Genetic diversity of Spanish *Cucurbita pepo* landraces: an unexploited resource for Summer Squash breeding.

Gesolmina Formisano[1*], Cristina Roig[2*], Cristina Esteras[2], Maria Raffaella Ercolano[1], Fernando Nuez[2], Antonio Jose Monforte[3], Maria Belen Picó[2]

[1]Department of Soil, Plant, Environmental and Animal Production Sciences, University of Naples 'Federico II', Via Università 100, 80055 Portici, Italy
[2]Institute for the Conservation and Breeding of Agricultural Biodiversity (COMAV-UPV), Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, Spain
[3]Instituto de Biología Molecular y Celular de Plantas (IBMCP), Universitat Politècnica de València (UPV)-Consejo Superior de Investigaciones Científicas (CSIC), Ciudad Politécnica de la Innovación (CPI), Ed. 8E, C/Ingeniero Fausto Elio s/n, 46022 Valencia, Spain

*These two authors contributed equally

Corresponding author
Maria Belén Picó Sirvent
COMAV, Ciudad Politécnica de la Innovación, CPI, Edificio 8E, Escalera J. Universitat Politècnica de València (UPV), Camino de Vera s/n, 46022, Valencia. Spain
E-mail: mpicosi@btc.upv.es
Phone: 34 96 3879415 Fax: +34 96 3879422

Abstract

Cucurbita pepo is a world-wide cultivated vegetable of American origin. Most of the widely grown comercial types are known as Summer squashes and belong to the elongated forms of *C. pepo* ssp. *pepo* (Cocozelle, Vegetable marrow and Zucchini Groups). These forms were developed in Europe after the arrival of the first american landraces through a process of selection and fixation that led to the loss of genetic diversity. Part of the genetic variability included on the first American cultigens remains in diverse landraces, still cultivated for self-consumption and sale in local markets. Using the first colecction of genomic and EST derived microsatellites that has just become available for the species, we compared the natural variation present in a collection of Spanish landraces with that of a set of commercial varieties and hybrids, representing the current Summer squash market offer. A total of 194 alleles allowed us to distinguish all genotypes, even those closely related. In general, Cocozelle and Vegetable marrow, Groups with considerably long histories, were more variable than the Zucchini Group, of more recent origin. We found significant genetic diversity among landraces. Variation present in the local accessions belonging to the Zucchini Group was larger than that of commercial cultivars. Cluster, principal coordinates and population structure results suggested that the variation of the Spanish landraces has not been extensivelly used in breeding. Commercial Summer squashes can then benefit from this underexploited variability, specially from some landraces that already displayed favourable commercial traits.

Key words: breeding, *Cucurbita pepo*, landraces, microsatellites, Summer squash.

Introduction

Cucurbita pepo L. is the most economically important species of the genus *Cucurbita* L. (Cucurbitaceae). Cultivated *C. pepo* has been traditionally considered to comprise two subspecies (Decker 1988; Sanjur et al. 2002; Nesom 2011), each one encompassing several cultivar-groups: *C. pepo* ssp. *pepo* L. (including Pumpkin Group, Vegetable Marrow Group, Cocozelle Group, and Zucchini Group) and *C. pepo* ssp. *texana* (Scheele) Filov (syn ssp. *ovifera* (L.) Decker) (including Acorn Group, Scallop Group, Crookneck Group, and Straightneck Group) (Ferriol et al. 2003; Paris 1986; Paris et al. 2003). The major economic value of this species is based mainly on the culinary use of immature fruits often referred to collectively as "Summer squash". Only cultivars of the Pumpkin and Acorn Groups display a major use as "Winter squashes" grown for consumption of their mature fruits (Paris 2008).

This species originated in North America (Smith 1997), although nowadays its distribution is worldwide, being one of the most phenotypically variable species in the plant kingdom. Wild relatives and ancient cultigens of the species, mostly round or nearly round fruited, are still found in USA, Mexico and Central America, where some variable morphotypes were selected by native Americans in pre-Columbian times (Paris, 2000, 2008). A number of these accessions are housed in facilities of Mexico, USA, and Costa Rica (Lira and Montes 1994).

