Liu, L.L., Yu, W-B., Li, D. Z., Mill, R. R., and Wang, H. 2013. Seed morphological diversity of *Pedicularis* (Orobanchaceae) and its taxonomic significance. Plant Syst. Evol. 299:1645-1657.

Lootens, P., Chaves, B., Baert, J., Pannecoucque, J., Van Waes, J., and Roldan-Ruiz, I. 2013. Comparison of image analysis and direct measurement of UPOV taxonomic characteristics for variety discrimination as determined over five growing seasons, using industrial chicory as a model crop. Euphytica 189: 329-341.

Luna-Paez, A., Valadez-Moctezuma, E., Barrientos-Priego, A. F., and Gallegos-Vázquez, C. 2007. Characterization of *Opuntia* spp. by means of seed with RAPD and ISSR markers and its possible use for differentiation. J PACD 9: 43-59.

Mebatsion, H. K., Paliwal, J., and Jayas D. S. 2012. Evaluation of variations in the shape of grain types using principal components analysis of the elliptic Fourier descriptors. Comput. Electron Agri. 80: 63-70.

Orozco-Segovia, A., Márquez-Guzmán, J., Sánchez-Coronado, M. E., Baskin, J. M., and Baskin, C.C. 2007. Seed anatomy and water uptake in relation to seed dormancy in *Opuntia tomentosa* (Cactaceae, Opuntioideae). Ann. Bot. 99: 581-592.

Rebman, J. P., and Pinkava, D. J. 2001. *Opuntia* cacti of North America-an overview. Fla Entomol. 84: 474-483.

Reyes-Agüero, J. A., Aguirre, J. R., and Flores, J. L. 2005a. Variación morfológica de *Opuntia* (*Cactaceae*) en relación con su domesticación en la Altiplanicie Meridional de México. Interciencia 30: 476-484.

Reyes-Agüero, J. A., Aguirre, J. R., and Hernández, H. M. 2005b. Systematic notes and a detailed description of *Opuntia ficus-indica* (L.) Mill. (Cactaceae). Agrociencia 39: 395-408.

Sassone, A., Giussani, L. M., Guaglianone, E. R. 2013. Multivariate studies of *Ipheion* (Amaryllidaceae, Allioideae) and related genera. Plant Syst. Evol.299: 1561-1575.

Scheinvar, L., Filardo-Kerstupp, S., Olalde-Parra, G., and Zavaleta-Beckler, P. 2009. Diez especies mexicanas productoras de xoconostles: *Opuntia* spp. y *Cylindropuntia imbricada* (Cacteceae), UNAM, UAEH, UAM, México.

Scheinvar, L. 1995. Taxonomy of utilized opuntias. In: Barbera, G., Inglese, P., Pimienta, E., Arias, E., (Eds): Agro-ecology, Cultivation and Uses of Cactus Pear. FAO. Roma.

Stuppy, W. 2002. Seed characters and the classification of the Opuntioideae. Succ. Plant Res. 6: 25-58.

Valadez-Moctezuma, E., Ortiz-Vásquez, Q., and Samah, S. 2014. Molecular based assessment of genetic diversity of xoconostle accessions (*Opuntia* spp.). Afr. J. Biotechnol. 13: 202-210.

Valdez-Cepeda, R. D., Blanco-Macías, F., and Gallegos-Vázquez, C. 2003. Ordering and numerical classification in prickly pear cactus using fruit attributes. Rev. Chap. Hortic. 9: 81-95.

Wu, B., Quilot, B., Kervella, J., Génard, M., and Li, L. 2003. Analysis of genotypic variation of sugar and acid contents in peaches and nectarines through the Principle Component Analysis. Euphytica. 132:375-384.

Yada, B., Tukamuhabwa, P., Alajo, A., and Mwanga R. O. M. 2010. Morphological Characterization of Ugandan Sweetpotato Germplasm. Crop Sci. 50: 2364-2371.

Chapter 3

Biochemical marker

FRACTIONATION AND ELECTROPHORETIC PATTERNS OF SEED PROTEIN OF *OPUNTIA* GENUS. A PRELIMINARY SURVEY AS A TOOL FOR ACCESSION DIFFERENTIATION AND TAXONOMY

Samir Samah, Elsa Ventura-Zapata and Ernestina Valadez-Moctezuma

Article submitted: Biochemical Systematics and Ecology. Manuscript Number: BSE-D-14-00463R2

Fractionation and electrophoretic patterns of seed protein of *Opuntia* genus. A preliminary survey as a tool for accession differentiation and taxonomy

ABSTRACT

At present, little is known about Opuntia seed proteins and their contribution to the characterization and taxonomy of genotypes that belong to this genus. The variation among 102 accessions of Mexican Opuntia was studied using electrophoretic patterns by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) of their seed total protein (STPs) and seed storage proteins (SSPs). Albumins and globulins, the saline-soluble proteins, were the most abundant protein fractions, with contents that varied from 2.6 to 11.9 mg/mL and 2.6 to 9.5 mg/mL, respectively. These were followed by glutelins (2.3-8.5 mg/mL) and prolamins as the lowest (1.1-7.9 mg/mL) of the four protein factions. On the other hand, STPs content varied between 1.13 and 7.12 mg/mL. Therefore, the total protein content and the different protein fractions were not found to be associated with any of the seed morphological traits. However, regardless of variations in protein content estimated in seeds, the electrophoretic patterns of STPs and SSPs, as outcome of their SDS-PAGE, were not so variable. Furthermore, the individual analysis of each STPs or the SSPs analyses, separately, were not enough to discriminate all accessions, since it was necessary to combine the data resulting from all protein profiles together to differentiate all Opuntia genotypes. The UPGMA and PCoA analyses indicated that there is no separation between accessions of species of the prickly pear (sweets fruits) and xoconostle (acidic fruits), even though the latter were grouped together. Based on biochemical markers analyzed herein, the need for revision of taxonomic assignment of genotypes belonging to the genus Opuntia is suggested.

keywords: Nopal, SDS-PAGE, biochemical markers, characterization.

INTRODUCTION

The genus *Opuntia* (Cactaceae, Opuntioideae) is native to arid and semiarid regions. It is widely distributed in Mexico and the arid regions of America, and also grows in Africa, Australia and the Mediterranean basin. Because Opuntia genotypes can withstand prolonged drought, they are considered a potential alternative crop and an interesting agricultural resource in marginal areas. It is one of the few crops that can be grown in areas that offer little opportunities of growth to produce fruits and vegetables. Opuntia or prickly pears, also known as "nopal" in Mexico, includes about 200 species, of which 66 to 83 species are reported in that country (Guzmán *et al.*, 2003). The Aztecs and other Mesoamerican civilizations used the cactus cladods as a vegetable or as fodder, the sweet ("tunas") or acidic fruits ("xoconostle") as seasonal fruits. Mexico possesses approximately 90 % of the world distribution of prickly pear production and it has been considered the most important producer with 79 % of the total world production (Chavez *et al.*, 2009). The interest for foods derived from these cacti has been increased nationally and internationally due to their potential nutraceutical effects. Furthermore, the *Opuntia* genus is known for its difficult taxonomy, due mainly to the wide morphological variation, synonyms and different ploidy levels (Caruso *et al.*, 2010; Valadez-Moctezuma *et al.*, 2014a).

In Mexico, Opuntia accessions with agronomic and economic importance are safeguarded in several germplasm banks. Information about the structure of germplasm collections is of utmost importance for both conservation and utilization of genetic

resources, considering the diverse nature of the genetic material that they protect (wild accessions, clones, released varieties and genetic stocks of different areas of origin). All this material provides relevant allelic diversity necessary for developing breeding programs (Fufa *et al.,* 2005). The new accessions should include additional intra- or inter- specific variability, in terms of morphological characteristics. Therefore, these new collections need to be evaluated to preserve those that are different (Kumar *et al.,* 2012).

The identification and characterization of accessions and their correct taxonomic assignment are a major challenge faced by collectors. Traditionally, genetic diversity and taxonomy are studied through the analysis of morphological traits. However, morphological characters are limited in number, modified by the environment, and may be controlled by epistatic and pleiotropic effects (Mirali *et al.,* 2007). Another way to estimate the genetic diversity is using molecular markers; these overcome many limitations of morphologically-based genetic diversity analysis and provide information that can help for discrimination between accessions, classification and phylogenetic position. Several molecular markers viz. RFLP, RAPD, ISSRs, AFLP and SSR, are presently available to assess the variability and diversity at the molecular level in Opuntia (Labra *et al.,* 2003; Caruso *et al.,* 2010; Valadez-Moctezuma *et al.,* 1014a; b). Another molecular technique that has proven to be useful in typing crop genotypes is the application of biochemical markers viz. Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis (SDS-PAGE) of seed proteins (Ladizinsky and Hymowitz, 1979). Moreover, the analysis of proteins, such as seed storage protein (SSPs), is of great interest (Kumar et *al.,* 2012). The SSPs are synthesized during seed development and their accumulation is specific of the embryogenesis (Ibl and Stoger, 2012). They provide sources of nitrogen and sulfur required for germination and the establishment of the new plant (Shewry *et al.,* 1995). The

SSPs are classified based on their solubility in albumins (water-soluble), globulins (saline solution-soluble), prolamin (alcohol-soluble) and glutelins (acid or alkaline solution-soluble) (Osborne, 1924). These proteins are expressed by specific genes, are physiologically stable and their composition is little affected by environmental conditions or seasonal fluctuations; they also remain unchanged in dry mature seeds. Consequently, the mature seeds of different ages have a similar protein profile (Galvez *et al.,* 2009). Furthermore, intrinsic changes in the plant, such as rearrangements and even doubling of the chromosomes number, have none or very little effect on these profiles (Ladizinsky and Hymowitz, 1979). For these reasons, SSPs have been used as genetic biomarkers in genetic diversity studies and to resolve taxonomic and evolutionary problems (Shewry *et al.,* 1995; Kumar *et al.,* 2012). They have been used to differentiate mutants from their parents (Hameed *et al.,* 2012), and also used to trace the lineages of *Pisum sativum* resulting from breeding programs (Baranger *et al.,* 2004). Likewise, they have been used in intra-specific variation studies and to discriminate cultivars (Vaz *et al.,* 2004, Stoilova *et al.,* 2006). Other proteins of interest to describe the genetic structure of plant germplasm are the seed total proteins (STPs), which were also analyzed by differences in electrophoretic profiles (Galvez *et al.,* 2009; Galván *et al.,* 2011).

In Opuntia, the edible part of the fruit contains a relatively large number of seeds which vary from 30 % to 40 % of the dry weight of the fruit. The seeds of *Opuntia* sp. contain oils that are rich in polyphenols, flavonoids and tannins, whose concentrations are higher than in the fruit pulp (Morales *et al.,* 2012; Chougui *et al.,* 2013). Nevertheless, the use of seed total proteins as biochemical markers has received little attention. In fact, Galvez *et al.* (2009) used proteins profiles to detect the inter-specific variation in six species of *Opuntia* grown in Argentina. However, there is no information about inter- or

intra- specific diversity using seed protein profiles in *Opuntia* species from Mexico. Similarly, there is no report on the solubility of Opuntia SSPs and no information is available on preferentially accumulated SSPs or their usefulness in systematics and germplasm characterization.

The present research was undertaken to explore the Opuntia seeds' protein profile, total protein and protein storage fractions, to study the genetic variability, and to determine their potential use for characterization of genotypes and taxonomy.

MATERIALS AND METHODS

Germplasm collection

A total of 105 accessions were analyzed in this study. Fruit samples of 102 Opuntia accessions were obtained from the CRUCEN-UACh (Zacatecas) and Nopalera-UACh (Texcoco) germplasm banks, in Mexico. Ten fruits from at least three individual plants of each accession were harvested at maturity. The fruits were washed with running water to remove glochids and impurities, and all mature seeds were removed manually and then dried in the open air; any remaining pulp was cleaned off, and only viable seeds were stored in paper bags at room temperature until use. Three cacti (one sample of *Cylindropuntia* sp., one pitahaya (*Hylocereus undatus*) and one pitaya (*Stenocereus thurberi*)) were included as out-groups. Some Opuntia accessions are classified in delimited species, while others have no specific assignation. The passport data of the 105 accessions are presented in Table 1.

Table 1 List of genotypes used in the present study and their oil percentage and protein content

N	Accession name	Species name	% of oil	Total protein content (mg/mL)	Protein fractions content (mg/mL)			
					Albumin	Globulin	Glutelin	Prolamin
1	Alfajayucan	*O. albicarpa* Scheinvar	10.0	7.1	4.7	3.7	3.8	1.5
2	Alteña Blanco	*Opuntia* sp.	14.6	3.6	2.9	4.6	4.4	2.2
3	Alteña Rojo	*Opuntia* sp.	13.6	4.6	4.2	4.7	4.8	2.2
4	Amarilla 2289	*Opuntia* sp.	11.4	5.0	5.6	7.8	5.9	2.9
5	Amarilla 3389	*Opuntia* sp.	6.1	5.7	4.0	5.2	7.8	2.5
6	Amarilla China	*Opuntia* sp.	7.1	4.1	4.8	4.7	6.4	1.5
7	Amarilla Jalpa	*O. ficus-indica* (L.) Mill.	5.3	6.0	4.3	3.4	4.5	2.8
8	Amarilla Jarro	*O. megacantha* Salm-Dyck.	10.3	4.1	4.0	2.6	7.0	1.7
9	Amarilla Milpa Alta	*O. ficus-indica* (L.) Mill.	12.3	1.1	3.4	5.6	7.0	2.4
10	Amarilla Montesa	*O. megacantha* Salm-Dyck	12.4	5.5	7.2	5.7	6.3	2.9
11	Amarilla Oro	*O. albicarpa* Scheinvar	12.7	2.8	4.4	4.5	6.3	2.4
12	Amarilla San Elías	*Opuntia* sp.	7.4	2.5	4.0	7.0	5.3	2.6
13	Amarilla Zacatecas	*O. megacantha* Salm-Dyck	5.9	2.8	5.6	4.9	3.4	2.8
14	Amarillo	*O. megacantha* Salm-Dyck	15.8	3.1	4.4	5.1	4.6	4.3
15	Amarillo Aguado	*Opuntia* sp.	9.2	5.6	4.7	5.4	6.1	4.0
16	Amarillo Plátano	*O. megacantha* Salm-Dyck.	14.0	3.1	4.3	5.1	6.6	3.1
17	Atlixco	*O. ficus-indica* (L.) Mill.	8.9	5.5	6.8	7.5	6.5	2.7
18	Bam	*Opuntia* sp.	12.8	4.0	5.5	6.4	2.6	2.8
19	Blanca de Castilla	*Opuntia* sp.	14.2	3.6	5.7	6.6	5.0	3.1
20	Blanca del cerro	*Opuntia* sp.	7.7	3.4	3.9	6.1	3.0	3.0
21	Blanca San José	*O. albicarpa* Scheinvar	10.1	4.3	6.8	4.9	3.9	4.3
22	Blanco Atlacomulco	*Opuntia* sp.	7.8	4.0	9.9	5.7	7.1	3.7
23	Blanco Huexotla	*Opuntia* sp.	15.0	5.2	6.8	5.2	6.3	4.1
24	Bola de Masa	*O. albicarpa* Scheinvar	13.4	5.7	5.3	4.6	5.2	4.2
25	Burrona	*O. albicarpa* Scheinvar	6.6	5.7	4.3	5.2	6.6	4.3
26	Cacalote	*O. cochinera* Grif.	12.2	3.1	6.3	7.3	6.2	2.2
27	Camuezo	*O. megacantha* Salm-Dyck	13.1	3.6	6.7	5.5	3.1	1.4
28	Cardón Blanco	*O. streptacantha* Lem.	11.6	3.4	6.2	5.1	6.0	1.1
29	Cardona	*O. streptacantha* Lem.	11.3	3.3	5.8	4.6	5.1	1.8
30	Cardona Castilla	*O. streptacantha* Lem.	12.0	2.5	8.5	5.8	7.2	1.6
31	Cascarón	*O. chaveña*	18.6	4.3	7.2	5.6	4.3	2.4
32	Chapeada	*O. albicarpa* Scheinvar	7.6	5.8	7.0	6.6	6.6	1.9
33	Charola Tardía	*O. streptacantha* Lem.	15.0	3.1	5.8	4.9	4.1	2.1
34	Chicle	*O. ficus-indica* (L.) Mill.	10.6	2.3	5.7	4.8	3.9	2.1
35	Col. Bar. Chica	*Opuntia* sp.	9.2	4.4	7.8	6.2	5.1	2.5
36	Col. Bar. Grande	*Opuntia* sp.	10.0	6.7	8.2	6.0	6.5	2.6
37	Color de Rosa	*O. albicarpa* Scheinvar	13.0	6.0	6.0	6.0	4.2	2.5
38	Colorada	*Opuntia* sp.	10.7	5.9	9.0	6.8	6.4	1.8
39	Copena CEII	*O. ficus-indica* (L.) Mill.	8.8	6.6	4.2	7.7	8.5	3.0
40	Copena F1	*O. ficus-indica* (L.) Mill.	7.4	4.1	4.7	6.2	5.3	2.6
41	Copena T12	*O. ficus-indica* (L.) Mill.	10.6	5.7	4.4	6.3	6.3	2.6
42	Copena T5	*O. ficus-indica* (L.) Mill.	8.3	5.4	2.6	5.5	4.9	3.5
43	Copena V1	*O. ficus-indica* (L.) Mill.	14.3	4.4	7.4	4.2	5.1	3.0
44	Copena Z1	*O. albicarpa* Scheinvar	12.3	6.4	5.9	4.9	3.7	3.5
45	Cristalina	*O. albicarpa* Scheinvar	12.1	2.5	3.5	6.2	4.3	3.1
46	Fafayuca	*O. albicarpa* Scheinvar	10.5	3.7	5.6	6.2	6.0	2.3
47	Gavia	*O. albicarpa* Scheinvar	9.1	5.6	5.4	6.6	4.8	4.3
48	Green de Guanajuato	*Opuntia* sp.	5.5	6.0	4.1	8.1	5.0	3.4
49	Huatusco	*Opuntia* sp.	8.8	4.1	3.7	9.2	5.2	2.8
50	INIFAP	*Opuntia* sp.	10.1	4.1	3.8	7.3	5.1	3.2
51	Jade	*Opuntia* sp.	12.9	4.5	4.2	6.4	5.6	3.5
52	Jarilla Grande	*Opuntia* sp.	9.9	2.9	7.9	5.8	5.8	3.6
53	Laltus	*Opuntia* sp.	5.6	5.9	3.4	8.2	3.0	4.4
54	Larreguin	*O. ficus-indica* (L.) Mill.	13.8	5.4	7.2	6.9	6.1	3.1
55	Liso Amarillo	*Opuntia* sp.	11.1	5.5	7.4	6.1	6.8	4.6
56	Liso Forrajero	*Opuntia* sp.	7.3	3.0	9.9	5.3	4.3	4.6
57	Mango	*O. albicarpa* Scheinvar	8.5	3.2	7.8	6.8	5.5	5.7
58	Mansa Amarilla	*Opuntia* sp.	12.0	4.1	3.1	3.5	4.1	4.6
59	Memelo	*O. affinis hyptiacantha*	10.9	3.3	8.7	5.7	4.0	4.1
60	Montesa	*Opuntia* sp.	8.6	4.3	10.1	7.5	5.9	4.4
61	Morada	*O. megacantha* Salm-Dyck.	6.0	3.8	6.4	5.3	4.4	3.5
62	Morada T10	*O. megacantha* Salm-Dyck.	10.0	4.3	8.4	8.5	4.2	2.8
63	Naranjón Legítimo	*O. albicarpa* Scheinvar	9.5	3.7	8.2	7.1	4.2	3.8
64	Naranjona	*Opuntia* sp.	12.1	4.2	9.8	6.2	3.9	2.6
65	*O. cochillinifera*	*O. cochillinifera*	7.5	5.9	7.8	6.1	4.9	2.3
66	Oreja de elefante	*O. undulata* Grif.	7.5	3.9	11.9	6.6	6.1	2.9
67	Pabellón	*O. ficus-indica* (L.) Mill.	8.0	2.8	3.9	6.8	4.4	3.5
68	Pachon	*Opuntia* sp.	10.5	5.0	7.0	5.6	4.9	3.9
69	Pico Chulo	*O. megacantha* Salm-Dyck.	6.7	5.3	5.3	5.3	5.3	3.1
70	Pico de Oro	*Opuntia* sp.	10.2	5.0	8.1	6.4	6.9	3.1
71	Platano	*Opuntia* sp.	7.6	6.0	7.2	6.5	5.4	3.0
72	Red Villa Puebla	*Opuntia* sp.	4.8	6.3	8.9	7.2	5.0	7.9
73	Reyna	*O. albicarpa* Scheinvar	9.4	5.1	6.0	4.4	4.7	4.9
74	Reyna Crucen	*Opuntia* sp	9.1	5.1	6.8	4.9	4.4	6.0
75	Roja Azteca	*O. megacantha* Salm-Dyck.	6.1	5.2	7.0	8.0	6.0	5.3
76	Roja San Martín	*O. megacantha* Salm-Dyck.	7.0	3.9	8.5	6.0	5.9	5.0

77	Rojo 3589	*Opuntia* sp.	9.5	4.2	6.2	7.7	6.8	4.6
78	Rojo Lirio	*O. megacantha* Salm-Dyck.	9.1	4.6	6.5	9.5	6.1	5.9
79	Rojo Pelón	*O. ficus-indica* (L.) Mill.	8.0	3.7	5.4	5.6	4.9	4.9
80	Rojo UACh	*Opuntia* sp.	11.6	4.4	7.5	6.6	6.3	4.6
81	Rojo Vigor	*O. ficus-indica* (L.) Mill	9.7	3.8	6.4	7.4	5.0	4.5
82	Rosa de Castilla	*O. megacantha* Salm-Dyck.	5.9	3.8	6.7	6.5	4.3	4.4
83	Rubí Reyna	*O. megacantha* Salm-Dyck.	10.5	5.3	8.8	7.3	7.2	4.0
84	San Juan	*Opuntia* sp.	20.2	4.5	7.4	7.4	3.5	3.7
85	Sangre de Toro	*Opuntia* sp.	10.7	5.2	6.4	8.0	7.3	5.6
86	Sanjuanera	*O. lasiacantha* Pfeiffer	9.9	3.8	6.1	7.4	7.1	4.7
87	Solferino	*Opuntia* sp.	11.4	3.8	7.0	7.5	3.4	4.1
88	Tapon rojo	*O. robusta* H.L.	12.1	3.3	5.5	7.4	6.1	3.9
89	Tapona de Mayo	*O. robusta* H.L.	10.2	4.8	8.7	6.6	6.9	4.9
90	Tobarito	*Opuntia* sp.	9.5	4.6	7.1	7.2	5.3	5.6
91	Toluca	*Opuntia* sp.	7.1	3.6	7.5	7.5	5.6	4.4
92	Torreoja	*O. megacantha* Salm-Dyck.	8.4	2.7	7.9	6.0	4.5	2.6
93	Trompa de Cochino	*Opuntia* sp.	10.5	2.7	8.5	8.7	3.1	3.3
94	Tuna Mansa	*O. albicarpa* Scheinvar	8.1	5.1	5.2	4.9	2.6	5.2
95	Tuna Rosa	*O. albicarpa* Scheinvar	8.1	3.1	5.2	6.8	2.3	5.4
96	Tuna Sandia	*Opuntia* sp.	10.0	3.9	5.2	6.8	3.8	4.8
97	Var S/I	*Opuntia* sp.	11.1	4.8	5.5	8.6	4.8	5.9
98	Villanueva	*O. albicarpa* Scheinvar	8.4	5.0	6.8	7.9	4.4	5.9
99	X. Colorado	*O. joconostle* F.A.C. Weber	10.6	4.8	6.6	6.8	4.1	7.0
100	X. Cuaresmero	*O. matudae* Scheinvar	9.6	3.6	6.6	4.4	4.5	6.3
101	X. Blanco	*O. joconostle* F.A.C. Weber	10.2	3.3	5.9	5.1	5.3	7.3
102	X. Manzano	*O. joconostle* F.A.C. Weber	14.6	3.8	4.3	7.9	4.5	6.6
	Mean		**10.1**	**4.4**	**6.2**	**6.2**	**5.2**	**3.7**
	Max		**20.2**	**7.1**	**11.9**	**9.5**	**8.5**	**7.9**
	Min		**4.8**	**1.1**	**2.6**	**2.6**	**2.3**	**1.1**
103	Cylindropuntia	*Cylindropuntia* sp.	8.9		3.5	4.8	6.1	6.4
104	Pitahaya	*Hylocereus undatus*	27.3		16.5	13.4	15.2	12.4
105	Pitaya	*Stenocereus thurberi*	15.8		9.1	6.6	5.7	6.9

Defatted flour preparation

The seeds were ground to a fine powder with a Coffee/Spice Grinder GX4100 (Krups, Mexico), and thereafter the powdered sample was de-oiled with hexane (10 % w/v) (Chougui *et al.*, 2013) in Falcon tubes for 12 h with constant agitation. The defatted flour was air-dried at room temperature and subsequently kept in air-tight plastic containers at 4 ºC prior to use. The oil weight was determined as follows: Oil weight (%) = $[(M_1-M_0)/M_2]*100$, where M_0 is the weight of the empty flask (g), M_1 the weight of the flask after evaporation (de-oiled sample) (g) and M_2 the weight of the seeds powder (g).

Seed total protein (STPs) preparation

STPs were extracted according to Galvez *et al.* (2009) with some changes. Three grams of seed flour were treated with 0.025 M Tris-0.192 M glycine, pH 8.3 buffer (10 v/w) in continuous agitation for 2 h at room temperature (20 ºC). The suspension was centrifuged at 15 000 rpm for 10 min and the supernatant was stored at -20 ºC.

Seed protein sequential fractionation based on solubility

Fractionation of Opuntia seed protein was carried out as described by Santos *et al.* (2013). The method was based on the classical Osborne protein fractionation procedure. For albumin extraction, 300 mg of defatted flour sample were suspended in 1.5 mL of 10 mM L^{-1} Tris-HCl solvent (pH 7.5) and 1 mM L^{-1} EDTA. The mixture was shaken for 12 h and centrifuged at 15 000 rpm for 15 min at 4 °C. The supernatant (albumin) was collected and precipitated with 1.5 mL of cold acetone, homogenized by inversion, and stored for 12 h in a freezer at -20°C. The solution was centrifuged at 15 °C for 15 min, and the supernatant (acetone) was discarded. The pellet (albumin) was stored in a freezer for subsequent analysis on SDS-PAGE. For globulin extraction, 1.5 mL of 10 mM L^{-1} Tris-HCl (pH 7.5), 1 mM L^{-1} EDTA, and 0.5 mM L^{-1} NaCl were added to the initial defatted flour. This solution was shaken and centrifuged, and the supernatant (globulin) was collected, precipitated with acetone, homogenized and stored overnight in a freezer. The solution was centrifuged and the supernatant (acetone) was discarded. The pellet (globulin) was stored in a freezer for subsequent analysis on SDS-PAGE. For prolamin extraction, 1.5 ml of 60 % isopropanol (v/v) was added to the Opuntia flour. The solution was shaken and centrifuged, and the supernatant (prolamin) was precipitated, homogenized and stored in a freezer overnight. Then, the solution was thawed and centrifuged, and the supernatant (acetone) was discarded. The pellet (prolamin) was stored in a freezer. For glutelin extraction, 1.5 mL of glacial acetic acid (v/v) was added to the flour. This solution was shaken vigorously for two hours and centrifuged for 15 min at 4 °C. The supernatant (glutelin) was collected in a new tube and precipitated with 1.5 mL of acetone, homogenized by inversion and stored in a freezer overnight. The solution was centrifuged and the supernatant (acetone) was discarded. The pellet (glutelin) was stored in a freezer for subsequent analysis on SDS-PAGE.

Protein sample preparations and content determination

Total protein and protein storage fractions were suspended in 2X SDS-PAGE sample buffer for continuous gel system [0.5 M sodium phosphate buffer pH 7 (0.64 mL), water (1.36 mL), glycerol (1.6 mL), 10 % SDS (3.2 mL), 2-mercaptoethanol (0.8 mL), 0.1 % bromophenol blue (0.4 mL)] (Hames, 1998). The samples were heated at 95 ºC for 5 min to denature proteins prior to being loaded on gel. The content of total protein and protein fractions in the samples was estimated by spectrophotometry (Thermo Scientific Nanodrop ND-1000, USA) with absorbance at 280 nm (Goldring, 2012).

Protein electrophoresis

Electrophoresis was carried out in continuous SDS-PAGE system (Hames, 1998), using 10 % polyacrylamide solution (30 % acrylamide:bisacrylamide, 0.5 M sodium phosphate pH 7, 10 % SDS, TEMED, 10 % ammonium persulfate). The 5X running buffer was composed of 0.5 M sodium phosphate, pH 7, and 0.5 % SDS. Then, 50 µg of each protein sample (40 µl) were loaded into each well. Gels were subjected to electrophoresis at 90 V for approximately six hours using Dual MGV-216-33 vertical electrophoresis gel system (CBS, USA). The gels were stained with Coomassie brilliant blue R-250 (BIO RAD, UK) containing 45 % methanol and 10 % acetic acid and de-stained in the same solution but without the dye. SDS-PAGE assays were performed in duplicate and only bands that were clearly obtained twice were considered for further analyses.

Data analysis

The electrophoresis pattern obtained was analyzed in two parts. First, the molecular weight of each protein band was estimated using the PageRuler^{TM} Prestained Protein

Ladder (Thermo Scientific, Lithuania). Second, the observed band patterns for total protein and proteins fractions were assigned to a binary system of 0 (absence) and 1 (presence). Then, the Nei and Li/Dice coefficient was used to calculate genetic similarity and, after that, cluster analysis was performed using the Unweighted Pair Group Method with Arithmetic Averages (UPGMA) in FreeTree software (v0.9.1.50) and Tree View (v1.6.6). In addition, Principal coordinate analysis (PCoA) was used to further confirm cluster analysis results with the GenAIEx 6.5 software (Peakall and Smouse, 2012).

RESULTS AND DISCUSSION

Experiments conditioning

A necessary step prior to the study of protein profiles of Opuntia seeds was the defatting of flour to avoid any effect of fat on the protein extraction and electrophoresis, as reported by Wang *et al.* (2007). Preliminary tests showed the presence of fat interference on protein profiles. The use of hexane to remove the oil showed that Opuntia seeds contain an average of 10.1 % of fat, ranging from 4.8 % to 20.2 %. Half of the genotypes showed oil content of more than the total average, while only 1 % of the genotypes had fat content below to 5 % (Table 1). This variation reflects the high variability present in Mexican genotypes in terms of oil content. Quantities of 5.5 % of oil were reported by Tlili *et al.* (2011) in Tunisian varieties of *O. ficus-indica* and 7.3-9.3 % in 5 varieties of Opuntia in Algeria (Chougui *et al.*, 2013).

The effect of seed age on protein content and on the total protein profiles was determined, since the accessions studied were collected from 2007 and 2013. The results of ten accessions showed a reduction in total protein content by an average of 23.2 %, ranging

from 1.3 to 51 % (Fig. 1). In contrast, seed total protein profiles of the same genotypes showed high stability over the years (data not shown). These results agree with those reported by Carreras *et al.* (1997) and Galvez *et al.* (2009), who concluded that the age of Cactaceae seeds did not affect protein profiles. Therefore, protein profiles can be used as genetic markers because they remain stable over years and have little environmental effect (Ladizinsky and Hymowitz, 1979; Shewry *et al.*, 1995; Kumar *et al.*, 2012).

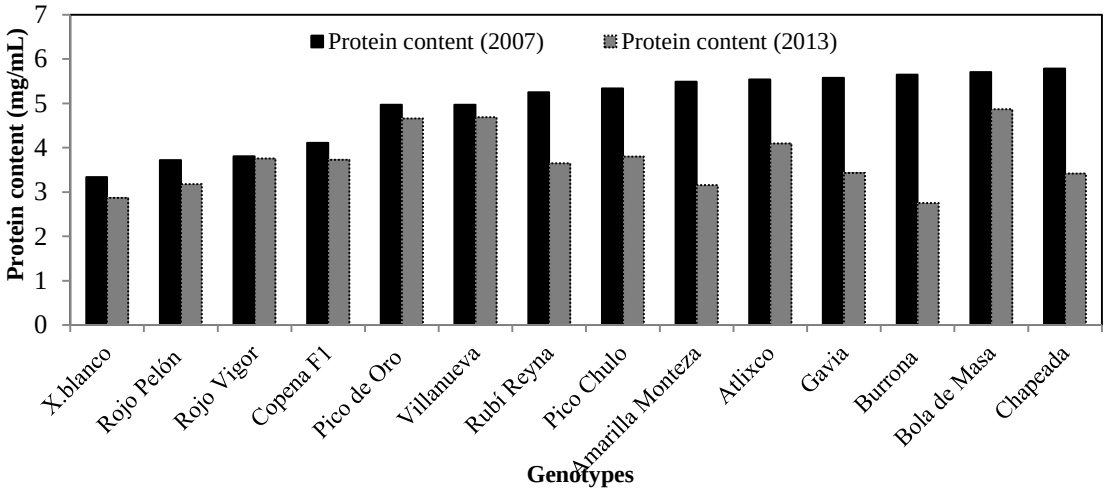


Fig. 1. Total protein content in seeds of 10 Opuntia accessions collected in different years (2007 and 2013), from two localities in Mexico.

Opuntia accessions protein characterization

The quantity of total protein in Opuntia seeds showed high variation, ranging between 1.13 and 7.12 mg/mL with a total average of 4.3 mg/mL (Table 1). The STPs levels were divided into three groups (Fig. 2): high (≥ 5 mg/mL), medium (3-5 mg/mL) and low (≤ 3 mg/mL). According to this classification, out of 102 Opuntia accessions studied here, 33 % of the genotypes showed high protein content and 60 % had intermediate protein content (Table 1). El-Guizani *et al.* (2012) informed that the amount of STPs in cultivars of *O. ficus-indica* was about 9.6-11.8 %; while Tlili *et al.* (2011) reported a quantity of 6.0 %. Such variation might reflect differences in environmental conditions. Furthermore, the

STPs content was not correlated with any of the morphological characteristics, such as weight and area, length and width of seeds.

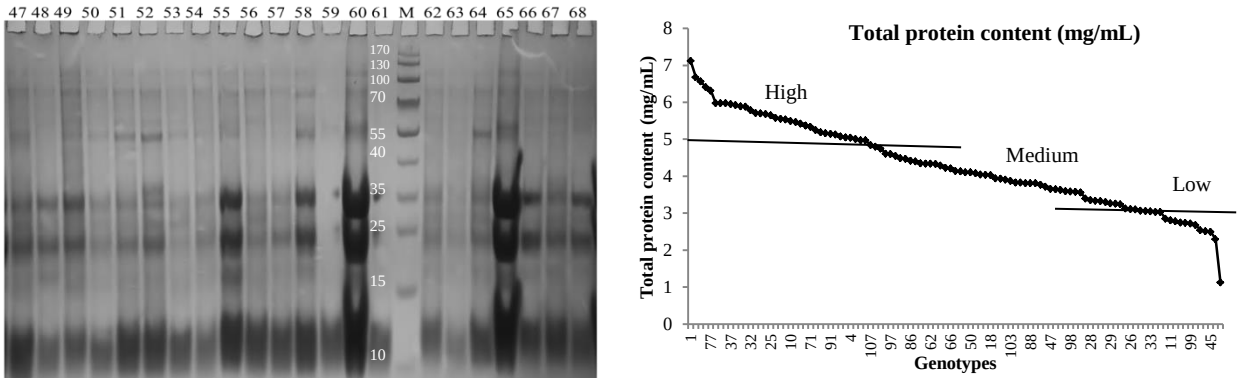


Fig. 2. SDS-PAGE analysis and contents of total proteins of Opuntia accessions. Proteins were resolved under reducing conditions on the basis of equal volume of protein. Left: Representative total protein profiles of Opuntia genotypes. M, protein weight marker (kDa). lanes 47-68: Opuntia genotypes listed in Table 1. Right: Graphical representation of total seed protein variation in the 102 genotypes of Opuntia from Mexico.

The SSPs were separated into four fractions according to their solubility differences. The SSPs sequential analysis revealed the presence of the four fractions in Opuntia seeds vis. albumins, globulins, prolamins and glutelins (Table 1; Fig. 3). The albumin and globulin fractions were the most predominant fractions in the Opuntia seeds with a concentration average of 6.2 mg/mL (Table 1). Albumin variation ranged from 2.6 to 11.9 mg/mL. A third of genotypes (34.3 %) had an albumin amount higher than 7 mg/mL (Fig. 3). A small proportion of the genotypes (11.4 %) exhibited amounts less than 3 mg/mL. However, half of the genotypes (54.3 %) displayed intermediate amounts of albumins (3-7 mg/mL). The globulin content varied between 2.6 and 9.5 mg/mL. A quart (26.9 %) of the genotypes had an amount of this protein fraction higher than 7 mg/mL. A minor percentage

of the genotypes (5.8 %) showed amounts lower than 3 mg/mL, while 37.3 % of the genotypes exhibited intermediate amounts of the globulin fraction (3-7 mg/mL) (Fig. 3). The prolamin fraction was also present in the Opuntia seeds but with quantities lesser than the other fractions (Table 1). The prolamin fraction showed an average of 3.7 mg/mL, and varied from 1.1 to 7.9 mg/mL. Similarly, the studied genotypes could be grouped based on the amount of prolamins (Fig. 3). A small percentage (4.8 %) of the genotypes had high levels (< 6 mg/mL), 37.5 % of genotypes low levels (> 3 mg/mL), and 57.7 % of genotypes presented intermediate levels of prolamins (3-6 mg/mL). The glutelin fraction exhibited an intermediate amount of the SSPs (average of 5.2 mg/mL). The variation of glutelin amounts oscillated between 2.3 and 8.5 mg/mL. Most of the genotypes (87.5 %) were found to have intermediate levels of glutelin (3-7 mg/mL), while only 8.7 % of genotypes had levels higher than 7 mg/mL and 3.9 % of the genotypes showed low levels of glutelins (> 3 mg/mL) (Fig. 3).

The majority of agriculturally important SSPs are albumins, globulins or prolamins. While albumins are found in all seeds, prolamins and glutelins are most abundant in monocotyledon seeds and globulins are prevalent in dicotyledon seeds (Higgins, 1984; Miernyk and Hajduch, 2011). Our results revealed that SSPs solubility in Opuntia is similar to the water-salt soluble proteins (albumins and globulins). However, glutelin and prolamin were also present, which were somewhat different from the proportions reported in dicot seeds. There are different soluble fractions in various plants studied so far; they are highly heterogeneous and are composed of number polypeptides of different molecular weights. The proportion of these fractions in the seed can vary and is regulated by a number of distinct genetic and environmental factors (Panozzo *et al.*, 2001). The final status of a particular fraction in the seed is also determined by the extent of its degradation,

if any, by specific proteinases during seed development. Unlike cereals, where prolamins and glutelins are the major storage proteins, albumins and globulins comprise the major storage proteins of legumes (Mirali *et al.,* 2007). Santos *et al.* (2013) reported that the glutelin fraction was the predominant protein in rice endosperm, accounting for 60-80 % per weight of total seed protein, and prolamin makes up about 20-30 %. A similar value was recorded in ornamental ecotypes of *Ebenus cretica* (Syros *et al.* 2003). In the case of *Mucuna pruriens* (Fabaceae), the globulin fraction was the greatest (62.5 %), followed by albumin (18.3 %), glutelin (15.8 %) and prolamin (3.4 %) (Bellani *et al.,* 2013).

The broad range of protein content in Opuntia seeds (STPs and SSPs) described above may be explained as the result of self-pollination inherent in the floral structure, reproduction system that is largely frequent in Opuntia (Reyes-Agüero *et al.,* 2006). A wide genetic base for protein content reported in self-pollinated crops such as legumes and cereals should be homozygously maintainable at low and high extremes of the range in natural accessions of individual species of these crops. In contrast, cross-pollination leads to mixing-up of genomes with different levels of protein content favoring the stabilization of lines with the most frequent range of protein contents (Singh and Matta, 2010).

There was no significant correlation between protein contents of STPs and SSPs or between the distinct storage proteins fractions contents (Fig. 3). This fact shows that genotypes with a greater amount of one fraction are not necessarily the same genotypes that have greater amounts of the remaining fractions. As seen in the case of STPs content, the proportion of different fractions, i.e. albumins, globulins, glutelins and prolamins, was also not found to be associated with any of the seed traits. Similar results were reported by Singh and Matta (2010) in *Citrullus* and *Praecitrullus* accessions.

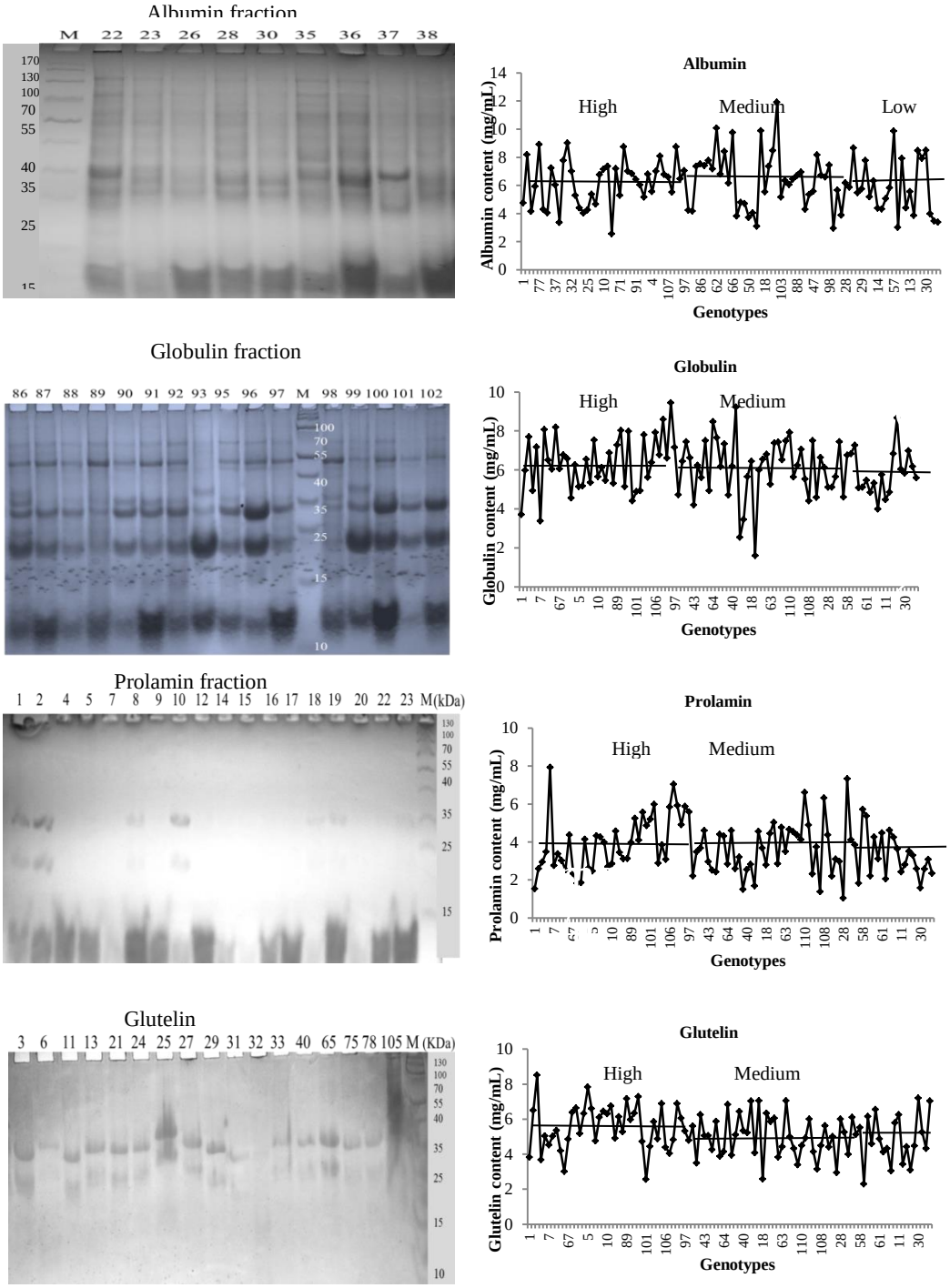


Fig. 3. SDS-PAGE analysis and SSPs contents of Opuntia accessions. Proteins were resolved under reducing conditions on the basis of equal volume of protein. Left: Representative SSPs profiles (albumin, globulin, prolamin and glutelin fractions) of Opuntia genotypes (lane numbers are listed in Table 1). M, protein marker (kDa). Right: Graphical representation of SSPs variation in the 102 genotypes of Opuntia from Mexico. The arrangement of the genotypes in the graph was based on the total protein content, from high to low, previously illustrated in Fig. 2.

