Scientia Horticulturae 192 (2015) 441–452

Contents lists available at ScienceDirect

Scientia Horticulturae

journal homepage: www.elsevier.com/locate/scihorti

Seeds morpho-colourimetric analysis as complementary method to molecular characterization of melon diversity

Diego Sabato[a], Cristina Esteras[b], Oscar Grillo[a,b,c], Belén Picó[b,*], Gianluigi Bacchetta[a]

[a] Centro Conservazione Biodiversità (CCB), Dipartimento di Scienze della Vita e dell'Ambiente, Università degli Studi di Cagliari, v.le Sant'Ignazio da Laconi 11-13, 09123 Cagliari, Italy
[b] Instituto de Conservación y Mejora de la Agrodiversidad Valenciana (COMAV), Universitat Politécnica de Valéncia, Camino de Vera s/n, 46022 Valencia, Spain
[c] Stazione Consorziale Sperimentale di Granicoltura per la Sicilia (SSGS), via Sirio 1, 95041 Borgo Santo Pietro, Caltagirone, Italy

ARTICLE INFO

Article history:
Received 19 September 2014
Received in revised form 27 April 2015
Accepted 2 June 2015
Available online 7 July 2015

Keywords:
Cucurbitaceae
Cucumis melo
Genetic characterization
Old landraces
Seed image analysis
Wild crop relatives

ABSTRACT

Melon has undergone an intense process of selection and crossbreeding, resulting in many landraces distributed all over Europe, Africa and Asia. Due to this huge variability, the systematic position of this *taxon* has been reviewed many times in the last two decades. The goal of this article is to compare the phenotypic characterization achieved by seed features with the molecular analysis on melon genotypes. A set of 124 accessions of *Cucumis melo* has been selected for molecular and morpho-colourimetric analyses plus an additional selection of accessions of *Cucumis sativus*, *Citrullus lanatus* and *Citrullus colocynthis* used to highlight seed morphology distances among genus and species. Genotyping was performed on the basis of 211 polymorphic SNPs and was executed using the iPLEX® Gold MassARRAY Sequenom technology. A total of 137 parameters were specifically designed to evaluate seed colour, size, shape and texture. Both molecular and seed morpho-colourimetrical analyses confirm the existence of two melon subspecies while an intermediate group has also been found. A non random allelic distribution in SNPs located in specific genomic regions suggests that some of these regions may account for a part of the observed variation in seed size. Six major groups of varieties can be discriminated on the basis on seed traits.

1. Introduction

The Cucurbitaceae family includes about 130 genera and 800 *taxa* (Jeffrey, 2005; Jeffrey and De Wilde, 2006). Among them, the most economically important species are *Cucumis melo* L. (melon), *Cucumis sativus* L. (cucumber), *Citrullus lanatus* (Thunb.) Matsum & Nakai (watermelon) and *Cucurbita* L. spp. (gourds and squashes).

Melon is worldwide diffused and comprises wild, feral and cultivated varieties, including sweet melons used for dessert and non-sweet ones consumed raw, pickled or cooked (Kirkbride, 1993; Bates and Robinson, 1995). Africa has been traditionally thought to be the centre of origin of this species. However, due to the high level of variation found in Asia, especially in India, melon could have originated there and then reached Africa (Renner et al., 2007; Sebastian et al., 2010). Other theories suggest that two independent domestications took place (Jeffrey, 1980; Esquinas-Alcázar and Gulick, 1983; Mallick and Masui, 1986; Bates and Robinson, 1995).

In the Mediterranean area the presence of melon is recorded in Egypt since the third millennium BC (Zohary et al., 2012), and in Greece and Italy since at least the late Bronze age (Megaloudi, 2006; Sabato et al., 2015). First representations show fruits likely belonging to the non-sweet melon varieties *chate* and *flexuosus* (with long cucumber-like fruits) (Janick et al., 2007). The presence of round sweet melons in the Mediterranean basin till the Classical age is uncertain, but it is well proven since the 11th century AD by Arabian trade with Central Asia (Paris et al., 2012).