The ssp. *texana* morphotypes were developed in America and today are more popular in this continent, being hardly found in Europe. Since the arrival to Europe of some American cultigens, an extraordinary variability of new phenotypes has been generated through hybridisation and recombination. Selection for long-fruitedness in the subspecies *pepo* apparently was first conducted in Italy in the 16th century (Paris 2008). New cultivars were then developed, particularly the elongated forms of this subespecies, first the Vegetable Marrow Group (short, tapered cylindrical fruits), followed soon afterward by the Cocozelle Group (long or very long, bulbous cylindrical) and much more recently, probably in the late 19th century, by the Zucchini Group (long, uniformly cylindrical fruits). The Zucchini morphotype is the one of most recent origin, and one of the less variable (Paris 2000). Nowadays, cultivars of the Zucchini Group dominate the squash market and breeding efforts of seed companies.

Summer squashes of the Vegetable Marrow, Cocozelle and Zucchini Groups are popular in the Middle East, North Africa and European countries bordering the Mediterranean Sea, with cultivars meeting particular consumers preferences in each area. Diverse collections of landraces can be still found in Italy, Turkey (Paris, 2008) and Spain (Ferriol and Pico 2008). In the last century, breeders have make a big

effort in Summer squash breeding by performing selection and fixation from some of these landraces to develop new improved cultivars and hybrids, that are now replacing old landraces.

In Spain, the Genebank of the Institute for the Conservation and Breeding of Agricultural Diversity (COMAV), located at the Polytechnic University of Valencia, is dedicated to preserve this genetic heritage. This Genebank maintains seed collections of old Winter and Summer squash Spanish cultigens of *C.pepo* ssp. *pepo* (Ferriol et al. 2003). These were mainly collected from the farmers that still use them for self consumption and/or sale in local markets. One of the aims of the COMAV Institute is to characterize these landraces in order to use it in Summer squash breeding.

The genetic variability within *C. pepo* has been previously assesed using allozymes and different DNA marker systems (Restriction Fragment Length Polymorphism, RFLP, Random Amplified Polymorphic DNA, Amplified Fragment Lenght Polymorphism, AFLP, and Inter Simple Sequence Repeats, ISSRs) (reviewed in Lebeda et al. 2007 and in Esteras et al. 2011). Most of the studies have been focused on the assessment of the genetic and evolutionary relationships between the wild types and the domesticates, between the two subespecies, and among the cultivar groups, including only a few representatives of European landraces (Ferriol et al. 2003; Paris et al. 2003).

Most of the markers systems used to date have limitations associated to their dominant and/or unreliable nature. Microsatellite markers (Simple Sequence Repeats, SSRs), if available, are preferred for being reliable and codominant, but also multiallelic and highly polymorphic, appropriated for detecting variation among closely related varieties. However, these markers are difficul to obtain in species for which genetic and genomic tools are lacking.

Cucurbits are becoming primary models for the study of several biological processes (Boualem et al. 2008; Ezura and Fukino 2009; Li et al. 2009). The availability of molecular tools, including the whole genome sequence, has significantly increased in the last years, mainly for melon and cucumber (Huang et al. 2009; Gonzalez et al. 2010), but more recently also for the *Cucurbita* genus.

Gong et al. (2008) developed the first collection of genomic SSRs (gSSR), about 400, using an enriched genomic library from an oil-seed Pumpkin cultivar. Recent advances in next-generation sequencing technologies (Metzker, 2010) has allowed us to generate the first *Cucurbita* transcriptome (Blanca et al. 2011), with 49,610 *Cucurbita* unigenes *de novo* assembled from 512,751 high quality ESTs, sequenced using Roche GS/454. These unigenes were screened for SSRs motifs, leading to the discovery of a collection of 1,882 unigenes with SSR motifs (EST-SSR), what could be considered as the first genomic

resource in this genus. EST-SSRs have the advantatge over gSSRs of being in the functional fraction of the genome.

In this paper, we evaluate the utility of these new genomic resources for assesing genetic variation within *Cucurbita pepo*, even among closely related cultivars. Boh gSSRs and EST-SSRs have proven to be useful for genotyping a selected set of Spanish landraces, belonging to the three Summer squash morphotypes with high commercial value. These varieties have been also phenotyped for the presence of traits of agronomic interest. Most of the genetic **v**ariation that still remains in these landraces is not present in a set of commercial varieties and hybrids, representative of the current offer of Summer squashes in the European market. Their diversity and the presence of some unique alleles make the Spanish landraces of *C. pepo* an unvaluable resource with potential in summer squash breeding, mainly the most widely grown and less variable Zucchini Group.