Electrophoretic patterns

Protein profiles of STPs and SSPs resolved by SDS-PAGE were analyzed qualitatively. For STPs, Opuntia genotypes were found to have an average of 10.5 bands per accession, and varied in number between 8 and 14 bands/accession. For albumins, this range oscillated from 7 to 15 bands, showing an average of 12.2 bands per accession. Moreover, the average of globulin bands per accession was 8.3, with a variation level from 4 to 11 bands per accession.

For STPs, a total of 16 polypeptides (bands) were generated by 102 genotypes in a range of molecular weight oscillating from 12 to 110 kDa. Five of these bands showed no qualitative differences between the genotypes studied, revealing a percentage of polymorphism of 68.8 %. Galvez *et al.* (2009) reported lower proportion of polymorphisms (23.7 %), perhaps because they analyzed only six species of Opuntia with few populations. In our case, monomorphic bands had extreme molecular weights (12 kDa, 60, 65, 80 and 110 kDa), while polymorphic bands showed intermediate molecular weights. In particular, a 14 kDa band was found in six accessions only (Amarilla Jalpa, Amarillo, Burrona, Cardona de Castilla, Pico Chulo and Pico de Oro). In turn, the three out-groups generated additional bands (two from Cylindropuntia, two from pitaya and three from pitahaya).

For the albumin fraction, Opuntia genotypes generated a total of 16 polypeptides with molecular weights ranging from 12 to 110 kDa. Five bands with extreme molecular weights (12 kDa, 35, 85, 90 and 110 kDa) were monomorphics; whereas bands with intermediate molecular weight were polymorphics, revealing a percentage of polymorphism of 68.8 %. It is noteworthy that the 40 kDa band was detected in four

accessions only (Villanueva, X. Colorado, X. Cuaresmero and X. Manzano). Another 55 kDa band was absent in 4 genotypes (Alfajayucan, Alteña Blanco, Amarilla Jalpa and Amarilla Jarro). The 63 kDa band was absent in two accessions (Amarilla Jalpa and Amarilla Jarro). In turn, the three out-groups accessions generated additional bands (three from Cylindropuntia, four from pitaya and five from pitahaya).

For the globulin fraction, a total of 14 bands were counted and their molecular weights varied between 12 and 75 kDa. Three of these bands (12, 23 and 75 kDa) showed no qualitative differences between the 102 Opuntia genotypes, displaying a percentage of polymorphism of 78.6 %. The 17 and 44 kDa bands were detected in Atlixco and Copena CEII genotypes, respectively. Another 25 kDa band was present in four genotypes (Rojo Lirio, Rojo UACH, Trompa de Cochino and X. Manzano). Two out-group accessions showed additional bands with respect to Opuntia genotypes (five from pitahaya and six from pitaya).

As for the prolamin and glutelin fractions, just three bands were obtained for each fraction; two of them were polymorphics for the prolamin fraction and only one polymorphic band was found in the glutelin fraction.

Five bands (12, 20, 35, 38 and 60 kDa) were detected indifferently in the two albumin and globulin fractions. This indicates different solubility of the same components of the protein. This result indicated that the cross-contamination of protein fractions is an Osborne fractionation problem, as it has been reported for seeds of other plants (Gazzola *et al.*, 2014; Rubio *et al.*, 2014). Applying the SDS-PAGE system under reducing conditions (without 2-mercaptoethanol) may avoid overlap of these two fractions, since differences

have been demonstrated in the number of bands for the same protein fraction depending on the presence or absence of 2-mercaptoethanol (Wang *et al.,* 2007). The dialyses of each fraction against the corresponding solution can also resolve the presence of cross-contamination (Gazzola *et al.*, 2014). Moreover, the sequence and the structure of the common polypeptide can confirm or not the cross-contamination among albumins and globulins. Nevertheless, qualitative differences observed in the two protein fractions were detected in different accessions; for this reason, both fractions were used separately to differentiate Opuntia genotypes.

Genetic diversity and genotypes differentiation

Seed protein polymorphism has been used in genetic studies in many plant species (Karihaloo *et al.,* 2002; Stoilova *et al.,* 2006; Galván *et al.,* 2011). For example, the most commonly used molecular markers in wheat have historically been SSPs (glutenins and gliadins). They have been recommended as reliable genetic markers to differentiate wheat genotypes for bread-making (Gupta *et al.,* 1999; Fufa *et al.,* 2005). Genes visualized as precise protein bands or spots, which reflect physiological status, are good candidates for assessing variability and establishing genetic distances and phylogenetic relationships between different species and individuals (Fufa *et al.,* 2005). At present, there have been few studies that used seed protein markers to determine the genetic diversity in the *Opuntia* genus. Carreras *et al.* (1997) reported the usefulness of these protein patterns to separate nine species belonging to the genera *Stetsonia*, *Cereus*, *Harrisia*, *Opuntia*, *Tephrocactus* of the Cactaceae family, while Galvez *et al.* (2009) differentiated 6 species of *Opuntia*.

In our study, the UPGMA analysis based on STPs profiles separated the studied accessions into three groups. The first cluster located the out-groups pitahaya and pitaya,

while the remaining two groups were formed by 42 and 61 genotypes of Opuntia. Within these two groups, several genotypes were grouped together without discrimination and presented the same profile of STPs. Only 18 genotypes (2, 3, 7, 14, 25, 41, 43, 48, 58, 69, 70, 74, 81, 87, 96, 103-105) clustered in a single clade each (supplementary data S1. A). The UPGMA analysis based on the profiles of the albumin fraction characterized only 28 genotypes (1, 2, 4, 7-10, 14-16, 18, 22, 27, 35, 37, 41, 43, 44, 50, 51, 53, 60, 98, 100, 101, 103-105). However, genotypes belonging to xoconostles (99-102) were grouped together in a sub-cluster (Supplementary data S1, B). Moreover, the globulin fraction profiles discriminated 13 genotypes only (7, 17, 37, 39, 40, 45, 48, 58, 66, 70, 103-105) (Supplementary data S1. C). UPGMA analysis of the prolamin and glutelin fractions separated the Opuntia accessions into two groups (Supplementary data S1, D and E), without the ability to differentiate any genotype from each other within each group. These results are consistent with the small amounts of these two fractions in the Opuntia seeds.

Whereas the genotypes discriminated in STPs and SSPs analyses were not the same, therefore the data resulting from all protein profiles were combined and the UPGMA analysis was carried out. Smith and Desborough (1987) and Syros *et al.* (2003) recommended the use of the pooled data from both types of electrophoresis for genetic diversity analysis and taxonomic comparisons in *Solanum* and *Ebenus* species. In our case, thirty-six polymorphic polypeptides resulting from STSs and SSPs analyses were combined to differentiate the 102 Opuntia accessions. The Dice coefficient revealed a similarity range that oscillated between 7 and 99 % when the out-groups were included, and between 63 % and 99 % within Opuntia accessions. This result showed the close genetic distance between the genotypes studied. Similar results were reported by Carreras *et al.* (1997) of Argentina's species recording a range of similarity oscillating from 65 % to

84 %, while closer similarities (81-97 %) were estimated by Galvez *et al*. (2009). Mirali *et al.* (2007) reported that SDS-PAGE analyses showed that the abundance of polymorphisms in cross-fertilized species is much higher than those of self-fertilized ones. This is in line with the generally accepted concept that in populations that have high levels of selfing, the heterozygosity level is much lower than in cross-fertilized species (Maquet *et al.*, 1996). This is reflected in the close genetic distance estimated by the analyses of seed proteins found here, despite these genotypes belonging to different species (Table 1).

The cophenetic correlation coefficient between the original similarity matrix and the cophenetic matrix derived from the UPGMA dendrogram was very high (r = 0.97), indicating a good fit between the dendrogram and the similarity matrix. UPGMA analysis separated the genotypes into 6 groups, three of which were formed with a single genotype corresponding to the out-groups Cylindropuntia, pitahaya and pitaya. Another group was formed with the "Amarilla Jalpa" accession. The remaining two groups were composed of 39 and 62 Opuntia genotypes (Fig. 4). Again, xoconostle genotypes were clustered together within a large clade of the remaining prickly pears. Likewise, the genotypes belonging to the Copena set were grouped together with the "Cristalina" accession. These genotypes were the result of selection programs in the 80s in Mexico. Generally, this assay allowed discrimination of all genotypes, thus showing the utility of protein markers for differentiation of Opuntia genotypes. Singh and Matta (2010), Kumar *et al.* (2012) and Gazzola *et al.* (2014) recognized that protein analyses should be sufficiently informative and inexpensive for molecular distance estimation.

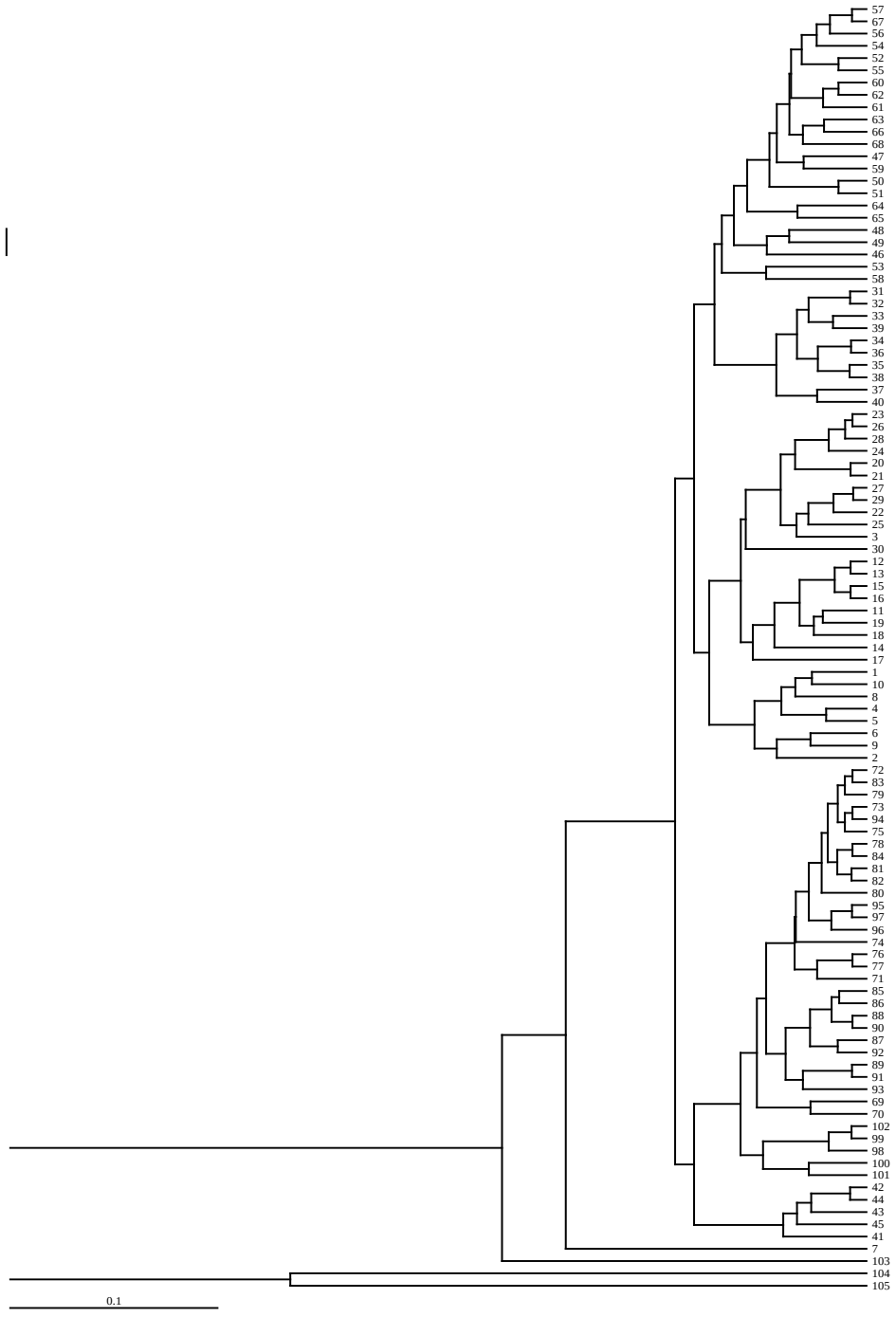


Fig. 4. The dendrogram based on the qualitative analysis of total protein and SSPs resolved in 102 Opuntia accessions contrasted with three out-groups obtained from UPGMA analysis. Protein bands were resolved under reducing conditions on SDS-PAGE continuous system. Accessions numbers are listed in Table 1.

Usefulness of the seed proteins SDS-PAGE analysis for taxonomy

Analysis of the proteins on SDS-PAGE gels is widely suggested as additional and important method for taxonomy studies, because the profiles patterns are very stable, besides this analysis is simple and of low cost (Galvan *et al.,* 2011; Kumar *et al.,* 2012).

To illustrate the taxonomic status of the genotypes studied using protein markers, data from 44 accessions, whose assignment of species is documented (Table 1), were subjected to cluster and PCoA analyses. The tree obtained by the UPGMA grouping methodology using Dice coefficient of combined data from STPs and SSPs is illustrated in Figure 5 (r = 0.98). The dendrogram showed 6 clusters, three of which (Grp. 1, Grp. 2 and Grp. 3) correspond to the genotypes included as out-groups. Once again, the sweet prickly pear "Amarilla Jalpa" was located in a separated cluster (Grp. 4); this accession may be an interesting genotype for further protein analysis and genetic expression. The remaining two clusters (Grp. 5 and Grp. 6) were composed with 22 and 21 accessions, respectively. The clustering of accessions was not in accordance with the current assignment of genotypes in their respective species. This is probably related to the high level of phenotypic plasticity and ploidy levels, and also due to the morphological diversity of these accessions. These genotypes had several end uses, as fruits, vegetables and/or forage (Samah and Valadez-Moctezuma, 2014). For these reasons, many studies have suggested the revision of the classification of the *Opuntia* genus (Helsen *et al.,* 2009; Caruso *et al.,* 2010; Valadez-Moctezuma *et al.,* 2014a). Typically, the location of an accession into a species is arbitrary due to lack of solid descriptors. For this reason, many genotypes considered in this study have not yet been assigned taxonomically (Table 1). However, the xoconostle genotypes were grouped together within the large Grp. 5 that also contained the sweet prickly pear genotypes. This behavior is also reported by Valadez-Moctezuma *et al.* (2014a) using

DNA markers (RAPD and ISSR). According to Morales *et al.* (2012) and Samah and Valadez-Moctezuma (2014), the absence of the pulp and the presence of an edible pericarp and the small seeds are the most significant morphological differences between prickly pears and xoconostles. Another overview of our results was the distribution across the tree of the accessions included in *O. ficus-indica*, *O. albicarpa* and *O. megacantha* species. The accessions documented in each species were not well-defined into separate clades (Fig. 5), since they were dispersed throughout the tree, suggesting that these species are closely related (Labra *et al.*, 2003; Valadez-Moctezuma *et al.*, 2014a). PCoA was used to further confirm cluster analysis results. This analysis revealed that the first three coordinates explained 60.8 % of the total variability. The first two coordinates accounted for 50.8 % of the variability, of which the first component contributed half of the variability (26.2 %). The projection of accessions on the first two coordinates showed a grouping of accessions very similar to that found in the UPGMA analysis (Fig. 6).

From the above, it can be deduced that biochemical markers, like the seed proteins studied here, are consistent with molecular DNA markers. Similar conclusions were reported by Fufa *et al.* (2005) and El Rabey *et al.* (2014) who found a highly significant correlation between seed storage proteins with SSR and AFLP markers; these methods are believed to be the most efficient to estimate genetic diversity markers. For this reason, protein markers have been suggested as an additional tool for taxa identification and to support systematic studies.

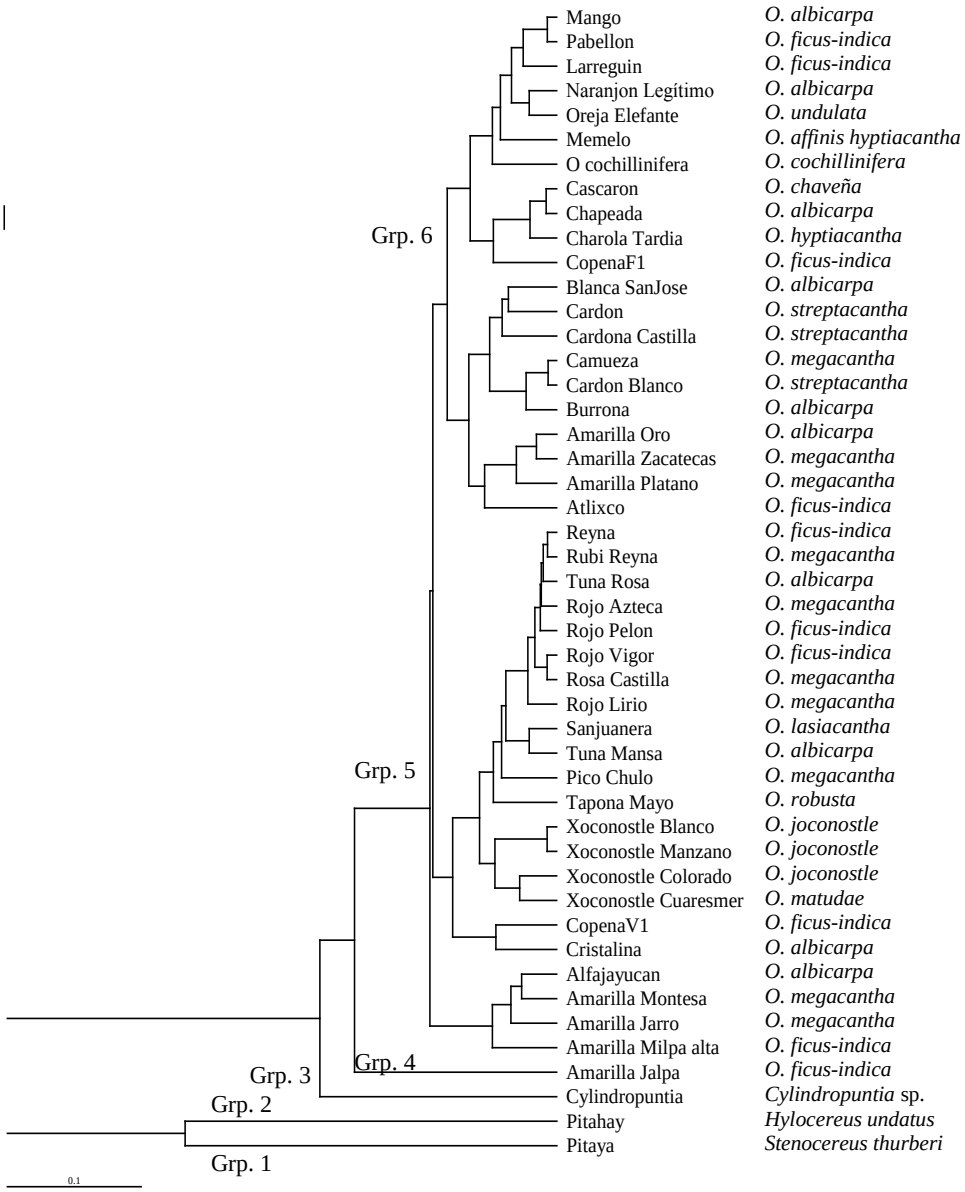


Fig. 5. The dendrogram based on the presence or absence of seed total and fractions proteins bands in 44 Opuntia accessions contrasted with three out-groups obtained from UPGMA analysis. Proteins were resolved under reducing conditions on SDS-PAGE continuous system.

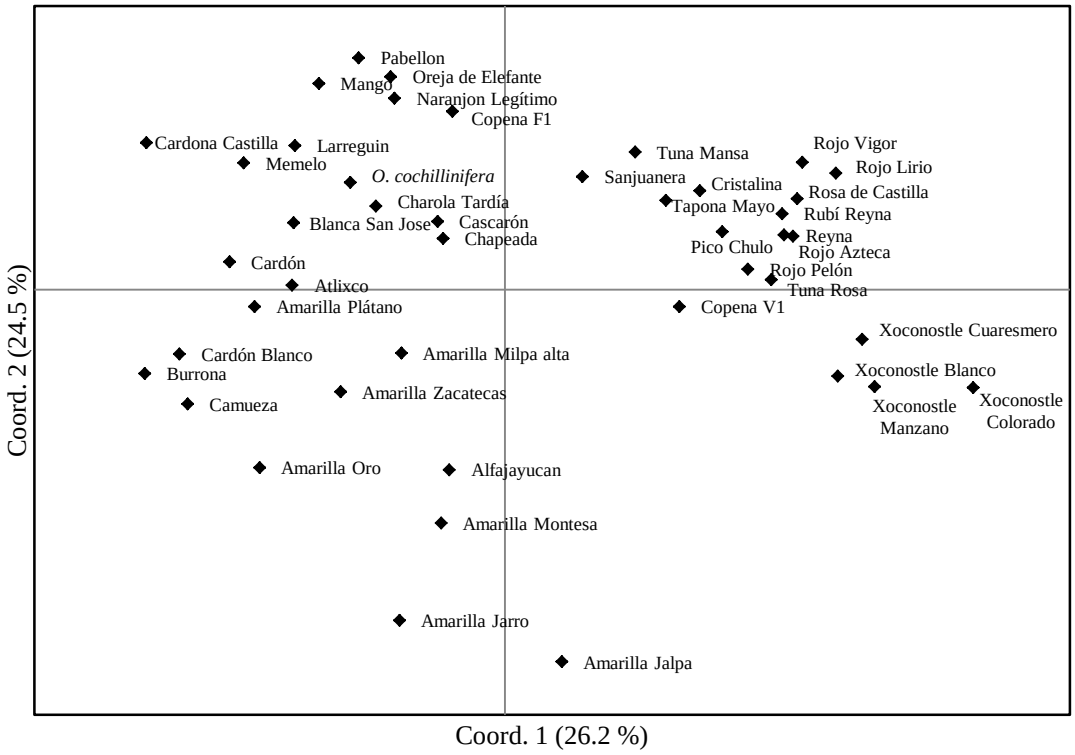


Fig. 6. Graph of distribution of 44 Opuntia accessions resulting by PCoA (first two factors, 50.7 % of the variance) from total proteins and SSPs profiles. Proteins were resolved under reducing conditions on SDS-PAGE continuous system.

CONCLUSIONS

In this research, we studied for the first time the relationships among Mexican Opuntia accessions using seed proteins as biochemical markers resolved by SDS-PAGE system under reducing conditions. The results obtained provide evidence of the broad variability of protein content in the seeds. The seeds contain about 4.3 mg/mL of total protein, with a variation ranging from 1.13 to 7.12 mg/mL. Similarly, the four protein fractions (SSPs) are all present in the Opuntia seeds with considerable amplitude; albumin and globulins are the most abundant protein fractions, while prolamins are present in small amounts. Contrary to the high variation in the relative contents of proteins, the composition of the STPs and SSPs, as showed by banding patterns resolved by the SDS-PAGE system, was shown to be similar among the different accessions studied. Moreover, the need to combine both STPs

and SSPs profiles data to differentiate all Opuntia accessions was demonstrated. However, the time of seed storing affected the amount of proteins, although not the protein profiles. The clustering of accessions by UPGMA and PCoA revealed no concordance with the current taxonomic status. However, the genotypes that produce acidic fruits (xoconostles) were clustered in a sister-no-separate group from prickly pears. Finally, protein profiles are solid markers that can be taken into account for genetic diversity studies and for the taxonomic revision of the genus *Opuntia*.

REFERENCES

Baranger, A., Aubert, G., Arnau, G., Lain, A.L., Deniot, G., Potier, J., Weinachter, C., Lejeune-Hnaut, I., Lallemand, J., and Burstin, J. 2004. Genetic diversity within *Pisum sativum* using protein and PCR-based markers. Theor. Appl. Genet. 108: 1309-1321.

Bellani, L., Giglioni, S., and Muccifora, S. 2013. Protein Characterization of Protein Bodies from Cotyledons of *Mucuna pruriens* (L.) DC. Plant Foods Hum. Nutr. 68: 47-51.

Carreras, M.E., Fuentes, E., and Merino, E.F. 1997. Seed protein patterns of nine species of Cactaceae. Bioch. Syst. Ecol. 25: 43-49.

Caruso, M., Currò, S., Las Casas, G., La Malfa, S., and Gentile, A. 2010. Microsatellite markers help to assess genetic diversity among *O. ficus-indica* cultivated genotypes and their relation with related species. Plant Syst. Evol. 290: 85-97.

Chavez, S. R. A., Gutierrez, U. J. A., and Serna, S. S. O. 2009. Phenolic Composition, Antioxidant Capacity and In Vitro Cancer Cell Cytotoxicity of Nine Prickly Pear (*Opuntia* spp.) juices. Plant Foods Hum. Nutr. 64: 146-152.

Chougui, N., Tamendjari, A., Hamidj, W., Hallal, S., Barras, A., Richard, T., and Larbat, R. 2013. Oil composition and characterization of phenolic compounds of *Opuntia ficus-indica* seeds. Food Chem. 139: 796-803.

El Rabey, H. A., Al-Malki, A. L., Abulnaj, K. O., Ebrahim, M. K., Kumosani, T., and Khan, J. A. 2014. Phylogeny of ten species of the genus *Hordeum* L. as revealed by AFLP markers and seed storage protein electrophoresis. Mol. Biol. Rep. 41: 365-372.

El-Guizani, T., Tlili, N., and Triki, S. 2012. Fruit characteristics and chemical proprieties of juice and seeds of three *Opuntia ficus-indica* cultivars. Afr. J. Biotechnol. 11: 1460-1464.

Fufa, H., Baenziger, P. S., Beecher, B. S., Dweikat, I. Graybosch, R. A., and Eskridge, K.M. 2005. Comparison of phenotypic and molecular marker-based classifications of hard red winter wheat cultivars. Euphytica 145, 133-146.

Galván, J.V., Valledor, L., Cerrillo, R.M.N., Pelegrín, E.G., and Jorrín-Novo, J.V. 2011. Studies of variability in Holm oak (*Quercus ilex* subsp. Ballota [Desf.] Samp.) through acorn protein profile analysis. J. Proteomics 74: 1244-1255.

Galvez, M.J., Castro, H.A., and Villamil, C.B. 2009. Antigenic patterns of seed proteins in Opuntioideae (Cactaceae). Biochem. Syst. Ecol. 37: 91-97.

Gazzola, D., Vincenzi, S., Gastaldon, L., Tolin, S., Pasini, G., and Curioni, A. 2014. The proteins of the grape (*Vitis vinifera* L.) seed endosperm: Fractionation and identification of the major components. Food Chem. 155: 132-139.

Goldring, J. P. D. 2012. Protein Quantification Methods to Determine Protein Concentration Prior to Electrophoresis. In: Biji, T. Kurien, Hal Scofield, R. (eds.), Protein Electrophoresis: Methods and Protocols. Meth. Mol. Biol. 869: 29-35.

Gupta, P.K., Varshney, R.K., Sharma, P.C., and Ramesh, B. 1999. Molecular markers and their applications in wheat breeding. Plant Breed. 118: 369-390.

Guzmán, U., Arias, S., and Dávila, P. 2003. Catálogo de cactáceas mexicanas. UNAM and CNCUB, D.F. México.

Hameed, A., Saddiqa, A., Nadeem, S., Iqbal, N., Atta, B. M., and Shah T. M. 2012. Genotypic variability and mutant identification in *Cicer arietinum* l. By seed storage protein profiling. Pak. J. Bot. 44: 1303-1310.

Hames, B.D. 1998. Gel electrophoresis of proteins. A practical approach. Third edition. Oxford University Press. UK.

Helsen, P., Verdyck, P., Tye, A., and Van Dongen, S. 2009. Low levels of genetic differentiation between *Opuntia echios* varieties on Santa Cruz (Galapagos). Plant Syst. Evol. 279: 1-10.

Higgins, T.J.V. 1984. Synthesis and regulation of major proteins in seeds. Ann. Rev. Plant Physiol. 35: 191-221.

Ibl, V., and Stoger, E. 2012. The formation, function and fate of protein storage compartments in seeds. Protoplasma 249: 379-392.

Karihaloo, J. L., Kaur, M., and Singh, S. 2002. Seed protein diversity in *Solanum melongena* L. and its wild and weedy relatives. Genet. Res. Crop Evol. 49: 533-539.

Kumar, A., Sharma, N., Panwar, P., and Gupta, A. K. 2012. Use of SSR, RAPD markers and protein profiles based analysis to differentiate *Eleusine coracana* genotypes differing in their protein content. Mol. Biol. Rep. 39: 4949-4960.

Labra, M., Grassi, F., Bardini, M., Imazio, S., Guiggi, A., Citterio, S., Banfi, E., and Sgorbati, S. 2003. Relationships in *Opuntia* Mill. Genus (Cactaceae) detected by molecular marker. Plant Sci. 165: 1129-1136.

Ladizinsky, G., and Hymowitz, T. 1979. Seed Protein Electrophoresis in Taxonomic and Evolutionary Studies. Theor. Appl. Genet. 54: 145-151.

Maquet, A., Zoro, I., Rocha, O., and Baudoin, J. 1996. Case studies on breeding systems and its consequences for germplasm conservation: I. Isozyme diversity in wild Lima beans populations in Central Costa Rica. Genet. Resour. Crop Evol. 43: 309-318.

Miernyk, J.A., and Hajduch, M. 2011. Review: Seed proteomics. J. Proteomics 74: 389-400.

Mirali, N., El-Khouri, S., and Rizq, F. 2007. Genetic diversity and relationships in some *Vicia* species as determined by SDS-PAGE of seed proteins. Biol. Plantarum 51: 660-666.

Morales, P., Ramírez-Moreno, E., Sanchez-Mata, M. C., Carvalho, A. M., and Ferreira, I. C. F. R. 2012. Nutritional and antioxidant properties of pulp and seeds of two xoconostle cultivars (*Opuntia joconostle* F.A.C. Weber ex Diguet and *Opuntia matudae* Scheinvar) of high consumption in Mexico. Food Res. Int. 46: 279-285.

Osborne, T.B. 1924. The vegetable proteins. Longmans. 2nd ed. New York.

Panozzo, J. F., Eagles, H. A., and Wootton, M. 2001. Changes in protein composition during grain development in wheat. Aust. J. Agric. Res. 52: 485-493.

Peakall, R., and Smouse, P.E. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28: 2537-2539.

Reyes-Agüero, J. A., Aguirre, J. R. and Valiente-Banuet, R. A. 2006. Reproductive biology of *Opuntia*: A review. J. Arid Envir. 64: 549-585.

Rubio, L. A., Pérez, A., Ruiz, R., Guzmán, M. A., Aranda-Olmedo, I., and Clemente, A. 2014. Characterization of pea (*Pisum sativum*) seed protein fractions. J. Sci. Food Agric. 94: 280-287.

Samah, S., and Valadez-Moctezuma, E. 2014. Morphological Seeds Descriptors for Characterize and Differentiate Genotypes of *Opuntia* (Cactaceae, Opuntioideae). Annual Research and Review in Biology 4: 3791-3809.

Santos, D. N. K. F., Silveira, D. R. D., Martin-Didonet G. C. C., and Brondani, C. 2013. Storage protein profile and amino acid content in wild rice *Oryza glumaepatula.* Pesq. Agropec. Bras. 48: 66-72.

Shewry, P.R., Johnathan A. Napier, J.A. and Tatham, A.S. 1995. Seed Storage Proteins: Structures and Biosynthesis. Plant Cell 7: 945-956.

Singh, N.P., and Matta, N.K. 2010. Levels of seed proteins in *Citrullus* and *Praecitrullus* accessions. Plant Syst. Evol. 290: 47-56.

Smith, J.A., and Desborough, S.L. 1987. The endosperm seed protein solin: biochemical characterization, introduction by ABA and species-specific subunitis. Theor. Appl. Genet. 74: 739-745.

Stoilova, T., Cholakova, N., and Markova, M. 2006. Variation in seed protein and isoenzyme patterns in *Cucurbita* cultivars. Biol. Plantarum 50: 450-452.

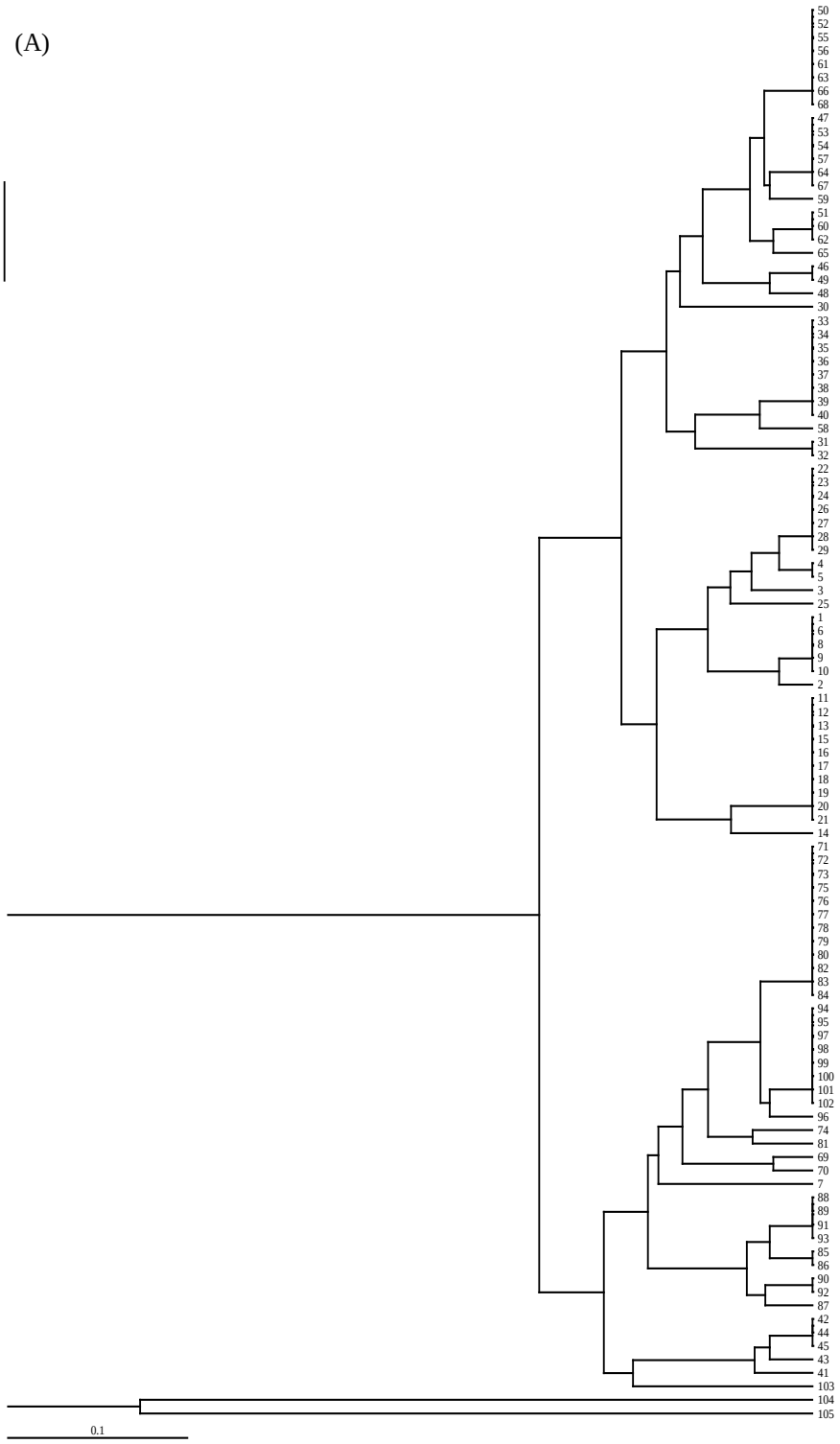
Syros, T., Yupsanis, T., and Economous, A. 2003. Fractionation and electrophoretic patterns of storage proteins of *Ebenus cretica*. A preliminary survey as a tool in taxonomy. Biol. Plantarum 46: 435-443.

Tlili, N., Bargougui, A., Elfalleh, W., Triki, S., and Nasri, N. 2011. Phenolic compounds, protein, lipid content and fatty acids compositions of cactus seeds. J. Med. Plant. Res. 5: 4519-4524.

Valadez-Moctezuma, E., Samah, S., and Luna-Paez, A. 2014a. Genetic diversity *of Opuntia* spp. varieties assessed by classical marker tools (RAPD and ISSR). Plant Syst. Evol. DOI: 10.1007/s00606-014-1112-y.

Valadez-Moctezuma, E., Ortiz-Vásquez, Q., and Samah, S. 2014b. Molecular based assessment of genetic diversity of xoconostle accessions (*Opuntia* spp.). Afr. J. Biotechnol. 13: 202-210.

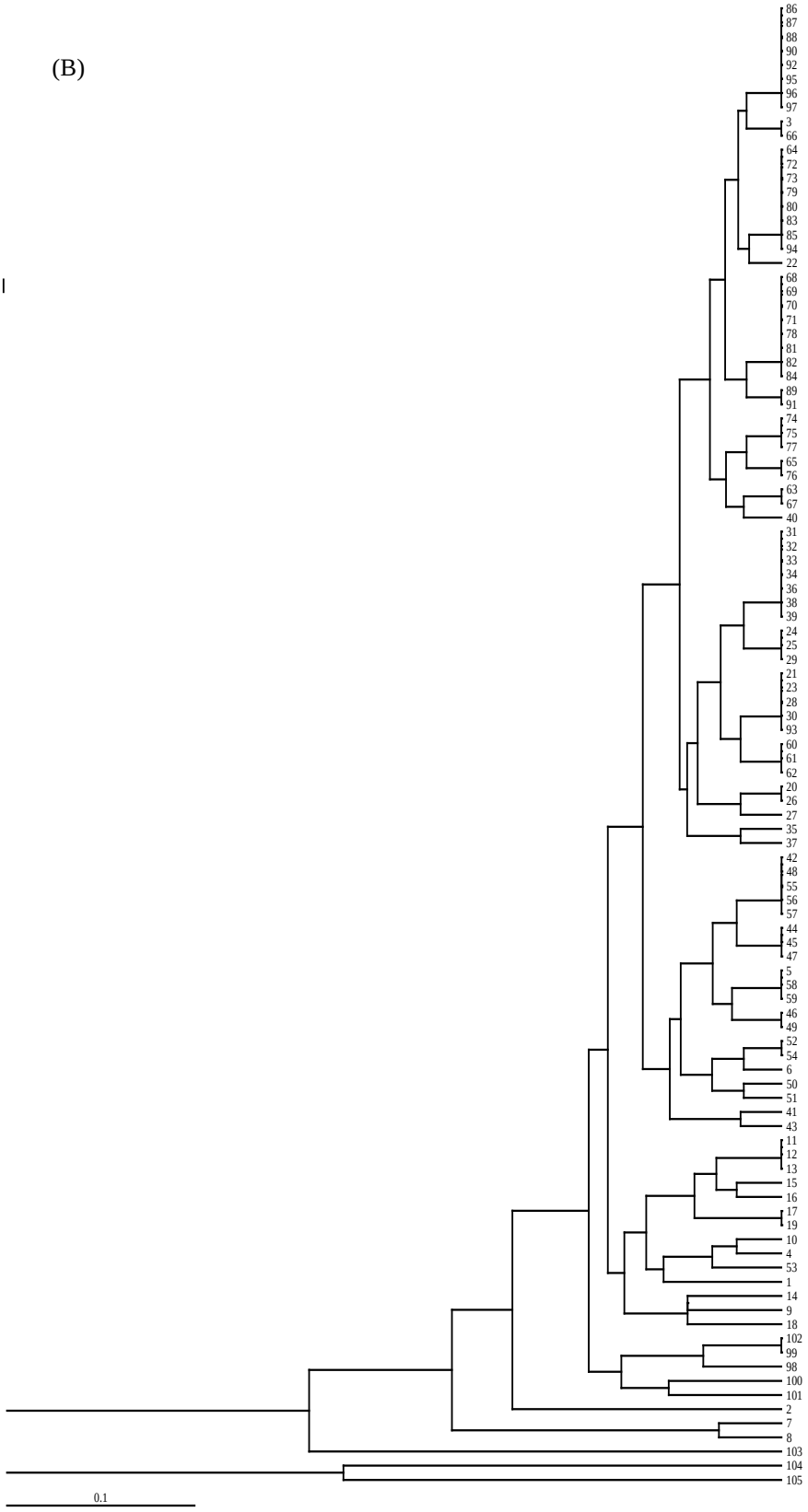
Wang, W., Alché, J.D., and Rodríguez-García, M.I. 2007. Characterization of olive seed storage proteins. Acta Physiol. Plant 29: 439-444.



Supplementary data S1. The dendrogram based on the presence or absence of seed total and fractions proteins bands in 102 Opuntia accessions contrasted with three out-groups obtained from UPGMA analysis. Proteins were resolved under reducing conditions on SDS-PAGE continuous system, (A) Total protein analysis, (B) Albumins analysis, (C) Globulins analysis, (D) Prolamins analysis, (E) Glutelins analysis.