C. melo has been traditionally separated into two subspecies, *melo* and *agrestis* (Naudin, 1859), each one including different varieties (Munger and Robinson, 1991). Pitrat et al. (2000) recognized 16 varieties: *cantalupensis* Naudin, *reticulatus* Ser. (cantaloupes, muskmelons), *inodorus* H. Jac. (winter melons, casaba melons), *flexuosus* L. (snake melons), *chate* Hasselq. (cucumber melons), *adana* Pangalo, *chandalak* Gabaev, *ameri* Pangalo (Asian melons), *chito* C. Morren (American melons), *dudaim* L. (pocket melons), and *tibish* Mohamed within the subsp. *melo* L. and *acidulus* Naudin, *conomon* Thunb., *makuwa* Makino and *chinensis* Pangalo (pickling melons),

* Corresponding author.
E-mail address: mpicosi@btc.upv.es (B. Picó).

http://dx.doi.org/10.1016/j.scienta.2015.06.006
0304-4238/

and *momordica* Roxb. (snap melons) within subsp. *agrestis* Naudin. In later revisions, Pitrat (2008) merged some varieties and Esteras et al. (2009, 2013), after further molecular studies, moved *tibish* and *chito* into the subspecies *agrestis*. Some of these varieties are quite heterogeneous, and accessions displaying intermediate features are difficult to classify. The wild forms of melon, usually referred to as *C. melo* subsp. *agrestis* var. *agrestis*, are mainly distributed in North and Eastern Africa and in Asia, while free-living forms of small size fruited melons have been found in Northern Australia, Southern USA and Central America (Roy et al., 2012). The sweet *cantalupensis*, *reticulatus* and *inodorus* melons are the ones with the most commercial interest worldwide (Pitrat, 2008).

In order to establish the genetic relationships among subspecies and varieties, several molecular studies have been carried out in melon, employing different markers (reviewed in Esteras et al., 2012). Most of them support the division at the sub-specific level and have contributed to better reclassify some of the varieties (Stepansky et al., 1999; Deleu et al., 2009; Esteras et al., 2009, 2013). SNPs (single nucleotide polymorphisms) are high-quality markers mostly used for genome-wide surveys in high to medium-throughput genotyping platforms (Fan et al., 2006; Steermers and Gunderson, 2007; Gabriel et al., 2009). The number of SNPs available in melon has largely increased in the last few years (Blanca et al., 2011, 2012; http://melogene.net/; Garcia-Mas et al., 2012). Esteras et al. (2013) reported the first application of a GoldenGate genotyping platform to analyze a melon core collection with SNPs distributed throughout the genome, demonstrating their usefulness for genetic diversity and population structure studies.

Also phenotypic variability has been studied with different core collections (Stepansky et al., 1999; Esteras et al., 2009; Leida et al., 2015), and with germplasm from specific centres of diversity (reviewed in: Esteras et al., 2012; Raghami et al., 2014). These phenotyping assays have basically focused on fruit traits and on the responses to biotic and abiotic stress, and many QTLs controlling these traits have been mapped in the melon genome (Diaz et al., 2011). In contrast to other species for which extensive efforts have been made in mapping QTLs for seed properties (Cai et al., 2012), and even in cloning the underlying genes (Orsi and Tanksley, 2009), only some studies have included seed traits in melons. Some of them report a significant correlation between seed traits and botanical classification (Stepansky et al., 1999; Yashiro et al., 2005; Tanaka et al., 2007). Fujishita and Nakagawa (1973) pointed out that seed size is one important trait for the identification of melon varieties. Fujishita (1980) described *makuwa* and *conomon* varieties with seeds smaller than 9 mm, *reticulatus* with seeds larger than 9 mm and *momordica* with intermediate seeds. Also in a recent study Tanaka et al. (2013) associated the variation in seed length and weight to chloroplast genome variation. However, seed traits have not been extensively analysed in large collections, representing the whole diversity of the species. In other cucurbits like *Cucurbita pepo*, seed traits have been used as discriminating factors since early studies (Decker and Newsom, 1988), and correlation between seed and fruit traits has been reported (Paris and Nerson, 2003).