Materials and Methods

Germplasm

A set of twenty three Spanish landraces, included in the core collection of *C. pepo* held and characterized by the Cucurbits breeding group at COMAV, representing the three more important morphotypes of *C. pepo ssp*. *pepo* (five belonging to the Zucchini Group, eleven to the Vegetable Marrow Group, and seven to the Cocozelle Group), and the variability found in all the agro-ecological regions in the country (Table 1, Figure 1), were selected for this study. All the selected accessions are primitive landraces, cultivated in small orchards used for self-consumption or sale in local markets. Original germplasm, maintained through sibing, was used for the molecular study.

Twelve commercial varieties (three belongind to the Zucchini Group, seven to the Cocozelle and two Summer squash Pumpkins) and six commercial hybrids (four, one and one belonging to the Zucchini,Vegetable marrow and Coccozelle Groups, repectivelly) were included, as representatives of the main commercial varieties currently offered in the market (Table 1). Other accessions were included as controls, one South American accession of the Zucchini Group, one North African accession of the Vegetable marrow Group, two Pumpkin accesions, used as parentals of mapping populations: the oil-seed styrian pumpkin (Gong et al. 2008) and the Turkish variety PI171678, that is being used for increasing the levels of carotenoids in *C. pepo* (Ferriol et al. 2003), and three accessions belonging to 3 morphotypes of the ssp. *texana*, one of each Group Acorn, Scallop and Crookneck.

Morphological characterization

Three plants per accession were grown at the greenhouse and characterized for different vine, flower and fruit traits of commercial interest (Table 2): plant growth habit (bushy (B), compact with short internodes, viney (V), spreading growth habit, with long thin internodes, and intermediate (I), semi-bushy; lateral branch development (no-branching (0), moderately branched (0.5), highly branched (1)); presence of spines in stems and leaves (spineless (0), sparse and small spicules (0.5), dense, large and sharp spicules (1)); days to male and female flowering from transplant (DMF and DFF); node in which the first male/female flower appears (NMF and NFF); and number of male/female flowers 7 days after the opening of the first female flower (NºMF and NºFF) and the femaleness ratio NºFF/ NºMF. Mature fruits were also characterized (Figure 1).

SSR analysis

Genomic SSRs

A set of 30 genomic SSRs (gSSRs) were used for the analysis (Table 3). These were selected among those developed by Gong et al. (2008), using a SSR-enriched partial genomic library prepared from the Austrian oil-pumpkin variety Gleisdorfer Ölkürbis (*C. pepo* ssp. *pepo*). We selected at least one marker in each of the 20 linkage groups of *C.pepo* ssp. *pepo* (oil-pumpkin variety) x *C. pepo* ssp. *texana* (vriety belonging to the Crookneck Group) map developed by these authors (Table 3).

EST-SSRs

We used a set of 27 ESTs-SSRs selected among those *in silico* identified by Blanca et al. (2011) after screening for SSR motifs the 49,610 unigenes of the first *C. pepo* transcriptome (assembled using ESTs from the cultivars ZU-MU16 (Zucchini Group) and UPV-196 (Scallop Group)) (Table 4). EST sequences containing perfect repetitions of more than seven units were selected. Most of the SSRs are within genes with a putative known function and were mainly located in ORFs (open reading frames), with only afew of them in UTRs (untranslated regions). Most of the ESTs-SSRs selected were in unigenes that showed homology to annotated protein sequences (Blanca et al. 2011). Some were involved in catabolism, transport, regulation of transcription, stress and defense response (Table 4). Primers used for each SSRs *locus* are listed in Blanca *et al*. (2011).