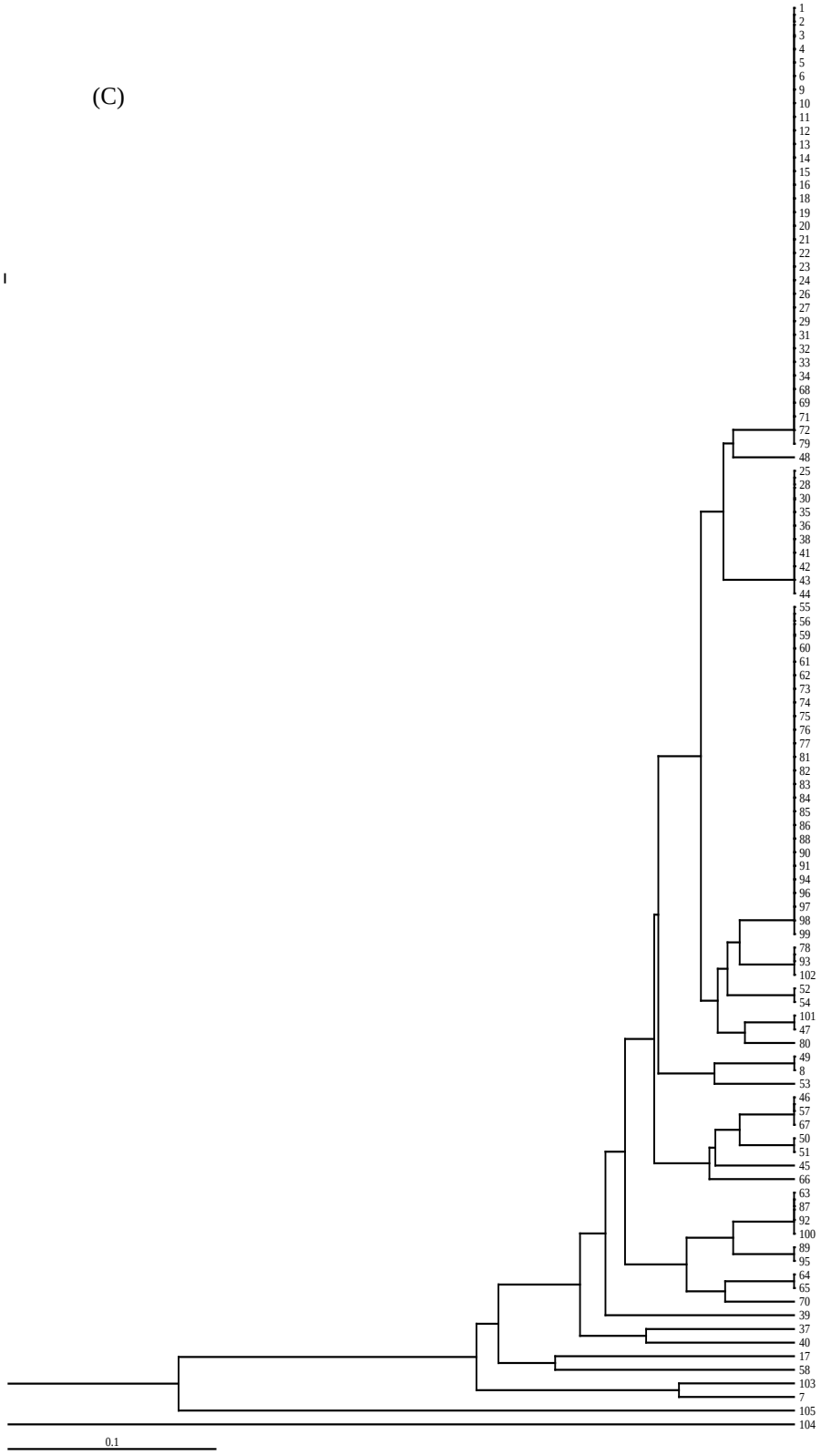
Supplementary data S1. Continued.

(B)



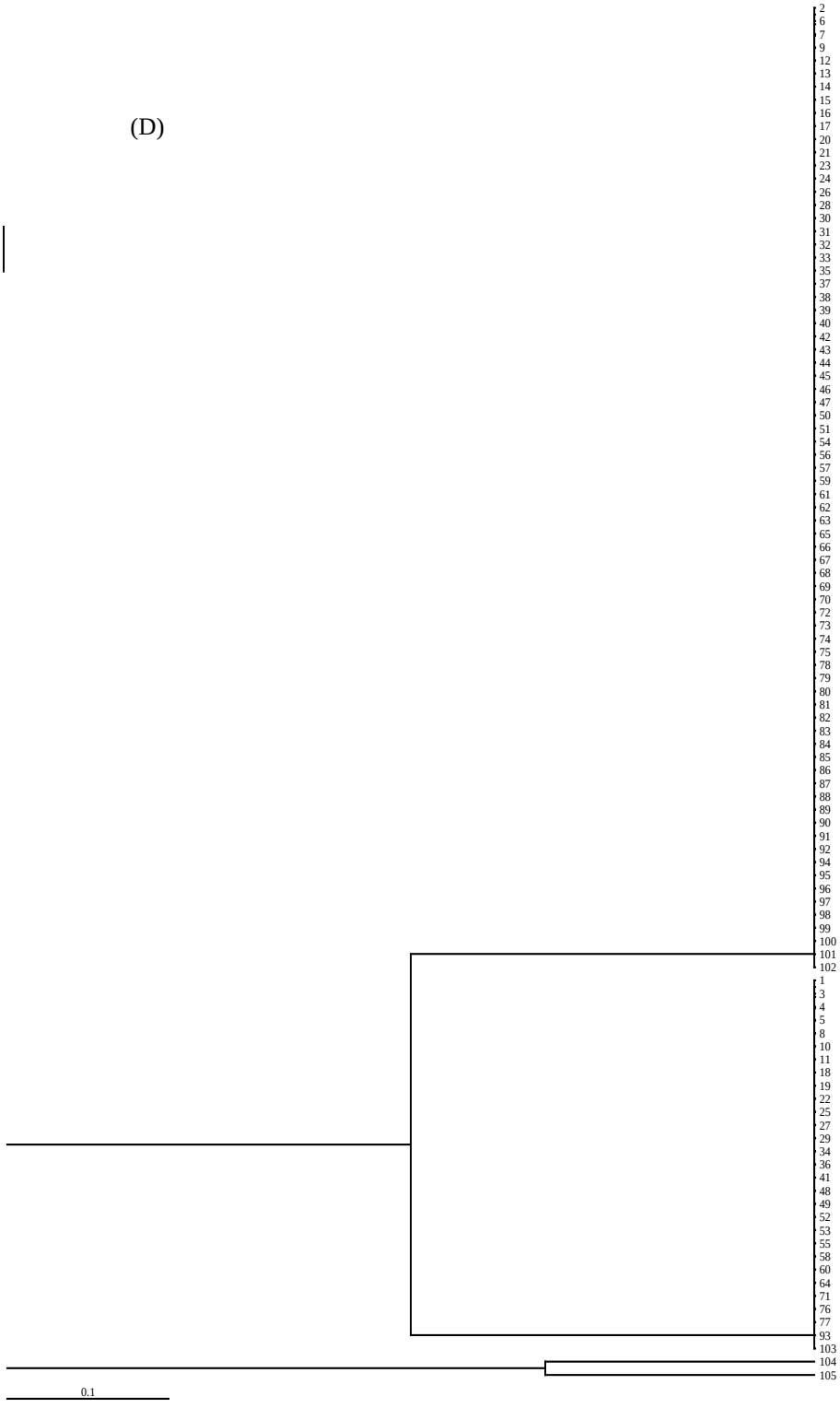
Supplementary data S1. Continued.

(C)



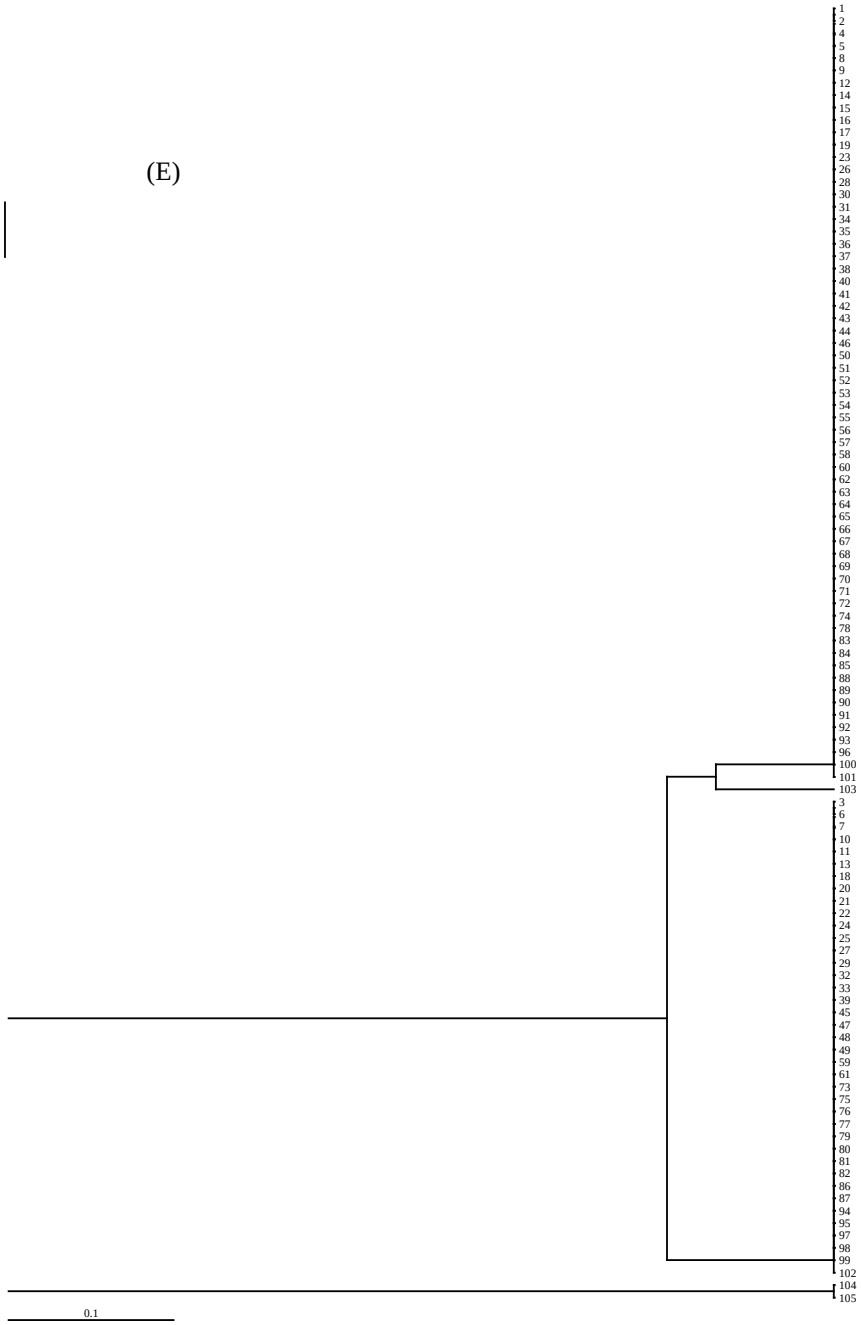
Supplementary data S1. Continued.

(D)



Supplementary data S1. Continued

(E)



Chapter 4

Part 1

DNA marker (Microsatellites)

GENETIC DIVERSITY, GENOTYPE DISCRIMINATION AND POPULATION STRUCTURE OF MEXICAN *OPUNTIA* SP., DETERMINED BY SSR MARKERS

Samir Samah, Claudia Verónica De Teodoro Pardo, Miguel Angel Serrato Cruz, Ernestina Valadez-Moctezuma

Article submitted: Plant Molecular Biology Reporter. PMBR-D-14-00501

GENETIC DIVERSITY, GENOTYPES DISCRIMINATION AND POPULATION STRUCTURE OF MEXICAN *OPUNTIA* SP., DETERMINED BY SSR MARKERS

ABSTRACT

The *Opuntia* genus, an important horticultural crop in Mexico, is essentially a fruit crop with two variants: sweet ("tunas") or acid ("xoconostles") fruits; it is also a source of vegetables "nopalitos" or fodder for livestock, among other uses. However, few studies are available on the genetic structure of Mexican Opuntia, and genetic differences between the two types of fruits are unknown. Opuntia genotype identification is still mainly based on morphological characters, and homonymies and synonymies are frequent. In this study, the genetic diversity of extensive Mexican Opuntia germplasm (88 accessions) was revealed, using 13 SSR markers in an attempt to explore the genetic relationships among them. A total of 159 alleles were detected ranging from 7 to 23 per locus with an average of 12.2. The SSR markers generated unique fingerprints for each Opuntia accession confirming its usefulness for genetic analysis of Opuntia germplasm. The accessions' grouping was defined by several complementary clustering methods and the moderate incongruences between the different methods did not influence the overall clustering. UPGMA and STRUCTURE analyses grouped the accessions studied into 7 and 5 clusters, respectively; confirming, thus, the incorrect delimitation of species in this genus. Median-Joining and Neighbor Net simulations classified all genotypes into a complex network; both linear and reticular ties between Mexican Opuntia genotypes were revealed. The genetic distance shows the importance of Mexican accessions for conservation and use in breeding programs.

Keywords Nopal, xoconostles, differentiation, Microsatellite markers, genetic structure

INTRODUCTION

Crop germplasm diversity can be exploited by numerous techniques such as analyses of morphological traits, total seed protein, isozymes, cytological and biochemical characteristics, and numerous types of DNA molecular markers (Helsen *et al.,* 2009; Caruso *et al.,* 2010). Morphology based characterization has some constraints in the accurate identification of the accessions, such as limited number of traits, and trait expression is subjected to strong environmental influence. Molecular markers can serve as powerful and reliable tools for screening variations and for studying genetic diversity and evolutionary relationships, determining genetic relatedness (Chae *et al.*, 2014). Likewise, molecular markers are not affected by physiology or the environment. Accurate classification and characterization of accessions based on molecular markers is a prerequisite for a successful breeding program and for evolutionary studies among different crops species, Opuntia genotypes included.

Opuntia sensu stricto (nopal, prickly pear) is the largest genus in Opuntioideae, the second largest subfamiliy of Cactaceae (Anderson, 2001). There are 150 (Stuppy, 2002) to 180 recognized species within the genus (Hunt 2006), of which, 66-83 are reported in Mexico. This genus is suggested to have originated as recently as 5-6 million years ago (Arakaki *et al.,* 2011). Members of *Opuntia* s.s. are cultivated worldwide as fruit and vegetable crops and are increasingly used as forage and fodder for livestock in arid areas of the world, mainly parts of Brazil and Mexico, and to a lesser extent in western Asia, northern and southern Africa. In Mexico, *Opuntia* species have been cultivated for at least 14 000 years (Casas and Barbera, 2002); Mesoamerican civilizations used the cladodes (cactus pads) as a vegetable or fodder and primarily as a seasonal fruit crop; with sweet (prickly pear or “tunas” in Spanish) or acid fruits (“xoconostle”). They are also considered

an alternative natural medicine due to their antihypoglycemic, oxidative stress and cancer prevention effects (Chavez-Santoscoy *et al.,* 2009; Morales *et al.,* 2012). Nevertheless, the misclassification in *Opuntia* species is reported. The continuous morphological variation, the synonyms, and the inadequate and limited morphological descriptors for cultivar discrimination are the most difficult obstacles to achieve a stable classification (Labra *et al.,* 2003; Caruso *et al.,* 2010; Valadez-Moctezuma *et al.,* 2014a; b). The difficulties in morphological interpretation have led to publication of a large number of binomials, many of which are synonyms, homonyms or false attributions (Gibson and Nobel, 1986). Moreover, species limits are still poorly understood as a result of the high frequency of polyploid taxa. *Opuntia* species have been recorded as diploid to enneaploid (1n = 1x = 11), with 60.4 % of reported counts showing different levels of polyploidy and 13.4 % representing taxa with both diploid and polyploidy cytotypes (Majure *et al.,* 2012).

In recent years, molecular markers based on polymerase chain reaction (PCR) amplification of specific genomic sequences have been proposed as a direct and effective tool to estimate inter-generic and inter-specific relationships among different taxa of Opuntia (Caruso *et al.,* 2010; Valadez-Moctezuma *et al.,* 2014a). Wang *et al.* (1998) showed that RAPD analyses can be useful in cultivar differentiation and identification of *Opuntia* duplicate accessions. Labra *et al.* (2003) documented the usefulness of molecular markers like AFLP and cpSSR in *Opuntia* species characterization and to study the relationships among different species. Similarly, the RAPD technique was used to assess the genetic diversity of *O ficus-indica* genotypes in Tunisia (Zhoghlami *et al.,* 2007; Bendhifi *et al.,* 2013). Recently, SSR markers have been developed to study the genetic diversity in the Galapagos Islands *Opuntia* (Helsen *et al.,* 2007), and they were applied to estimate the genetic relationship in some species from the Mediterranean region and

Mexico (Caruso *et al.,* 2010). Few molecular based studies have been carried out in Mexico, region of origin and diversification of many *Opuntia* species with agronomic and economic interests. Molecular characterization of some genotypes from central Mexico was reported by Mondragon-Jacobo *et al.* (2003). Luna-Paez *et al.* (2007) characterized 22 *Opuntia* sp. cultivars with RAPD and ISSR markers. García-Zambrano *et al.* (2009) differentiated 12 accessions using the AFLP technique. Recently, Valadez-Moctezuma *et al.* (2014 a; b) applied RAPD and ISSR markers to differentiate 52 cultivated accessions of Opuntia form Mexico revealing a great genetic diversity and pointing to the taxonomic misclassification. Thus, genetic relationships and population structure of Opuntia germplasm are still scarcely studied.

Microsatellites, also known as simple sequence repeats (SSRs), are noncoding, repetitive DNA regions consisting of tandem repeated small motifs (1-6 bp); they are present throughout the genome of an individual, both in coding and non-coding regions (Gao *et al.,* 2005). In recent years, microsatellites have been demostrated to have many important biological functions (e.g. the regulation of chromatin organization, DNA metabolic processes, gene activity, and RNA structure), and have therefore emerged as the third major class of genetic variations, alongside single nucleotide polymorphisms (Ekué *et al.,* 2009). In comparison to other molecular markers (e.g. RFLP, RAPD, ISSR, AFLP, SRAP and SNP), microsatellites are the most informative molecular markers due to their reliability and abundant multi-allelic forms. They exhibit higher mutation rates than the rest of the genome (Gao *et al.,* 2005). They can be easily analyzed by PCR-based methods, including fluorescent automated genotyping and multiplexing. They can be isolated from an SSR-enriched genomic library using oligonucleotide probes complementary to the repeated sequences, or from sequences in the public domain. Therefore, SSR markers have

been the preferential choice for various applications, such as variety identification, genetic diversity evaluation, phylogenetic relationship analysis, genetic map construction, linkage/association mapping of gene/QTL, marker-assisted selection and comparative mapping (Shi *et al.*, 2014).

Few studies have been conducted to identify/characterize genomic/genic microsatellites and to develop markers in *Opuntia* species. Helsen *et al.* (2007) developed 16 SSR markers from *O. echios* through a probe hybridization (containing a repeated motif) against genomic/cDNA clones and sequencing; and Erre *et al.* (2011) obtained, with the same method, ten SSR markers from *O. ficus-indica*. Moreover, two expressed sequence tag sequences were reported by Caruso *et al.* (2010). Therefore, the pattern of microsatellite distribution has remained ambiguous, and the development/utilization of SSR marker has still been limited in Opuntia, which is mostly because the lack of genome sequences and the initial development of SSR primers is rather costly and time consuming. Alternatively, in cases where primers have been developed for related taxa, cross-amplification can be attempted, and there are now numerous studies reporting successful interspecific transferability of SSR primers (Ducarme *et al.*, 2008; Ekué *et al.*, 2009).

As has been mentioned above, Opuntia accessions are several levels of polyploidy, so SSR analysis presents some drawbacks and it is difficult to determine exact heterozygosity. For lower polyploids like tetraploids, it may sometimes be possible to determine allelic configurations, for example, by using specific techniques like microsatellite DNA allele counting-peak ratios (Esselink *et al.*, 2004) or other mathematical methods (Bruvo *et al.*, 2004). In high polyploids, however, allele dosage of SSRs cannot easily be determined and alleles are not easily attributable to potentially diploidized loci. Nevertheless, polyploids

may offer an important redeeming feature for population genetic analysis using microsatellites if their interpretation as dominant markers can be tolerated (Pfeiffer *et al.*, 2011; Moscoe and Emshwiller, 2014), and SSR profiles are typically much more diverse, a fact that allows for fingerprinting with fewer markers (Andreakis *et al.*, 2009).

A crucial role attributed to the solving ability of SSR technique in Opuntia will be the clarification of the relationship between "xoconotles" and "tunas" since the separation between these two groups has beed controversial. Morales *et al.* (2010) indicated these two types are well differenced morphologically; while the molecular results (RAPD and ISSR) (Luna-Paez *et al.*, 2007; Valadez-Moctezuma *et al.*, 2014b) and storage proteins (Samah *et al.*, in press) suggests close relationship between these two groups.

In the present research, we used a set of previously described SSR markers to define the genetic relationships between 88 previously untested (except for four accessions) Mexican accessions, belonging to 17 species. Two types of SSR markers: gSSRs and EST-SSRs were used to (1) determine the genetic relationships between 88 accessions/species of Opuntia native to Mexico (2) elucidate the evolution type and population structure and (3) separate genotypes with acid fruits ("xoconostles") from sweets ones ("tunas").

MATERIALS AND METHODS

Plant material and microsatellite analysis

A total of 91 accessions/species were analyzed in the present study. Several accessions are cultivated in different growing regions, but little is known about their ancestries and levels of genetic diversity. Samples of 88 Opuntia accessions were obtained from the

germplasm banks of Crucen-UACh (Zacatecas) and Nopalera-UACh (Texcoco) in Mexico. Three cacti (one sample of *Cylindropuntia* sp., one pitahaya (*Hylocereus undatus*) and one pitaya (*Stenocereus thurberi*) were included as outgroups. Some of the *Opuntia* accessions are classified in delimited species but others have no taxonomic assignation (Table 1).

Table 1. List of Opuntia accessions tested with 13 SSRs and their corresponding species.

No.	Accession	Taxonomic classification	No. of total alleles	Average no. of alleles	No. of unique alleles
1	Alfajayucan	*O. albicarpa* Scheinvar	29	2.23	0
2	Alteña Blanco	*Opuntia* sp.	43	3.31	0
3	Alteña Rojo	*Opuntia* sp.	44	3.38	0
4	Amarilla Miquihuana	*O. lasiacantha* Pfeiffer	37	2.85	0
5	Amarilla Montesa	*O. megacantha* Salm-Dyck.	40	3.08	0
6	Amarilla San Elías	*Opuntia* sp.	33	2.54	0
7	Amarilla Zacatecas	*O. megacantha* Salm-Dyck.	35	2.69	0
8	Amarillo Aguado	*Opuntia* sp.	33	2.54	1
9	Amarillo Plátano	*O. megacantha* Salm-Dyck.	31	2.38	0
10	Atlixco	*O. ficus-indica* (L.) Mill.	36	2.77	0
11	Blanca de Castilla	*Opuntia* sp.	39	3.00	0
12	Blanca San José	*O. albicarpa* Scheinvar	39	3.00	0
13	Blanco de Atlacomulco	*Opuntia* sp.	35	2.69	0
14	Blanco Huexotla	*Opuntia* sp.	34	2.62	0
15	Bola de Masa	*O. megacantha* Salm-Dyck.	30	2.31	0
16	Burrona	*O. albicarpa* Scheinvar	32	2.46	0
17	Cacalote	*O. cochinera* Griffits	31	2.38	0
18	Camuezo	*O. megacantha* Salm-Dyck.	30	2.31	0
19	Cardon Blanco	*O. streptacantha* Lem.	29	2.23	0
20	Cardona de Castilla	*O. streptacantha* Lem.	33	2.54	0
21	Cascarón	*O. chaveña*	30	2.31	0
22	Chapeada	*O. albicarpa* Scheinvar	33	2.54	0
23	Charola Tardia	*O. hyptiacantha* Lem.	33	2.54	0
24	Chicle	*O. ficus-indica* (L.) Mill.	30	2.31	0
25	Col. Barr.	*Opuntia* sp.	27	2.08	0
26	Color de Rosa	*O. albicarpa* Scheinvar	33	2.54	0
27	Colorada	*Opuntia* sp.	33	2.54	0
28	Copena F1	*O. ficus-indica* (L.) Mill.	34	2.62	0
29	Copena V1	*O. ficus-indica* (L.) Mill.	30	2.31	0
30	Copena Z1	*O. albicarpa* Scheinvar	29	2.23	0
31	Fafayuca	*O. albicarpa* Scheinvar	30	2.31	0
32	Gavia	*O. albicarpa* Scheinvar	33	2.54	1
33	Green Guanajuato	*Opuntia* sp.	25	1.92	0
34	Huatusco	*Opuntia* sp.	39	3.00	1
35	Jarilla Grande	*O. megacantha* Salm-Dyck.	41	3.15	1
36	Laltus	*Opuntia* sp.	36	2.77	0
37	Larreguin	*O. ficus-indica* (L.) Mill.	28	2.15	0
38	Liso Amarillo	*Opuntia* sp.	42	3.23	1
39	Liso Forrajero	*O. ficus-indica* (L.) Mill.	34	2.62	0
40	Mango	*O. albicarpa* Scheinvar	36	2.77	0
41	Memelo	*O. affinis hyptiacantha*	32	2.46	0
42	Milpa Alta	*O. ficus-indica* (L.) Mill.	34	2.62	0
43	Montesa	*Opuntia* sp.	33	2.54	0
44	Morada	*O. megacantha* Salm-Dyck.	32	2.46	0
45	Naranjón Legítimo	*O. albicarpa* Scheinvar	35	2.69	0
46	Naranjona	*O. megacantha* Salm-Dyck.	35	2.69	0
47	O. cochillinifera	*O. cochillinifera*	20	1.54	0
48	Oreja de Elefante	*O. undulata* Griffiths *(lindheimeri)*	34	2.62	1
49	Pabellón	*O. ficus-indica* (L.) Mill.	35	2.69	0

50	Pachon	*O. hyptiacantha* Lem.	35	2.69	0
51	Pico Chulo	*O. megacantha* Salm-Dyck.	39	3.00	1
52	Pico de Oro	*Opuntia* sp.	40	3.08	0
53	Platano	*Opuntia* sp.	38	2.92	0
54	Princesa	*Opuntia* sp.	32	2.46	0
55	Reyna	*O. albicarpa* Scheinvar	33	2.54	0
56	Reyna Crucen	*Opuntia* sp.	30	2.31	0
57	Roja Azteca	*O. megacantha* Salm-Dyck.	32	2.46	0
58	Roja San Martín	*O. megacantha* Salm-Dyck	31	2.38	0
59	Rojo Lirio	*O. megacantha* Salm-Dyck.	35	2.69	0
60	Rojo Liso	*Opuntia* sp.	33	2.54	0
61	Rojo Pelón	*O. ficus-indica* (L.) Mill.	17	1.31	0
62	Rojo Vigor	*O. ficus-indica* (L.) Mill.	36	2.77	0
63	Rosa de Castilla	*O. megacantha* Salm-Dyck.	31	2.38	0
64	Rubí Reyna	*O. megacantha* Salm-Dyck.	30	2.31	0
65	San Juan	*Opuntia* sp.	37	2.85	0
66	Sangre de Toro	*O. megacantha* Salm-Dyck.	36	2.77	0
67	Sanjuanera	*O. lasiacantha* Pfeiffer	39	3.00	0
68	Solferino	*O. ficus-indica* (L.) Mill.	28	2.15	0
69	Tapón Aguanoso	*O. robusta* H.L. Wendland	30	2.31	0
70	Tapon Rojo	*O. robusta* H.L. Wendland	25	1.92	0
71	Tapona de Mayo	*O. robusta* H.L. Wendland	31	2.38	0
72	Tobarito	*Opuntia* sp.	36	2.77	0
73	Toluca	*Opuntia* sp.	33	2.54	0
74	Torreoja	*O. megacantha* Salm-Dyck.	25	1.92	0
75	Trompa de Cochino	*O. streptacantha* Lem.	33	2.54	0
76	Tuna Mansa	*O. albicarpa* Scheinvar	30	2.31	0
77	Tuna Rosa	*O. albicarpa* Scheinvar	30	2.31	0
78	Tuna Sandia	*O. streptacantha* Lem.	28	2.15	0
79	Villanueva	*O. albicarpa* Scheinvar	29	2.23	0
80	X. Colorada	*O. joconostle* F.A.C. Weber	27	2.08	0
81	X. Cuaresmero	*O. matudae* Scheinvar	27	2.08	0
82	X. Blanco	*O. joconostle* F.A.C. Weber	34	2.62	1
83	X. Chivo	*Opuntia* sp.	26	2.00	0
84	X. Manzano	*O. joconostle* F.A.C. Weber	29	2.23	0
85	X. Rojo Dulce	*Opuntia* sp.	31	2.38	1
86	O. matudae	*O. matudae* Scheinvar	19	1.46	0
87	O. leucotricha	*O. leucotricha*	24	1.85	0
88	O. rzedowskii	*O. rzedowskii*	29	2.23	0
89	Cylindropuntia	*Cylindropuntia* sp.	18	1.38	1
90	Pitahaya	*Hylocereus undatus*	10	0.77	0
91	Pitaya	*Stenocereus thurberi*	8	0.62	0

Total genomic DNA was extracted using the CTAB method (Luna-Paez *et al.*, 2007). The DNA quantification was estimated by spectrophotometry (ND-1000 Thermo scientific, USA) and the DNA quality was determined in 1 % agarose gels. Microsatellite primers were obtained from Helsen *et al.* (2007), Caruso *et al.* (2010) and Erre *et al.* (2011) (Table 2). A total of 18 SSR markers were tested. The PCR was carried out in a final volume of 25 μL containing nuclease-free water, 500 mM dNTPs, 1 x Taq buffer, 25 mM $MgCl_2$, 20 pmol primers, 1.5 U Taq DNA polymerase (Promega) and 100 ng templates DNA. The thermo-cycling conditions (MaxyGene Thermel Cycler, Applied Biosystem, USA) were: one 4 min cycle at 94 °C, 35 cycles [94 ºC for 30 s; an annealing step for 45 s; 72 ºC for 2

min] and one final extension cycle at 72 ºC for 10 min. Primers information and annealing temperature were described in Table 2. Amplifications were carried out separately for each primer pair. The PCR products were mixed with 5 µl loading buffer (98 % formamide, 0.05 % bromophenol blue, 0.05 % xylene cyanol, and 10 mM NaOH), denatured at 95 °C for 5 min and then separated on 8 % polyacrylamide gel for 1.5 h. Markers 25 pb DNA step ladder (Promega, USA) and 100 pb DNA ladder (Promega, USA) were used to estimate molecular weight of the amplified bands. The voltage applied was 220 volts for 1.5 h in 1 x TBE buffer (89 mM Tris-borate, 2 mM EDTA pH 8) using Dual MGV-216-33 vertical electrophoresis gel system (CBS, USA). Silver nitrate solution (AgNO3) 0.2 % was used to stain DNA fragments. To confirm the presence of bands, electrophoresis was replicated twice.

Table 2. List of SSR markers (eleven gSSRs and two EST-SSRs) used in the analysis, numbers of alleles, PIC and MI values obtained from each primer pair.

Primer	Repeat motif	AT (ºC)	No. of alleles	No. of unique alleles	Average PIC value	MI
Opuntia5[1]	$(TAC)_5$	65	07	0	0.290	2.028
Opuntia11[1]	$(CT)_{13}$ TT $(CT)_2$	64	14	1	0.253	3.539
Opuntia13[1]	$(AG)_{12}$	65	13	1	0.248	2.975
Ops.9[2]	$(TGA)_9$	67	12	3	0.231	2.777
Ops.24[2]	$(CT)_{24}$	65	14	1	0.277	3.884
Opufic01[3]	$(CT)_{16}$	68	07	0	0.332	2.324
Opufic03[3]	$(TG)_{12}$	67	07	1	0.203	1.419
Opufic04[3]	$(TG)_{12}$	68	11	1	0.217	2.386
Opufic09[3]	$(AAG)_{16}$	67	23	7	0.157	3.604
Opufic10[3]	$(GT)_{12}$	66	18	3	0.183	3.298
Opufic15[3]	$(CT)_{22}$	65	12	0	0.275	3.294
Opufic16[3]	$(CT)_{17}$	65	12	0	0.257	3.088
Opufic17[3]	$(AG)_{13}$	66	09	0	0.281	2.531

[1] Helsen *et al.* (2007)
[2] Caruso *et al*. (2010)
[3] Erre *et al*. (2011)
AT annealing temperature
PIC polymorphism information content
MI marker index (No. of alleles x PIC)
Number of unique alleles refers to rare alleles (alleles with a frequency lower than 5 %)

Data analysis

The quantification of genetic diversity for organisms with polyploidy genomes can be difficult because these genomes, unlike diploids, transmit more than two alleles per individual and locus, potentially including multiple copies of a given allele. Although some methods for measuring the allele dosage of each individual have been successfully applied, these calculations are often unreliable in high order polyploids (i.e. hexaploids and higher ploidy levels, Obbard *et al.*, 2006). In addition, segregation of alleles in allopolyploids typically follow a disomic inheritance, in which the alleles at a given locus on homeologous chromosomes segregate independently and it is not often clear which alleles are associated with which of the duplicate loci, i.e. isoloci. As a consequence of such independent segregation, a typical phenomenon in allopolyploids is the detection of fixed heterozygosity, i.e. homozygote individuals are not observed due to the presence of at least two allele variants associated with different isoloci. Due to these deviations from diploid meiotic behavior, standard summary statistics routinely used for diploids, such as expected heterozygosity, cannot be used to quantify genetic diversity in allopolyploids (Bruvo *et al.*, 2004; Obbard *et al.*, 2006); and partial heterozygosity makes it impossible to score genotypes exactly (Dufresne *et al.*, 2014). The most often cited technique to counter this problem is to score alleles as presence/absence data (Helsen *et al.*, 2009a; Caruso *et al.*, 2010).

The bands of each SSR marker were recorded as qualitative characters for their presence (1) or absence (0) in the 91 accessions and a binary matrix was then created. Number of alleles and number of unique alleles (alleles with a frequency lower than 5 %) were also calculates for each locus. PowerMarker 3.25 (Liu and Muse, 2005) was used to determine the average polymorphism information content (PIC value) for each primer pair and Marker index (MI) was calculated as: MI = No. of alleles x PIC.

Genetic similarity between accessions was calculated based on the Dice coefficient using the SIMQUAL subprogram of NTSYSpc 2.21o (Rohlf, 2002). Cluster analysis was performed using the unweighted pair group arithmetic mean method (UPGMA) in the SAHN subprogram of NTSYSpc. Principal coordinate analysis (PCoA) based on the genetic similarity matrix was performed using GenAlEx 6.501 (Peakall and Smouse, 2012). PCoA was used to visualize genetic distances among accessions and species. Genotypic variations across clusters resolved in UPGMA tree and across species were assessed by means the non-parametric analysis of molecular variance (AMOVA) using GenAlEx 6.501.

In addition, a model-based Bayesian analysis was performed to evaluate the genetic structuring of the 88 accessions of *Opuntia* (outgroups excluded) and to identify admixed individuals using STRUCTURE 2.3.4 (Pritchard *et al.,* 2000). Structure analysis of the data (dominant markers), admixture and independent allele frequencies model was used in estimating of the proper subgroups, without using species assignation of individuals as priors. The number of subgroups (K) was set from 2 to 17 with 20 runs independent simulations per k performed separately. For each run, 70,000 Markov chain Monte Carlo repetitions were carried out after a burn-in periods of 30,000 iterations. The mean estimate across of the 20 runs of the log posterior probability of the data for a given k, Pr(X|k), called L(k), was plotted for each k category on a graph to determine the k value of the population as the value of k for which the distribution of L(k) plateaus or continues to increase, but much more slowly. Because this point is known to be difficult to determine, we also used Δk, an ad hoc quantity proposed by Evanno *et al.* (2005) related to the second-order rates of change of the likelihood function with respect to k that is supposed to

show a clear Δk at a true value of k. The accessions were subdivided into different subgroups using a membership probability threshold of 0.80. Those varieties with less than 0.80 membership probabilities were identified. Also, microsatellite data were used to perform a network analysis using the software SplitsTree version 4.11.3 (Huson and Bryant, 2006). Specifically a Median-Joining and NeighborNet networks were computed using the default settings.

RESULTS AND DISCUSSION

In the present study, the genetic diversity of 88 accessions/species of Mexican *Opuntia* was analyzed based on 18 SSR markers; 16 genomic markers previously designed by Helsen *et al.* (2007) and Erre *et al.* (2011), and two expressed sequence tag-derived SSR markers (EST-SSR) reported by Caruso *et al.* (2010). Out of 18 SSR markers, 13 markers generated reproducible, clear, distinct and polymorphic amplification products in all loci. Meanwhile, one marker (Opuntia12) showed no amplification, and the remaining four markers (Opuntia3, Opuntia9, Opufic13 and Opufic14) were excluded from further analysis because they showed a large proportion of missing data among accessions (>5 %). This failure might be due to different mutation events at the level of the microsatellite and/or of the primer binding sites. The other 13 primers successfully amplified fragments in all of the genotypes, showing a high degree of cross-transferability among the analyzed species. Of the 16 primers designed from Galapagos' *O. echios* by Helsen *et al.* (2007), only six loci generated amplification in 62 Opuntia genotypes studied by Caruso *et al.* (2010); from these 6 markers, three were discarded in the present study due to lack of amplification. This was probably due to the high divergence between the Mexican species of Opuntia and those from the Galapagos Islands. The Galapagos Opuntia ranges from 12 m high tree-like plants with hard dimorphic spines and long pendant branches, to 2 m high

scrubby plants without branches and with soft bristly spines (Helsen *et al.* 2009a). On the other hand, Mexican species included here reach a maximum height of 5 to 5.2 m for the species *O. hyptiacantha* and *O. cochinera*, respectively (Reyes-Agüero *et al.* 2009). Furthermore, two markers designed specifically from *O. ficus-indica* (Table 2) were discarded due to the lack of amplification in at least 5 % of the accessions, possibly due to the higher diversity in the Mexican germplasm compared to that found in the Mediterranean basin (Caruso *et al.* 2010). In addition, the two SSR markers excluded interrupted the motifs sequences (Erre *et al.* 2011).

A total of 159 alleles were detected, ranging from 7 (Opuntia5, Opufic01 and Opufic03) to 23 (Opufic09), with mean allelic wealth of 12.2 alleles per marker. Out of these, 18 were regarded as rare alleles with a low frequency in the total accessions, representing 11.3 % of the total alleles. The SSR loci exhibited a high level of polymorphism (100 %). The average PIC and MI values are a reflection of allelic diversity and frequency among the genotypes analyzed. In the present study, the average PIC values ranged from 0.157 (Opufic09) to 0.332 (Opufic01), with an average of 0.246 for all the analyzed loci. The MI values for SSR loci ranged from 1.419 (Opufic03) to 3.884 (Ops.24), with an average of 2.857. Similar results were reported by Caruso *et al.* (2010), who recorded an average PIC of 0.191 (0.11 to 0.25); these authors recorded a value of the rare alleles of 25 % (of a total of 133 alleles), and an average of alleles per locus of 16.63 (ranging from 7 to 33) were slightly different. Therefore, the higher average number of alleles per SSR marker reported in the previous study may be primarily attributed to the electrophoretic system. However, genotypes may also influence the number of alleles detected at each SSR locus, and also the selection of locus SSR sets (five SSR markers studied here were previously used by Caruso *et al.* (2010)). The results showed that the

Opuntia germplasm was rich in genetic diversity but contained a small percentage of rare alleles. Unlike the results reported by Labra *et al.* (2003), where chloroplast SSRs were not able to define genetic differences between closely related species. Considering both MI values and allele numbers, about half of the markers (6 loci) could be considered informative because they showed MI values higher than 3 and generated more than 12 alleles per locus (Table 2).

The SSR markers used in this study consist of both genomic SSR (g-SSR) and expressed sequence tag-derived SSR (EST-SSR). A total of two EST-SSRs and 11 g-SSRs have been used to estimate the genetic diversity in Opuntia germplasm. Generally, EST-SSRs display lower levels of polymorphism than genomic SSRs, since expressed sequences are more likely to be more conserved than non-coding regions (Shepherd *et al.* 2002; Ekué *et al*. 2009). However, in this study, the levels of polymorphisms of EST-SSRs are higher than those observed for g-SSRs. For example, the average number of alleles per EST-SSR locus (13) is higher than that of alleles per g-SSR locus (12.09). The average PIC value and the MI value were higher in the loci EST-SSRs (0.254 and 0.331) than in the g-SSRs loci (0.245 and 2.771). The EST-SSR Ops.24 marker displayed the high level of MI. Thus, the findings in this study have demonstrated that highly polymorphic SSR markers could be derived from expressed sequences. Similar results are reported in other plant species by Martin *et al.* (2010) and Zhang *et al.* (2012). However, both genomic and genic microsatellite markers can provide an accurate indication of genetic diversity in several plant species.

It is noteworthy that the same 13 SSRs markers were analyzed in three cacti included as outgroups: Cylindropuntia, Pitahaya and Pitaya. Ten of these SSRs markers (except

Opuntia5, Opuntia13 and Opufic10) generated amplicons in Cylindropuntia amplifying a total of 18 bands. However, only four SSR markers (Opuntia11, Opufic15, Opufic16 and Opufic17) were cross-transferred to Pitaya and Pitahaya, generating a total of bands of 10 and 8, respectively. These results were as expected since Cylindropuntia is phylogenetically as close to Opuntia as the other two outgroups. Opuntias are often divided into cylindropuntias and platyopuntias (Gibson and Nobel 1986). Cylindropuntias are shrubby species with cylindrical stems (or joints). Platyopuntias have flattened stems called cladode. Cross-species amplification and utility of molecular markers depend on the conservation of priming sites within flanking sequences, and on the maintenance of arrays long enough to promote polymorphisms (Fitz Simmons *et al.* 1995). Generally, such attempts are the most successful for closely related species, although even then the amplifying loci are not always polymorphic. Several studies have demonstrated the utility of using primer pairs designed from one species with other species of the same genus (Mottura *et al.* 2005; Takayama *et al.* 2008) and even with species of other genera (Barreneche *et al.* 2004; Ekué *et al.* 2009).

Genotype differentiation, genetic diversity and cluster analysis

The 13 SSR loci amplified different numbers of alleles ranging from 17 to 44, with an average of 32.4 per accession. The lowest number (17) was recorded in the cultivar "Rojo Pelón", while the highest (44) was registered in the accession "Alteña Rojo". Each individual accession had a unique microsatellite pattern, suggesting that none of them are a vegetative clone or a duplicate of any other sampled Opuntia. Although a relatively low average PIC value was found in this study (0.246), the diversity of SSR markers here proved to be a reliable tool for accession discrimination. The level of polymorphism

detected by SSRs in this study was higher than that detected by Caruso *et al.* (2010). This indicates a great genetic diversity of Mexican Opuntia germplasm.

The estimated Dice coefficient among accessions ranged from 0.033 (the most distanced accessions were "Laltus" and "*O. leucotricha*") to 0.956 (the most similar accessions were "Blanco de Atlacomulco" and "Blanco Huexotla"), and these relationships were supported by cluster analysis (Fig. 1), indicating high genetic diversity. The relationship among 88 individuals of Opuntia was further determined by UPGMA analysis; the tree cactus outgroups were also included. Cluster analysis showed a good fit with distance matrix as reflected by the cophenetic correlation coefficient (r=0.72). This analysis clearly discriminated all accessions according to their SSR profiles. These results contradict that was reported by Caruso *et al.* (2010), who didn't differentiate some Opuntia accessions using 8 SSRs markers. This discrepancy could be due to the number of markers tested and the backround of the germplasm studied. Three pairs of accessions shared the minimal genetic distances; "Blanco Huexotla" and "Banco Atlacomulco", "Naranjón Legítimo" and "Naranjona", the latter accession pairs are very similar and the small genetic differences (probably somatic) were conserved through the vegetative propagation. Likewise, the two accessions "Alteña Blanco" and "Alteña Rojo" are very similar genetically and differentiate only from fruit color, white in the first accession and red in the second. The phenotypical differences might have been determined by relatively few genes which could not be reflected in the molecular results.

The 88 accessions were grouped into nine clusters (A-I), from which clusters H and I correspond to the accessions included as outgroups (Fig. 1). The number of genotypes in each cluster was different: 12, 9, 18, 20, 13, 14 and 2 for groups A to G, respectively (Fig.

1). The G cluster was formed by the species *O. cochillenifera* and the variety "Rojo Pelón" at the similarity coefficient of 0.28, two accessions with the minimum amounts of alleles detected.