Since the inception of the taxonomy, hierarchical classifications have been constructed on the basis of morphology, and molecular analyses can support those parts of phylogeny for which morphological data is lacking (Scotland et al., 2003). Morpho-colourimetric evaluations are commonly employed as tools to assess shape, size and colour of objects (Bacchetta et al., 2008; Venora et al., 2009; Grillo et al., 2010). Several works about the application of image analysis to the diaspores of wild vascular flora have been carried out, providing excellent results of classification within taxonomic units close to infra-generic and infra-specific levels (Bacchetta et al., 2011a,b; Grillo et al., 2012; Pinna et al., 2014). Many studies have been focused also on crop wild relatives and landraces (Venora et al., 2007a,b; Smykalova et al., 2011; Smykalova et al., 2013), and recently many authors focused on the *Vitis vinifera* complex (Rivera et al., 2007; Terral et al., 2010; Orrù et al., 2013a,b).

The knowledge of the existing diversity in melon is important, not only for its conservation, but also for its exploitation in commercial breeding, as this species displays crossability problems with other species of the genus *Cucumis*.

The goals of this research are to:

- Compare the groups established using molecular analyses with those achieved by seed characters.
- Analyze the variability of morpho-colourimetric seed features.
- Implement statistical classifiers able to discriminate among the studied varieties.
- Increase the knowledge about the variation of the current extant melon seed collections.

2. Methods

2.1. Seed lots

The whole melon collection was established on the framework of a previous project (MELRIP 2007–2010, Esteras et al., 2009, 2013; Leida et al., 2015). Accessions were characterized for vine and fruit traits, multiplied and conserved at the COMAV Genebank (Institute for the Conservation and Breeding of the Agrobiodiversity). We selected 124 seed lots (103 for both molecular and seed analyses plus 21 only for morpho-colourimetric analysis of the seeds). Seeds belonged to accessions from 48 countries and represented all melon varieties. Fruits were collected at the optimum maturity stage, corresponding to the complete morphologic and chromatic seed development. To avoid over-representation of single plants and/or fruit features, seeds from the highest number of plants and fruits available for each accession were analyzed. Undeveloped, deformed and sterile seeds were excluded. For further details about the composition of the analyzed collection see Supplementary data.

With the purpose to evidence morphological distances at the genus and species levels, a small set of close relatives of melon were used: twenty-one accessions of *C. sativus*, 18 of *C. lanatus* and 9 of *C. colocynthis*, were selected (see Supplementary data). All accessions were supplied by COMAV Genebank and represent mainly Mediterranean, African and Asian landraces.

2.2. Molecular analysis

The DNA was extracted from young leaves using the CTAB method with minor changes (Esteras et al., 2013). Genotyping was done with a total of 211 polymorphic SNPs, evenly distributed throughout the genome, that were selected from the SNP melon collection available in the Melogene database (http://www.melogene.net/) and *in silico* identified in two previous re-sequencing analysis (Blanca et al., 2011, 2012). Genotyping was performed using the iPLEX® Gold MassARRAY Sequenom technology at the Epigenetic and Genotyping Unit of the University of Valencia (Unitat Central díInvestigació en Medicina UCIM). The basis of this technology is described in Gabriel et al. (2009).

The genotyping results were employed to perform a cluster analysis using the PowerMarker software (Liu and Muse, 2005). Nei's genetic distance (Nei et al., 1983) was used, and the support values for the degree of confidence at the nodes of the dendrogram were analyzed by bootstrap re-sampling 1000 times. Phylip 3.69 software (Felsenstein, 1997) was employed to construct the consensus tree and TreeView32 (Page, 1996) to visualize it. Genotyping summary statistics such as the number of alleles, the frequency of the most common allele (MAF) and the polymorphism informa-