PCR amplification and detection of SSR loci

DNA was extracted using a modified CTAB method (Doyle and Doyle 1990) from young leaves of three plants per accession. PCR reactions were performed in a final volume of 15 µL with 15 ng of genomic

DNA, 1.5U Taq DNA polymerase (Biotools), 1X PCR buffer (75mM Tris-HCl pH 9, 20mM (NH4)2SO4, 50mM KCl), 3 mM MgCl2, 200 µM of each dNTP, 0.15 µM each forward and reverse primers, 0.02µM/0.2µM dye-labelled primer (FAM, VIC, PET or NED for ABI sequencer and IRDye® 700 or 800 for LI-COR). Forward primers were designed with an M13 tail sequence added to its 5' end (5'-cacgacgttgtaaaacgacc-3'). The cycling conditions were as follows: denaturation at 95ºC for 3 min, followed by 10 cycles of 30 s at 95ºC, 30 s at 65ºC (with each cycle the annealing temperature decreasing 1ºC), and of 30 s at 72ºC. Products were subsequently amplified for 20 cycles at 95ºC for 30 s, 55ºC for 30 s and 72ºC for 30 s, with a final extension at 72ºC for 5 min. SSR markers were scored either on the LI-COR®4300 DNA Analyzer (Li-Cor Bioscience, Lincoln, Nebraska, USA) and on the ABI 3130XL capillar sequencer (Applied Biosystems, Carlsbad, California, USA).

Genetic variability analysis

Due to the bulking of samples during DNA extraction, the observation of two or more SSR alleles in a single reaction could have resulted from the presence of heterozygous plants, homozygous plants for the alternative alleles, or a combination of both. All the detected alleles were assumed to have a frequency of 1/n (n = number of alleles). Microsatellite allele sizes were estimated comparing their migration with the IRDye® 700 or 800 50–350 bp size standards (Li-Cor Bioscience, Lincoln, Nebraska, USA). Number of alleles, and the polymorphism information content (PIC) were calculated for each *locus* for the 48 genotypes, and also for groups of morphotypes, using the PowerMarker software (Liu and Muse 2005). ANOVA was used to compare mean allele number and PIC between EST-SSRs and gSSRs. Correlation between the number of repeated units and PIC was also studied. All statistical analyses were performed with Statgraphics plus v.5.0 (Statistical Graphics. Corporation, Inc., Rockville, MD, USA).

Genetic similarity based on the presence or absence of SSR alleles between accessions was calculated using the Dice (Nei and Li 1979) similarity coefficient (S_{ij}) Genetic distance (GD) values among genotypes were calculated as $1 - S_{ij}$. From the genetic distance matrix a neighbor-joining (NJ) tree was obtained using Powermarker. Distance matrices obtained with gSSRs and EST-SSRs were compared by the Mantel test (Mantel 1967). The reliability and robustness of the tree were tested by bootstrap analysis with 1,000 replications to assess branch support using PHYLIP 3.6 software. The population structure of our collection was estimated using STRUCTURE version 2.3 (Pritchard et al., 2000, Falush et al. 2003, 2007, Hubisz et al., 2009). Accesions CR-NSL5227, AC-PI615111 and SC-UPV196 were excluded from STRUCTURE analysis because they belong to ssp. *texana* while the rest of genotypes belong to ssp.

pepo. Admixture and independent allele frequencies options were chosen. A burn-in period of 500,000 Markov Chain Monte Carlo iterations and the 1,000,000 iterations after burn were performed to estimate the parameters. Twenty runs were done for each K (number of populations) from 1 to 8. The K optimum was defined according Evanno et al. (2005). Principal coordinate analysis (PCoA) was carried out using the Dice genetic distance matrix and the DCENTER and EIGEN procedures of NTSYSpc 2.02 (Rohlf 1998).

RESULTS AND DISCUSSION

Phenotypic variation

The phenotypic variability found among the Summer squash genotypes used in the present study is provided in Table 2. The landraces morphological characterization allowed us to define the morphotype of each accession and to record the presence or the absence of traits of commercial interest.

The Zucchini Group is by far the most widely grown Summer squash and for this reason it is the morphotype in which the seed companies have mainly focused their breeding efforts. The Zucchini commercial hybrids and varieties were, as expected, bushy, unbranched, spineless and with uniformly cylindrical fruits. The bushy growth habit, conferred by a single gene (*Bu*) (Paris and Brown 2005), greatly facilitates multiple harvesting, and all the current commercial Zucchini types have this gene introgressed. The lack of branching improve an efficient picking of the fruits, and few or absent spicules is a desireable trait for any cultivar. Commercial cultivars of the Zucchini Group were also early flowering, with male and female flowers starting to appear from 30 to 36 DAP (Days after planting) and from 41 to 52 DAP respectivelly, and with a high female/male ratio (0.3-0.6). Selection for earliness and female tendency was another important achievement of pre-20th century squash breeding (Paris 2008).