The grouping of accessions was not in accordance with the current taxonomic classification of species, since defined clusters contain accessions of the different species. The multivariate PCoA analysis was used to further confirm this result (Fig. 2). PCoA revealed that the first three coordinates explained 24.69 % of the total variability. The first two coordinates accounted for a similar percentage of variability: 9.33 % and 8.08 %, respectively. As can be seen in Fig. 2, the distribution of genotypes was not according to the current taxonomic assignation of accessions. Similar results were found by Helsen *et al.* (2009a), who studied the genetic structure of Opuntia populations in the Galapagos Islands using SSRs markers. Overall, this provides little evidence for genetic structuring within the dataset. However, three accession representatives of the *O.robusta* species were grouped together in the D cluster at a similarity coefficient of 0.55. Similar results were reported by Valadez-Moctezuma *et al.* (2014a). *O. robusta* H.L.W. is large, polyploid, fleshy-fruited taxa from central and northern Mexico. *O.robusta* is cultivated throughout Mexico for its large edible fruits (Anderson, 2001) or for ornamental purposes (Griffith, 2004); it is a shrub with height from 0.61 to 2 m and width from 1 to 3.7 m, and the branches are born from the base (hence the name in Spanish "robusta") (Reyes-Agüero *et al.* 2009).

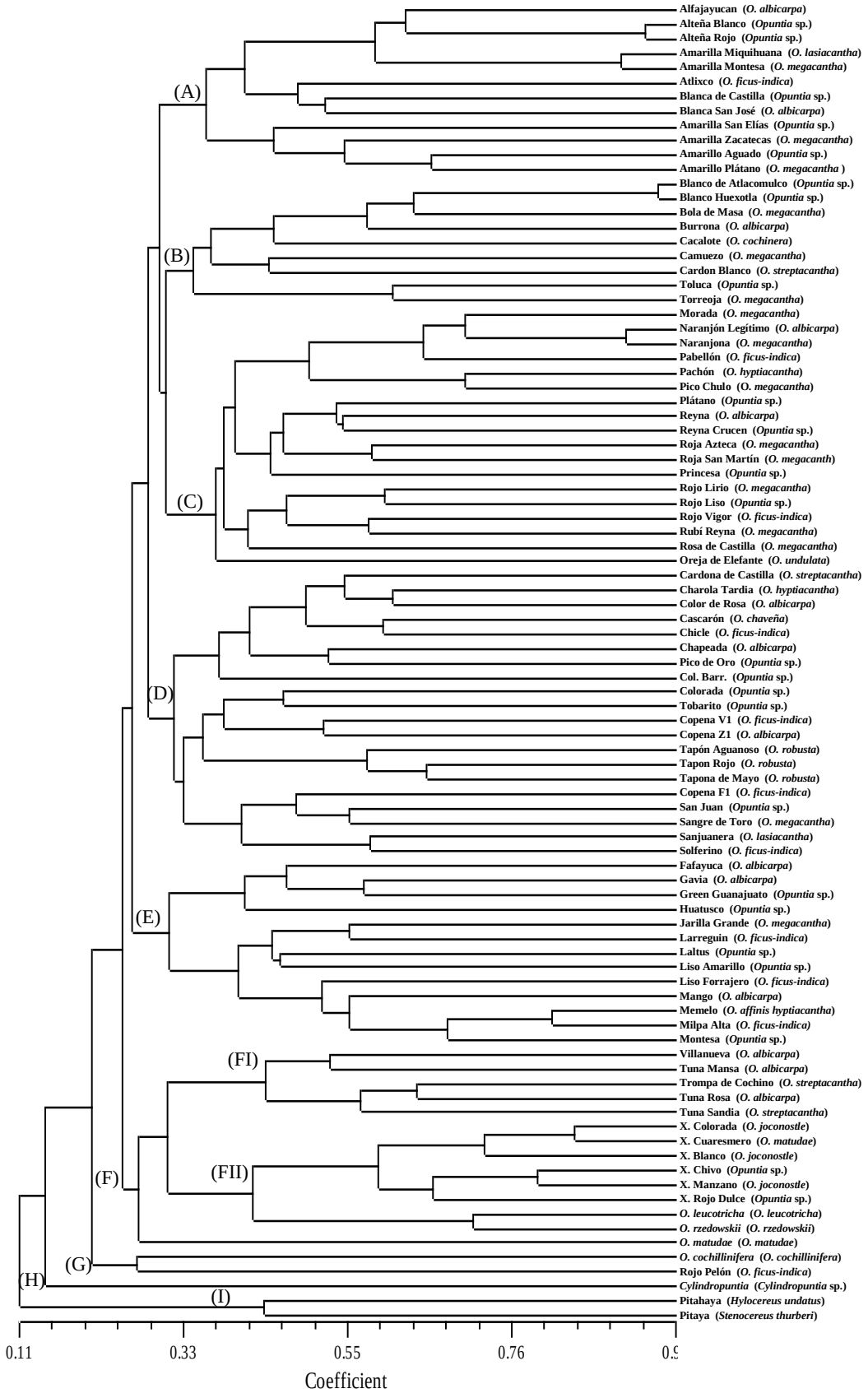


Fig. 1. Dendrogram of 88 Opuntia accessions obtained from SSR markers based on Dice coefficient using UPGMA as clustering method with 1,000 bootstraps. Cylindropuntia, Pitahaya and Pitaya were included as out-groups.

Meanwhile, the accession "Oreja de Elefante" was situated as outgroup of cluster C at a similarity coefficient of 0.4. This accession is well-known for its peculiar shaped cladodes from which comes the Spanish name "Oreja de Elefante" or elephant ear; also, it is characterized by the absence of spines, reason why it is a variety of extreme importance to livestock feed, specifically in Brazil. This accession has been classified as *O. undulata* due to the presence of cladodes with wavy margin; while Reyes-Agüero et al. (2009) classified this accession as *O. lindheimeri*. The F cluster (14 accessions) was divided into two sub-clusters at a similarity coefficient of 0.3; sub-cluster FI consists of five accessions that produce sweet fruit ("tunas") and sub-cluster FII is made up of accessions that produce acid fruits ("xoconostles") with the exception of *O. rzedowskii*; *O. matudae*, a wild species producing acid fruits, was located as out-cluster of the F cluster. According to Morales *et al.* (2012) and Samah and Valadez-Moctezuma (2014), the absence of pulp and the presence of an edible pericarp and small seeds are the most significant morphological differences between prickly pears and "xoconostles".

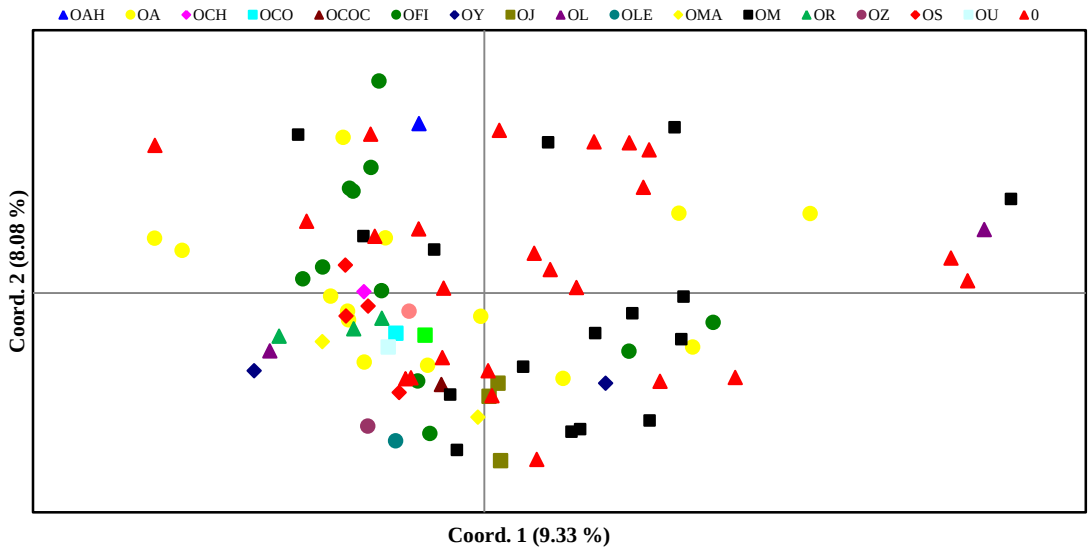


Fig. 2 Principal coordinate analysis of the 88 accession of Opuntia based on the genetic distance obtained by SSR analysis. OAH: *O. affinis hyptiacantha*, OA: *O. albicarpa* Scheinvar, OCH: *O. chaveña*, OCO: *O. cochillinifera*, OCOC: *O. cochinera* Griffits, OFI: *O. ficus-indica* (L.) Mill., OY: *O. hyptiacantha* Lem., OJ: *O. joconostle* F.A.C. Weber, OL: *O. lasiacantha* Pfeiffer, OLE: *O. leucotricha*, OMA: *O. matudae* Scheinvar, OM: *O. megacantha* Salm-Dyck, OR: *O. robusta* H.L. Wendland, OZ: *O. rzedowkii*, OS: *O. streptacantha* Lem., OU: O. *undulata* Griffiths, 0: *Opuntia* sp.

In this study, 41 accessions (47 % of studied Opuntias) are located in the three species with higher agronomic and economic importance in Mexico (14, 11 and 16 accessions, classified as *O. albicarpa*, *O. ficus-indica* and *O. megacantha*, respectively). The distribution of these accessions was over a dendrogram without any tendency to group according to taxonomic assignation. The accessions documented in each species were not well-defined into separate clades (Fig. 1), since they were dispersed throughout the tree; similar results were reported by related studies (Labra *et al.* 2003; Valadez-Moctezuma *et al.* 2014b; Samah and Valadez-Moctezuma, 2014), suggesting that these species are closely related, and many studies have suggested the revision of the classification of the *Opuntia* genus (Helsen *et al.*, 2009a; Caruso *et al.*, 2010; Valadez-Moctezuma *et al.*, 2014a). However, the increased differentiation reported in recent molecular studies (Labra *et al.*, 2003; Griffith, 2004; Helsen *et al.*, 2009b; Caruso *et al.*, 2010; Valadez-Moctezuma *et al.*, 2014a; b) combined with the putative plasticity of morphological characters used to distinguish species (Gibson and Nobel, 1986), raises questions as to the correctness of the current taxonomic classification. Since overlapping and continuous morphological characters are used to discriminate taxa, taxonomic classification is problematic and is currently based primarily on locality (Helsen et al. 2009b). It is noteworthy that many of the genotypes considered in this study have not yet been assigned taxonomically (Table 1); the location of an accession in a species is arbitrary due mainly to the lack of solid descriptors. This is probably related to the high level of phenotypic plasticity and ploidy levels, and also due to the morphological diversity of these accessions. These genotypes had several end uses, as fruits, vegetables and/or forage (Samah and Valadez-Moctezuma 2014).

Some accessions were reported as synonyms (Caruso et al. 2010): Fafajuca (syn. Blanca de San José or Blanca de Castilla), Naranjona (syn. Pico Chulo) and Reyna (syn. Alfajayucan). Nevertheless, the UPGMA analysis presented here showed that these accessions have different genetic profiles and are not considered synonyms. The accessions named "Blanca de San José" and "Blanca de Castilla" were grouped in the same cluster A at a similarity coefficient of 0.5. While the accession "Fafajuca" was located in cluster E, the two accessions "Naranjona" and "Pico Chulo" were grouped into cluster C at a similarity coefficient of 0.5. In turn, the two accessions "Reyna" and "Alfajayucan", reported as synonyms (Caruso et al. 2010), were grouped into different clusters, A and C, respectively (Fig. 1).

The genetic difference between the groups formed by the UPGMA analysis was estimated by the molecular analysis of variance (AMOVA). This was due to the mismatch between the results obtained in our study and the assigned genotypes in their current species (see above). AMOVA attributed 84 % of variance to the genotypes (within groups) and 16 % between groups (Table 3). For comparison, the same analysis (AMOVA) was conducted among species represented by more than one accession; the results revealed an occurrence of the genetic variance of 95 % between accessions and only 5 % of the variability among species (Table 3). Labra et al. (2003) found that the AMOVA analysis attributed 98.4 % of variance to the individuals within populations and 1.6 between *O. ficus-indica* and *O. megacantha* species. Moreover, the degree of genetic differentiation among the different species represented with more than one accession was quantified with the PCoA analysis. The results showed that species *O. joconostle* and *O. matudae* were distanced from the remaining species on axis 1 (Fig. 3). These two species are characterized by producing acid fruits. Also, *O. robusta* was separated from the rest of

species on axis 2; this species, as mentioned above, is characterized by its robust plants. *O. ficus-indica*, *O. albicarpa*, *O. megacantha*, *O. streptacantha*, *O. lasiacantha* and *O. hyptiacantha* were distributed in the center of the graph without any clear separation and showed a common genomic constitution. These results confirmed the findings obtained by the UPGMA analysis and support the idea about the need to revise the limits between *Opuntia* species.

Table 3. Analysis of molecular variance of Opuntia germplasm using SSR markers.

Source of variation	df	SS	MS	Est. Var.	Value (%)	*P*-value
Cluster groups						
Among groups	6	386.9	64.5	3.62	16	0.0001
Within groups	81	1603.9	19.8	19.8	84	
Species assignation						
Among species	8	225.3	28.2	1.13	5	0.002
Within species	48	1039.9	21.7	21.7	95	

df, degree of freedom; SS, sum of square; MS, mean square; Est. Var., estimated variation.

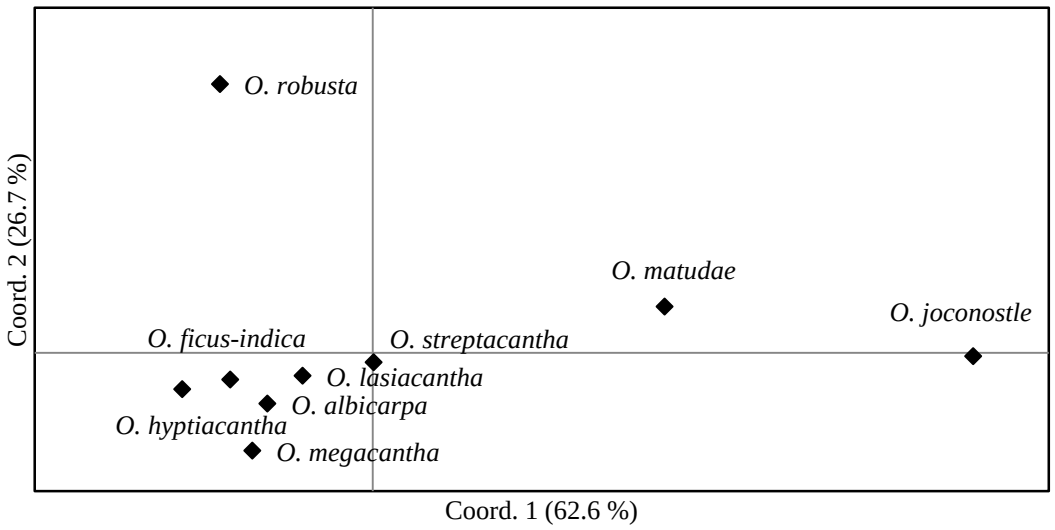


Fig. 3. Principal coordinate analysis of the nine *Opuntia* species based on SSR analysis

Assignment of individual accessions using STRUCTURE

Because only 24.69 % of the variation in the present study was included in the first three coordinate components, the Opuntia germplasm was also analyzed using an alternative model-based method implemented in the software STRUCTURE. Applications of model-based clustering methods and Bayesian assignment include demonstrating the presence of population structure, identifying distinct genetic populations, assigning individuals to populations, and identifying admixed individuals (Pritchard *et al.*, 2000). Bayesian clustering yielded estimated Ln probabilities for K=2 to K= 17 that ranged from -7878 to -7054. LnP(D) values decreased with increasing K values, but did not show an abrupt change for the true K, since the differences in log [P(X|K)] for K = 5, K = 6, and K = 7 were too small to make any decisions. Thus, Evanno's correction method (Evanno et al. 2005) was applied. Evanno's approach indicated the presence of five groups, since ΔK =5 was larger than all other ΔK values (data not shown), where each group has a unique set of allele frequencies. This method is clearly a simplification of the observed data; however, it can be used to compare with other methods of clustering and to test models of association analysis that would account for genetic associations arising from structure presence (Montilla-Bascón *et al.*, 2013).

The number of genotypes in each group was varied; 47, 5, 9, 7 and 20 accessions in groups G1 to G5, respectively (Fig. 4). Certain congruence between STRUCTURE grouping and UPGMA clusters was detected. In fact, the group G1 contains accessions of clusters B, 7 genotypes of cluster A, the subcluster FI, and cluster D except the accession "Pico de Oro". The group G2 is composed of 5 accessions out of the 12 accession from cluster A. The group G3 consists of the most accessions from cluster E. The group G4 coincides with subcluster FII except for the *O. leucotricha* species. The group G5 contains

genotypes from clusters G and C. Unlike the UPGMA grouping, where the "xoconostles" accessions were grouped in cluster F together with other "tunas", Bayesian model clustered the "xoconostles" accessions in an isolated group (G4), except for *O. matudae* and *O. leucotricha*, species that also produce "xoconostle" fruit types and were assigned to the group G1 with a percentage of membership of 85 % and 60 %, respectively. It is noteworthy that the species *O. rzedowskii* (wild) was grouped with "xoconostles" in both analyses; we cannot make a conclusion from this fact as it can be attributed to incidents of mislabeling during collection.

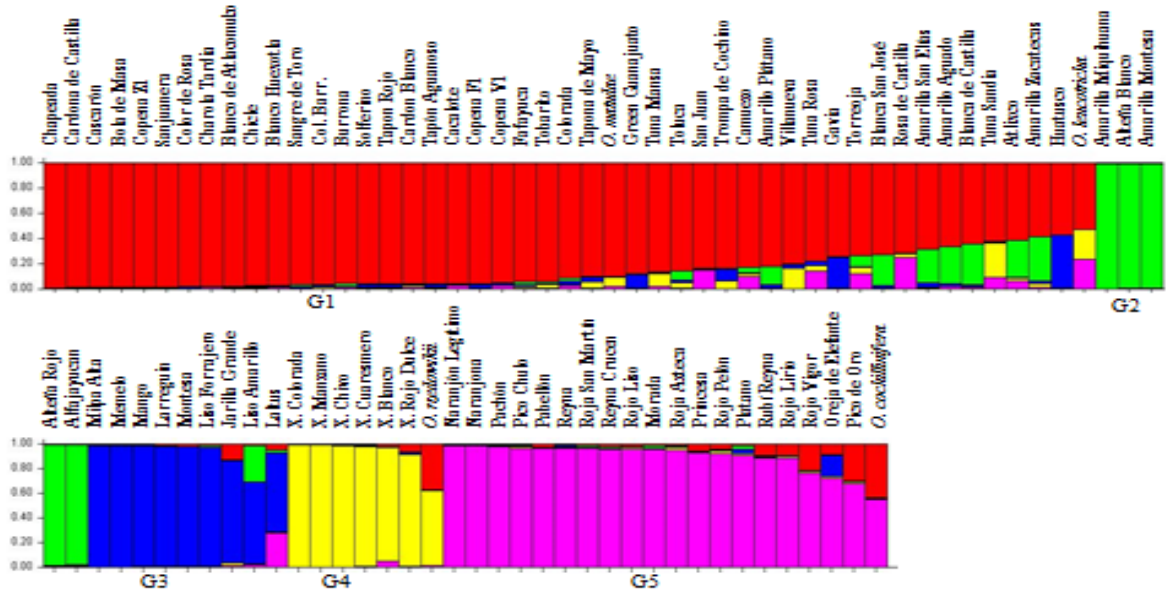


Fig. 4. STRUCTURE bar plot assigning the 88 accessions/species of *Opuntia* five groups (K=5).

An assignment of individual accessions in STRUCTURE at K = 5 (Fig. 4) shows admixture (membership coefficient less than 80 %) in the genetic background of 21 accessions (24 % from the total accessions): 14/47, 0/4, 2/9, 1/7, 4/20 in the groups G1 to G5, respectively. Particularly, the accessions with lower coefficient of membership decreased the congruence between STRUCTURE and UPGMA analyses; for example, the

accessions “Pico de Oro”, *O. cochillinifera*, *O. leucotricha* and “Rosa de Castilla” had membership percentages to their respective groups lower than 70 %. The different degrees of admixture among the analyzed genotypes might have been caused by natural hybridization (Caruso *et al.,* 2010) or probably due to shared ancestry.

Once again, accession clustering was not according to their taxonomic classification, as different accessions assigned in distinct species were commonly confused in the same group. This phenomenon reflected the complex genetic backgrounds of Mexican Opuntia germplasm. The congruence of patterns obtained with Bayesian and multivariate analyses suggests that the estimates of these admixture proportions are reasonably reliable. Because the relative impact of human activity on the distribution and population structure of crops cannot be concluded easily from genetic data (Hunt *et al.,* 2011), we cannot exclude that human intervention may have influenced the population structure of Opuntia.

Networks analysis

To overcome the limits of the UPGMA tree, where not all the branches were fully supported and the cophenetic correlation coefficient was moderate (r=0.72), we performed a network analysis based on the Median-Joining and NeighborNet networks implemented in SplitsTree version 4.11.3 (Huson and Bryant, 2006).

The clustering obtained by applying the NeighborNet networks method resulted in a similar UPGMA clustering (Fig. 5). However, the cluster B that resulted from the UPGMA analysis (Fig. 1) was divided into two subclusters in the NeighborNet analysis (Fig. 5). While cluster D in the UPGMA analysis was divided into three subclusters (Cluster D1,

D2 and D3). However, the NeighborNet analysis separated the representatives of *O. robusta* into an isolated group (cluster D3) at contrast of the UPGMA analysis.

The grouping obtained by the Median-Joining networks method resulted in a reticulate distribution with the presence of some linear formations (Fig. 6). These results agree with the findings of Caruso *et al.* (2010) who revealed a reticular evolution in Opuntia. The location of the genotypes in this analysis was similar to that resulting from the STRUCTURE analysis. The members of groups G2 and G4 agreed in both analyses, while the constituents of the groups G3 and G5 were partially similar in both analyses. The G1 accessions resulted from the STRUCTURE analysis were distributed throughout the Median Joining tree, presenting a reticular formation.

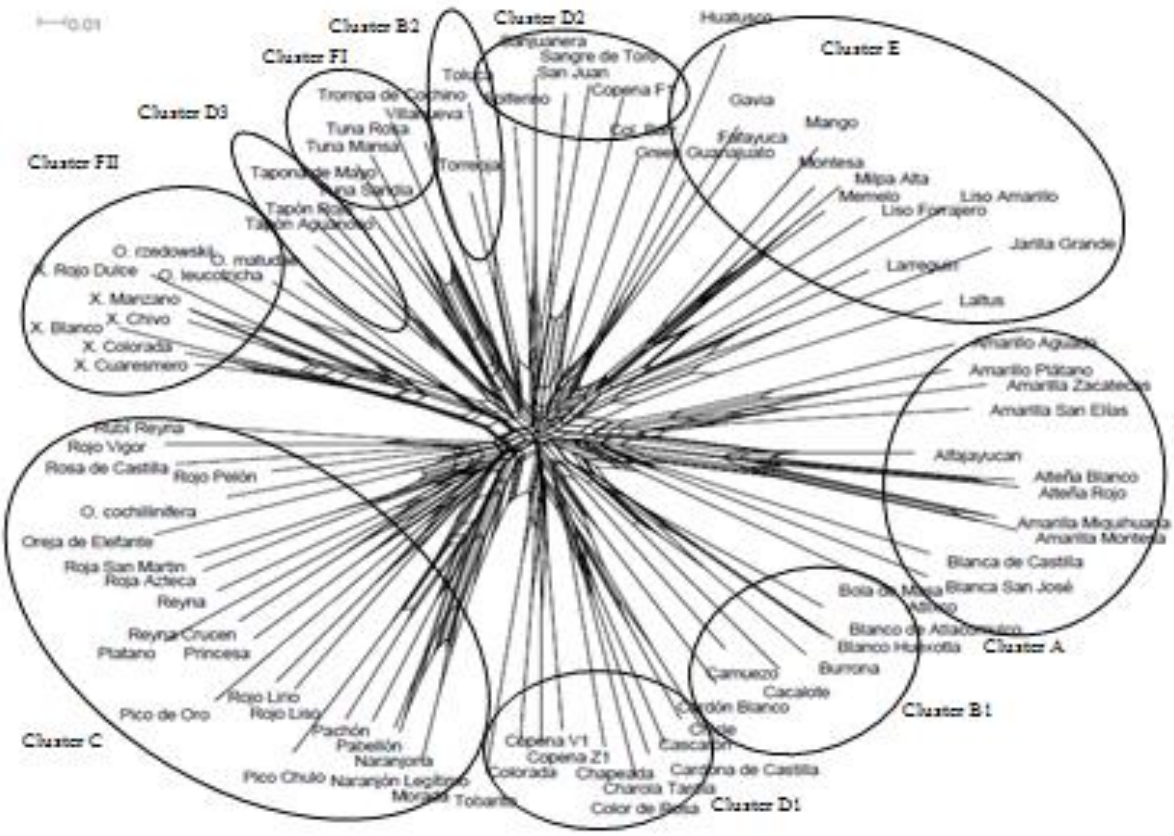


Fig. 5 NeighborNet tree of SSR data obtained from 88 Opuntia genotypes

The accessions in the groups G5 and G4 showed linear ties; the latter group contained all "xoconostle" accessions and was distinguished by grouping the *O. matudae* accession, which had been distanced from the "xoconostles" types in the previous analyses. This analysis revealed the relationship between the "xoconostles" accessions, where the cultivated genotypes are found at the end of the linear tie while the wild ones are located at its base (Fig. 6).

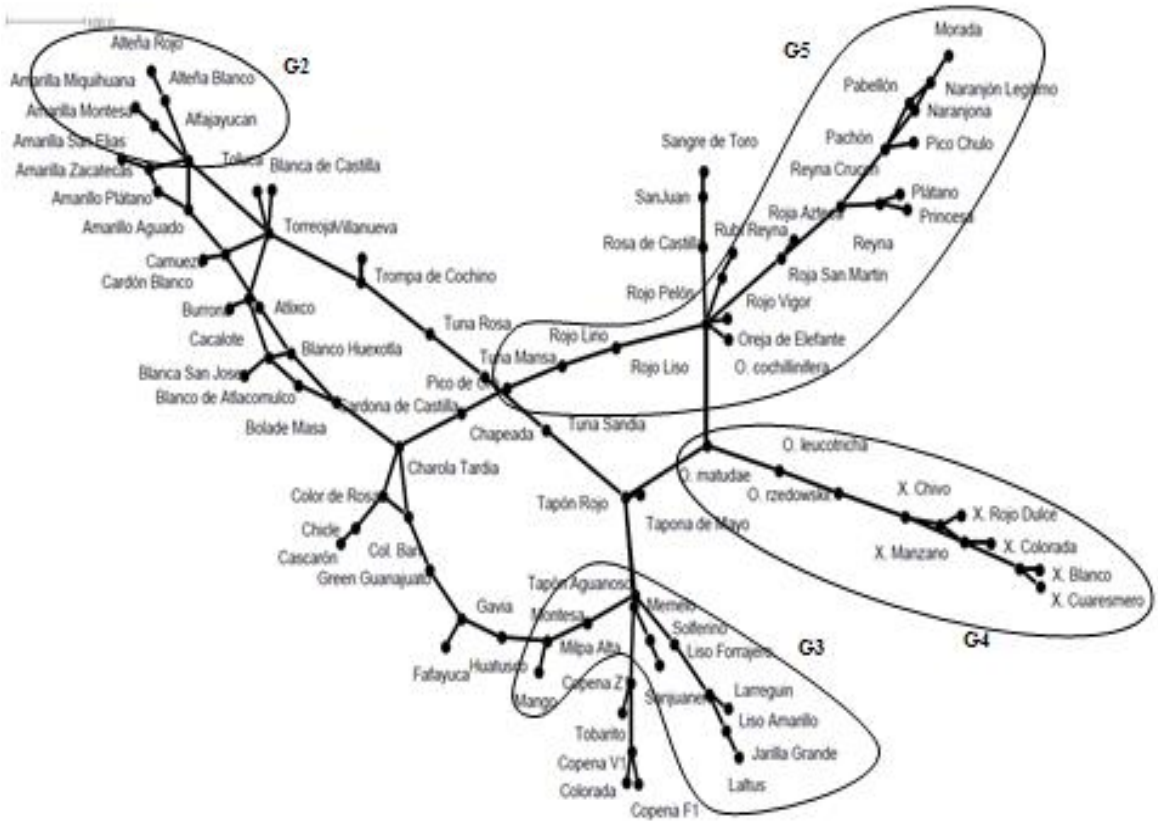


Fig. 6. Median-Joining tree of SSR data obtained from 88 Opuntia genotypes.

CONCLUDING REMARKS

This study represents the first report on the genetic diversity and population structure of Opuntia in Mexico analyzed by SSR markers. We used several complementary methods to estimate the number of clusters at both hierarchical and non-hierarchical levels. The

congruence of patterns obtained with Bayesian, networks and multivariate analyses suggests that the obtained groupings are reasonably reliable.

The results revealed a reticular evolution of the *Opuntia* species and they confirmed the incorrect definition of the species limits of this genus in Mexico. The accessions currently defined in *O. ficus-indica*, *O. megacantha*, *O. albicarpa*, *O. lasiacantha*, *O. hyptiacantha* and *O. streptacantha* were not separated genetically. Representative accessions of *O. robusta* species differ from other species as revealed by NeighborNet networks analysis. For their part, genotypes that produce acid fruits ("xoconostles") were partially separated from sweets ones when UPGMA and STRUCTRUE analyses were carried out. However, the results obtained by the Networks analyses (Median-Joining and NeighborNet) clearly separated the "xoconostles" accessions, and even revealed a linear evolution from wild species.

Three scenarios are suggested for the Mexican germplasm structure. Firstly, a scattering pattern of genotypes related to human exchange. Secondly, a potential regional selection of wild *Opuntia* sp. with desirable agronomic traits. Finally, the effect of end-use attributed to distinct genotypes in Mexico: fruits ("xoconostles" or "tunas"), vegetables or livestock feed, among other uses. The above scenarios are in accordance to the network analysis presented in the current study. Both linear and reticulate ties were revealed among the Mexican Opuntia accessions, suggesting mainly vegetative propagation of Opuntia accessions and, at a lesser extent, sexual heredity through selection of the possible natural crosses. Nonetheless, the domestication of Opuntia was likely a complex process not yet fully understood. The results presented here serve as a base for evolution studies under domectication concept.

Overall, the present results confirm the usefulness of SSR markers for characterization and for genetic diversity analysis of Mexican Opuntia accessions. According to the present and previous results, there is a high degree of genetic diversity. Clearly, these findings improved our knowledge about the situation of Opuntia diversity in Mexico and lead to appropriate information, which is useful for the successful management of germplasm and the prevention of its loss. Many efforts have been conducted to collect morphological variants of this crop and interesting genotypes for the production of fruits or vegetables have been selected, for cultivation at the household level or the local scale. However, the lack of a reliable characterization and lack of breeding programs have not allowed the development of improved cultivars.

In extension, a complete characterization of cultivars is needed in order to remove incidents of mislabeling or synonymy. Until now, only few genotypes have been fingerprinted genetically, while in the present study the genetic structure of an extensive Mexican Opuntia collection was exploited and relationships among entries were revealed. Such information will aid the selection of cultivars for germplasm conservation and implementation in breeding programs, providing information of diverse genetic backgrounds in crops, while monitoring the trade of plant material.

Hence, further studies are required to develop new SSR markers derived from genome sequence information of Opuntia rather than that of relative species, which will provide better insight and an understanding of the genetic diversity of Opuntia, and can be used for marker-assisted breeding of new cultivars.

REFERENCES

Anderson, E. F. 2001. The cactus family. Timber press, Portland. USA 776 p.

Arakaki, M., Christin, P. A., Nyffeler, R., Lendel, A., Eggli, U., Ogburn, R. M., Spriggs, E., Moore, M. J., and Edwards, E. J. 2011. Contemporaneous and recent radiations of the world's major succulent plant lineages. PNAS 108:8379-8384. doi: 10.1073/pnas.1100628108

Barreneche, T., Casasoli, M., Russell, K., Akkak, A., Meddour, H. 2004. Comparative mapping between *Quercus* and *Castanea* using simple-sequence repeats. Theor. Appl. Genet. 108: 558-566. doi:10.1007/s00122-003-1462-2

Bendhifi, M., Baraket, G., Zourgui, L., Souid, S., Salhi-Hannachi, A. 2013. Assessment of genetic diversity of Tunisian Barbary fig (*Opuntia ficus indica*) cultivars by RAPD markers and morphological traits. Sci. Hortic. 158: 1-7. doi: 10.1016/j.scienta.2013.04.015

Bruvo, R., Michiels, N. K., D'Souza, T. G., and Schulenburg, H. 2004. A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. Mol. Ecol. 13: 2101-2106. doi: 10.1111/j.1365-294X.2004.02209.x

Caruso, M., Currò, S., Las Casas, G., La Malfa, S., Gentile, A. 2010. Microsatellite markers help to assess genetic diversity among *O. ficus-indica* cultivated genotypes and their relation with related species. Plant Syst. Evol. 290: 85-97. doi: 10.1007/s00606-010-0351-9

Casas, A., and Barbera, G, 2002. Mesoamerican domestication and diffusion. In: Nobel, P. S. (ed) Cacti: Biology and uses. University of California Press, USA, pp: 143-162.

Chae, W. B., Hong, S. J., Gifford, J. M., Lane Rayburn, A., Sacks, E. J., and Juvik, J. A. 2014. Plant morphology, genome size, and SSR markers differentiate five distinct taxonomic groups among accessions in the genus *Miscanthus*. GCB Bioenergy 6: 646-660. doi: 10.1111/gcbb.12101

Chavez-Santoscoy, R. A., Gutierrez-Uribe, J. A., and Serna-Saldívar, S. O. 2009. Phenolic Composition, Antioxidant Capacity and In Vitro Cancer Cell Cytotoxicity of Nine Prickly Pear (*Opuntia* spp.) juices. Plant Foods Hum Nutr 64: 146-152. doi:10.1007/s11130-009-0117-0

Ducarme, V., Risterucci, A. M., Wesselingh, R. A. 2008. Permanent genetic resources: Development of microsatellite markers in *Rhinanthus angustifolius* and cross-species amplification. Mol. Ecol. Res. 8: 384-386. doi: 10.1111/j.1471-8286.2007.01966.x

Dufresne, F., Stift, M., Vergilino, R., and Mable, B. k. 2014. Recent progress and challenges in population genetics of polyploid organisms: an overview of current state-of-the-art molecular and statistical tools. Mol. Ecol. 23: 40-69. doi: 10.1111/mec.12581

Ekué, M., Gailing, O., and Finkeldey, R. 2009. Transferability of Simple Sequence Repeat Markers Developed in *Litchi chinensis* to *Blighia sapida* (Sapindaceae). Plant Mol. Biol. Rep. 27: 570-574. doi: 10.1007/s11105-009-0115-2

Erre, P., Nieddu, G., Chessa ,I. 2011. Identification of microsatellite loci in *Opuntia* spp. and their characterization in cultivars and species. Acta Hort. 918: 327-332.

Esselink, G. D., Nybom, H., Vosman, B. 2004. Assignment of allelic configuration in polyploids using the MAC-PR (microsatellite DNA allele counting-peak ratios) method. Theor. Appl. Genet. 109: 402-408. doi: I 10.1007/s00122-004-1645-5

Evanno, G., Regnaut, S., and Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14: 2611-2620. doi: 10.1111/j.1365-294X.2005.02553.x

Falush, D., Stephens, M., and Pritchard, J. K. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. Mol. Ecol. notes 7: 574-578. doi: 10.1111/j.1471-8286.2007.01758.x

FitzSimmons, N. N., Moritz, C., and Moore, S. 1995. Conservation and dynamics of microsatellite loci over 300 million years of marine turtle evolution. Mol. Biol. Evol. 12: 432-440.

Gao, L. Z., Zhang, C. H., and Jia, J. Z. 2005. Cross-species transferability of rice microsatellites in its wild relatives and the potential for conservation genetic studies. Genet. Resour. Crop Evol. 52: 931-940. doi: 10.1007/s10722-003-6124-3

García-Zambrano, E. A., Zavala-García, F., Gutiérrez-Diez, A., Ojeda-Zacarías, M. C., and Cerda-Hurtado, I. 2009. Estimación de la diversidad genética del Nopal, usando marcadores moleculares tipo AFLP. FYTON 78: 117-120.

Gibson, A. C., and Nobel, P. S. 1986. The cactus primer. Harvard University Press, Cambridge.

Griffith, M. P. 2004. The origins of an important cactus crop, *Opuntia ficus-indica* (Cactaceae): New molecular evidence. Am. J Bo.t 91: 1915-1921. doi: 10.3732/ajb.91.11.1915

Helsen, P., Browne, R., Anderson, D., Verdyck, P., and Van Dongen, S. 2009a. Galápagos' *Opuntia* cacti: extensive morphological diversity, low genetic variability. Biol. Linn. Soc. 96: 451-461. doi: 10.1111/j.1095-8312.2008.01141.x

Helsen, P., Verdyck, P., Tye, A., Desender, K., Van Houtte, N., and Van Dongen, S. 2007. Isolation and characterization of polymorphic microsatellite markers in Galapagos prickly pear (*Opuntia*) cactus species. Mol. Ecol. Notes 7: 454-456. doi: 10.1111/j.1471-8286.2006.01615.x

Helsen, P., Verdyck, P., Tye, A., and Van Dongen, S. 2009b. Low levels of genetic differentiation between *Opuntia echios* varieties on Santa Cruz (Galápagos). Plant Syst. Evol. 279: 1-10. doi: 10.1007/s00606-008-0064-5

Hunt, D. R. 2006. The New Cactus Lexicon. dh books, The Manse, Milbourne, UK. 526 p.

Hunt, H. V., Campana, M. G., Lawes, M. C., Park, Y. J., Bower, M. A., Howe, C. J,, Jones, M. K. 2011. Genetic diversity and phylogeography of broomcorn millet (*Panicum miliaceum* L.) across Eurasia. Mol. Ecol. 20: 4756-4771. doi: 10.1111/j.1365-294X.2011.05318.x

Huson, D. H., and Bryant, D. 2006. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. 23: 254-267. doi:10.1093/molbev/msj030

Labra, M., Grassi, F., Bardini, M., Imazio, S., Guiggi, A., *et al.* 2003. Relationships in *Opuntia* Mill. Genus (Cactaceae) detected by molecular marker. Plant Sci. 165: 1129-1136. doi:10.1016/S0168-9452(03)00321-2

Liu, K., and Muse S. V. 2005. PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21: 2128-2129. doi: 10.1093/bioinformatics/bti282

Luna-Paez, A., Valadez-Moctezuma, E., Barrientos, P. A. F., Gallegos, V. C. 2007. Characterization of *Opuntia* spp. by means of seed with RAPD and ISSR markers and its possible use for differentiation. J PACD 9: 43-59.

Majure, L. C., Judd, S., Soltis, S., Soltis D. 2012. Cytogeography of the *Humifusa* clade of *Opuntia* s.s. Mill. 1754 (Cactaceae: Opuntioideae): Correlations with geographic distributions and morphological differentiation of a polyploid complex. Comp Cytogenet 6: 53-77. Doi: 10.3897/CompCytogen.v6i1.2523

Martin, M. A., Mattioni, C., Cherubini, M., Taurchini, D., and Villani F. 2010. Genetic diversity in European chestnut populations by means of genomic and genic microsatellite markers. Tree Genet. Genomes 6: 735-744. doi: 10.1007/s11295-010-0287-9

Mondragón-Jacobo, C. 2003. Molecular characterization using RAPDs of a cactus (*Opuntia* spp. Cactaceae) collection from central Mexico as a basis for plant breeding. Rev. Chap. Hortic. 9: 97-114.

Montilla-Bascón, G., Sánchez M., J., Rispail, N., Rubiales, D., Mur, L. *et al.* 2013. Genetic Diversity and Population Structure among Oat Cultivars and Landraces. Plant Mol. Biol. Rep. 31: 1305-1314. doi: 10.1007/s11105-013-0598-8

Morales, P., Ramírez-Moreno, E., Sanchez-Mata, M. C., Carvalho, A. M, and Ferreira, I. 2012. Nutritional and antioxidant properties of pulp and seeds of two xoconostle cultivars (*Opuntia joconostle* F.A.C. Weber ex Diguet and *Opuntia matudae* Scheinvar) of high consumption in Mexico. Food Res. Int. 46: 279-285. doi:10.1016/j.foodres.2011.12.031

Moscoe, L. J., Emshwiller, E. 2014. Diversity of *Oxalis tuberosa* Molina: a comparison between AFLP and microsatellite markers. Genet. Resour. Crop Evol. doi: 10.1007/s10722-014-0154-x

Mottura, M. C., Finkeldey, R., Verga, A. R., and Gailing, O. 2005. Development and characterisation of microsatellite markers for *Prosopis chilensis* and *Prosopis flexuosa* and cross-amplification. Mol. Ecol. Notes 5: 487-489. doi:10.1111/j.1471-8286.2005.00965.x

Obbard, D. J., Harris, S. A., and Pannell, J. R. 2006. Simple allelic phenotype diversity and differentiation statistics for allopolyploids. Heredity 97: 296-303. doi:10.1038/sj.hdy.6800862

Peakall, R., and Smouse, P. E. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28: 2537-2539. doi: 10.1093/bioinformatics/bts460

Pfeiffer, T., Roschanski, A. M., Pannell, J. R., Korbecka, G., Schnittler, M. 2011. Characterization of Microsatellite Loci and Reliable Genotyping in a Polyploid Plant, *Mercurialis perennis* (Euphorbiaceae). J. Hered. doi: 10.1093/jhered/esr024

Pritchard, J. K., Stephens, M., and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. Genet. 155: 945-959.

Reyes-Agüero, J. A., Aguirre-Rivera, J. R., Carlín, F. C., and González, D. A. 2009. Catálogo de las principales variantes silvestres y cultivadas de *Opuntia* en la Altiplanicie Meridional de México. UASLP, México.

Rohlf, F. J. 2002. NTSYS-PC, numerical taxonomy system for the PC, ExeterSoftware, Ver. 2.2. Exeter Software, Setauket, New York, USA.

Samah, S., and Valadez-Moctezuma, E. 2014. Morphological Seeds Descriptors for Characterize and Differentiate Genotypes of *Opuntia* (Cactaceae, Opuntioideae). Annual Research and Review in Biology 4: 3791-3809. doi: 10.9734/ARRB/2014/11606

Shepherd, M., Cross, M., Maguire, T. L., Dieters, M. J., Williams, C. G., Henry, R. J. 2002. Transpecific microsatellites for hard pines. Theor. Appl. Genet. 104: 819-827. doi:10.1007/s00122-001-0794-z

Shi, J., Huang, S., Zhan, J., Yu, J., Wang, X., Hua, W., Liu, S., Liu, G., and Wang, H. 2014. Genome-Wide Microsatellite Characterization and Marker Development in the Sequenced Brassica Crop Species. DNA Res. 21: 53-68. doi: 10.1093/dnares/dst040.

Stuppy, W. 2002. Seed characters and the generic classification of the Opuntioideae (Cactaceae). In Hunt, D. R., and Taylor, N. P. (eds) Studies in the Opuntioideae (Cactaceae). The Manse, Sherborne, UK, pp 25-58.

Takayama, K., Tamura, M., Tateishi, Y., and Kajita, T. 2008. Isolation and characterization of microsatellite loci in the red mangrove *Rhizophora mangle* (Rhizophoraceae) and its related species. Conserv. Genet. 9: 1323-1325. doi:10.1007/s10592-007-9475-z

Valadez-Moctezuma, E., Ortiz-Vásquez, Q., Samah, S. 2014b. Molecular based assessment of genetic diversity of xoconostle accessions (*Opuntia* spp.). Afr. J. Biotech. 13: 202-210. doi: 10.5897/AJB2013.13350

Valadez-Moctezuma, E., Samah, S., and Luna-Paez, A. 2014a. Genetic diversity *of Opuntia* spp. varieties assessed by classical marker tools (RAPD and ISSR). Plant Syst. Evol. DOI: 10.1007/s00606-014-1112-y

Wang, X., Felker, P., Burow, M. D., and Peterson A. H. 1998. Comparison of RAPD marker whit morphological and physiological data in the classification of *Opuntia* accessions. JPACD 3: 1-5.

Zoghlami, N., Chrita, I., Bouamama, B., Gargouri, M., Zemni, H., Ghorbel, A., and Mliki, A. 2007 Molecular based assessment of genetic diversity within Barbary fig (*Opuntia ficus indica* (L.) Mill. in Tunisia. Sci. Hortic. 113: 134-141. doi: 10.1016/j.scienta.2007.02.009

Chapter 4

Part 2

DNA marker (PCR-RFLP)

GENETIC DIVERGENCE BETWEEN MEXICAN OPUNTIA ACCESSIONS INFERRED BY PCR-RFLP ANALYSIS

Samir Samah, Ernestina Valadez-Moctezuma, Kenia Socorro Peláez-Luna, Sonia Morales-Manzano, Pablo Meza-Carrera, Roberto Cid-Contreras

Article submitted: Plant Systematics and Evolution. Manuscript Number: PLSY-D-14-00388.

GENETIC DIVERGENCE BETWEEN MEXICAN OPUNTIA ACCESSIONS INFERRED BY PCR-RFLP ANALYSIS

ABSTRACT

Molecular tools offer alternatives for better characterization and determination of phylogenetic relationships between plants. The aim of this research was to study the genetic divergence between 103 accessions/species of Mexican Opuntia. To accomplish this, the polymerase chain reaction-restriction fragment length polymorphism analysis of three chloroplast intergenic spacers (*atp*B-*rbc*L, *trn*L-*trn*F and *psb*A-*trn*H), one chloroplast gene (*ycf*1), two nuclear genes (*ppc* and *Phy*C), and one mitochondrial gene (*cox*3), was carried out. The amplified product from all samples had very similar molecular size, and there were very small variations among the undigested PCR amplicons for all regions, with the exception of the gene *ppc*. We obtained a total of 5,850 pb from the seven regions. A total of 136 fragments were detected with 8 enzymes, of which 37 fragments (27.2 %) were polymorphic. A total of 40 % polymorphic fragments were obtained in the chloroplast regions, while only 9.8 % of polymorphic bands were detected in the nuclear genes, and 20 % of the polymorphic fragments were identified in the mitochondrial locus *cox*3. The intergenic spacers *trn*L-*trn*F and *psb*A-*trn*H were the most variables regions. The genetic distance was very narrow, ranged from 0 to 0.12. Indeed, 77 out of the 103 genotypes had the same genetic profile. Moreover, all *xoconostle* accessions (acidic fruits) were grouped together without being separate from three genotypes of prickly pears (sweet fruits). We assume that the genetic divergence between prickly pears and *xoconostles* is very narrow and the number of *Opuntia* species currently considered in Mexico is questioned.

Keywords prickly pears; *xoconostle*; phylogenetic relationships; mitochondria; chloroplast; nucleus.

INTRODUCTION

Cactaceae is a major New World plant family and popular in semiarid horticulture. Cactaceae is comprised of 127 genera and 1438 species divided in four subfamilies: Cactoideae, Opuntioideae, Pereskiodeae and Maihuenioideae (Hunt, 2006). The two latter are each comprised of a single genus. Cactoideae, represented with seven tribes, is the largest subfamily, and Opuntioideae is the next largest subfamily; it includes two tribes, Opuntieae and Cylindropuntieae, comprised by ten and seven genera, respectively, and 192 species, of which 75 species are placed in the largest genus *Opuntia* (Hunt, 2006).

The subfamily Opuntioideae is widespread throughout the Americas, from Canada to southern Patagonia. It has traditionally been recognized as a monophyletic taxonomic entity (Griffith and Porter, 2009; Hernández-Hernández *et al.*, 2011). It is characterized by a number of synapomorphies: (1) presence of glochids: small and deciduous barbed spines; (2) woody funicular tissue surrounding the seed (funicular envelope); (3) high amounts of calcium oxalate monohydrate druses and monoclinic cluster crystals in the outer hypodermis of stems; and (4) polyporate pollen grains with peculiar exine structures (Anderson, 2001; Stuppy, 2002; Hunt, 2006). Opuntioideae include *Opuntia* Mill. *sensu lato* (s.l.) and four associated genera (*Cumulopuntia* F. Ritter s.l., *Maihueniopsis* Speg. s.l., *Pterocactus* K. Schum. and *Puna* R. Kiesling s.l). Molecular phylogenetic studies demonstrated that *Opuntia* s.l. was shown to be polyphyletic (Griffith and Porter, 2009). Thus, *Opuntia sensu stricto* (s.s.) has been reduced drastically in size with many segregate genera now recognized (e.g., *Austrocylindropuntia* Backeb., *Brasiliopuntia* (K. Schum.) A.

Berger, *Cylindropuntia* (Engelm.) F. M. Knuth) (Hunt, 2006; Griffith and Porter, 2009). *Opuntia* s.s. (nopal, prickly pears) is the largest genus in Opuntioideae (Anderson, 2001). This genus is suggested to have originated as recently as 5.6 (± 1.9) million years ago (mya) (Arakaki *et al.,* 2011). There are 150 (Stuppy, 2002) to 180 recognized species within the genus including *Nopalea* (Anderson, 2001; Hunt, 2006), of which there are 66-83 reported only in Mexico.

Members of *Opuntia* s.s. are cultivated worldwide as fruit and vegetable crops and are increasingly used as forage and fodder for livestock in arid areas of the world, such as parts of Brazil, Mexico, western Asia, and northern and southern Africa. In Mexico, where species of *Opuntia* have been cultivated for at least 14 000 yr (Casas and Barbera, 2002), they represent an iconic national figure, illustrated on the country's flag. The Aztecs and other Mesoamerican civilizations used the cactus pads as a vegetable or fodder and the prickly pear as a seasonal fruit: sweet (cactus pear) or acid fruits (*xoconostle*). They are also considered an alternative natural medicine due to their antihypoglycemic, oxidative stress and cancer prevention effects (Chavez-Santoscoy *et al.,* 2009). They also have antihyperlipidemic, antinflammatory, antidiuretic, hypocholesterolemic, immunostimulatory and antiulcerogenic activity, including weight-loss effects (Morales *et al.,* 2012).

Also, *Opuntia* is known for its difficult taxonomy. The continuous morphological variation in the genus, the synonyms and the insufficient and inadequate morphological descriptors, all have led to misclassifications (Labra *et al.,* 2003; Caruso *et al.,* 2010; Valadez-Moctezuma *et al.,* 2014). Moreover, species limits are still poorly understood as a result of the high frequency of polyploid taxa. Thus, it seems that chromosomal data are of

little value for infrageneric classification (Majure *et al.*, 2012a). Opuntia has been recorded as diploid to enneaploid (2n = 9x = 99), with 60.4 % of reported counts thus far pertaining to polyploidy individuals, and another 13.4 % representing taxa with both diploid and polyploid cytotypes (Majure *et al.,* 2012a). Furthermore, there is no comprehensive phylogeny of *Opuntia* s.s., so limits of major clades are largely unknown. Numerous morphological and cytological studies have been conducted on large groups of taxa and species complexes (e.g. Majure *et al.*, 2012a). Griffith and Porter (2009) included 28 species of *Opuntia* s.s. in their molecular phylogeny of Opuntioideae but were unable to resolve relationships within *Opuntia* s.s. using the nuclear internal transcribed spacer (ITS) and the plastid intergenic spacer *trn*L-F. Hernández-Hernández *et al.* (2011) and Bárcenas *et al.* (2011) studied the phylogenetic relationship in South American *Opuntia* species. However, these last studies only surveyed seven species of *Opuntia* s.s., or had no resolution among clades, respectively. The most complete phylogeny available for the subfamily Opuntioideae is that of Griffith and Porter (2009). They sampled 110 specimens using ITS and the chloroplast *trn*L-*trn*F regions. Majure *et al.* (2012b) studied 98 species of Opuntioideae, they sequenced the plastid intergenic spacers *atp*B-*rbc*L, *ndh*F-*rpl*3, *psb*J-*pet*A, and *trn*L-*trn*F, the plastid genes *mat*K and *ycf*1, and the nuclear gene *ppc* and ITS to reconstruct the phylogeny of the Opuntieae tribe, including *Opuntia* s.s., which is a well-supported clade but includes *Nopalea* (Majure *et al.*, 2012b). On the other hand, *Opuntia* s.s. shows the signature of a clade that has undergone a rapid radiation (i.e., broad distribution, high morphological and species diversity, and low molecular marker divergence) (Helsen *et al.*, 2011). However, small Mexican species were included in the molecular studies of the Opuntioideae subfamily, viz. *O. ficus-indica* and *O. lasiacantha* (Bárcenas *et al.*, 2011); *O. ficus-indica* and *O. megacantha* (Griffith and Porter, 2009); *O. ficus indica, O. megacantha*, *O. streptacantha, O. hyptiacantha, O. robusta* (Griffith,

2004); *O. ficus-indica, O. hyptiacantha, O. lasiacantha, O. robusta, O. streptacantha* (Yesson *et al.,* 2011); *O. lasiacantha* (Hernández-Hernández *et al.,* 2011); *O. megacantha* (Majure *et al.,* 2012b). Also, these species were represented with one specimen of each species in all studies mentioned above except for *O. ficus indica* in Griffith (2004). Other Mexican *Opuntia* species with high agronomic and economic importance like *O. xoconostle, O. albicarpa, O. matudae, O. undulata, and O. chaveña* have never been included in molecular phylogenetic studies. Furthermore, there are no phylogenetic studies for Mexican Opuntia and the relationships between these species are still unclear. Also, the genetic divergence between *xoconostles* (Opuntia with acid fruits) and other *Opuntia* species (with sweet fruits) is quite unknown.

Knowledge of genetic variation and phylogenetic relationship among genotypes is an important consideration for classification, utilization of germplasm resources, and plant breeding. Traditionally, cultivar identification has relied on morphological and agronomic characteristics of plant materials. Although there is substantial intraspecific variation in vegetative traits, especially leaf (spine in Opuntia) and fruit characters, it is difficult to distinguish genotypes based on their external morphology alone (Khan *et al.,* 2014). Furthermore, phenotypic characters are generally influenced by environmental factors and the growth stage of the plant. In *Opuntia* species, this requires a lengthy and expensive evaluation during the whole vegetative growth (Labra *et al.,* 2003). However, molecular markers overcome many limitations of morphologically-based genetic analysis and provide information that can help to differentiate accessions, classification and phylogenetic position. Molecular markers are classified into two types' viz.: one DNA marker system based on hybridization blotting and another based on PCR, principally. The hybridization blotting system is represented for the Restriction Fragment Length Polymorphism (RFLP)

technique (the first molecular marker technique reported), where the DNA polymorphism is detected through a labelled DNA probe on a southern blot containing DNA digested by restriction endonucleases. The polymorphism is generated due to nucleotide substitutions or DNA rearrangements like insertion, deletion or single nucleotide polymorphisms (Agarwal *et al.,* 2008). DNA marker systems based on PCR have been frequently used for the last two decades and include techniques like RAPD, ISSR, AFLP, SSR, DArT and SNP (Khan et al. 2014). PCR-RFLP technique, sometimes also known as cleaved amplified polymorphic sequences (CAPS) (Konieczny and Ausubel, 1993), is a variant of AFLP. PCR-RFLP are performed by digesting locus-specific PCR amplicon with one or more restriction enzymes, followed by separation of the DNA fragments on agarose or polyacrylamide gels (Agarwal *et al.,* 2008). The method is considered to be an easy and advantageous tool for detecting DNA variations very fast. The polymorphisms were detected to have stable genetic attributes and thus can be used as molecular markers to distinguish individuals or genotypes within or among species (Bertea *et al.,* 2007; Lin *et al.,* 2012; Alvarez *et al.,* 2013).

Plants have two cytoplasmic genomes, chloroplast cpDNA and mitochondrial mtDNA. Mitochondria are involved in basic metabolic functions (respiration) and chloroplasts contain the entire enzymatic machinery necessary for photosynthesis and other biochemical pathways (Dong *et al.,* 2012; Liu *et al.,* 2013). Both genomes are generally uniparentally inherited in most flowering plants; they do not undergo recombination, so they act as a single locus. Also, the evolutionary rate of mtDNA and cpDNA has been frequently assumed to be slow, so that DNA divergence levels should roughly reflect divergence times (Galtier *et al.,* 2009; Nock *et al.,* 2011). Chloroplast/mitochondrial-encoded genes have been considered as less likely than other genes to be involved in

adaptive processes. These specific biological properties should make it an appropriate marker of molecular biodiversity and taxonomy. For the purpose of resolving relationships at and/or below species level, researching rapidly evolving DNA spacers and intron sequences is usually required, and many of such regions are readily amplifiable from both organelles (Shaw *et al.,* 2005) and the nucleus (Steele *et al.,* 2008) using standard PCR protocols. When PCR is used to retrieve discrete chloroplast regions for phylogenetic analysis, the target sequences need to be carefully selected. Sequences with relatively low mutation rates are required for higher-level phylogenetic comparisons, while higher mutation rates are needed to discriminate among closely related species (Nock *et al.,* 2011). Amplification of non-chloroplast regions is often challenging due to lack of specific primers, and limiting general use to the easily amplifiable regions such as the internal transcribed spacers (ITS) of the 18S-26S nrDNA tandem repeats, and the closely related external transcribed spacer (ETS) adjacent to the 5' end of the 28S nrDNA. Both have an extensive tradition of use in various phylogenetic studies but are known to occur in multiple, although not necessarily identical, copies (Álvarez and Wendel, 2003; Schlüter *et al.,* 2005). The number of copies is often dependent on species-specific ploidy levels and is likely to increase with an increasing number of chromosome sets. Nuclear data can also be subject to stochastic processes such as recombination and lineage sorting which may result in topologies differing from those obtained from other sources of data (Poczai and Hyvönen, 2010).

The detection of RFLP in specific regions of cpDNA and mtDNA amplified by PCR has been developed as a method for detecting variations and is a source of original markers potentially useful for studies of evolution, phylogeny and population genetics (Poczai *et al.,* 2011; Alvarez *et al.,* 2013). PCR-cpRFLP and PCR-mtRFLP have been used to

separate two complex species of *Leonurus cardiaca* (*L. cardiaca* and *L. quinquelobatus*) (Marciniuk *et al.,* 2014). PCR-cpRFLP has been used to reveal genetic diversity among apples (Khadivi-Khub *et al.,* 2014), to analyze genetic composition and differentiation of *Prunus spinosa* L. populations (Leinemann *et al.*, 2014), as well as to fingerprint, identify and authenticate varieties and species (Bertea *et al.,* 2006; Karaca *et al.,* 2008; Melgarejo *et al.,* 2009; Zhu *et al.,* 2011; Lin *et al.,* 2012; Sarin *et al.,* 2013), in order to infer biogeographic and evolutionary relationships, origin and domestication history of plants (Dane and Liu, 2007).

Based on previous information, the objectives of this study were: (1) to infer phylogenetic relationships among 103 agronomic important accessions/species of Mexican Opuntias by PCR-RFLP analysis of seven loci, and (2) to estimate the genetic divergence among prickly pears and *xoconostle* genotypes. The present paper represents the first report from our ongoing research investigating the phylogeny of Mexican Opuntia.

MATERIALS AND METHODS

Plant materials and DNA extraction

A total of 106 accessions/species were analyzed in the present study. Samples of 103 Opuntia accessions were obtained from the germplasm banks of Crucen-UACh (Zacatecas) and Nopalera-UACh (Texcoco), in Mexico. Three cacti (one sample of *Cylindropuntia* sp., one *pitahaya* (*Hylocereus undatus*) and one *pitaya* (*Stenocereus thurberi*) were included as outgroups. Some *Opuntia* accessions are classified in delimited species but others have no specific assignation (Supplementary data S1). Total genomic DNA was extracted using the CTAB protocol according to Luna-Paez *et al.* (2007). The DNA quantification was

estimated by spectrophotometry (ND-1000 Thermo scientific, USA) and its quality was determined in 1 % agarose gels.

PCR-RFLP analysis

A pilot study including 9 markers (four loci of the chloroplast genome (intergenic spacers *atp*B-*rbc*L, *trn*L-*trn*F, *psb*A-*trn*H and *pet*A-*psb*J) and one chloroplast gene (*ycf*1), two nuclear genes (*ppc* and *Phy*C), one nuclear internal transcribed spacer region (ITS) and one mitochondrial gene (*cox*3)) was conducted to evaluate the suitability of various markers for the present study. Primers information and molecular weight of the PCR product were described in Table 1.

Table 1. DNA regions, molecular weight (pb) and associated primers used in this study.

Genome	Region	Primer name	Source	annealing temperature (°C)	PCR-product Size (pb)
chloroplast	*trnL-trnF*	TrnL (UAA) 5' exon TrnF (GAA)	Taberlet *et al.* (1991) Taberlet *et al*. (1991)	60	1100
	atpB-rbcL	atpB.Op rbcL.Op	Majure *et al.* (2012b) Majure *et al*. (2012b)	65	900
	psbJ-petA	psbJ petA.Op	Shaw *et al.* (2007) Majure *et al.* (2012b)	60	-
	*psb*A-*trn*H	trnH psbA	Azuma *et al*. (2001) Sang *et al*. (1997)	62	500
Nucleus	*ycf1*	ycf1.Op118F ycf1.Op1330R	Majure *et al*. (2012b) Majure *et al.* (2012b)	65	900
	nrITS	ITS4 ITS5	White *et al.* (1990) White *et al.* (1990)	58	700
	ppc	ppc.Op.19F ppc.Op.569R.2	Majure *et al.* (2012b) Majure *et al.* (2012b)	60	550
	*Phy*C	PhyC F PhyC R	Helsen *et al.* (2009) Helsen *et al.* (2009)	60	1150
Mitochondrie	*cox*3	cox3 F cox3 R	Duminil *et al.* (2002) Duminil *et al.* (2002)	52	750

The PCRs were carried out in a final volume of 25 µL containing nuclease-free water, 500 mM dNTPs, 1 x Taq buffer, 25 mM $MgCl_2$, 20 pmol primers, 1.5 U Taq DNA polymerase (Promega) and 100 ng templates DNA. The thermo-cycling conditions

(MaxyGene Thermel Cycler, Applied Biosystem, USA) were: one 4 min cycle at 94 °C, 35 cycles [94 ºC for 30 s; annealing step (temperature for each primer pair are listed in Table 1) for 30 s; 72 ºC for 1.5 min], and one final extension cycle at 72 ºC for 3 min. After that, PCR products were separated into 1.2 % agarose gels in 1 x TAE buffer (40 mM Tris-acetate, 1 mM EDTA pH 8.0). DNA staining was performed with ethidium bromide 0.5 µg mL^{-1} and photographed using a Molecular Imager® ChemiDoc™ XRS system (BioRad, Japan). PCR products were digested with *Alu*I, *Ava*II, *Eco*RI, *Hae*III, *Hha*I, *Hinf*I, *Mbo*I, and *Tru*9I restriction enzymes according to specifications by the manufacturers. Digested PCR products were loaded on 8 % polyacrylamide gels. Marker Gene Ruler 1 kb DNA ladder MBI (Fermentas, USA) and 100 pb DNA ladder (Promega, USA) were used as reference to molecular weight. The electrophoresis was run using 220 volts for 1.5 h in 1 x TBE buffer (89 mM Tris-borate, 2 mM EDTA pH 8.0), using Dual MGV-216-33 vertical electrophoresis gel system (CBS, USA), and silver nitrate solution (AgNO3) 0.2 % was used to stain DNA fragments.

Data analysis

DNA band from each marker-endonuclease enzyme combination was considered as a qualitative character and coded as absent ‘0’ or present ‘1’ in each accession/locus/enzyme combination. Only consistent and reproducible DNA bands of the two replicates were used for the corresponding analysis. Binary matrices were constructed and the similarity between varieties was calculated using the Nei and Li/Dice coefficient. The dendrograms were constructed independently for each genome marker and for the combined data from all genome markers. The cophenetic correlation coefficient was computed for the dendrogram after the construction of a cophenetic matrix to measure the goodness of fit between the original similarity matrix and the dendrogram, using NTSYSpc version 2.2.

Bootstrap support values were obtained from 1,000 replicates. The dendrograms were obtained using the FreeTree V0.9.1.5 software, and then the TreeView V1.6.6 software was used to display the trees.

RESULTS

The phylogenetic relationships between the 103 accessions, belonging to more than 13 species of the *Opuntia* genus collected in Mexico, was studied using three chloroplast loci (intergenic spacers *atp*B-*rbc*L, *trn*L-*trn*F and *psb*A-*trn*H), the *ycf*1 chloroplast gene, the *ppc* and *Phy*C nuclear genes, and the *cox*3 mitochondrial gene. The seven genomic regions studied were amplified using universal and specific primers (Table 1) and the size of amplified fragments varied between 500 and 1,550 bp. A total of 5,850 bp were amplified, of which 3,400 bp were obtained from the chloroplast genome, 1,700 bp from the nucleus and 750 bp from the mitochondria. There was no observed change in length between the non-digested PCR products in the 103 accessions. However, the amplified fragment from the *ppc* gene showed small length variations ranging from 550 to 600 bp. The *psb*J-*pet*A locus was not considered in this study because the PCR product obtained wasn't unique.

PCR-RFLP polymorphisms and genetic diversity

Seven pair primers used in the present study were successfully amplified from the corresponding DNA genome regions in all Opuntia accessions (Fig.1; Table 2). These PCR products were digested with *Alu*I, *Ava*II, *Eco*RI, *Hae*III, *Hha*I, *Hinf*I, *Mbo*I and *Tru*9I restriction endonucleases having different recognition sites, of which four enzymes (*Eco*RI, *Hha*I, *Hinf*I, *Tru*9I) had restriction sites in the *atp*B-*rbc*L locus, five endonucleases (*Alu*I, *Hha*I, *Hinf*I, *Mbo*I, *Tru*9I) had restriction sites in the *trn*L-*trn*F region, five endonucleases (*Alu*I, *Hae*III, *Hha*I, *Hinf*I, *Mbo*I) had restriction sites in the *psb*A-*trn*H

region, five endonucleases (*Alu*I, *Hae*III, *Hinf*I, *Mbo*I, *Tru*9I) had restriction sites in the *ycf*1 locus, seven (*Alu*I, *Ava*II, *Hae*III, *Hha*I, *Hinf*I, *Mbo*I, *Tru*9I) had restriction sites in the *ppc* region, seven (*Alu*I, *Ava*II, *Hae*III, *Hha*I, *Hinf*I, *Mbo*I, *Tru*9I) had restriction sites in the *Phy*C locus, and four (*Alu*I, *Hae*III, *Hinf*I, *Mbo*I) had restriction sites in the *cox*3 region.

Table 2. Digested DNA fragments of seven loci in 103 *Opuntia* accessions/species

Restriction enzymes		*Alu*I		*Ava*I		*Eco*RI		*Hae*III		*Hha*I		*Hinf*I		*Mbo*I		*Tru*9I		Total	
Genome	Locus	TF	PF	TF	PF	TF	PF	TF	PF	TF	PF	TF	PF	TF	PF	TF	PF	TF	PF
Chloroplast	*atp*B-*rbc*L	//	//	//	//	3	0	//	//	2	0	6	0	//	//	6	0	17	0
	*trn*L-*trn*F	2	2	//	//	//	//	//	//	2	2	7	5	5	5	9	1	26	15
	*psb*A-*trn*H	3	2	//	//	//	//	2	2	2	2	2	2	2	2	//	//	12	10
	*ycf*1	4	2	//	//	//	//	3	0	//	//	2	0	5	0	6	3	20	5
Subtotal																		**75**	**30**
Nucleus	*ppc*	2	0	2	0	//	//	2	0	4	4	2	0	3	1	4	0	19	5
	*Phy*C	3	0	2	0	//	//	3	0	3	0	6	0	7	0	8	0	32	0
Subtotal																		**51**	**5**
Mitochondria	*cox*3	4	2	//	//	//	//	2	1	//	//	1	0	3	0	//	//	**10**	**2**
Total																		**136**	**37**

TF total fragments, PF polymorphic fragments
// No restriction site for the enzyme

Digestion of amplified products totally detected 136 fragments of which 37 (27.2 %) were polymorphic (Table 2). A total of 40 % polymorphic fragments were obtained in the chloroplast genomic regions (*atp*B-*rbc*L (0 %), *trn*L-*trn*F (57.7 %), *psb*A-*trn*H (83.3 %) and *ycf*1 (25 %)), and only a total of 9.8 % polymorphic bands in nuclear genes (*ppc* (26 %) and *Phy*C (0 %)), as well as 20 % of the polymorphic fragments in the mitochondrial gene *cox*3. The digestion products of the *atpB*-*rbc*L, *ycf*1 and *cox*3 regions in some Opuntia accessions are shown in Figs. 2, 3 and 4. Similarly, Fig. 5 illustrates the few polymorphisms revealed between species. The estimated genetic distance using the Dice coefficient showed the close relationship between the studied accessions/species of *Opuntia*. In the chloroplast regions this distance ranged between 0 and 0.13, and in the nuclear genes it varied from 0 to 0.04, and from 0 to 0.11 in the mitochondrial gene. The estimated distance using all combined data ranged between 0 and 0.12.

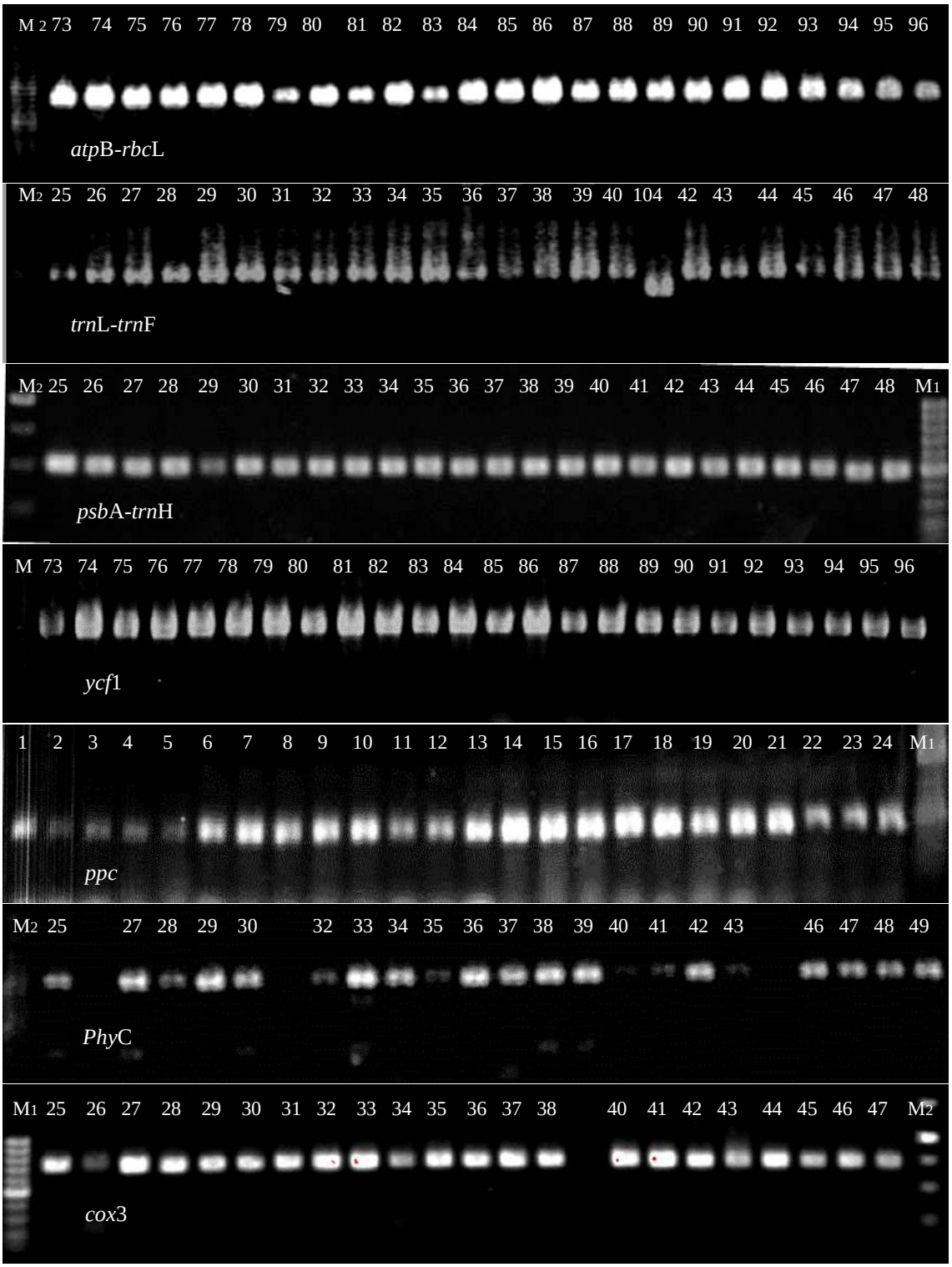


Fig. 1 Representative samples of PCR amplifications of the seven loci resolved in agarose gels 1.2 %. The numbers in each lane correspond to the accessions listed in Supplementary data S1. Molecular size: *atp*B-*rbc*L (900 pb), *trn*L-*trn*F (1100 pb), *psb*A-*trn*H (500 pb), *ycf*1 (900 pb), *ppc* (550-600 pb), *Phy*C (1150) and *cox*3 (750 pb). *M1* and *M2* are 100 pb DNA ladder and 1 kb DNA ladder, respectively.

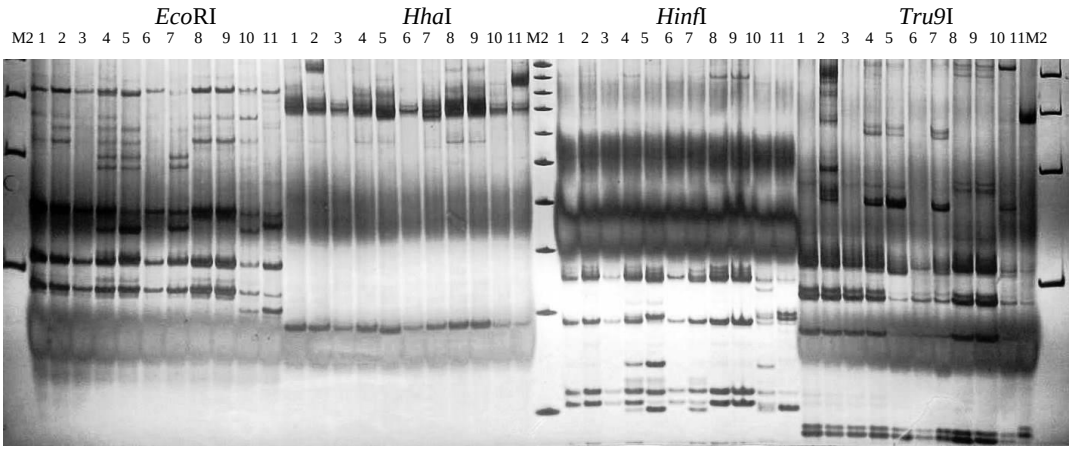


Fig. 2. Digested products of *atp*B-*rbc*L region with four enzymes (*Eco*RI, *Hha*I, *Hinf*I and *Tru9*I). 1-10 are accessions, 1 Alfajayucan, 2 Amarilla Jarro, 3 Amarillo Aguado, 4 Atlixco, 5 Cylindropuntia, 6 Naranjo Legitimo, 7 Xoconostle Colorado, 8 Xoconostle Cuaresmero, 9 Xoconostle Blanco, 10 Pitahaya, 11 Pitaya. *M1* and *M2* are 100 pb DNA ladder and 1 kb DNA ladder MBI, respectively.

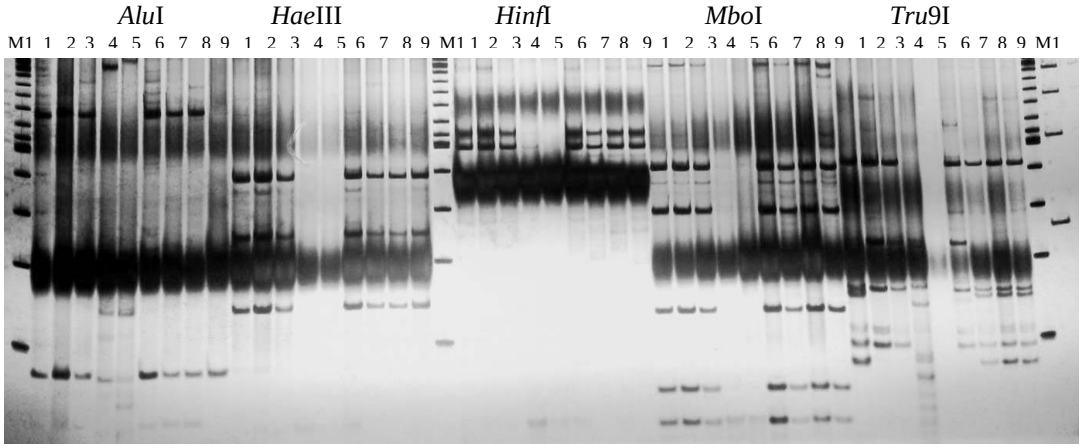


Fig. 3. Digested products of *ycf*1 gene with five enzymes (*Alu*I, *Hae*III, *Hinf*I, *Mbo*I and *Tru9*I). 1-9 are accessions, 1 Cascaron, 2 Copena VI, 3 Cylindropuntia, 4 Pitahaya, 5 Pitaya, 6 Tuna rosa, 7 Xoconostle colorado, 8, Xoconostle Cuaresmero, 9 Xoconostle Blanco. *M1* and *M2* are 100 pb DNA ladder and 1 kb DNA ladder MBI, respectively.

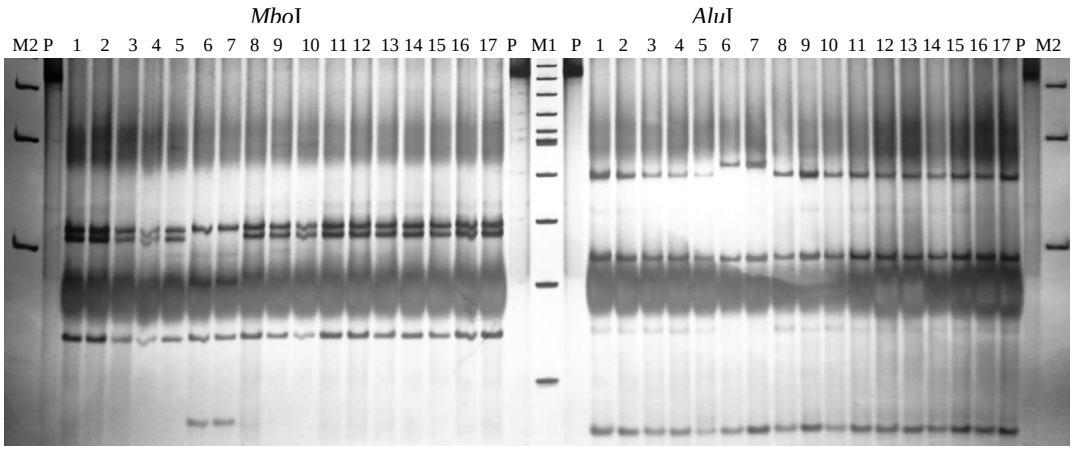


Fig. 4. Digested products of *cox*3 gene with two enzymes (*Mbo*I and *Alu*I). *P* PCR-product. 1-17 are accessions, 1 Alteña Blanco, 2 Alteña Rojo, 3 Amarilla Milpa Alta, 4 Amarilla Montesa, 5 Montesa, 6 Pitahaya, 7 Pitaya, 8 Reyna, 9 Toluca, 10 Plátano, 11 Cardona de castilla, 12 Atlixco, 13 Blanco de Atlacomulco, 14 Xoconostle Colorado, 15 Xoconostle Cuaresmero, 16 Xoconostle Blanco, 17 Xoconostle Manzano. *M1* and *M2* are 100 pb DNA ladder and 1 kb DNA ladder, respectively.

Genetic relationship based on neighbor-joining tree

A total of 26 different genotype profiles (the three outgroups included) were obtained with the 106 accessions studied. Based on the results, obtained with all the PCR-RFLP markers, a method applying a similarity matrix based on the Dice coefficient was used to generate a Neighbor-Joining (N-J) dendrogram (Fig. 6). The cophenetic correlation coefficient between the original similarity matrix and the cophenetic matrix derived from the tree was high (0.998), indicating a good fit between the tree and the similarity matrix.

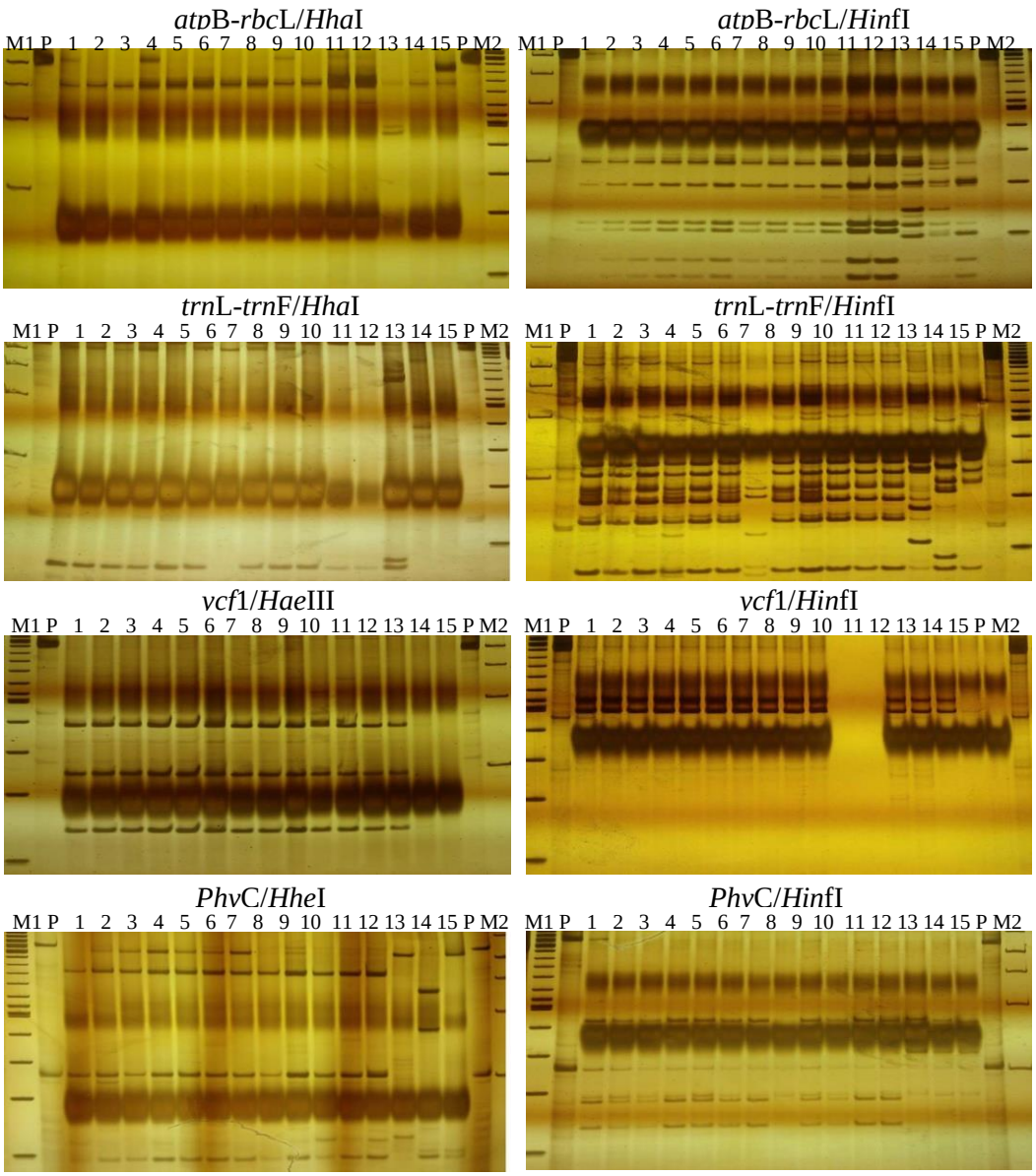


Fig. 5. Amplified (P) and digested products of primer/enzyme combination in *Opuntia* accessions. *1* Memelo (*O.affinis hyptiacantha*), *2* Reyna (*O.albicarpa*), *3* Cascarón (*O. chaveña*), *4* Atlixco (*O. ficus-indica*), *5* Charola Tardía (*O. hyptiacantha*), 6. Sanjuanera (*O. lasiacantha*), *7* Pico Chulo (*O. megacantha*), *8* Tapón Aguanoso (*O. robusta*), *9* Cardona de Castilla (*O. streptacantha*), *10* Oreja de Elefante (*O. undulata*), *11* Xoconostle Blanco (*O. joconostle*), 12 Xoconostle cuaresmero (*O. matudae*), *13 Cylindropuntia* sp., *14* Pitahaya (*Hylocereus undatus*), *15* Pitaya (*Stenocereus thurberi*). *M1* and *M2* are 100 pb DNA ladder and 1 kb DNA ladder, respectively.

Fig. 6. Phylogenetic relationships in Mexican Opuntia. Dendrogram constructed using the NJ method applying Nei and Li/Dice coefficient. Data were resolved from PCR-RFLP analysis of the seven regions representing the three plants genomes. The numbers at the nodes are bootstrap values as a percentage of 1,000 replications.

The close genetic distance mentioned previously was reflected in the N-J tree for the genomic regions of the chloroplast, in which 81 of the 103 genotypes had a similar genetic profile. The “Pico Chulo” accession was the most distanced. Accessions “Tapón Aguanoso”, “Tobarito”, “Tapón Rojo”, “Oreja de Elefante”, “Sanjuanera”, “Rojo Lirio”, “Roja Azteca”, “Reyna Crucen”, “Rojo Vigor”, “Blanca del Cerro” and “Copena Z1” were also distanced. Moreover, the *xoconostle* accessions (“Blanco”, “Colorado”, “Cuaresmero”, “Chivo”, “Manzano” and “Rojo”) were grouped together with three prickly pear accessions (“San Martin”, “Cascarón” and “Colorada” (Supplementary data S2-A)).

The tree obtained from the N-J analysis of nuclear genes showed that 88 genotypes had the same profile. The “Amarilla San Elias” accession was the most separated. The genotypes “Oreja de Elefante”, “Alteña Rojo”, “Alteña Blanco”, and “X. Rojo” were slightly distanced from the other ones (supplementary data S2-B). The dendrogram derived from the mitochondrial gene revealed the formation of two groups of 11 and 92 genotypes each. No genetic variation between accessions within each group was registered (Supplementary data S2-C).

The close genetic relationship described separately in three genomes was similar to that resulting from the combined data. In fact, 77 of 103 genotypes showed identical profiles to the seven genomic loci studied (Fig.6). The “Pico Chulo” accession was the most isolated from other Opuntia accessions. All genotypes belonging to the *xoconostle* group were grouped together but didn’t diverge from the prickly pear accessions, “Cascarón” (with accuracy of 48 %), “Colorada” and “Roja San Martin” (with accuracy of 18 %). Moreover, this grouping showed no concordance with the current taxonomic

classification. It is noteworthy that the three outgroups were separated in discrete clusters from the *Opuntia* accessions/species with an accuracy of 100 % (Fig. 6).

DISCUSSION

Generic limits within the genus *Opuntia* s.s. have historically been determined by morphological characters. Therefore, this study included only DNA data through PCR-RFLP analysis of seven loci, covering all three plant genomes, in order to provide an independent test of morphologically-determined classification. The amplification of the seven genomic regions was successful and the presence of one band per PCR product during gel electrophoresis reflected the appropriate primers used. The locus *psb*J-*pet*A presented an exception and it is required to design more specific primers for Mexican Opuntia genotypes.