In the Cocozelle commercial Group, bushy growth habit was not so common. Vine and semibushy more branched and spiny varieties were found. A more delayed flowering (male and female flowers appearing from 32 to 45 DAP, and 43 to 61 DAP) associated to a decreased female/male ratio (0.1-0.5) was also displayed. The commercial Vegetable marrow and Pumpkin Groups displayed intermediate traits.

Spanish landraces are mostly primitive cultivars no subjected to intensive selection. They harbor a great deal of phenotypic variation. Most of the Zucchini landraces were bushy rather than viney, and develop uniformaly cylindrical dark-green fruits (Figure 1). Some landraces showed desired commercial trais, being unbranched, spineless and early flowering, and even more positive female/male ratio than some commercial varieties (ZU-E27 and ZU-MU16). Others have some traits commercially undesireable,

related to more ancestral *C. pepo* (ZU-E10, ZU-MU20, ZU-A13), also found in the landrace from Ecuador (ZU-ECU227). Our results are in agreement with the primitive traits reported in Italian and Turkish landraces (Paris 2008).

Phenotypic variation was more apparent among Spanish landraces of the Vegetable marrow and Cocozelle Groups, being indeterminate or viney, much branched from medium to highly spiculate foliage, and from early to late flowering (male and female flowers appearing from 30 to 63 and DAP, and 45 to 77 DAP, respectivelly), with some accesions with a predominant male sexuality (female/male ratio 0.05-0.2), mainly in the Cocozelle group. Only one Cocozelle variety was bushy and no branched (CO-V185). High variability in fruit color, color pattern, rind texture and fruit size was found in Vegetable marrow and Cocozelle Spanish landraces, ranging from yellow to cream, white, light, medium and dark green, striped, dotted, smooth, ribed, wrinkled and/or warted (Figure 1).

Molecular variation

gSSR and EST-SSRs

A summary of results obtained with genomic- and EST-SSRs are shown in tables 3 and 4. Amplification products were obtained for 29 (97%) gSSRs, being 27 (90%) polymorphic among the *Cucurbita* genotypes. A total of 109 alleles were found across the full set of accessions, ranging from 2 to 9 (average 3.8) alleles per SSR. The average allele number in our collection was similar to that reported by Gong et al. (2008) (3.7), who validate these gSSRs using a set of 8 genotypes of both subespecies of *C. pepo* (one cultivar of each Group: Zucchini, Cocozelle, Vegetable marrow, oil-Pumpkin, Acorn, Scallop, Croockneck and Straighneck), but also including 3 accessions of the related species *C. moschata* and 1 of the wild species *C. ecuadorensis*. Despite we have a larger number of accessions, we assesed here only intraspecific variability, so we expected the SSR polymorphism to be lower in our collection. Nevertheless, our results demonstrate that this germplasm sample is highly variable, noteworthy 12 of these markers (40%) detected more alleles in our population than in that of Gong et al. (2008).

Twenty-six ESTs-SSRs (96%) amplified polymorphic fragments in the set of *Cucurbita* genotypes. The total (85) and average allele number (3.2), was similar, no significant differences (p=0.11), to that obtained with gSSRs, and also similar to that previously reported for this set of ESTs-SSRs, validated using 9 genotypes representative of the diversity within *C. pepo* (4 Zucchini, 1 Vegetable marrow, 1 Pumpkin, 1 Styrian Pumpkin, 1 Croockneck and 1 Scallop) and 1 *C. moschata* accession (Blanca et al., 2011).

The allele diversity was similar for both SSRs sets, with PIC values ranging from 0.1 to 0.74 (mean, 0.42) in gSSRs and from 0.1 to 0.66 (mean, 0.36) in EST-SSRs, no significant differences (p=0.26). The correlation between the number of SSR repeats and PIC was positive and significant both for gSSR ($r = 0.49$, $p = 0.02$) and EST-SSRs (0.62, $p < 0.001$). Also in melon this correlation was positive , but with a considerable higher mean PIC value (mean 0.58 and 0.54 for EST and gSSRs respectively) (Fernandez-Silva et al. 2008; Fergany et al. 2011). Differences could be due to the fact that in melon analysis both subespecies are well represented and here we use only a few accessions of the subspecies *texana* as control.