Plant molecular evolutionary and systematic biologists have traditionally used ribosomal/nuclear (rDNA/ncDN) and organelle DNA (cpDNA and mtDNA) markers to reconstruct phylogenies at or below the genus level (Zimmer and Wen, 2012). The cpDNA and mtDNA markers are considered to be an ideal system in phylogeny and population genetics because of their uniparental mode of inheritance and low mutation rate related to the nuclear genome (Galtier *et al.*, 2009; Nock *et al.*, 2011). Tsumura *et al.* (2000) and Dane and Liu (2007) reported that the nature of specific DNA polymorphism detectable using PCR-RFLP is typically limited to restriction site changes and indel mutations. In this study, the few polymorphisms detected (total of three genomes; 27.2 %) were higher in the chloroplast genome (40 %) followed by the mitochondrial genome (20 %), while the nucleus genome reveled the lowest percentage of polymorphism (9.8 %). Chloroplasts contain both highly conserved genes fundamental to plant life and more variable regions,

which are informative over broad time scales (Nock *et al.*, 2011). However, the mitochondrial genome evolves considerably more slowly at the nucleotide sequence level than the nuclear or the chloroplast genomes, although the rate of rearrangements is extraordinarily faster in plant mtDNA than in cpDNA due to the presence of repeated regions, source of recombination within and between mtDNA genomes. PCR-based markers useful at low taxonomic levels are therefore difficult to obtain, due to the presence in higher plant mtDNA of introns, intergenic sequences, duplicate sequences and sequences of plastid and nuclear origin (Galtier *et al.*, 2009). In turn, nuclear ribosomal RNA encoding repeats are not always completely homogenized (Álvarez and Wendel, 2003) and do not always tracks both parents' genomes in hybrids and polyploids (Zimmer and Wen, 2012). This can be compared with the results of polymorphism in each genome displayed in this study.

The spacers *psb*A-*trn*H (83.3 %) and *trn*L-*trn*F (57.7 %) were the most polymorphic regions and the *ppc* (26 %), *ycf*1 (25 %) and *cox*3 (20 %) genes had an intermediate polymorphic percentage. Korotkova *et al.* (2011) suggested that the spacers are more effective but the introns still perform better compared to the coding regions. The *psb*A-*trn*H spacer is among the most variable plastid spacers for genome across a wide range of groups of plants. However, there are some problems limiting its use in phylogenetics, such as frequent indels, microsatellites, inversions, and a high degree of homoplasy, and the presence of inversions in the middle of the sequence, which can lead to incorrect alignments (Whitlock *et al.*, 2010); however, *psb*A-*trn*H may still be a successful marker due to its high interspecific variability (Kress and Erickson, 2007; Korotkova *et al.*, 2011). Similarly, Korotkova *et al.* (2011) reported that the *psb*A-*trn*H spacer was the most successful individual region for operational taxonomic unit identification (at the

intraspecific level). Likewise, the intergenic spacer *trn*L-*trn*F has been the most frequently used locus in phylogenetic studies in Opuntioideae (Griffith and Porter, 2009; Hernández-Hernández *et al.,* 2011; Majure *et al.,* 2012b). In contrast, the *ppc* and *ycf*1 genes have been previously sequenced in the Opuntieae tribe and demonstrated to be moderately informative (Hernández-Hernández *et al.,* 2011; Majure *et al.,* 2012b); nevertheless, the amplification of the mitochondrial gene *cox*3 in Opuntia is reported here for the first time. Moreover, the phytocrome C gene (*Phy*C) showed to be invariable in all genotypes studied here. Phytochromes are proteins involved in the sensing of light cues in green plants. They are encoded by a small gene family (*Phy*A-*Phy*E) in *Arabidopsis*. Most species trees use phytochrome genes for lower taxonomic levels, and they employ *Phy*A, *Phy*B or *Phy*C markers (Mathews *et al.,* 2010; Zimmer and Wen, 2012). Helsen *et al.* (2009) found that the *Phy*C gene was phylogenetically informative in four species of *Opuntia* grown in the Galapagos Islands. Similarly, the variability of *Phy*C markers was high within the ingroup of the cacti of tribe Tephrocacteae (Opuntioideae) (Ritz *et al.,* 2012), which indicates no variability in the sequence of this gene in the Mexican *Opuntia* species studied here. Similarly, Majure *et al.,* (2012b) defined 20 informative characters in a fragment of 861 bp of the *atp*B-*rbc*L region in *Opuntia* species. However, the results found here didn't reveal any polymorphism site in this region with any of the eight enzymes used, suggesting that Mexican accessions preserve the same sequence in this region. It is noteworthy that the PCR-RFLP analysis was carried out for the ITS region in 70 genotypes (all included in the material studied here), and digested with four restriction enzymes (*Eco*RI, *Hae*III, *Hha*I and *Hinf*I). However, no polymorphism was detected among accessions either (data not shown).

The phylogenetic relationship among accessions/species of *Opuntia* was visualized in the phylogenetic trees derived from the N-J method (Fig. 6 and Supplementary data S2-A, B and C). The cophenetic correlation coefficient was very high (0.998), indicating a consistent dendrogram. Also, the accuracy of 100 % separated the three outgroups from the *Opuntia* accessions/species, confirming the high divergence between these cacti taxa. In total, three quarters of the studied genotypes did not reveal differences in the profiles of seven loci digested with 8 endonuclease enzymes. This reflects the low genetic divergence among accessions/species of Mexican *Opuntia* studied here, although they differ morphologically and have been assigned in more than 13 taxonomic species (Supplementary data S1). Helsen *et al.* (2009) indicated that *Opuntia* species in Galápagos differ morphologically but not genetically. The discrepancy may be due, first, to the existence of a small set of unsampled loci determining morphology. Second, to the fact that adaptive radiation, a process attributed to these cacti, is characterized by rapid phenotypic evolution with respect to comparatively little genetic variation (Helsen *et al.*, 2011). Many plant species evolved via adaptive radiations or explosive patterns of speciation, and have evolutionary histories of only a few million years. Arakaki *et al.* (2011) suggested that many of the major radiations within Cactaceae were initiated at the end of the Miocene (ca. 10-5 mya). These very short evolutionary histories result in low sequence divergence. And third, the production of allopolyploid species is mentioned in *Opuntia* and has led to the origin of new species (Pinkava, 2002). However, these new genomic combinations often result in morphologically distinct entities, which may propagate themselves indefinitely via agamospermy, vegetative or apomixis to conserve the new phenotypes (Rebman and Pinkava, 2001). Also, phenotypic differences could be due to somatic mutations; traditional cultivars have been vegetatively propagated and could have accumulated somatic mutations with accompanying phenotypic consequences in crop morphology and

agronomic performance. Therefore, caution must be exercised when evaluating only morphological or only genetic data (Helsen *et al.,* 2009).

The clustering of accessions was not in accordance with the current assignment of genotypes in their respective species. This is probably related to the high level of phenotypic plasticity and ploidy levels, and also due to the morphological diversity of these accessions (Valadez-Moctezuma *et al.,* 2014). For these reasons, many studies have suggested the revision of the classification of the *Opuntia* genus (Helsen *et al.,* 2009; Caruso *et al.,* 2010; Valadez-Moctezuma *et al.,* 2014). It should be noted that many of the genotypes considered in this study have not yet been assigned taxonomically. On the other hand, the non-divergent accessions/species established here can respond to several hypotheses. First, the low efficiency of the genomic regions used to differentiate accessions/species of Mexican *Opuntia*. However, these same loci were shown to be efficient to separate species and genera in the subfamily Opuntoideae and the Cactaceae family in general (Griffith and Porter, 2009; Helsen *et al.,* 2009; Calvente *et al.,* 2011; Demaio *et al.,* 2011; Hernández-Hernández *et al.,* 2011; Majure *et al.,* 2012). Another reason to discard this hypothesis is the fact that all these genomic regions are known to be highly variable at low taxonomic levels (specifically *ycf*1 and *trn*H-*psb*A; Dong *et al.,* 2012; Franck *et al.,* 2012). Another hypothesis is the absence of many *Opuntia* species currently recognized, suggesting the need to define more precisely the limits among species of this genus in Mexico. Based on cpSSR and AFLP molecular markers, Labra *et al.* (2003) hypothesized that *O. ficus-indica* (the most important agronomic species) should be considered to be a domesticated form of *O. megacantha*. However, in this study there was a lack of comparisons to other *Opuntia* species. Griffith (2004), through the use of Bayesian phylogenetic analyses of nrITS DNA sequences, assumed that the taxonomic

concept of *O. ficus-indica* may include clones derived from multiple lineages and, therefore, be polyphyletic. This hypothesis seems to be supported by SSRs analyses (Caruso *et al.,* 2010) because the *O. ficus-indica* accessions did not cluster separately from other arborescent cactus pear species, such as *O. amyclaea, O. megacantha, O. streptacantha*, *O. fusicaulis*, and *O. albicarpa*. Valadez-Moctezuma *et al.* (2014), using RAPD and ISSR markers, supposed the existence of a smaller number of *Opuntia* species in Mexico, but with high intraspecific genetic variation. All these suggestions support our second hypothesis.

Other information that can be displayed in the tree (Fig. 6) is the incomplete divergence between genotypes producing sweet fruits (prickly pears) and the genotypes that produce acid fruits (*xoconostles*). The separation between the two groups was poorly supported by bootstrap value (48 %). According to Morales *et al.* (2012) and Samah and Valadez-Moctezuma (2014), the absence of pulp and the presence of an edible pericarp and small seeds are the most significant morphological differences between prickly pears and *xoconostles*. However, the separation between *xoconostles* and prickly pears has not been clear when the RAPD and ISSR markers were used (Valadez-Moctezuma *et al.,* 2014), or when biochemical markers (seed storage proteins) were applied (Samah *et al.*, chapter 2). Despite the clear morphological differences between these two types of Opuntia, these differences are not yet well defined at the DNA level and perhaps need much more time for it.

To differentiate genotypes morphologically as highly divergent but not genetically, using genes that are related with the high phenotypic variation is recommended (e.g. color, size of fruit or acidity of fruit), as well as developing molecular techniques such as cleaved

amplified polymorphic sequences (CAPS), which would be an alternative to differentiate and authenticate varieties of Opuntia. We have evaluated the CAPS technique on a related fruit size gene (Fruitfull), but genetic differences were negligible (unpublished data).

CONCLUSIONS

In this study, and for the first time, the genetic divergence among accessions/species of Mexican *Opuntia* was investigated by applying the PCR-RFLP analysis of seven regions used for molecular taxonomy. The results showed little genetic divergence despite the great morphological variability in the Mexican genotypes. No complete separation between genotypes producing sweet fruits (prickly pears) and genotypes producing acidic fruits (*xoconostles*) was demonstrated either. Moreover, the number of *Opuntia* species considered in Mexico is questioned, and the need to find other methods to delimit species of this genus is of great urgency. Species delimitation will require development of more appropriate markers to allow for the discovery of intraspecific variation, using multiple accessions from each potential species.

REFERENCES

Agarwal, M., Shrivastava, N., and Padh, H. 2008. Advances in molecular marker techniques and their applications in plant sciences. Plant Cell Rep. 27: 617-631.

Álvarez, I., and Wendel, J. F. 2003. Ribosomal ITS sequences and plant phylogenetic inference. Mol. Phylogenet. Evol. 29: 417-434.

Alvarez, J. B., Carmona, S., and Caballero, L. 2013. Identification and characterization by PCR-RFLP analysis of the genetic variation for the Glu-A1x and Glu-B1x genes in rivet wheat (*Triticum turgidum* L. ssp. *turgidum*). J. Cereal Sci. 57: 253-257.

Anderson, E. F. 2001. The Cactus Family. Timber, Portland. USA, 776 p.

Arakaki, M., Christin, P. A., Nyffeler, R., Lendel, A., Eggli, U., Ogburn, R. M., Spriggs, E., Moore, M. J., and Edwards, E. J. 2011. Contemporaneous and recent radiations of the world's major succulent plant lineages. PNAS 108: 8379-8384.

Azuma, H., García-Franco, J. G., Rico-Gray, V., and Thien, L. B. 2001. Molecular phylogeny of the Magnoliaceae: the biogeography of tropical and temperate disjunctions. Am. J. Bot. 88: 2275-2285.

Bárcenas, R. T., Yasson, T., and Hawkins, J. A. 2011. Molecular systematics of the Cactaceae. Cladistics 27: 470-489.

Bertea, C. M., Luciano, P., Bossi, S., Leoni, F., Baiocchi, C., Medana, C., Azzolin, C. M. M., Temporale, G., Lombardozzi, M. A., and Maffei, M. E. 2006. PCR and PCR-RFLP of the 5S-rRNA-NTS region and salvinorin A analyses for the rapid and unequivocal determination of *Salvia divinorum.* Phytochem. 67: 371-378.

Calvente, A., Zappi, D. C., Forest, F., and Lohmann, L. G. 2011. Molecular phylogeny of tribe Rhipsalideae (Cactaceae) and taxonomic implications for *Schlumbergera* and *Hatiora*. Mol. Phylogen. Evol. 58: 456-468.

Caruso, M., Currò, S., Las Casas, G., La Malfa, S., and Gentile, A. 2010. Microsatellite markers help to assess genetic diversity among *O. ficus indica* cultivated genotypes and their relation with related species. Plant Syst. Evol. 290: 85-97.

Casas, A., and Barbera, G. 2002. Mesoamerican domestication and diffusion. In: Nobel PS (ed) Cacti: Biology and uses. University of California Press, USA, pp 143-162.

Chavez-Santoscoy, R. A., Gutierrez-Uribe, J. A, and Serna-Saldívar, S. O. 2009. Phenolic Composition, Antioxidant Capacity and In Vitro Cancer Cell Cytotoxicity of Nine Prickly Pear (*Opuntia* spp.) juices. Plant Foods Hum. Nutr. 64: 146-152.

Dane, F., and Liu, J. 2007. Diversity and origin of cultivated and citron type watermelon (*Citrullus lanatus*). Genet. Resour. Crop Evol. 54: 1255-1265.

Demaio, P. H., Barfuss, M. H. J., Kiesling, R., Till, W., and Chiapella, J. O. 2011. Molecular phylogeny of *gymnocalycium* (Cactaceae): assessment of alternative infrageneric systems, a new subgenus, and trends in the evolution of the genus. Amer. J. Bot. 98: 1841-1854.

Dong, W., Liu, J., Yu, J., Wang, L., and Zhou, S. 2012. Highly Variable Chloroplast Markers for Evaluating Plant Phylogeny at Low Taxonomic Levels and for DNA Barcoding. PLoS One 7:e35071.

Duminil, J., Pemonge, M-H., and Petit, R. J. 2002. A set of 35 consensus primer pairs amplifying genes and introns of plant mitochondrial DNA. Mol. Ecol. Notes 2: 428-430.

Franck AR, Cochrane BJ, Garey JR (2012) Low-copy nuclear primers and *ycf*1 primers in Cactaceae. Am J Bot 99:405-407

Galtier, N., Nabholz, B., Glémin, S., and Hurst, G. D. D. 2009. Mitochondrial DNA as a marker of molecular diversity: a reappraisal. Mol. Ecol. 18: 4541-4550.

Griffith M. P. 2004. The origins of an important cactus crop, *Opuntia ficus-indica* (Cactaceae): New molecular evidence. Am. J. Bot. 91: 1915-1921.

Griffith, M. P., and Porter, M. 2009. Phylogeny of Opuntioideae (Cactaceae). Int. J. Plant Sci. 170: 107-116.

Helsen, P., Verdyck, P., Tye, A., and Van Dongen, S. 2009. Low levels of genetic differentiation between *O. echios* varieties on Santa Cruz (Galapagos). Plant Syst. Evol. 279: 1-10.

Helsen, P., Verdyck, P., and Van Dongen, S. 2011. The Influence of Historical Geneflow, Bathymetry and Distribution Patterns on the Population Genetics of Morphologically Diverse Galápagos' *Opuntia echios*. J. Mol. Evol. 72: 315-325.

Hernández-Hernández, T., Hernández, H., De Nova, J. A., Puente, R., Eguiarte, L. E., and Magallón, S. 2011. Phylogenetic relationships and evolution of growth form in Cactaceae (Caryophyllales, Eudicotyledoneae). Am. J. Bot. 98: 44-61.

Hunt, D. R. 2002. That's Opuntia, that was! Succ. Plant. Res. 6: 245-248.

Hunt, D. R. 2006. The New Cactus Lexicon. dh books, The Manse, Chapel Lane, Milbourne Port, UK. 526 p.

Karaca, M., Ince, A. G., Ay, S. T., Turgut, K., and Onus, A. N. 2008. PCR-RFLP and DAMD-PCR genotyping for Salvia species. J Sci Food Agric. 88: 2508-2516.

Khadivi-Khub, A., Jahangirzadeh, S., Ahadi, E., Aliyoun, S. 2014. Nuclear and chloroplast DNA variability and phylogeny of Iranian apples (*Malus domestica*). Plant Syst Evol DOI 10.1007/s00606-014-1007-y.

Khan, M. K., Pandey, A., Choudhary, S., Hakki, E. E., Akkaya, M. S., and Thomas, G. 2014. From RFLP to DArT: molecular tools for wheat (*Triticum* spp.) diversity analysis. Genet. Resour. Crop Evol. 61: 1001-1032.

Konieczny, A., and Ausubel, M. 1993. Procedure for mapping *Arabidopsis* mutations using codominant ecotype-specific PCR-based markers. Plant J. 4: 403-410.

Korotkova, N., Borsch, T., Quandt, D., Taylor, N. P., Müller, K. F., and Barthlott, W. 2011. What does it take to resolve relationships and to identify species with molecular markers? An example from the epiphytic rhipsalideae (cactaceae). Amer. J. Bot. 98: 1549-1572.

Kress, W. J., and Erickson, D. L. 2007. A two-locus global DNA barcode for land plants: The coding *rbcL* gene complements the noncoding *trn*H-*psb*A spacer region. PLoS One 2:e508.

Labra, M., Grassi, F., Bardini, M., Imazio, S., Guiggi, A., Citterio, S., Banfi, E., and Sgorbati, S. 2003. Relationships in *Opuntia* Mill. Genus (Cactaceae) detected by molecular marker. Plant Sci. 165: 1129-1136.

Leinemann, L., Kleinschmit, J., Fussi, B., Hosius, B., OKuchma, L., Arenhovel, W., Lemmen, P., Katzel, R., Rogge, M., Finkeldey, R. 2014. Genetic composition and differentiation of sloe (*Prunus spinosa* L.) populations in Germany with respect to the tracing of reproductive plant material. Plant Syst. Evol. DOI 10.1007/s00606-014-1027-7.

Lin, T. C., Yeh, M. S., Cheng, Y. M., Lina, L. C., Sung, J. M. 1012. Using ITS2 PCR-RFLP to generate molecular markers for authentication of *Sophora flavescens* Ait. J. Sci. Food Agric. 92: 892-898.

Liu, Y., Huo, N., Dong, L., Wang, Y., Zhang, S. 2013. Complete Chloroplast Genome Sequences of Mongolia Medicine *Artemisia frigida* and Phylogenetic Relationships with Other Plants. PLoS One 8:e57533.

Luna-Paez, A., Valadez-Moctezuma, E., Barrientos-Priego, F., and Gallegos-Vázquez, C. 2007. Characterization of *Opuntia* spp. by means of seed with RAPD and ISSR markers and its possible use for differentiation. J PACD 9: 43-59.

Majure, L. C., Judd, W. S., Soltis, P. S., and Soltis, D. E. 2012a. Cytogeography of the *Humifusa* clade of *Opuntia* s.s. Mill. 1754 (Cactaceae: Opuntioideae): Correlations with geographic distributions and morphological differentiation of a polyploid complex. Comparative Cytogenetics 6: 53-77.

Majure, L. C., Puente, R., Griffith, M. P., Judd, W. S., Soltis, P. S., and Soltis, D. E. 2012b. Phylogeny of *Opuntia* s.s. (Cactaceae): clade delineation, geographic origins, and reticulate evolution. Am. J. Bot. 99: 847-864.

Majure, L. C., Puente, R., and Pinkava, D. J. 2012c. Miscellaneous Chromosome Numbers in Opuntieae Dc. (Cactaceae) with a Compilation of Counts for the Group Source: Haseltonia 18: 67-78.

Marciniuk, P., Gawronska, B., Marciniuk, J., and Joachimiak, A. J. 2014. Taxonomic individuality of *Leonurus cardiaca* and *Leonurus quinquelobatus* in view of morphological and molecular studies. Plant Syst. Evol. 300: 255-261.

Mathews, S. 2010. Evolutionary studies illuminate the structural-functional model of plant phytochromes. Plant Cell 22: 4-16.

Melgarejo, P., Martínez, J. J., Martínez, R., Legua, P., Oncina, R., and Martínez-Murcia, A. 2009. Cultivar identification using 18S-28S rDNA intergenic spacer-RFLP in pomegranate (*Punica granatum* L.). Sci. Horti. 120: 500-503.

Morales, P., Ramírez-Moreno, E., Sanchez-Mata, M., Carvalho, M., and Ferreira, I. 2012. Nutritional and antioxidant properties of pulp and seeds of two xoconostle cultivars (*Opuntia joconostle* F.A.C. Weber ex Diguet and *Opuntia matudae* Scheinvar) of high consumption in Mexico. Food Res. Int. 46: 279-285.

Nock, C. J., Waters, D. L. E., Edwards, M. A., Bowen, S. G., Rice, N., Cordeiro, G. M., and Henry, R. J. 2011. Chloroplast genome sequences from total DNA for plant identification. Plant Biotech. J. 9: 328-333.

Pinkava, D. J. 2002. On the evolution of the North American Opuntioideae. In Hunt, D. R., and Taylor N. P. (eds) Studies in the Opuntioideae (Cactaceae). The Manse, Sherborne, UK, pp 59-98.

Poczai P, Cseh A, Taller J, Symon DE (2011) Genetic diversity and relationships in *Solanum* subg. *Archaesolanum* (Solanaceae) based on RAPD and chloroplast PCR-RFLP analyses. Plant Syst. Evol. 291: 35-47.

Poczai, P., and Hyvönen, J. 2010. Nuclear ribosomal spacer regions in plant phylogenetics: Problems and prospects. Mol. Biol. Rep. 37: 1897-1912.

Rebman, J.P., and Pinkava, D. J. 2001. *Opuntia* of North America: An overview. The Florida Entomologist 84: 474-483.

Ritz, C.M., Reiker, J., Charles, G., Hoxey, P., Hunt, D., Lowry, M., Stuppy, W., and Taylor, N. 2012. Molecular phylogeny and character evolution in terete-stemmed Andean opuntias (Cactaceae, Opuntioideae). Mol. Phylogenet. Evol. 65: 668-681.

Samah, S., and Valadez-Moctezuma, E. 2014. Morphological Seeds Descriptors for Characterize and Differentiate Genotypes of *Opuntia* (Cactaceae, Opuntioideae). Annual Research and Review in Biology 4: 3791-3809.

Sang, T., Crawford, D. J., Stuessy, T.F. 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *paeonia* (paeoniaceae). Am J, Bot. 84: 1120-136.

Sarin, B., Clemente, J., Mohanty, A. 2013. PCR-RFLP to distinguish three *Phyllanthus* sp., used in herbal medicines. South Afr. J. Bot. 88: 455-458.

Schlüter, P. M., Stuessy, T. F., Paulus, H. F. 2005. Making the first step: practical considerations for the isolation of low-copy nuclear markers. Taxon 54: 766-770.

Shaw, J., Lickey, E. B., Beck, J. T., Farmer, S. B., Liu, W., *et al.* 2005. The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. Am. J. Bot. 92: 142-166.

Steele, P. R., Guisinger-Bellian, M., Linder, C. R., and Jansen, R. K. 2008. Phylogenetic utility of 141 low-copy nuclear regions in taxa at different levels in two distantly related families of rosids. Mol. Phylogenet. Evol. 48: 1013-1026.

Stuppy, W. 2002. Seed characters and the generic classification of the Opuntioideae (Cactaceae). In Hunt, D.R., Taylor, N.P. (eds) Studies in the Opuntioideae (Cactaceae). The Manse, Chapel Lane, Milborne Port Sherborne, UK, pp 25-58.

Taberlet, P., Gielly, L., Pautou, G., Bouvet, J. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol. Biol. 17: 1105-1109.

Tsumura, Y., Suyama, Y., Yoshimura, K. 2000. Chloroplast DNA inversion polymorphism in populations of *Abies* and *Tsuga*. Mol. Biol. Evol. 17:1302-1312.

Valadez-Moctezuma, E., Samah, S., and Luna-Paez, A. 2014. Genetic diversity *of Opuntia* spp. varieties assessed by classical marker tools (RAPD and ISSR). Plant Syst Evol DOI: 10.1007/s00606-014-1112-y.

Wallace, R. S., and Dickie S. L. 2002. Systematic implications of chloroplast DNA sequences variation in the Opuntioideae. In Hunt, D. R., and Taylor N. P. (eds) Studies in the Opuntioideae (Cactaceae). The Manse, Sherborne, UK, pp 9-24.

White, T. J., Bruns, T., Lee, S., and Taylor, J. W. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ (eds) PCR protocols: A guide to methods and applications. Academic Press, San Diego, California, USA, pp 315-322.

Whitlock, B. A., Hale, A. M., and Groff, P. A. 2010. Intraspecific inversions pose a challenge for the *trn*H-*psb*A plant DNA barcode. Plos One 5: e11533.

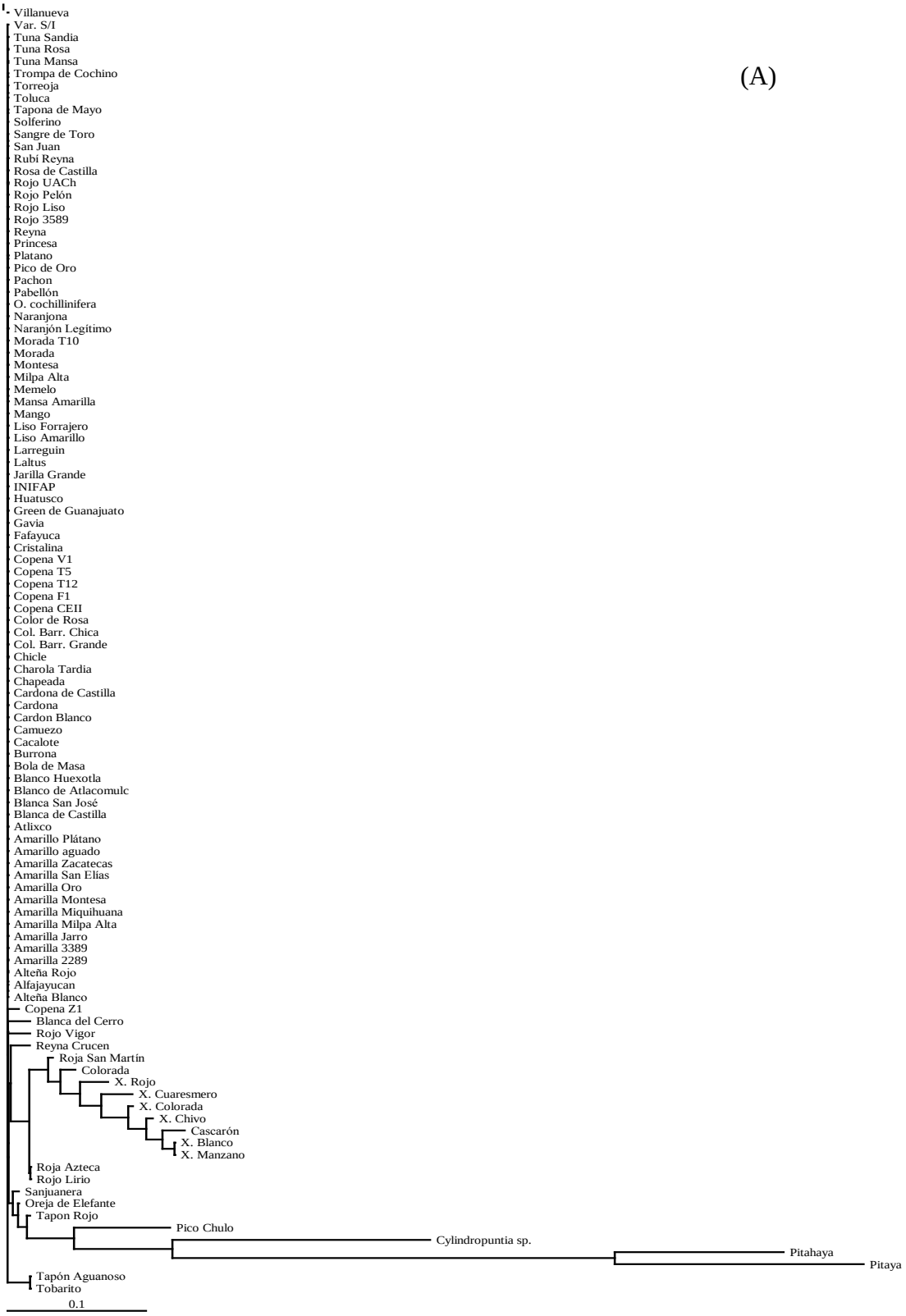
Yesson, C., Bárcenas, R. T., Hernández, H. M, Ruiz-Maqueda, M., Prado, A., Rodríguez, V.M., and Hawkins, J. A. 2011. DNA barcodes for Mexican Cactaceae, plants under pressure from wild collecting. Mol. Ecol. Resour. 11: 775-783.

Zhu, S., Bai, Y., Oya, M., Tanaka, K., Komatsu, K., Goda, Y., Kawasaki, T., Fujita, M., Shibata, T. 2011. Genetic and chemical diversity of *Eleutherococcus senticosus* and molecular identification of Siberian ginseng by PCR-RFLP analysis based on chloroplast trnK intron sequence. Food Chem. 129: 1844-1850.

Zimmer, E. A., and Wen, J. 2012. Review. Using nuclear gene data for plant phylogenetics: Progress and prospects. Mol. Phylogenet. Evol. 65: 774-785.

Supplementary data S1 Opuntia accessions and their corresponding species used in this study

N	Accession	Species	N	Accession	Species
1	Alfajayucan	*O. albicarpa* Scheinvar	54	Mansa Amarilla	*Opuntia* sp.
2	Alteña Blanco	*Opuntia* sp.	55	Memelo	*O. affinis hyptiacantha*
3	Alteña Rojo	*Opuntia* sp.	56	Milpa Alta	*O. ficus-indica* (L.) Mill.
4	Amarilla 2289	*Opuntia* sp.	57	Montesa	*Opuntia* sp.
5	Amarilla 3389	*Opuntia* sp.	58	Morada	*O. megacantha* Salm-Dyck.
6	Amarilla Jarro	*O. megacantha* Salm-Dyck.	59	Morada T10	*O. megacantha* Salm-Dyck.
7	Amarilla Milpa Alta	*O. ficus-indica* (L.) Mill.	60	Naranjón Legítimo	*O. albicarpa* Scheinvar
8	Amarilla Miquihuana	*O. lasiacantha* Pfeiffer	61	Naranjona	*Opuntia* sp.
9	Amarilla Montesa	*O. megacantha* Salm-Dyck.	62	*O. cochillinifera*	*O. cochillinifera*
10	Amarilla Oro	*O. albicarpa* Scheinvar	63	Oreja de Elefante	*O. undulata* Griffiths
11	Amarilla San Elías	*Opuntia* sp.	64	Pabellón	*O. ficus-indica* (L.) Mill.
12	Amarilla Zacatecas	*O. megacantha* Salm-Dyck.	65	Pachon	*Opuntia* sp.
13	Amarillo Aguado	*Opuntia* sp.	66	Pico Chulo	*O. megacantha* Salm-Dyck.
14	Amarillo Plátano	*O. megacantha* Salm-Dyck.	67	Pico de Oro	*Opuntia* sp.
15	Atlixco	*O. ficus-indica* (L.) Mill.	68	Platano	*Opuntia* sp.
16	Blanca de Castilla	*Opuntia* sp.	69	Princesa	*Opuntia* sp.
17	Blanca del Cerro	*Opuntia* sp.	70	Reyna	*O. albicarpa* Scheinvar
18	Blanca San José	*O. albicarpa* Scheinvar	71	Reyna Crucen	*Opuntia* sp.
19	Blanco de Atlacomulco	*Opuntia* sp.	72	Roja Azteca	*O. megacantha* Salm-Dyck.
20	Blanco Huexotla	*Opuntia* sp.	73	Roja San Martín	*O. megacantha* Salm-Dyck
21	Bola de Masa	*O. albicarpa* Scheinvar	74	Rojo 3589	*Opuntia* sp.
22	Burrona	*O. albicarpa* Scheinvar	75	Rojo Lirio	*O. megacantha* Salm-Dyck.
23	Cacalote	*O. cochinera* Griffits	76	Rojo Liso	*Opuntia* sp.
24	Camuezo	*O. megacantha* Salm-Dyck.	77	Rojo Pelón	*O. ficus-indica* (L.) Mill.
25	Cardon Blanco	*O. streptacantha* Lem.	78	Rojo UACh	*Opuntia* sp.
26	Cardona	*O. streptacantha* Lem.	79	Rojo Vigor	*O. ficus-indica* (L.) Mill.
27	Cardona de Castilla	*O. streptacantha* Lem.	80	Rosa de Castilla	*O. megacantha* Salm-Dyck.
28	Cascarón	*O. chaveña*	81	Rubí Reyna	*O. megacantha* Salm-Dyck.
29	Chapeada	*O. albicarpa* Scheinvar	82	San Juan	*Opuntia* sp.
30	Charola Tardia	*O. hyptiacantha* Lem.	83	Sangre de Toro	*Opuntia* sp.
31	Chicle	*O. ficus-indica* (L.) Mill.	84	Sanjuanera	*O. lasiacantha* Pfeiffer
32	Col. Barr. Grande	*Opuntia* sp.	85	Solferino	*Opuntia* sp.
33	Col. Barr. Chica	*Opuntia* sp.	86	Tapón Aguanoso	*O. robusta* H.L. Wendland
34	Color de Rosa	*O. albicarpa* Scheinvar	87	Tapon Rojo	*O. robusta* H.L. Wendland
35	Colorada	*Opuntia* sp.	88	Tapona de Mayo	*O. robusta* H.L. Wendland
36	Copena CEII	*O. ficus-indica* (L.) Mill.	89	Tobarito	*Opuntia* sp.
37	Copena F1	*O. ficus-indica* (L.) Mill.	90	Toluca	*Opuntia* sp.
38	Copena T12	*O. ficus-indica* (L.) Mill.	91	Torreoja	*O. megacantha* Salm-Dyck.
39	Copena T5	*O. ficus-indica* (L.) Mill.	92	Trompa Cochino	*Opuntia* sp.
40	Copena V1	*O. ficus-indica* (L.) Mill.	93	Tuna Mansa	*O. albicarpa* Scheinvar
41	Copena Z1	*O. albicarpa* Scheinvar	94	Tuna Rosa	*O. albicarpa* Scheinvar
42	Cristalina	*O. albicarpa* Scheinvar	95	Tuna Sandia	*Opuntia* sp.
43	Fafayuca	*O. albicarpa* Scheinvar	96	Var._S/I	*Opuntia* sp.
44	Gavia	*O. albicarpa* Scheinvar	97	Villanueva	*O. albicarpa* Scheinvar
45	Green de Guanajuato	*Opuntia* sp.	98	X. Blanco	*O. joconostle* F.A.C. Weber
46	Huatusco	*Opuntia* sp.	99	X. Chivo	*Opuntia* sp.
47	INIFAP	*Opuntia* sp.	100	X. Colorada	*O. joconostle* F.A.C. Weber
48	Jarilla Grande	*Opuntia* sp.	101	X. Cuaresmero	*O. matudae* Scheinvar
49	Laltus	*Opuntia* sp.	102	X. Manzano	*O. joconostle* F.A.C. Weber
50	Larreguin	*O. ficus-indica* (L.) Mill.	103	X. Rojo	*Opuntia* sp.
51	Liso Amarillo	*Opuntia* sp.	104	*Cylindropuntia* sp.	*Cylindropuntia* sp.
52	Liso Forrajero	*Opuntia* sp.	105	Pitahaya	*Hylocereus undatus*
53	Mango	*O. albicarpa* Scheinvar	106	Pitaya	*Stenocereus thurberi*



(A)

Supplementary data S2-A, B and C Phylogenetic relationships in Mexicans Opuntia. Dendrograms constructed using the NJ method applying Nei and Li/Dice coefficient. Data were resolved from PCR-RFLP analyses, (A) chloroplast regions, (B) mitochondrial regions and (C) nucleus gene.

Supplementary data S2. Continued

Supplementary data S2. Continued

(C)

Chapter 4

Part 3

DNA marker (CAPS)

CAPS ANALYSIS OF FRUITFULL-HOMOLOGUE GENE IN OPUNTIA ACCESSIONS

Samir Samah and Ernestina Valadez-Moctezuma

Article will submit: Fruits

CAPS ANALYSIS OF FRUITFULL-HOMOLOGUE GENE IN OPUNTIA ACCESSIONS

ABSTRACT

Opuntia species, commonly known as "nopal" in Mexico, are interesting economic and agronomic fruit and vegetable crop. Opuntia fruits ("xoconostles" and "tunas") to have a great variability of sizes and colors; but the genetic base of this morphologic variability still unstudied. In this study, the genetic diversity among 91 Opuntia genotypes was studied by CAPS technique using the FRUITFULL gene with the primer pair Fruitfull-R1218 and Fruitfull-F131 developed previously in *Ficus carica*. The PCR products were digested with *Alu*I, *Ava*II, *Hae*III, *Mbo*I, *Hha*I, *Hinf*I and *Rsa*I restriction enzymes and separated on 6 % polyacrylamide gels. A UPGMA dendrogram was constructed based on Dice similarity coefficient. As results, the primers used were considered adequate for the detection of the presence of the FRUITFULL gene in Opuntia cultivars. PCR products were approximately of 850 pb and the number of fragments digested varied between four and six depending on the enzyme. Only the enzymes *Alu*I, *Hae*III and *Mbo*I generated different band patterns; thus, they were considered informative and polymorphic. UPGMA analysis revealed that the most of the genotypes contain similar genetic profile; thus, 69 % of the accessions included were shown to have the same genetic profile of the FRUITFULL gene, indicating the lower rate of mutation occurred in this region of the genome. The CAPS method was successfully applied since it is simple, has a low cost, and is quicker than other molecular techniques and/or morphological differentiation.

Keywords: Nopal, fruit size, genetic diversity, differentiation.

INTRODUCTION

Mexico is known for its natural richness of several types of plants, of which the "nopal" (*Opuntia* sp.) reflects a great generic and phenotypic diversity; thus, Mexico is considered the center of origin and diversification of this crop (Griffith, 2004). Opuntia genotypes play an important role in agricultural economies throughout arid and semi-arid areas. This genus is becoming an alternative fruit crop for semi-arid areas of the world. The main producing of Opuntia are Mexico, Spain, Italy, Morocco, Algeria, Egypt, Saudi Arabia, Palestine, Chile, Brazil, Turkey and South of Africa countries. It takes advantage of their fresh fruits ("tunas" or sweet fruits and "xoconostles" or acidic fruits) and their tender cladodes ("nopalitos") as a vegetable. Fresh and processed fruits are considered as an important source of vitamins, minerals and carbohydrates; moreover, derived juices, jams and syrups among other byproducts (Rosas-Cardenas *et al*., 2007). The "tunas" and "xoconostles" show ample morphological variations in size, color, sugar levels, skin and pulp size. However, the genetic basis of this high morphological variation quiets unknown.

Generally, Opuntia collections with commercial agronomic value are described only from a morphological and/or anthropocentric point of view (Reyes-Agüero *et al.*, 2013). Nevertheless, the morphological, anatomical and chemical analyses used to detect and quantify the variability are affected by environmental factors and/or growth factors (Rubiolo *et al.*, 2009; Bertea and Gnavi, 2012). Thus, strengthening the characterization of Mexican "nopal" with alternative tools would reveal more efficiently the existing phenetic and genetic diversity, contributing to breeding programs for the creation of new varieties and enrichment of the germplasm banks.

DNA molecular methods have several advantages over classical morphological and chemical analyses. The genetic method requires genotype instead than phenotype; therefore PCR-based techniques have been widely used for a rapid identification of genetic diversity. Detection of the PCR products is performed by agarose and/or acrylamide gel electrophoresis and the polymorphisms can be detected using restriction digestion and comparison of the products size with a DNA size marker (Rubiolo *et al.,* 2009). The bands on gels, which typically serve as molecular markers, may arise from cutting DNA at specific sites with restriction enzymes, in order to detect restriction fragment length polymorphisms (RFLPs). Cleaved Amplified Polymorphic Sequences (CAPS) deciphers the RFLPs caused by single base changes like SNPs, insertions/deletions, which modify restriction endonuclease recognition sites in PCR amplicons (Konieczny and Ausubel, 1993). The CAPS assays are performed by digesting locus-specific PCR amplicons with one or more restriction enzyme, followed by separation of the digested DNA on agarose or polyacrylamide gels (Abe *et al.,* 2013). The CAPS markers are codominants and locus specific and have been used to distinguish between plants. The method is simple, relatively inexpensive. This technique has been considered a useful technique for the distinction of closely related genotypes and considered as a rapid and precise method for plant identification (Bertea and Gnavi, 2012). With this method, great diversities can be found in the genes and non-coding regions; this allows the design of specific primers (Gnavi *et al.,* 2010).

For Opuntia, the use of DNA marker has successfully demonstrated to study the genetic diversity and to characterize genotypes; i.e. ISSR (Valadez-Moctezuma *et al.,* 2014a), RAPD (Bendhifi *et al.,* 2013), SSR (Helsen *et al.,* 2009; Caruso *et al.,* 2010).

Recently, the CAPS of the gene FRUITFULL has been used by to study genetic diversity in *Ficus carica* (García-Ruiz *et al.*, 2013).

The aim of this research was to amplify FRUITFULL-homologous gene, and to study the genetic diversity of this gene in ninety-one accessions of Opuntia using the CAPS technique.

MATERIALS AND METHODS

Ninety-one Opuntia genotypes; and two outgroups: pitaya (*Stenocereus stellatus*) and *Cylindropuntia* sp., were studied. The accessions were obtained from the germplasm banks of Crucen-UACh (Zacatecas) and Nopalera-UACh (Texcoco), Mexico. The *O. spunilifera*, *O. velutina*, *O. leucotricha*, *Nopalera karwinskiana* samples were provided by the botanical garden, UNAM, Mexico, which is sincerely grateful (Table 1).

The genomic DNA was extracted according to Luna-Paez *et al.* (2007). The DNA quantification was estimated by spectrophotometry (ND-1000, Thermo scientific, USA) and its quality was determined in 1 % agarose gel. The applied voltage was 90 volts for 60 min in 1X TAE buffer (Trisbase, glacial acetic acid, 0.5 M EDTA pH 8.0). DNA staining was performed with ethidium bromide 0.5 $\mu g.mL^{-1}$ and the gels were documented with Molecular Imager® ChemiDoc™ XRS system (BIO RAD, Japon).

One forward (Fruitfull-F131; 5’ ATGGGAAGAGGTAGGG 3’) and two reverse primers (Fruitfull-R1218; 5’ GATACTTGAACGCTATGAT 3’) and Fruitfull-R3056; 5’ GATTAAGGAGAGGGAGAAG 3’) were used. These primers were designed on the basis of *Arabidopsis thaliana* gene and were chosen from sequences highly conserved across

plant species (García-Ruiz *et al.*, 2013). PCR was carried out in thermal cycler (MaxyGene Gradient Axygen, Brazil).

Table 1. List of Opuntia accessions studied.

Nº	Accession	Nº	Accession	Nº	Accession
1	Alfajayucan	32	Copena Z1	63	Reyna
2	Alteña Blanco	33	*Cylindropuntia* sp.	64	Reyna Crucen
3	Alteña Rojo	34	Fafayuca	65	Roja Azteca
4	Amarilla 2289	35	Gavia	66	Roja San Martín
5	Amarilla Milpa Alta	36	Green de Guanajuato	67	Rojo 3589
6	Amarilla Miquihuana	37	INIFAP	68	Rojo dulce
7	Amarilla Monteza	38	Jarilla Grande	69	Rojo Lirio
8	Amarilla Oro	39	Laltus	70	Rojo Liso
9	Amarilla Plátano	40	Liso Amarillo	71	Rojo Pelón
10	Amarilla San Elias	41	Liso Forrajero	72	Rojo UACh
11	Amarillo Aguado	42	Mango	73	Rojo Vigor
12	Blanca Castilla	43	Memelo	74	Rosa Castilla
13	Blanco Atlacomulco	44	Milpa alta	75	Rubí Reyna
14	Blanco Huexotla	45	Montesa	76	San Juan
15	Bola Masa	46	Morada	77	Sangre Toro
16	Burrona	47	Morada T10	78	Sanjuanera
17	Cacalote	48	Naranjón Legítimo	79	Solferino
18	Camuezo	49	Naranjona	80	Tapón Aguanoso
19	Cardón Blanco	50	Nopalera karwinskiana	81	Tapón Rojo
20	Cardona	51	*O. cochillinifera*	82	Tapona Mayo
21	Cardona Castilla	52	*O. spunilifera*	83	Tobarito
22	Cascarón	53	*O. velutina*	84	Toluca
23	Chapeada	54	*O. leucotricha*	85	Torreoja
24	Charola Tardía	55	Oreja Elefante	86	Trompa Cochino
25	Chicle	56	Pabellón	87	Tuna Mansa
26	Colección Barrientos	57	Pachón	88	Tuna Rosa
27	Color Rosa	58	Pico Chulo	89	Tuna Sandia
28	Colorada	59	Pico Oro	90	Villanueva
29	Copena CE II	60	Pitaya	91	X. Blanco
30	Copena T12	61	Princesa	92	X. Chivo
31	Copena T5	62	Red Villa Puebla	93	X. Manzo

Components of the PCR mix (25 μL) were as follows: 100 ng genomic DNA, 1 mM $MgCl_2$, 0.25 mM dNTPs, 0.4 μM primer, 1X ThermoPol buffer and 1 U Taq DNA polymerase (Invitrogen). PCR cycling conditions were: 95°C for 5 min for initial denaturation; 35 cycles at 95 °C for 1 min, 40/45 °C for 45 s and 68 °C for 3 min and 15 s; and 68 °C for 5 min for final extension. Undigested PCR products were mixed with 10 μL 6X bromophenol-blue loading dye. A volume of 8 μL mix was loaded on a 1.2 % agarose

gel. A 1 Kb DNA ladder (Promega) was used as molecular weight markers. Gels electrophoresis and staining were realized as above.

PCR products were digested with *Alu*I, *Ava*II, *Hae*III, *Mbo*I, *Hha*I, *Hinf*I and *Rsa*I restriction enzymes at 37°C for overnight in a final volume of 20 µL. Digested PCR products were mixed with 10 µL 6X bromophenol-blue loading dye with formamide. A volume of 10 µL mix was loaded on 8 % polyacrylamide gel in 1X TBE buffer (89 mM Tris-borate, 2 mM EDTA pH 8.0) using a vertical Dual TMVG-216-33 gel rig (C.B.S. Scientific, USA), for 90 min at 240 V. A 100 pb DNA ladder and 1 Kb DNA ladder (Promega) were used as standard molecular weight markers. The gels were stained with silver nitrate solution (AgNO3) 0.2 % to reveal DNA fragments. Finally, the gels were photodocumented with a Kodak Digital Camera

The electrophoresis pattern obtained after restriction digestion was considered as qualitative character and visually coded as absent ("0") or present ("1") in each genotype. Binary matrice was constructed and the genetic similarity was calculated between pairs of accessions using the Dice coefficient. The cluster analysis was performed using the arithmetic averages with the unweighted pair group method (UPGMA), considering all enzymes restrictions together. The analysis was performed with FreeTree software Version 0.9.1.50. The cophenetic correlation coefficient was computed using NTSYSpc 2.2.

RESULTS AND DISCUSSION

To date, the characterization of Opuntia genotypes was based on morphological characters, mainly cladode and fruit (Gallegos-Vázquez *et al.* 2012; Reyes-Agüero et *al.*, 2013). Recently, molecular characterization has been has won more attention and it has

been successfully applied in Opuntia; mainly the RAPD and ISSR techniques were used to differentiate and characterize varieties of "nopal" in Mexico (Luna-Paez *et al.,* 2007; Valadez-Moctezuma *et al.,* 2014a; b). Analysis at the molecular level of ripening-related genes of *Opuntia ficus-indica* was reported by Collazo-Siqués *et al.* (2003), and Prickly pear polygalacturonase gene by Rosas-Cárdenas *et al.* (2007). The expression of the *Opacco-1* and *Opaccs-1* ripening genes showed correlation with the high content of sugars and flavonoids contents of prickly pear cultivars (El-Shehawi *et al.,* 2013). However, no information is available on the genetic diversity of genes related to fruit characteristics, neither other genes. However, in the model plant *Arabidopsis thaliana*, 27 genes are responsible for the development of fruit characteristics (Roeder and Yanofsky, 2006). FRUITFULL gene is a member of the extended family of MADS-box genes and it is strongly expressed in the fruit valves. In *arabidopsis*, loss of FRUITFULL expression results in a severe reduction in fruit size (Gu *et al.,* 1998). FRUITFULL is required for the expansion and differentiation of fruit valves after fertilization (Roeder and Yanofsky, 2006). FRUITFULL family genes are important regulators of ripening in fleshy fruits (Bemer *et al.,* 2012; Shima *et al.,* 2013). An FRUITFULL like gene is associated with genetic variation for fruit flesh firmness in apple (Cevik *et al.,* 2010). Interestingly, the gene FRUITFULL is highly conserved in plants (García-Ruiz *et al.,* 2013). This study was undertaken for purposes of determining the level of conservation of this gene between genotypes.

Thus, we hypothesized that Opuntia has an FRUITFULL homolog. To test this hypothesis, we successfully amplified fragments of 850 bp for all genotypes studied of Opuntia with the primer pair Fruitfull-F131-Fruitfull-R1218; while, for pitaya sample a fragment of 950 bp was obtained. The weight of the FRUITFULL gene in common fig

(*Ficus carica* L.) amplified with the same primer pair was approximately of 900 bp (García-Ruiz *et al.*, 2013). There was no length variation observed among the undigested PCR products from the 93 accessions. The amplified product from all the samples had a very similar molecular size (Fig. 1).

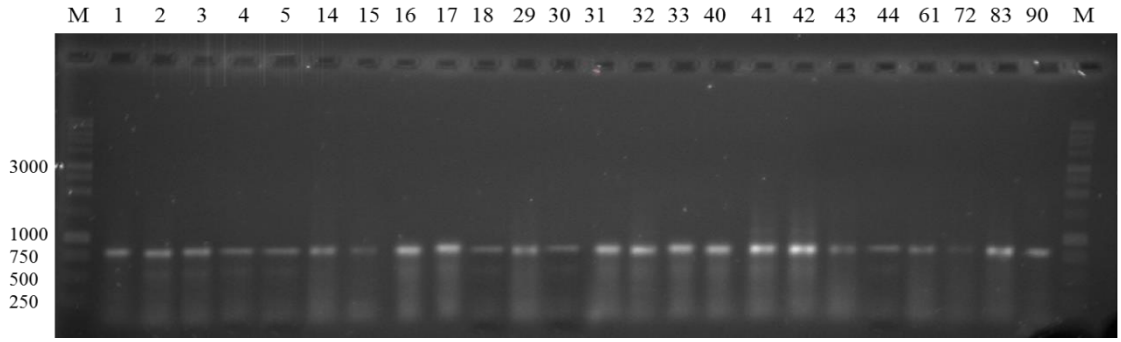


Fig. 1. Electrophoresis patterns of undigested PCR products run on 1.2 % agarose gel. M: 1 Kb DNA ladder markers. Lanes number corresponds to the genotypes referred in Table 1. The gene Fruitfull-homologue amplicons were approximately 850 pb.

The enzymes *Hinf*I and *Rsa*I didn't have cleavage site in the amplified fragment; whereas, the enzymes *Ava*II and *Hha*I didn't generated any polymorphism between the genotypes. However, *Alu*I, *Hae*III and *Mbo*I generated different band patterns; thus, they were considered informative. The endonuclease *Alu*I generated six distinct fragments (400 bp, 350 or 300 bp, 170 or 160 bp and 100 bp); while the endonuclease *Hae*III generated four bandas (500 or 520 bp, 400 bp and 330 bp); in turn, *Mbo*I generated five fragments 550 pb, 510 or 500 bp, 200 and 100 bp.

On the basis of the electrophoresis pattern observed from sets of digested PCR products (Fruitfull-F131 and Fruitfull-R1218 primer pair and the polymorphic endonuclease *Alu*I, *Hae*III, *Mbo*I), the Dice similarity coefficient using UPGMA was estimated. The assessed genetic distance ranged from 0.27 to 1 (outgroups included) and from 0.46 to 1 (outgroups excluded). However, the most genotypes studied showed narrow

genetic distance ranging from 0.75 to 1. The UPGMA dendrogram is illustrated in the Fig. 2. The cophenetic correlation coefficient between the original similarity matrix and the cophenetic matrix derived from the dendrogram was high (0.998), indicating a good fit between the dendrogram and the similarity matrix. This analysis revealed the most of the genotypes had a similar genetic profile; thus, 69 % (63 genotypes) of the accessions included in this study were shown to have the same genetic profile of the FRUITFULL gene (Fig. 2), indicating the lower rate of mutation occurred in this region of the genome, It is possible to infer that there is a great sequence similarity between these genotypes. However, the genotypes "Amarilla San Elías", "Colección Barrientos", "Roja Azteca", "Alteña Blanco", "Tapona de Mayo", "Oreja de Elefante", "Naranjona", "Torreoja" were grouped into separate clusters. This indicates that these genotypes present some variation in the sequence of the gene FRUITFULL. These genotypes can be characterized by their RFLP profiles generated by the combined enzymes *Alu*I, *Hae*III and *Mbo*I. Likewise, three clusters grouped two genotypes with the same genetic profile each one ("Amarilla Oro" and "Solferino"; "Amarilla Miquihuana" and "*O. leucotricha"*; "Rojo Vigor" and "Tobarito", respectively), indicating the high sequence similarity between these three pairs of genotypes. The genotypes "Xoconostle Blanco", "Rojo Lirio", "Amarilla Monteza","Amarilla Milpa Alta", "Amarilla 2289" and "Alfajayucan" presented the same profile of the FRUITFULL gene being located in a same cluster. Similar results from the genotypes "Cacalote", "Camuezo", "Jarilla Grande", "Laltus", "Liso Forrajero", "Pachón", "Pico Chulo", "Pico de Oro" were obtained.

In conclusion, molecular techniques are required to quickly and precisely characterize and certify different cultivated varieties of Opuntia. In this study, the procedure adopted was based on the restriction fragment length polymorphisms of the FRUITFULL gene. The

results proved that the most cultivars were maintained similar FRUITFULL-homologue gene according to their genetic profiles.

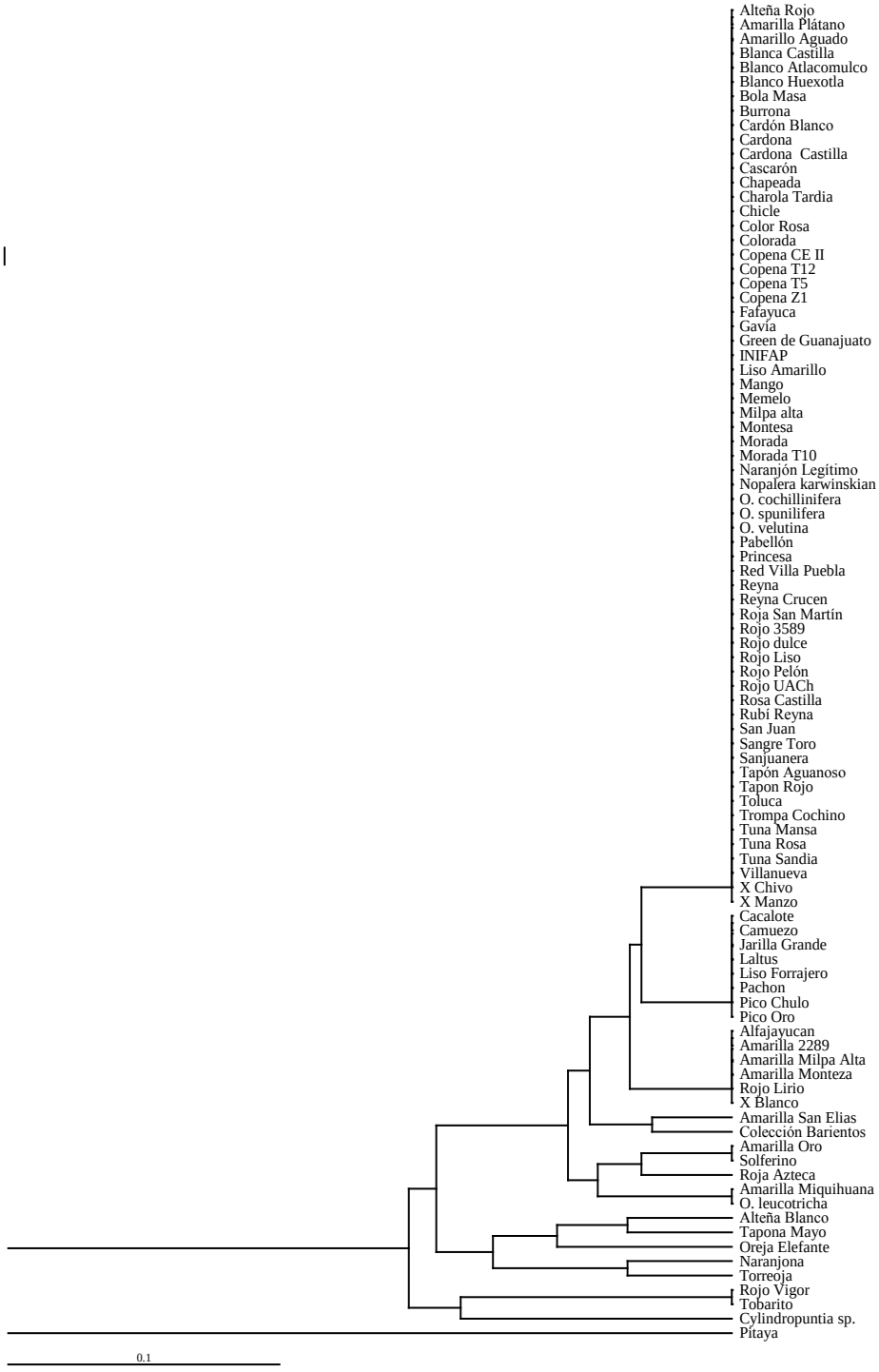


Fig. 2. UPGMA dendrogram demonstrating the relationships among 91 genotypes of Opuntia based on CAPS technique of the FRUITFULL-homologue gene. Cylindropuntia and Pitaya are outgroups.

REFERENCES

Abe, M., Kobayashi, M., Fujiyoshi, E., Tamaki, M., Kikuchi, N. and Murase, N. 2013. Use of PCR-RFLP for the discrimination of Japanese *Porphyra* and *Pyropia* species (Bangiales, Rhodophyta). J. Appl. Phycol. 25: 225-232.

Bendhifi, M., Baraket, G., Zourgui, L., Souid, S., Salhi-Hannachi, A. 2013. Assessment of genetic diversity of Tunisian Barbary fig (*Opuntia ficus indica*) cultivars by RAPD markers and morphological traits. Sci. Hortic. 158: 1-7.

Bertea, C. M. and Gnavi, G. 2012. Restriction Fragment Length Polymorphism of the 5S-rRNA-NTS Region: A Rapid and Precise Method for Plant Identification. In Nikolaus, J. S. *et al.* (eds.), Plant DNA Fingerprinting and Barcoding: Methods and Protocols. Methods in Molecular Biology 862: 89-101.

Caruso, M., Currò, S., Las Casas, G., La Malfa, S., Gentile, A. 2010. Microsatellite markers help to assess genetic diversity among *O. ficus-indica* cultivated genotypes and their relation with related species. Plant Syst. Evol. 290: 85-97.

Cevik, V., Ryder, C. D., Popovich, A., Manning, K., King, G. J., and Seymour, G. B. 2010. A FRUITFULL-like gene is associated with genetic variation for fruit flesh firmness in apple (*Malus domestica* Borkh.). Tree Genetics and Genomes 6: 271-279.

Collazo-Siqués, P., Valverde, M. E., Paredes-López, O., and Guevara-Lara, F. 2003. Expression of full mature-related genes in Prickly Pears (*Opuntia* sp.) fruits. Plant foods Hum. Nutr. 58: 317-326.

El-Shehawi, A. M., Nagaty, M. A., and Fahmi, A. I. 2013. Gene expression profiling and fruit quality during ripening stages of Prickly pear (*Opuntia ficus-indica*) in Taif. Life Sci 10: 1267-1278.

Gallegos-Vázquez, C., Scheinvar, L., Núñez-Colín, C. A., and Mondragón-Jacobo, C. 2012. Morphological diversity of xoconostles (*Opuntia* spp.) or acidic cactus pears: a Mexican contribution to functional foods. Fruits 67: 109-120.

García-Ruiz, M. T., Mendoza-Castillo, V. M., Valadez-Moctezuma, E., and Muratalla-Lúa, A. 2013. Initial assessment of natural diversity in Mexican fig landraces. Genet. Mol. Res. 12: 3931-3943.

Gnavi, G., Bertea, C. M., and Maffei, M. 2010. PCR, sequencing and PCR-RFLP of the 5S-rRNANTS region as a tool for the DNA fingerprinting of medicinal and aromatic plants. Flavour. Frag. J. 25:132–137.

Griffith, M. P. 2004. The origins of an important cactus crop, *Opuntia ficus-indica* (Cactaceae): New molecular evidence. Am. J. Bot. 91: 1915-1921.

Gu, Q., Ferrándiz, C., Yanofsky, M. F., and Martienssen, R. 1998. The FRUITFULL MADS-box gene mediates cell differentiation during Arabidopsis fruit development. Development 125: 1509-1517.

Helsen, P., Browne, R., Anderson, D., Verdyck, P., and Van Dongen, S. 2009. Galápagos' *Opuntia* cacti: extensive morphological diversity, low genetic variability. Biol. Linn. Soc. 96: 451-461.

Konieczny, A., and Ausubel, F. 1993. A procedure for mapping *Arabidopsis* mutations using co-dominant ecotypespecific PCR-based markers. Plant J. 4: 403-410.

Luna-Paez, A., Valadez-Moctezuma, E., Barrientos, P. A. F., Gallegos, V. C. 2007. Characterization of *Opuntia* spp. by means of seed with RAPD and ISSR markers and its possible use for differentiation. J PACD 9: 43-59.

Reyes-Agüero, J. A., Aguirre, R. J. R., Carlín, C. F., and González A., D. A. 2013. Diversity of wild and cultivated *Opuntia* variants in the meridional highlands plateau of Mexico. Acta Hortic. 995: 69-74.

Roeder, A. H. K., and Yanofsky, M. F. 2006. Fruit Development in *Arabidopsis.* The Arabidopsis Book 4. 51p.

Rosas-Cárdena, F. F., Valderrama-Cháirez, M. L., Cruz-Hernández, A. and Paredes-López, O. 2007. Prickly pear polygalacturonase gene: cDNA cloning and transcript accumulation during ethylene treatment, cold storage and wounding. Posth. Biol. Techn. 44: 254-259.

Rubiolo, P., Matteodo, M., Bicchi, C. *et al.* 2009. Chemical and biomolecular characterization of *Artemisia umbelliformis* Lam., an important ingredient of the alpine liqueur "Genepì". J. Agr. Food Chem. 57: 3436-3443.

Valadez-Moctezuma, E., Ortiz-Vásquez, Q., Samah, S. 2014b. Molecular based assessment of genetic diversity of xoconostle accessions (*Opuntia* spp.). Afr. J. Biotech. 13: 202-210.

Valadez-Moctezuma, E., Samah, S., and Luna-Paez, A. 2014a. Genetic diversity *of Opuntia* spp. varieties assessed by classical marker tools (RAPD and ISSR). Plant Syst. Evol. DOI: 10.1007/s00606-014-1112-y

Chapter 5

GENERAL DISCUSSION

GENERAL DISCUSSION

Until now, relationships within the Cactaceae are insufficiently understood, and fewer molecular phylogenetic studies have been conducted; thus, the unstable generic limits and constant movement of species between Cactaceae genera have resulted in instability of names. The Opuntioideae differs from all other cacti in having glochids (small, barbed, and deciduous spines) and seeds completely enwrapped by a funicular stalk, which becomes hard and bony. The largest genus in this subfamily is *Opuntia*, and in its very broad sense numbers perhaps 200 species. However, Opuntia has been reduced drastically in size with many segregate genera now documented. Members of *Opuntia* s.s. are cultivated worldwide as fruit and vegetable crops and are increasingly used as forage and fodder for livestock in arid areas of the world. In Mexico, where species of *Opuntia* have been cultivated for at least 14 000 yr, they represent an iconic national figure. The cactus pads ("nopalitos") and the prickly pear as a seasonal fruit: sweet ("tunas") or acid fruits ("xoconostle") were well expended. Moreover, Opuntia is known for its difficult taxonomy. The continuous morphological variation in the genus, the synonyms and the insufficient and inadequate morphological descriptors, high frequency of polyploidy levels, all have led to misclassifications (Labra *et al.*, 2003; Caruso *et al.*, 2010, Valadez-Moctezuma *et al.*, 2014). Also, the genetic divergence between "xoconostles" and other *Opuntia* species ("tunas") is quite unknown.

In extension, a complete characterization of cultivars is needed in order to remove incidents of mislabeling or synonymy. Before the present work, only few genotypes have been fingerprinted genetically; while in the present study the genetic structure of an extensive Mexican Opuntia collection was exploited and relationships among entries were revealed. Such information will aid the selection of cultivars for germplasm conservation

and implementation in breeding programs, providing information of diverse genetic backgrounds in cultivated, while monitoring the trade of plant material. Therefore, the three constituent parts of this thesis illustrate the usefulness of various types of markers (morphological, biochemical and molecular) to study the genetic diversity and to clarify the taxonomic status of Opuntia genotypes with greater economic and agricultural importance.

The currently accepted classification is based on morphology. A drawback to a morphology-based classification is the plasticity of morphological characteristics. In addition, the continuous nature of many of the morphological characters limits their usefulness for classification (Labra *et al.,* 2003). The organs most commonly used to differentiate species and plant variants of Opuntia are cladodes, spines, fruits and the flower in smaller scale (Reyes-Agüero *et al.,* 2005). However, the discriminative power of the seed characteristics has not been reported despite this organ is known to be very stable. Seed variables of Opuntia are little influenced by environmental pressure and are more affected by the genetic control, which is likely due (i) to the hardness of the seed; (ii) the protective effect offered by the pulp and seeds testa and (iii) the short period of exposure the fruits to environmental factors. The results presented in this thesis evidenced the utility of the seed variables as grouping characteristics such as weight, size and dimensions of seeds. We demonstrated the potential discriminatory of seed variables derived from image analysis, in order to their consideration in characterization studies and to assign new identified genotypes in their respective taxa. The seed features were successfully used to discriminate between the “xoconostle” and “tunas” genotypes. The variables responsible for the separation between genotypes were seed area, major axis length, minor axis length,

Feret diameter and seed weight. Likewise, these characters are potential candidates to be transferable as characters state useful for varieties register.

The second method applied was the biochemical marker and the objectives were to explore the Opuntia seeds protein profile, and to determine their potential use for characterization of genotypes and taxonomy. The obtained results provided evidence of the broad variability of protein content in the seeds. Similarly, the four protein fractions (SSPs) are all exist in the Opuntia seeds with considerable amplitude; albumin and globulins are the most abundant protein fractions, while prolamins are present in small quantities. Contrary to the high variation in the relative contents of proteins, the composition of the seed total proteins and SSPs, as showed by banding patterns resolved by SDS-PAGE system, was shown to be similar among the different accessions studied. Moreover, the need to combine both seed total proteins and SSPs profiles data to differentiate all Opuntia accessions was demonstrated. The clustering of accessions revealed no concordance with the current taxonomic status. However, the "xoconostles" genotypes were clustered in a sister-no-separate group from "tunas".

For the DNA markers, the genetic diversity of an extended Mexican Opuntia germplasm (88 accessions) was revealed using 13 SSR markers in an attempt to explore the genetic relationships among them. SSR markers generated unique fingerprints for each *Opuntia* accessions confirming its usefulness for genetic analysis of Opuntia germplasm. UPGMA and STRUCTURE analyses confirmed the incorrect delimitation of the species in this genus. Median-Joining and NeighborNet simulations classified all genotypes into a complex network; both linear and reticulate ties were revealed among the Mexican Opuntia accessions, suggesting mainly vegetative propagation of Opuntia accessions and at a lesser

extent sexual heredity through selected the possible natural crosses.. The accessions currently defined in *O. ficus-indica, O. megacantha, O. albicarpa. O. lasiacantha, O. hyptiacantha* and *O. streptacantha* species were not separated genetically. Representative accessions of *O. robusta* differ from other species as revealed by NeighborNet Networks analysis. For their part, "xoconostles" genotypes were partially separated from the "tunas" when UPGMA and STRUCTRUE analyses were carried out. However, the results obtained by the Networks analyses (MedianJoining and NeighborNet) clearly separated the "xoconostles" accessions, even revealed a linear evolution from wild species. Three scenarios are suggested for the Mexican germplasm structure. Firstly, a dispersion pattern of genotypes related to human exchange. Secondly, local and regional selection of wild *Opuntia* sp. with desirable agronomic traits. Finally, the effect of end-use attributed to distinct genotypes in Mexico: fruits ("xoconostles" or "tunas"), vegetables or livestock feed, among other uses. The above scenarios are in accordance to the network analysis presented in the current study. According to the current and previous results (Caruso *et al.,* 2010; Valadez-Moctezuma *et al.,* 2014), there is a high degree of genetic diversity. Hence, further studies are required to develop new SSR markers derived from genome sequence information of Opuntia rather than that of relative species, which will provide better insight and an understanding of the genetic diversity of Opuntia and can be used for marker-assisted breeding of new cultivars.

The polymerase chain reaction-restriction fragment length polymorphism analysis of the three plant genomes (chloroplast, mitochondria and nucleus; *atp*B-*rbc*L, *trn*L-*trn*F and *psb*A-*trn*H, *ycf*1, *ppc*, *Phy*C and *cox*3), was carried out to investigate the phylogenetic relationship among 103 accessions. The results showed that the genetic distance was very narrow, and the 75 % of genotypes had the same genetic profile. Moreover, all

"xoconostle" accessions were grouped together without being totally separated from three genotypes of "tunas". We assume the genetic divergence between "tunas" and "xoconostles" is very narrow.

The PCR-RFLP (CAPS) technique was therefore applied to study the genetic diversity in 91 accessions of Opuntia accessions based on a fruit size-related gene "FRUITFULL". The results demonstrated the narrow genetic divergence and estimated distance. Thus, the most cultivars (69 %) were maintained similar gene FRUITFULL-homologue according to their genetic profiles.

The various tools applied in this thesis confirmed the erroneous current taxonomic classification of Opuntia in Mexico. For the other hand, the techniques shown to be useful and effective for genetic diversity studies and for clarifying the taxonomic status. Similarly, the results found here serve as a solid foundation for further studies related to the evolution of these plants under the concept of domestication.

Throughout the present thesis, several statistical methods have been applied to achieve a subjective interpretation. ANOVA and Principal component analysis have been carried out for morphological data, while multivariate analysis such as PCoA and UPGMA, model-based Bayesian (STRUCTURE) and network analysis (Median-Joining and NeighborNet networks) were applied to molecular markers. The application of these statistical methods allowed us to affirm the great genetic diversity of the Mexican "nopal"; however the results disagree with the current taxonomic classification, as reported in previous studies (e.g. Labra *et al.*, 2003; Griffith, 2004; Caruso *et al.*, 2010; Valadez-Moctezuma *et al.*, 2014). Labra *et al.* (2003) suggest that *O. ficus-indica* should be

considered as a domesticated form of *O. megacantha*. Griffith (2004) hypothesized that the taxonomic concept of *O. ficus-indica* may include clones derived from multiple lineages and therefore be polyphyletic. Caruso *et al.* (2010) pointed to the inconsistencies of previous taxonomical genotype assignments and supported the Griffith' hypothesis. Valadez-Moctezuma *et al.* (2014) hypothesis about the existence of a smaller number of *Opuntia* species in accordance with those currently described, but with high intraspecific genetic variation. The results presented in this thesis confirmed the last hypothesis. The 110 accessions studied here can assigned to three complex species; the first complex correspond to the genotypes producing "xoconostles" (*O. matudae, O. joconostle* and *O. leucotricha*); the second complex correspond to *O. robusta* and the third complex resembles the remaining genotypes (*O. ficus indica, O albicarpa, O. megacantha, O. hypthiacanta, O. undulta, O. cochinera, O. lasiacantha, O. streptacantha, O. chaveña*).

There still much work to be done on the molecular systematic in the *Opuntia* genus. Species limits, genotypes specific assignation and relationships between the genera and species are poorly studied. The main obstacles in the way of achieving these ambitious goals are complexities that one encounters when trying to resolve relationships between very recently differentiated species, most notably low numbers of variable sites, incomplete lineage sorting and polyploids levels. In addition to increased efforts toward exploration in the field, various initiatives to promote and develop taxonomic expertise, generalization of collaborative work, and improved access to natural history collections and literature, major advances in technology provide new opportunities to facilitate delimiting of species.

Today, most recognized species have been delineated and described based on morphological evidence: in general, they have been delimited based on one or more qualitative or quantitative morphological characters that show no-or very little-overlap with other species. The use of molecular data in plant taxonomy has been era-splitting and highly successful in many instances, but we also highlighted some limits and cautions to consider when adopting this approach. Most importantly, a species description solely based on molecular evidence would obviously seem critically disconnected from the natural history of the species, e.g. its life-history traits, ecological requirements, co-occurring species, and biotic interactions. As sources of relevant characters, many fields of biology might contribute to taxonomic studies: they include morpho- anatomy which takes advantage of new techniques such as scanning electron microscopy, remotely operable digital microscopy, computer-assisted tomography, confocal laser microscopy, and automatic image processing for morphometry and cytogenetics but also palynology, physiology, chemistry (production of secondary compounds), breeding relationships, and ecological niche modeling. Other sources of information will also most probably be more widely used in the future, such as transcriptomics, metabolomics and proteomics.

It is also clear that end users of taxonomy such as conservation planners need an operational, character-based, and cheap way to discriminate species. This could tend to diminish the perceived potential of molecular taxonomy, but the molecular taxonomy obviously has a great role to play. DNA can aid to delimit taxa and to group specimens among which to find morphological-or other types of-affinities in further investigations. Taxonomic circumscriptions are scientific hypotheses, which are ideally validated by evidence from multiple sources, and molecular methods offer the opportunity to yield high-potential information. However, there is not a single, best method to be used in all

plant groups and the molecular taxonomist will have to face multiple questions: before anything, it is necessary to identify the optimal sampling strategy, the most suited genomic compartment (s) to examine, the right technique(s) to use, and the adequate method(s) of statistical analysis to extract the relevant information about species limits and relationships. In addition to the complementarily of "traditional" and genetic approaches, molecular taxonomy itself will often require to gather and compare patterns based on several types of data e.g., nuclear vs. cytoplasmic markers or markers with different rates of evolution.

REFERENCES

Caruso, M., Currò, S., Las Casas, G., La Malfa, S., Gentile, A. 2010. Microsatellite markers help to assess genetic diversity among *O. ficus-indica* cultivated genotypes and their relation with related species. Plant Syst. Evol. 290: 85-97. doi: 10.1007/s00606-010-0351-9

Labra, M., Grassi, F., Bardini, M., Imazio, S., Guiggi, A., Citterio, S., Banfi, E., and Sgorbati, S. 2003. Relationships in *Opuntia* Mill. Genus (Cactaceae) detected by molecular marker. Plant Sci. 165: 1129-1136.

Reyes-Agüero, J.A., Aguirre, J. R. and Hernández, H. M. 2005. Systematic notes and a detailed description of *Opuntia ficus-indica* (L.) Mill. (Cactaceae). Agrociencia 39: 395-408.

Valadez-Moctezuma, E., Samah, S., and Luna-Paez, A. 2014. Genetic diversity *of Opuntia* spp. varieties assessed by classical marker tools (RAPD and ISSR). Plant Syst Evol DOI: 10.1007/s00606-014-1112-y.

ANNEX 1

UTILITY OF SEED CHARACTERISTICS DERIVED FROM IMAGE ANALYSIS FOR NOPAL (*OPUNTIA* SPP.) VARIETIES DISCRIMINATION

Samir Samah, Elsa Ventura-Zapata and Ernestina Valadez-Moctezuma

UTILITY OF SEED CHARACTERISTICS DERIVED FROM IMAGE ANALYSIS FOR NOPAL (*OPUNTIA* SPP.) VARIETIES DISCRIMINATION

ABSTRACT

The novel varieties of prickly pear (*Opuntia* spp.) require registration to UPOV (International Union for the Protection of New Varieties of Plants). These varieties must complete the tests for distinctness, uniformity and stability (DUS). To carry out the DUS process, morphological characters of the plant, cladodes, fruits and phenology are used, which express high variation and instability due to environmental conditions. Therefore, the aim of this study was to assess the discriminatory potential of 19 seed variables based on analysis of variance, principal component analysis and conglomerations to differentiate and characterize 33 prickly pear varieties registered at SNICS (*Servicio Nacional de Inspección y Certificación de Semillas, México*). Analysis of variance showed highly significant differences between varieties for all studied and the coefficients were less than 20. Seeds have internal and external features that are stable with high discriminatory power. UPGMA and Principal component analysis discriminate the all studies varieties based only on seed traits. The variables: area and perimeter of the seed, major axis length, area and perimeter of the embryo, area and perimeter of the perisperm, seed weight and Feret diameter, differentiated varieties of Opuntia considered in the present study. This study provides additional characteristics, with good discriminating power, determined in an objective and standardized, so it could be considerate in test for DUS in Opuntia varieties.

Key words: *Opuntia* spp.; DUS; UPOV; external morphology; longitudinal section.

INTRODUCTION

Opuntia, also known as nopal in Mexico, is a genus plant used as livestock feed and as fruits and/or as vegetable for human in North of Africa, Middle East, Southern Europe and in America. Nopal is a potential crop for agricultural production and has an important role in ecosystem sustainability and to safeguard the natural heritage (Erre and Chessa, 2013). In Mexico, *O. ficus-indica* is grown on 70,000 ha; each year 430 t of pads are transplanted into commercial plantations (Gallegos *et al.,* 2013).

Tests for distinctness, uniformity and stability (DUS) are an essential component of variety registration and granting plant breeders' rights (Lootens *et al.,* 2013). Tests for DUS should be performed according to international guidelines established by the UPOV (International Union for the Protection of New Varieties of Plants), since the UPOV has published specific guidelines for several crops, even nopal (UPOV, 2006). These guidelines outline the relevant characteristics to be evaluated, along with the recommended procedures to conduct the experiments and the statistical analyses to undertake; also UPOV established the characteristics to be visually classified and with ones to be accurately measured. UPOV characteristics for DUS testing are primarily morphological and chosen to reflect general genetic differentiation among varieties (Lootens *et al.,* 2013). For the nopal (prickly pears and xoconostles) guidelines illustrate 61 characters to evaluate for DUS, of which 3 characters are for the plant, 16 for cladodes, 9 for spines and central spine, 5 for flower and 23 for fruit; also, the time of beginning of flowering, flowering type, time of maturity, duration of harvest and seed size (small, medium, large) (UPOV, 2006); since some of these characters are of environment conditions dependents. However, accuracy data from seed characteristics are missing.

Moreover, the selection of the reference varieties, the effect of environment on the expression of some distinguishing features and data collection by the experts, may also be prone to subjectivity and visual scores provide discontinuous data that are not easy to analyze using standard statistical methods (Lootens *et al.,* 2013). In this sense, the lack of ability to discriminate features for examining the distinctness can obstruct the progress improvement by not discriminating novel varieties (Lootens *et al.,* 2013). All these difficulties have led to the search for alternative procedures for DUS testing. Therefore, the use of molecular markers and image analysis are the subject of considerable attention (UPOV 2011).

The *Opuntia* genus displays a great morphological diversity in Mexico, so it is important to document and record the different varieties by a reliable classification procedure. Therefore, the aim of this study was to determine whether variables derived from image analysis of seed provides high discriminatory power between varieties of nopal and its possible use or consideration in the test for DUS.

MATERIALS AND METHODS

Fruits of 33 Opuntia varieties were collected at commercial maturity from two germplasm (CRUCEN, Zacatecas and Nopalera-UACh, Texcoco, Mexico). All seeds were extracted, and then dried at room temperature, after they were cleaned off any remaining pulp and only viable seeds were stored in paper-bags until use. Three replicates of 100 fully developed seeds were weighed using an analytical balance (ABS 220-4; Karn and GmbH) to determine seed weight. For external morphological: 36 randomly seeds/repetition with 3 replicates of each sample were selected, and then images were taken with an Olympus digital camera. For internal morphology: the technique developed

by Guerrero *et al.* (2006) was applied. Three replicates of 5 seeds were adhered to the surface of a glass slide. These seeds were worn symmetrically and in parallel to the median plane with fine sandpaper. Subsequently, they were viewed and photo-documented under a Leica EZ4 stereoscope (Leica Microsystems, Switzerland) with an integrated camera. Consequently, all images obtained were processed using Photoshop CS5 12.0 program to define the area of seed, embryo, perisperm and funicular seedcoat. The seed variables were then obtained by UTHSCSA ImageTool version 3.00. The methodology described by Mebatsion *et al.* (2012) was adopted to improve the contrast.

The variables obtained from entire seeds were: Area = the area of the object measured as the number of pixels in the polygon; Perimeter = the length of the outside boundary of the object; Major Axis Length = the length of the longest line that can be drawn through the object: Minor Axis Length = the length of the longest line that can be drawn though the object perpendicular to the major axis; Elongation = the ratio of the length of the major axis to the length of the minor axis (if the value is 1, the object is roughly circular or square, whereas it is more elongated when the ratio decreases from 1); Roundness = if the ratio is equal to 1, the object is a perfect circle, when the ratio decreases from 1, the object departs from a circular shape, calculated as $R = [(4\pi * \text{area})/\text{perimeter}^2]$; Feret Diameter = the diameter of a circle having the same area as the object, calculated with the formula: $FD=\sqrt{[(4 * \text{area})/\pi]}$; Compactness = provides a measure of the object's roundness: at 1 the object is roughly circular, when it decreases from 1, the object results less circular, calculated as C = FD/Major Axis Length. The variables obtained from the median section of the seeds (internal morphometric) were: Area and Perimeter of embryo, Area and perimeter of perisperm and funicular seedcoat. Ratios between variables were also calculated (Table 1).

Statistical analyses were performed using SPSS 20.0 for Windows (SPSS, USA). A basic descriptive statistical analysis, followed by an ANOVA. Tukey test ($P < 0.05$) was used to group the varieties for their averages. Analysis of bilateral Pearson correlation was also carried out to reveal possible relationships between seed characters. Furthermore principal component analysis (PCA) and cluster analysis based on Ward's method and squared Euclidean distance between variables were performed.

RESULTS AND DISCUSSION

Analysis of variance showed highly significant differences between varieties for all studied characteristics of seeds, indicating the existence of a high degree of morphological diversity. The coefficient of variation ranged from 1.7 and 28.3. However, most of the variables showed a coefficient of variation of less than 20 (Table 1). Tukey separated the varieties into different groups depending on the variable (data not shown). Thus, the "Oreja de Elefante" was separated from other varieties by its greater Seed Area (20.8 mm^2), Seed Perimeter (18.2 mm), Major Axis Length (5.78 mm), Minor Axis Length (4.71 mm) and Feret Diameter (0.52). "Amarilla Jalpa" was isolated by its low Seed Weight (1.03 g), while the "Bola de Masa" variety was characterized by its low MjA (3.66 mm). Therefore, the variables Perisperm Area/Seed Area, Perisperm Area/Embryo Area, Perisperm Area and Seed Weight were the most variable among the studied characteristics (values *F*, Table 1).

Table 1. ANOVA and descriptive analyses of the seed morphological characteristics of 33 prickly pear varieties.

Characteristics	Abbreviation	Minimum	Maximum	Mean	CV (%)	*F*-Values
100 seeds weight (g)	SW	1.01	2.41	1.70	17.9	64.3***
Seed Area (mm^2)	SA	10.9	21.9	14.1	12.9	35.0***
Seed Perimeter (mm)	SP	13.2	18.5	15.0	6.48	17.2***
Major Axis Length (mm)	MjA	3.68	5.97	4.76	7.43	36.9***
Minor Axis Length (mm)	MnA	3.35	4.81	3.89	6.63	21.7***
Elongation	Elg	1.13	1.41	1.24	4.01	06.4***
Roundness	R	0.66	0.84	0.79	3.65	02.5**
Feret Diameter	FD	0.37	0.53	0.42	6.24	31.6***
Compactness	C	0.84	0.92	0.89	1.70	06.4***
Embryo Area (mm^2)	EA	3.06	6.67	5.28	13.2	06.2***
Embryo Perimeter (mm)	EP	8.83	14.4	11.4	8.79	05.4***
Perisperm Area (mm^2)	PA	0.10	0.37	0.23	25.4	04.7***
Perisperm perimeter (mm)	PP	1.98	4.11	3.05	14.4	05.1***
Embryo Area/Seed Area	EA/SA	0.19	0.53	0.38	15.7	07.0***
Perisperm Area/Seed Area	PA/SA	0.01	0.03	0.02	28.2	05.3***
Perisperm Area/Embryo Area	PA/EA	0.02	0.08	0.04	28.3	03.3***
Embryo Perimeter/Seed Perimeter	EP/SP	0.55	0.99	0.76	10.1	07.1***
Perisperm Perimeter/Seed Perimeter	PP/SP	0.14	0.29	0.20	15.7	05.4***
Perisperm Perimeter/Embryo Perimeter	PP/EP	0.18	0.37	0.27	15.1	03.4***

CV: coefficient of variation; *F*: critical value test; g: gram; mm: millimeter; mm^2: millimeter square.

** and *** Indicates significant difference at 0.01 and 0.001 levels, respectively.

The high significance of the seeds variables and the low coefficients of variation indicate the high discriminatory potential of internal and external seed characteristics for the *Opuntia* genus. On the other hand, low values of CV suggest the stability discrimination of quantitative variables of seeds, as reported by Guerrero *et al.* (2006). According to Lootens *et al.* (2013), there are at least two ways in which image analysis can help to improve tests for DUS. First, this type of analysis increases the accuracy of measurements. In our case, the data demonstrated that the method of image analysis was used successfully to estimate the fundamental dimensions of seeds (length, width and elongation) with high accuracy, since as the manual measurements are difficult due to the small size of the seeds of Opuntia genotypes. Another advantage of image analysis is to provide additional features to be determined objectively and with a good power of

discrimination, such as FD, C and Elg (Table 1). Additionally, these are all continuous variables, which allow the use of ANOVA statistics Moreover.