These results confirm that both gSSRs and EST-SSRs discovered in a small germplasm sample can be transferred to different cultivar groups, being useful for depicting genetic relationships as well as for identification of closely related cultivars. Also as some of the tested SSRs are in functional regions, the study of their variability may give insight into functional variability and its possible relationship with phenotypic variability.

Molecular variability within the cultivar groups

In general, a lower genetic variability is found within commercial hybrids compared with commercial varieties and landraces (Table 5), which is consistent with the narrower origin of the hybrids and genetic erosion due to intensive breeding (Formisano et al 2010). On the other hand, commercial cultivars and Spanish landraces showed similar genetic diversity. When we analyze the genetic variability of the different groups, the Cocozelle Group show the highest genetic varibility among commercial cultivars, what is coherent with the more ancient characteristics found in this Group, more similar to ancient Mexican landraces, as this Group is thougth to have been developed by the 17th century (Paris 2000). The less variable Group among commercial cultivars, as expected, was the Zucchini. In fact 35 alleles were found exclusivelly in all the other morphotypes, Pumpkin, Cocozelle and Vegetable marrow, being absent in Zucchini, both the commercial and the landraces, even in the sudamerican control. However, no differences in genetic variability among cultivar groups was found for the Spanish landraces.

Twenty one alleles were detected exclusively within Spanish landraces, being absent in the commercial varieties of the types Zucchini, Cocozelle and Vegetable marrow. Spanish landraces belongiong to the Zucchini Group showed the highest number of exclusive alleles (27), followed by the 15 alleles exclusive of Cocozelle Spanish landraces. The number of alleles exclusive of the commercial cultivars was considerably lower (9 in each case). Therefore, despite the fact that the Zucchini Spanish landraces are

morphologically not very variable, they have a level of genetic diversity higher than the commercial types, so they can be considered a reservoir of alleles useful for breeding.

Relationships among varieties

The genetic relationships among accessions based on SSR polymorphism were investigated by cluster analysis. No significant differences were found in the average pair-wise distances based on gSSRs and ESTs-SSRs. The correlation between the two distance matrices was 0.83 ($P < 0.002$) according to Mantel's test (Mantel 1967), confirming that the new ESTs-SSRs set is as effective as the previously available gSSRs in establishing genetic relationships among summer squash accessions. Similar results were observed previously in melon when the polymorphism of gSSRs and EST-SSRs was compared (Fernandez-Silva et al., 2008) . Comparing different Groups, average intragroup pair-wise distances in the Cocozelle and Vegetable marrow (0.27 and 0.28) were significantly higher than in the zuchini group (0.20), consistently with a higher intragroup variability within the former. Intergroup distances were also significantly higher between the Cocozelle and the Vegetable marrow group (0.27), than between Zucchini and these two Groups (0.22).

The NJ dendrogram based on 109 gSSR and 85 ESTs-SSRs alleles shown in Figure 2 fits very well with previous classifications using different markers (Esteras et al. 2011; Ferriol et al. 2003; Paris et al. 2003) with two major clusters that clearly separated accessions of both subespecies (bootstrap = 1000). Two major patterns of genetic association were observed. Most the commercial accessions, except one Group of Cocozelle, and one Vegetable marrow and one Pumpkin variety, were separated from the Spanish landraces. Cocozelle commercial cultivars were distributed in several groups with some cutivars even being similar to more ancient Pumpkin types (CO-BDT and PI171675, bootstrap=548).

Despite the lower level of polymorphism, all Zucchini genotypes could be distinguished with the set of SSRs. Only one Zucchini landrace was molecularly similar to the group of Zucchini commercial cultivars (MU-16 was highly similar to ZU-GIO, bootstrap=895), whereas the other were genetically more similar to Vegetable marrow or Cocozelle types (ZU-A13 similar to VM-V21, bootstrap=629; ZU-E10 similar to CO-V74, bootstrap=690). Interestingly, the accession ZU-E27, phenotipically similar to commercial Zucchini cultivars, was molecularly similar to Cocozelle and Vegetable marrow landraces. This accession has potential value for introgressing molecular variation into the commercial Zucchini Group without introducing unfavorable commercial traits. Cocozelle landraces were highly variable intermingled with Vegetable marrow landraces spread across the dendrograme .