According to the Pearson correlation coefficients, a high positive correlation between several pairs of variables was obtained: SA *vs* SP, FD *vs* SA, SP *vs* FD, MjA *vs* MnA, SA *vs* MjA and SA *vs* MnA. These results suggest that developmental increases in seed size correspond to increases in the width thereof, as well as in its length. One aspect that should be considered before adding new features to the test for DUS is the fact that they are not correlated to avoid double counting (UPOV, 2002). This can be overcome by taking into account the easy and the automation to get the variables from the analysis of images of Opuntia seeds as it involves no additional costs or efforts (Guerrero *et al.,* 2006). Also the use of two highly correlated variables may be useful in replacing a missing variable.

The principal component analysis (PCA) was performed to study the combination of variables that best explain the variation. The first 4 principal components (CPs) explain the major variation. Their eigenvalues were 6.48, 4.71, 3.65 and 2.06 for PC 1, PC 2, PC 3 and PC 4, respectively. The PCA revealed that the first four components explained 88.95 % of the total variability. The first three accounted for 78.14 %, of which the first component contributed for 34.12 % and the second one for 24.79 %. The variables that defined, by its eigenvectors, the first component in the positive direction were SP (0.94), FD (0.94), SA (0.94), MnA (0.92), MjA (0.85) and SW (0.69) and in the negative direction EA/SA (0.77) and EP/SP (0.63). The second component was related to the variables PA (0.92), PP/PS (0.91), PA/EA (0.84), PP (0.84), PA/SA (0.76) and PP/SP (0.68), all in the positive sense. The variables EP (0.90), FD (0.87) and EP/SP (0.68) defined the third component in the positive direction. The fourth component was determined by variable Elg (0.81) in the positive direction and by C (0.82) in the negative one. The first component seems to be

associated with the external characteristics and the second component is correlated with the internal characteristics of seeds (perisperm traits); thus these traits are independent and offered very power discriminative variables.

The projection of the varieties on the first two components (PC1 and PC 2) showed high dispersion around the origin of the graph (Fig. 1) showing no clustering pattern; neither the grouping was according to the assignation of varieties in their respective species. This is probably due to the morphological diversity observed in the suited varieties. Also, these varieties have several end-uses; they were cultivated for fresh fruits, for vegetables uses and/or like forage (Gallegos *et al.*, 2013). However, the varieties "Oreja de Elefante" and "Memelo" were separated from the remain ones; "Oreja de Elefante" was separated on PC 1 to have large size of seeds (Area and Perimeter of the seed, Major and Minor Axis Length and Feret Diameter); and "Memelo" was taken place on the second component in the negative sense, mainly because it has smaller Area and Perimeter of perisperm.

Cluster analysis separated the 33 studied varieties into three main groups with a base of a Euclidean distance of 25 (Fig. 2). The first group was composed of a single variety (Oreja de Elefante). This variety was separated from the other defined groups to have the larger seeds. The other two groups contain 12 and 20 genotypes, respectively, as well as separated with the base of a Euclidean distance of 20. The second group was distanced from the third one mainly because it contained varieties with seeds that have more SW (18.1 %), SA (13.2 %), SP (7.4 %), MjA (8.4 %), MnA (6.0 %), FD (6.9 %) and lower EA/SA (20.7 %), EP/SP (12.2 %) and PP/SP (12.8 %). The latter group was divided into two subgroups composing with 7 and 13 varieties as well as defined at a Euclidean distance of 20 (Figure 2), whereby the separation of these subgroups was the result of

differences in SP (3.3 %), FD (3.8 %), EA (18.7 %), EP (13.4 %), EA/SA (11.2 %) and PS/SP (10.3 %).

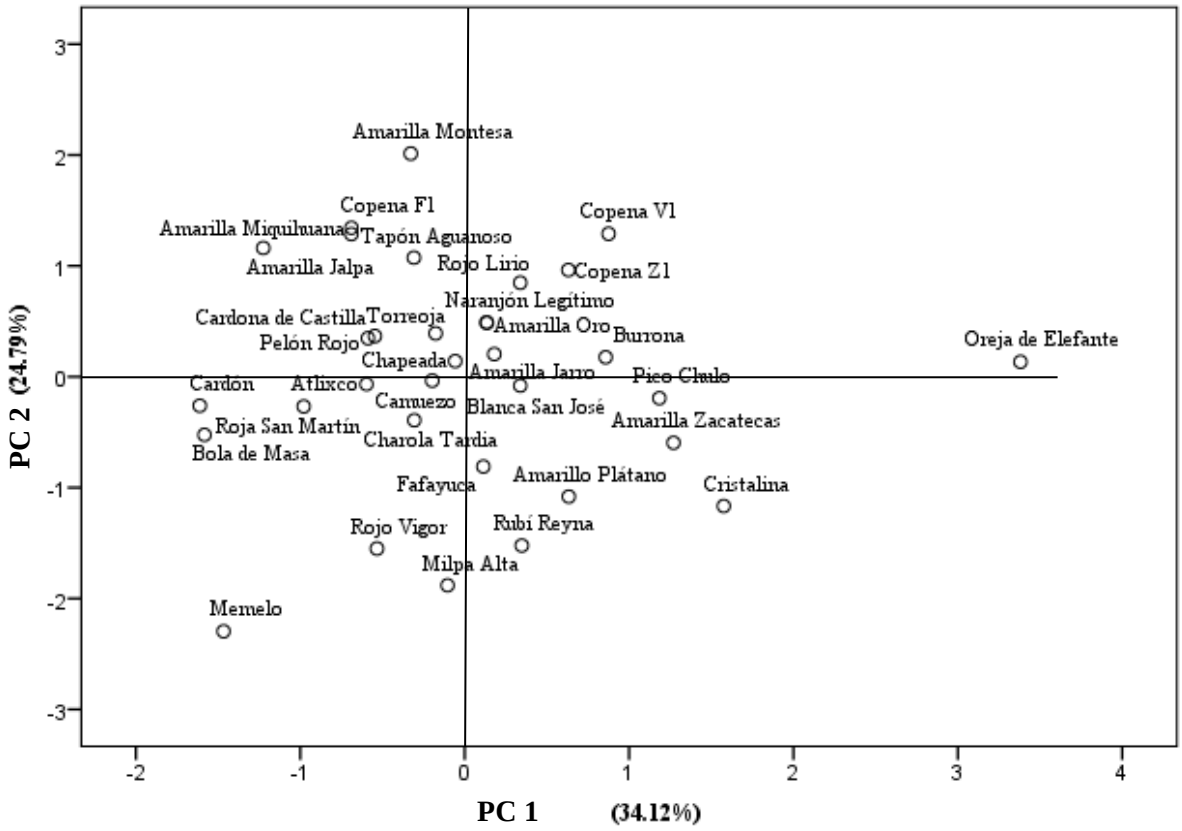


Fig. 1. Distribution of the 33 varieties of Opuntia on the first two principal components (PC 1 and PC 2) based on 19 external and internal quantitative traits from the seeds.

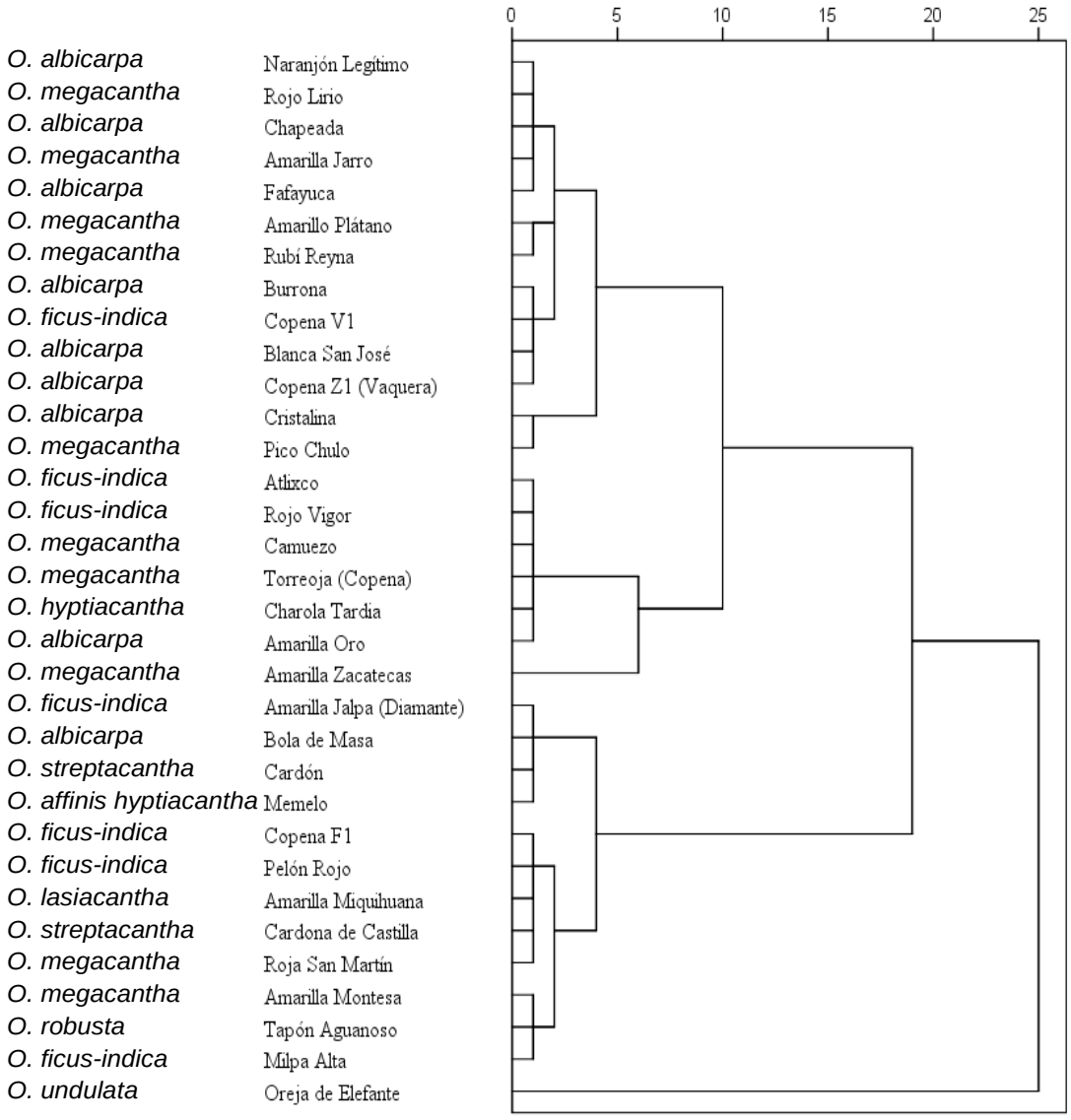


Fig. 2. Grouping of 33 varieties of Opuntia resulted from Ward method based on squared Euclidean distance of 19 morphological of seed.

For Opuntia varieties, the usefulness of vegetative characteristics such as the color of areoles in the cladode, the thickness of the fruit peel; and reproductive such as seed size, flower habit and time of harvest, was documented to group varieties (UPOV, 2006). Grouping characteristics contribute to select varieties of common knowledge to be grown in the trials with candidate varieties and to define the way in which these varieties are divided into groups to prove the distinctness. The results presented here demonstrated the

utility of the seed variables as features for varieties grouping such as weight, size and seed dimensions. Moreover, the test of stability is more difficult to accomplish because in practice it is not usual to perform tests of stability that produce certain results as those tests for distinctness and uniformity. To overcome this drawback, it is considered that when a variety has been shown to be uniform, it also is considered stable (UPOV, 2006). However, the characters considered in these tests are highly affected by the environment conditions, and the test of stability derived from one trail still questioned. Contrary, the additional use of variables derived from the seeds could be a good alternative. Quantitative variables of seeds are little influenced by environmental pressure, which is probably due to the following reasons: (i) the hardness of the seed, (ii) the protective effect offered by the pulp and the peel for the seeds (iii) the short period of exposure of fruits to environmental factors.

Finally, the image analysis processing has been found to be more useful over the past two decades; these trends were mainly to replace and/or to support the classification by visual examination (D'Imperio *et al.,* 2012). Image analysis of seeds was used to differentiate genotypes and crops (Viscosity and Fortini, 2011). In the present study, we demonstrated the potential discriminatory of the variables derived from image analysis of seed and their possible consideration in the evaluation and registration of Opuntia varieties. This would involve an initial investment in the development of the procedures, but it also would reduce costs at long-term of the evaluation, due to reductions in the time required for data collection and the size of the field experiments. In addition, the UPOV provides a suitable feature for DUS testing which should be "accurately defined and recognized." Therefore, it is important that any new features added to eventually DUS system are also recognizable. This case can be achieved with the image analysis technique.

CONCLUSIONS

In the present study, it was shown that the variables: area and perimeter of the seed, major axis length, area and perimeter of the embryo, area and perimeter of the perisperm, seed weight and the Feret diameter, all derived from analysis of images, have a high discriminatory potential and could be considered in tests for DUS. Moreover, the varieties "Oreja de Elefante" and "Memelo" showed to be candidates as to witness varieties in DUS tests, since they had extreme values of the variables evaluated derived from seeds.

ACKNOWLEDGEMENTS

The first author received a scholarship for doctoral studies from the Ministry of Foreign Affairs, Mexico. We are grateful to Dr. Rodríguez de La O J L for the stereomicroscope facilities.

REFERENCES

D'Imperio, M., Viscosi, V. Scarano, M.T. D'Andrea, M.. Zullo, B.A and Pilla, F. 2011. Integration between molecular and morphological markers for the exploitation of olive germoplasm (*Olea europaea*). Sci. Hortic. 130: 229-240.

Erre, P., and Chessa, I. 2013. Discriminant analysis of morphological descriptors to differentiate the *Opuntia* genotypes. Acta Horticulturae 995: 43-50.

Gallegos V., C., Mondragón, C. Núñez, C.A., and Scheinvar L. 2013. Germplasm diversity of xoconostles (*Opuntia* spp.) or acidic cactus pears from central Mexico. Acta Hortic. 995: 75-81.

Guerrero M., P., Zavaleta, A. Barrientos A. F., Gallegos C., Núñez, C. A, Valadez M., E. and Cuevas, J. A. 2006. Técnica para el estudio de la micromorfología interna de semillas duras en *Opuntia*. Rev Fitot Mex 29:37-43.

Linkies, A., Graeber, K. Knight, C. and Leubner-Metzger, G. 2010. The evolution of seeds. New Phytologist 186: 817-831.

Lootens, B., Baert, J., Pannecouc, J., VanWaes, J., and Roldan-Ruiz, I. 2013. Comparison of image analysis and direct measurement of UPOV taxonomic characteristics for

variety discrimination as determined over five growing seasons, using industrial chicory as a model crop. Euphytica 189: 329-341.

Mebatsion, H. Paliwal, K. J., and Jayas, D. S. 2012. Evaluation of variations in the shape of grain types using principal components analysis of the elliptic Fourier descriptors. Computers and Electronics in Agriculture 80: 63-70.

UPOV. 2002. General introduction to the examination of distinctness, uniformity and stability and the development of harmonized descriptions of new varieties of plants, TG/1/3.

UPOV. 2006. Chumbero, nopal tunero, tuna y xoconostles. Directrices para la ejecución del examen de la distinción, la homogeneidad y la estabilidad.

UPOV. 2011. Image analysis for DUS in the United Kingdom, TWC/29/19.

Viscosi, V., and Fortini, P. 2011. Leaf shape variation and differentiation in three sympatric white oak species revealed by elliptic Fourier analysis. Nordic Journal of Botany 29: 632-640.

ANNEX 2

MORPHOLOGICAL SEED DIFFERENTIATION BETWEEN CULTIVARS OF XOCONOSTLES AND TUNAS (*OPUNTIA* SPP.)

Samir Samah, Elsa Ventura-Zapata and Ernestina Valadez-Moctezuma

Submitted: Acta Horticulturae

VIII International Congress on Cactus Pear and Cochineal October 28-31, 2013

Palermo, ITALY

MORPHOLOGICAL SEED DIFFERENTIATION BETWEEN CULTIVARS OF XOCONOSTLES AND TUNAS (*OPUNTIA* SPP.)

ABSTRACT

The classification and description of cactus pear varieties, tunas (sweet cactus pear) and its wild relatives xoconostles (acidic cactus pear), has been based on morphological traits of the fruits and cladodes. The absence of pulp and the presence of an edible thick pericarp are the most significant morphological differences between xoconostles and tunas. This work was focused on the differentiation between tunas and xoconostles by analyzing the external and internal seeds characteristics. Twenty three tunas and six xoconostles accessions were compared using sixteen variables. The ANOVA was performed and significant variables were selected by correlation and principal component analysis (PCA), and subsequently a cluster analysis was performed too. All variables were found to be significant in order to discriminate between these types of plants: the perisperm area and the whole seed area were 50 % and 40 % bigger in tunas than in xoconostles, respectively. While the perimeter area, the major and minor axis length, the Feret diameter, the embryo area, and 100 seeds weight, all were 20 % smaller in xoconostles than in tunas. Two principal components (PC) explained 88.96 % of the variability; the PC1 (Seed area, major axis length, Feret diameter, embryo area and 100 seeds weight) explained 60.61 %, and the PC2 (minor axis length and elongation) explained 28.36 % of the variability. The PCA plots and the cluster analysis clearly separated the xoconostles accessions from the tunas ones, but both analysis did not match up with the actual taxonomy classification. The present study demonstrated that the internal and external seeds variables are an alternative to the cactus pear characterization and it showed a tendency to increase the seeds weight in relation to the domestication level.

Keywords: external morphology, longitudinal section, embryo, *Opuntia*

INTRODUCTION

In Mexico, the word "nopal" refers to each plant of most of the *Opuntia* species disregarding if they are used for fresh fruit, for vegetable or for animal feed. Opuntia (cactus pear) plants are closely associated to the cultural development of the Mexican people; they have been used as human food in the semi-arid regions of the south-west of Tamaulipas and in the Tehuacán valley from 9,000 to 11,000 years ago. The fruits of Opuntia are known as sweet cactus pears ("tunas") and acidic cactus pears ("xoconostles"). Xonocostle fruits differ from tunas because they have a very thick edible pericarp, the pulp is almost absent and highly acidic, and has a long shelf life (Gallegos-Vazquez *et al.,* 2013).

It is highly important to document and register different cactus pear cultivars under the basis of a reliable classification procedure. The classification of cultivars of tuna and its wild relative xoconostle has been based upon morphological features mainly from the fruits and cladodes (Reyes-Agüero *et al.,* 2013). All these studies were based on quantitative approaches that attempt to group cultivars using similarities detected in the measured characteristics. However, in any characterization study, the differences that may exist in this type of plants between the cactus pear seeds and their possible discriminatory potential have not been considered. Morphological characterization is needed to provide users valuable information on individual accessions, relationship among traits, and the structure of collections (Erre and Chessa, 2013).

The present study has the following objectives: to assess external and internal morphological seed variables of 23 tunas and 6 xoconostles accessions, to determine the most discriminant variables and grouping genotypes based on these characteristics. On the other hand, to quantify internal and external morphological seed differences between sweet and acidic cactus pear and to determine whether there are, in fact, clear morphological difference between these crops.

MATERIALS AND METHODS

Fruit samples of 23 tunas, 6 xoconostles accessions and one outgroup (Pitahaya: *Hylocereus undatus*, Cactaceae) were collected at two locations; CRUCEN-UACh germplasm bank, Zacatecas and "Nopalera" UACh, Texcoco (Table 1). Ten fruits from at least three individuals of each accession were harvested at commercial maturity, from which all mature seeds were removed manually, then dried in open air, cleaned of any remaining pulp and only viable seeds were stored in paper-bags at room temperature until use. Accession names, species designations, use, domestication level and maturity duration are cited in Table 1.

To determine the weight of seeds, 100 fully developed seeds for each of the three replicates were counted and weighed with analytic balance (220g/0.1mg). For external morphology; 36 seeds for each of the 3 repetitions of each sample were randomly chosen to take pictures of them with a digital camera.

Table 1. List of tunas and xoconostle accessions from Mexico, evaluated to study morphological seed diversity of *Opuntia* spp.

Label	Nopal	Accessions	*Opuntia* specie	Domestication level*	100 seeds weight (g)	Use*	Maturity duration*
1	Tunas	Amarilla Jalpa	*O. ficus-indica*	Low	1.03	Fruit	May-June
2		Amarilla Miquihuana	*O. lasiachanta*	Medium	1.53	Fruit	July-August
3		Amarilla Montesa	*O. megacantha*	High	1.82	Fruit	July-August
4		Amarilla Plátano	*O. megacantha*	High	2.08	Fruit	September
5		Burrona	*O. albicarpa*	High	2.17	Fruit	August-September
6		Cacalote	*O. cochinera*	Low	1.74	Fruit	August
7		Cardón	*O. streptacantha*	Low	1.34	Fruit	July-September
8		Cardona de Castilla	*O. streptacantha*	Low	1.54	Fruit	August
9		Chapeada	*O. albicarpa*	High	1.79	Fruit	June-July
10		Charola Tardía	*O. hyptiacancha*	Medium	1.52	Fruit	September-November
11		Chicle	*O. ficus-indica*	Medium	1.53	Fruit	May-June
12		Copena F1	*O. ficus-indica*	Low-Medium	1.41	Fruit	June- July
13		Cristalina	*O. albicarpa*	High	2.30	Fruit	August-October
14		Memelo	*O. sffinis-hyptiacantha*	Low	1.35	Fruit	August
15		Milpa Alta	*O. ficus-indica*	High	1.57	vege	All year
16		Naranjón Legítimo	*O. megacantha*	Medium	1.79	Fruit	July-August
17		Oreja de Elefante	*O. undulata*	Medium	1.75	Fruit	June
18		Rojo Pelón	*O. ficus-indica*	High	1.94	Fruit	July-August
19		Pico Chulo	*O. megacantha*	Medium	2.10	Fruit	July-August
20		Pitahaya**	*Hylocereus undatus*	Medium	0.13	Fruit	September
21		Reyna	*O. albicarpa*	High	1.50	Fruit	July-August
22		Roja San Martin	*O. megacantha*	Low-Medium	1.31	Fruit	June-September
23		Tapón Aguanoso	*O. robusta*	Medium	1.93	Fruit	July-August
24		Villanueva	*O. albicarpa*	Medium	1.79	Fruit	April-June
25	Xoconostles	X_Blanco	*O. joconostle*	Medium	1.30	Fruit	All year
26		X_Colorado	*O. joconostle*	Medium	1.34	Fruit	-
27		X_Cuaresmero	*O. matudae*	Medium	1.35	Fruit	All year
28		X_Chivo	*O. durangensis*	Low	1.12	Fruit	-
29		X_Manzano	*O. joconostle*	Low	1.54	Fruit	September-December
30		X_Rojo	*O.* sp.	Low	1.17	Fruit	-

*Data from Gallegos-Vazquez and Mondragón-Jacobo (2011); ** Outgroup

For internal morphology, the technique developed by Guerrero-Muñoz *et al.* (2006) was applied: 5 clean and viable seeds for each of the three repetitions were stuck to the surface of a glass slide and oriented in parallel to the median section. These seeds were polished symmetrically and parallel to the median section (longitudinal section) until reaching the embryo exposure, with fine sandpaper, and they were viewed and photographed individually under a Leica EZ4 stereoscope (Leica Microsystems, Switzerland) with an integrated camera. All images obtained were processed using Photoshop CS5 12.0 program to define the area of seed, the area of the embryo, the perisperm, and the funicular seedcoat (testa). The seed variables were obtained by UTHSCSA ImageTool 3.00 software. The variables measured from whole seeds were: Area (SA), Perimeter (SP), Major Axis Length (MjA), Minor Axis Length (MnA), Elongation (Elg), Feret Diameter (DF) and Compactness (C). The variables measured from

the median section of the seeds (internal morphology) were: Area (EA) and Perimeter of embryo (EP), Area (PA) and perimeter of perisperm (PP) and funicular seedcoat. Ratios were also calculated; embryo area/area seed (EA/SA), area perisperm/area seed (PA/SA) and Minor Axis Length/Major Axis Length (MnA/MjA) (Table 2).

The analysis of variance (ANOVA) was applied to detect discriminant variables among genotypes, and multiple comparisons (Duncan's Multiple Range Test) were computed to identify the difference between each pair of groups (alpha level was 0.05). Correlation matrix was built using Pearson Correlation Coefficients, and thereafter a Principal Component Analysis (PCA) was performed considering the dataset composed from 30 accessions and 10 variables. All variables were subjected to a cluster analysis using the squared Euclidean distance and Ward's minimal variance method. All calculations were done using SAS 9.0 software.

RESULTS AND DISCUSSION

The embryo is spirally enrolled, and bent around a strongly reduced perisperm. The cactus pear seed is small and ovoid. Seeds of *Opuntia* species have hard (to-the-touch) seed covers. The hardness of the Opuntia seed cover is related to the presence of a funicular envelope, which completely encloses the seed.

The analysis of variance showed highly significant differences ($p \leq 0.001$) among accessions for all variables studied indicating that there was a high degree of phenotypic diversity among the accessions (Table 2). The seed weight, the Feret diameter, the major axis length and the seed area were more variable among all the measured variables (*F* value at 142, 127, 125 and 105, respectively). The coefficient of variation ranged from

0.99 (C) to 19.3 (PA/SA). However, most of the variables showed a coefficient of variation of less than 10 (Table 2). Duncan's Multiple Range Test separated genotypes into different groups depending on the variable (data not shown). Nevertheless, the variable SA, SP, MjA, MnA, EA, EP, PA, PP and EA/SA ratio clearly separated tunas from xoconostles accessions. In this sense, the seed weight, seed perimeter, the major axis length, the minor axis length, the Feret diameter, and the embryo area, were on average about 20 % larger in tunas than in xoconostles. The perimeter and the area of perisperm were 31 % and 50 % larger respectively, comparing tunas with xoconostles. The seed area was in average 40 % greater in sweet cactus pear (Table 2). On the other hand, the Pitahaya has seeds with small sizes, small dimensions and small weight. When considering all the studied accessions, the embryo area represents 42 % and the perisperm 2 % of the total seed area. Portions of the embryo, of the perisperm and the funicular area, compared to the total area of seeds, were similar to those reported in the literature (Stuppy, 2002; Guerrero-Muñoz *el al.*, 2006). However, the seed size has been shown to be phylogenetically significant (Rojas-Aréchiga *et al.*, 2013). The results suggest that seed properties are related in a broad sense to the phylogenetic position of the plant.

The high significance of the analysis of variance and the low coefficients of variation of the variables indicate the high discriminative potential of the internal and external seed variables of cactus pear. On the other hand, low values of CV, like those reported by Guerrero-Muñoz *et al.* (2006), suggest discriminatory stability of quantitative seed variables.

Table 2. ANOVA and descriptive analysis of 23 tunas, 6 xoconostles and one Pitahaya accessions (Mean, mean value of the continuous variable; Max, maximum value; Min,

minimum value; CV, coefficient of variation; *F*, critical value from *F*-test; g, gram; mm; millimeter, mm^2: square millimeter).

Variables	Abreviation	All accessions					Cactus pear			Xoconosltes			Pitahaya		
		Mean	Max	Min	CV	*F* value	Mean	Max	Min	Mean	Max	Min	Mean	Max	Min
100 Seeds Weight (g)	SW	1.56	2.41	0.13	3.88	142.5***	1.69	2.41	1.01	1.30	1.58	1.07	0.13	0.13	0.13
Seed Area (mm2)	SA	12.7	21.9	2.54	4.48	105.3***	14.1	21.9	10.9	8.86	11.1	7.46	2.70	2.85	2.54
Seed Perimeter (mm)	SP	14.1	18.5	6.54	2.87	076.5***	14.9	18.5	13.2	12.0	13.7	11.0	6.68	6.85	6.54
Major Axis Length (mm)	MjA	4.49	5.97	2.15	2.34	125.1***	4.78	5.97	4.16	3.73	4.16	3.47	2.19	2.25	2.15
Minor Axis Length (mm)	MnA	3.65	4.81	1.47	2.80	091.5***	3.89	4.81	3.35	3.11	3.54	2.86	1.53	1.56	1.47
Elongation	Elg	1.24	1.46	1.13	2.41	010.1***	1.24	1.34	1.13	1.21	1.25	1.14	1.43	1.46	1.38
Roundness	R	0.79	0.84	0.66	2.84	04.53***	0.79	0.84	0.66	0.77	0.83	0.72	0.74	0.75	0.73
Feret Diameter	FD	0.40	0.53	0.18	2.32	126.5***	0.42	0.53	0.37	0.33	0.38	0.31	0.18	0.19	0.18
Compactness	C	0.88	0.92	0.83	0.99	011.2***	0.88	0.92	0.84	0.90	0.92	0.88	0.83	0.84	0.83
MinorAxisLength/MajorAxisLenght	MnA/MjA	0.81	0.89	0.68	2.26	10.04***	0.81	0.89	0.75	0.83	0.88	0.80	0.70	0.72	0.68
Embryo Area (mm2)	EA	5.15	6.67	1.05	8.92	13.07***	5.51	6.67	4.33	4.46	5.08	2.86	1.09	1.18	1.05
Embryo Perimeter (mm)	EP	11.4	14.5	5.16	7.79	06.65***	11.8	14.5	10.2	10.5	12.8	7.79	6.17	8.09	5.16
Perisperm Area (mm2)	PA	0.20	0.37	0.01	18.2	13.06***	0.23	0.37	0.10	0.12	0.20	0.08	0.02	0.02	0.01
Perisperm Perimeter (mm)	PP	2.81	5.08	0.54	13.9	09.03***	3.08	5.08	1.98	2.11	2.77	1.56	0.69	0.97	0.54
Embryo Area/Seed Area	EA/SA	0.42	0.63	0.27	11.3	05.04***	0.40	0.52	0.27	0.51	0.63	0.36	0.41	0.46	0.37
Perispem Area/Seed Area	PA/SA	0.02	0.03	0.01	19.3	07.98***	0.02	0.03	0.01	0.01	0.02	0.01	0.01	0.01	0.01

*** Indicates significant difference at 0.001 levels.

According to the Pearson Correlation Coefficients, a high positive correlation (0.99) was obtained between several pairs of variables: SA *vs* MjA, FD *vs* SA, SP *vs* MjA, SP *vs* FD, MjA *vs* FD, MnA *vs* FD. Likewise, significant correlations were obtained between other variables (Table 3). However, PA/SA and Roundness variables had no correlation with any variable. These results suggest that evolutionary increases in seed size corresponded with increases in seed width as well as seed length. Indeed, seed size increases when embryo size (area and perimeter) increases.

In order to study the combination of variables that better explained the existing variability, Principal Components Analysis (PCA) was conducted. Variables with less correlation coefficient were evaluated (Table 3), and data matrix was built using 10 variables and 30 genotypes, which are projected on the first two axes of CPA. The PCA revealed that the first three components explained 95.39 % of the total variability. The first two axes described 88.97 % of the variability, of which the first component explained most of the variability (60.61 %). The variables that had the greatest contribution in the positive direction of the first axis were: FD, EA, SA, and MjA and SW. The second component was

determined by the variable Elg in the positive sense and by C and MnA/MjA in the negative one. The third axis was determined by the greater weight of variable PA and PP in the positive direction.

Table 3. Matrix of Pearson Correlation Coefficients between 16 variables of internal and external seed of tunas, xoconostles and Pitahaya; N = 30.

Variables	SA	SP	MjAL	MnAL	Elg	R	FD	C	MnA/MjA	EA	EP	PA	PP	EA/SA	PA/SA	SW
SA	1															
SP	**0.980**	1														
MjA	**0.986**	**0.991**	1													
MnA	**0.972**	**0.984**	**0.975**	1												
Elg	-0.218	-0.289	-0.211	-0.414	1											
R	0.459	0.373	0.437	0.495	-0.354	1										
FD	**0.987**	**0.993**	**0.994**	**0.992**	-0.308	0.476	1									
C	0.112	0.16	0.083	0.287	**-0.955**	0.364	0.189	1								
MnA/MjA	0.198	0.266	0.188	0.395	**-0.998**	0.355	0.287	**0.960**	1							
EA	**0.855**	**0.903**	**0.898**	**0.898**	-0.347	0.439	**0.907**	0.251	0.325	1						
EP	**0.772**	**0.821**	**0.819**	**0.833**	-0.379	0.452	**0.828**	0.257	0.363	**0.886**	1					
PA	**0.652**	**0.664**	**0.699**	**0.643**	-0.003	0.364	**0.676**	-0.097	-0.021	**0.716**	0.615	1				
PP	**0.727**	**0.749**	**0.776**	**0.730**	-0.097	0.422	**0.763**	0.017	0.072	**0.826**	**0.727**	**0.933**	1			
EA/SA	**-0.691**	-0.592	-0.626	-0.588	-0.130	-0.307	-0.608	0.221	0.133	-0.251	-0.289	-0.318	-0.296	1		
PA/SA	0.066	0.081	0.079	0.104	-0.136	0.102	0.089	0.094	0.126	0.223	0.111	0.317	0.254	0.122	1	
W	**0.815**	**0.840**	**0.833**	**0.839**	-0.318	0.391	**0.843**	0.220	0.301	**0.793**	**0.788**	0.514	**0.666**	-0.443	0.024	1

Bold indicates level of significance at 0.001

The projection of genotypes into the two-dimensional plot in the first two principal components (88.36 % of the total variability) showed a clear separation between the two types of cactus pear, especially along PC1 (Fig. 1), and also genotypes that belong to the tuna displayed a greater dispersion than the xoconostle ones. This is probably due to a higher level of domestication and cultivation of tuna genotypes which were studied here, because their use is not exclusive to the fruit; it is also consumed as a vegetable and as cattle feed (Reyes-Agüero *et al.,* 2013). Among the sweet cactus pears, genotypes named “Oreja de Elefante” and "Memelo" were dispersed from other genotypes because they have maximum and minimum extremes as the internal and external seed variables, respectively. All these results suggest that the morphological variables of seed were less influenced by environmental pressure and were more affected by genetic control.

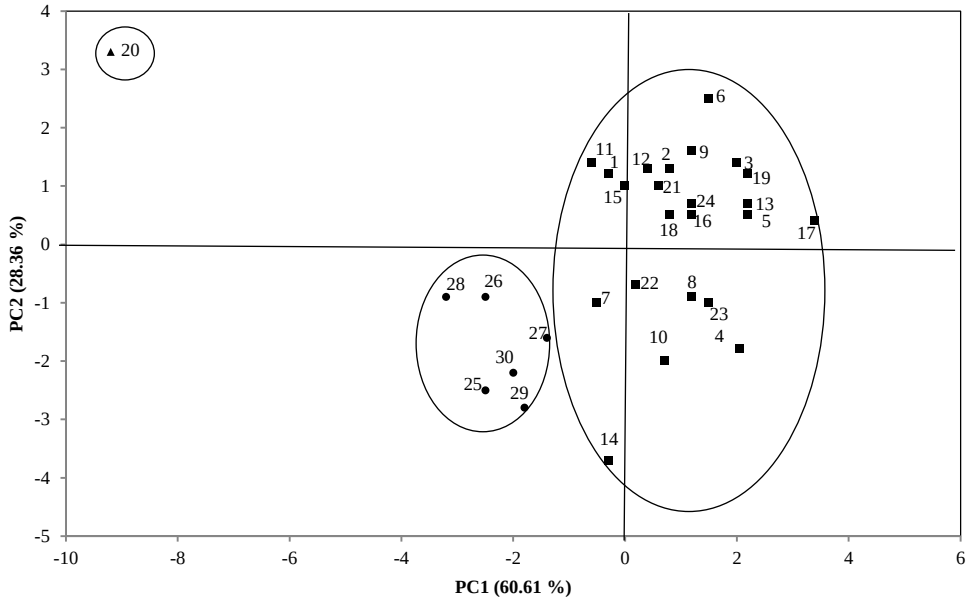


Fig. 1. Two-dimensional representation of 23 tunas (rectangular), 6 xoconostles (circular) and one Pitahaya (triangle) groups' dispersions as determined by PCA. Labels 1 to 30 are the accessions referred to in Table 1.

Cluster analysis revealed the separation of the 30 genotypes studied into three main clusters. The first cluster, compound of a single genotype (Pitahaya, outgroup), was separated from the two other clusters with a Semi-partial R^2 coefficient of 0.18 (Fig. 2), differing by the small size of their seeds. A second cluster, containing the 6 xoconostle accessions included in this study, along with group 1, was separated from the third group by a Semi-partial R^2 coefficient of 0.58. The genotypes located in this group are characterized by intermediate-size seeds between tunas and Pitahaya. The third cluster contained all 23 genotypes of tunas with a separation level of 0.13. The latter group can be divided into two subgroups; the first subgroup enclosed 6 genotypes (Oreja de elefante, Pico chulo, Cristalina, Villanueva, Burrona and Amarilla plátano) and the other one with the remnant 17 cactus pear (Figure 2). The separation of these two subsets is mainly due to the differences in level of SA, SP, MjA, MnA,FD, PA, PP and SW.

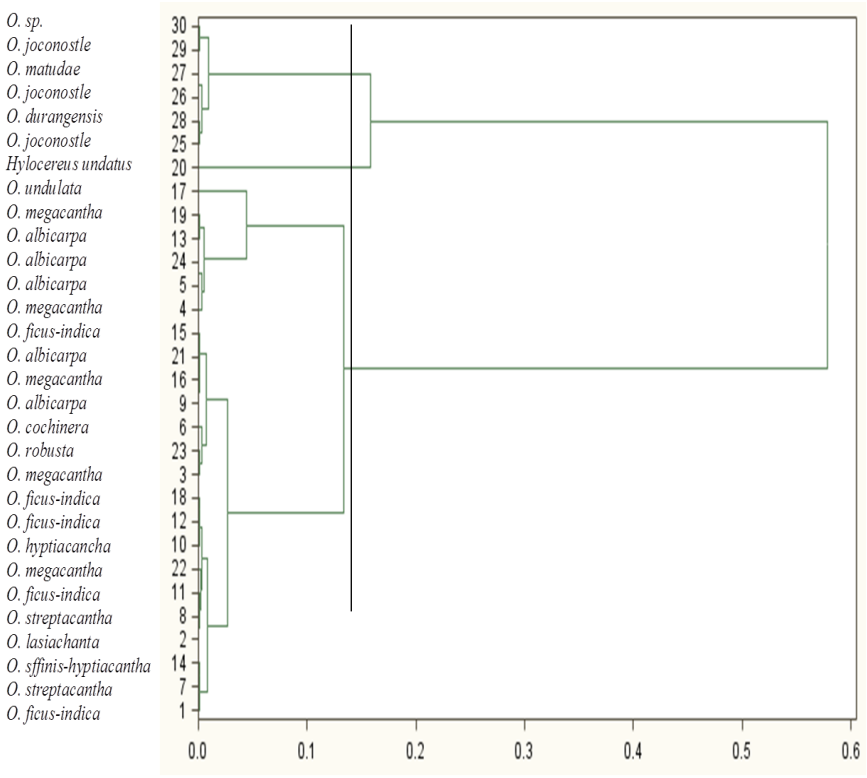


Fig. 2. Dendrogram obtained with UPGMA method by using internal and external morphological seed variables in the 23 tunas, 6 xoconostles and one Pitahaya accessions. Labels 1 to 30 are accessions mentioned in Table 1.

The cluster analysis confirms the separation between sweet and acidic cactus pears as it is revealed by ANOVA and PCA. The xoconostles were distinguished from tunas mainly for their small size and light seeds. Xoconostles are produced by specific cactus pear plants, bearing fruits which are prized for their fleshy and acidic mesocarp. They are morphologically different from their cousins, the cactus pears, which are recognized by their sweet, juicy and seedy endocarp. Both plants grow under the semiarid conditions of the highlands of Central Mexico.

It is worth mentioning that a correlation was obtained between the seed mass and the domestication level, since domesticated genotypes tend to have heavier abundant seeds. Increased seed mass could be related to the reduction of the number of seeds, which is a desirable characteristic in selection programs of cactus pear breeding. Even if we are not evaluating the number of seed per fruit in our study, Foster et al., (2013) found a negative correlation between seed number and seed mass in *Opunta littoralis*. This is consistent with a life history strategy in which an organism adapts aspects of its biology, in order to optimize size and age at reproductive maturity, by balancing the number of offspring. The existence of a negative correlation between these two components (mass and number of seeds) is reported in other crops by Dreccer *et al.* (2009) and Fang *et al*. (2011).

Finally, digital image processing has significantly gained in acceleration during the last two decades, this trend occurred mainly because classification by visual examination is intrinsically applied in the fields. Seed image analysis has been used to differentiate genotypes and crops (Mebatsion *et al.*, 2012), these authors recognized variables such as area, perimeter and major axis to differentiate genotypes. In the present study, we showed the discriminative potential of variables derived from image analysis. Cactus seeds therefore are highly potential for use in studies of identification and description of Opuntia variants.

CONCLUSIONS

In the present study, we used statistical procedures to assess the descriptive ability of seeds variables (internal and external) and its possible use for Opuntia characterization. The seed variables were obtained from digital images whose handling and processing are easier, faster and less expensive. The area and the perimeter of seed, the major axis length,

the minor axis length, the area and the perimeter of the embryo, the area and the perimeter of the perisperm, the seed weight, and the Feret diameter, all had a high discriminative potential for accessions description and separated the xoconostles and the tunas through multiple comparisons analysis, principal component analysis and cluster analysis.

ACKNOWLEDGEMENTS

The first author receives a doctorate scholarship from the "*Secretaría de Relaciones Exteriores*", Mexico. Sr. Cruz-Miranda, F.M. (UACh-"Facundo Barrientos Perez" Germplasm Bank) and Dr. Gallegos-Vazquez C. (CRUCEN-UACh Germplasm bank, Zacatecas) are greatly appreciated for providing plant material. We thank Dr. Rodriguez De La O J. L. for lending us stereoscope service.

REFERENCES

Dreccer, M.F., van Herwaarden, A.F., and Chapman, S.C. 2009. Grain number and grain weight in wheat lines contrasting for stem water soluble carbohydrate concentration. Field Crops Research 112: 43-54.

Erre, P., and Chessa, I. 2013. Discriminant analysis of morphological descriptors to differentiate the O*puntia* genotypes. Acta Hortic. (ISHS) 995: 43-50.

Fang, H. X., Li, Z., Jin, Z. Z., and Lai, W. J. 2011. Preliminary study on the relationship between yield and other agronomic traits in wheat. Chinese Seed 2(1): 37-40.

Foster, K., Hoey, C., and Carmona-Galindo, V. D. 2013. An examination of *Opuntia littoralis* fruit volume, sugar concentration, number of seeds and average seed mass in relation to fitness. BIOS 84: 89-91.

Gallegos-Vázquez, C., Mondragón-Jacobo, C., Núñez-Colín, C.A. and Scheinvar, L. 2013. Germplasm diversity of xoconostles (*opuntia* spp.) or acidic cactus pears from central Mexico. Acta Hort. (ISHS) 995:75-81.

Guerrero-Muñoz, P., Zavaleta-Mancera, H. A., Barrientos-Priego, A. F., Gallegos-Vaques, C., Núñez-Colín, C. A., Valadez-Moctezuma, E. and Cuevas-Sánchez, J. A. 2006. Técnica para el estudio de la micromorfología interna de semillas duras en *Opuntia*. Rev. Fitotec. Mex. 29: 37-43.

Mebatsion, H. K., Paliwal, J. and Jayas, D. S. 2012. Evaluation of variations in the shape of grain types using principal components analysis of the elliptic Fourier descriptors. Comp. Elect. Agri. 80: 63-70.

Reyes-Agüero, J. A., Aguirre-Rivera, J. R., Carlín-Castelán, F. and González-Durán, A. 2013. Diversity of wild and cultivated *Opuntia* variants in the meridional highlands plateau of Mexico. Acta Hortic. (ISHS) 995: 69-74.

Rojas-Aréchiga, M., Mandujano, M. C. and Golubov, J. K. 2013. Seed size and photoblastism in species belonging to tribe Cacteae. J Plant Res 126: 373-386.

Stuppy, W. 2002. Seed characters and the classification of the Opuntioideae. Succ. Plant Res. 6: 25-58.