The population structure estimated by STRUCTURE analysis (Pritchard et al. 2000) revealed that the genotypes beloging to *C. pepo* ssp. *pepo* can be separated in three populations, most of the genotypes could be assigned clearly to one of the two major populations (green or red in Figure 2), only accession CO-BDT could be separated clearly from the two previous populations, whereas the remaining five accessions could not be assigned to any of the former populations. The results are consistent with the NJ cluster analysis, showing two major populations, one of them consisting mainly in commercial cultivars and the second one in Spanish landraces. All cultivar groups are represented in both populations.

Figure 3 represents the distribution of the different a ccessions according to the two principal axes of variation using principal coordinates analysis (PCoA). We performed a first PCoA which explained the 31% of the variation. On the basis of the first coordinate, which accounted for 15.4% of the total variation, the accessions were clearly grouped according to subspecies. We performed a second PCoA analysis after removing the ssp. *texana* accessions. The firts coordinate, explaining 10% of the variation, separated all the commercial cultivars of the Zucchini Group from the other commercial groups and from the Spanish landraces, supporting the differentiation of this modern group. Paris et al. (2003) and previous studies performed with first generation molecular markers (Esteras et al., 2011), also found that the Zucchini group, the most recent of the subspecies, is the most distinct of the edible fruited cultivar-groups of ssp. *pepo*. However, European ancient landraces are not included in most of these previous studies, and our study reveals that this distinction only occur in commercial Zucchni, but not in Zucchini landraces that are more similar to the oldest groups. The second coordinate, explaining 8% of the variation, separated the remaining commercial cultivars, mostly Cocozelle, but also Pumpkin and Vegetable marrow from the Spanish landraces .

All previous three analysis support that the genetic structure of the current summer squash sampled genotypes is mainly due to the type of cultivar (commercial or landrace), i. e., Spanish landraces are clearly separated from commercial varieties, although within Spanish landraces there is no clear genetic structure due to cultivar type. The fact that the Spanish landraces are not close related to the commercial cultivars of the same cultivar types seems to suggest that the gene variation included in the Spanish landraces has not been used extensively to develop commercial cultivars, confirming that this gene pool is a reservoir of genetic variability underexploited in modern summer squash breeding.

Conclusions

This is the first study with codominant markers of the diversity of a wide collection of European landraces belonging to the elongated forms of *C.pepo* ssp. *pepo*. Our results indicate that these accessions retain traits common to ancient cultivars of this species. These germplasm resources could be useful for the enrichment of the current commercial cultivars. They still conserve phenotipic and molecular variation that has been lost during the breeding process. It is particularly noteworhty the existence of landraces with favourable commercial traits, such as the ZU-E27 or the CO-V185, but molecularly variable, that could provide new alleles useful for breeding without altering some required commercial characteristics.

This new-found genetic potential in Spanish landraces not present in the restricted gene pool of modern varieties could be used for example to adapt market types to regional preferences of color, secondary design, shapes, but could also provide other interesting traits such us flavor, vitamins and mineral content that have been associated with color variation in summer squash. Other traits that are current objectives of Summer squash breeding, such as fruit glossiness, parthenocarpy, response to diseases could be also variable in this collection.

Acknowledgements

This research was funded by the project of the Spanish Instituto Nacional de Investigación y Tecnología Agraria INIA (RTA2008-00035-C02-02). G. Formisano was supported by the Italian Doctorate School in Agrobiology and Agrochemestry University of Naples “Federico II”.

References

Blanca J, Cañizares J, Roig C, Ziarsolo P, Nuez F, Picó B (2011) Transcriptome characterization and high throughput SSRs and SNPs discovery in Cucurbita pepo (Cucurbitaceae). BMC Genomics 10;12:104.

Boualem A, Fergany M, Fernandez R, Troadec C, Martin A, Morin H, Sari MA, Collin F, Flowers JM, Pitrat M, Purugganan MD, Dogimont C, Bendahmane A (2008) A conserved mutation in an ethylene biosynthesis enzyme leads to andromonoecy in melons. Science, 321:836-838

Decker DS (1988) Origin(s), evolution, and systematics of *Cucurbita pepo* (Cucurbitaceae). Econ Bot 42:4-15.

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15

Esteras C, Nuez F, Picó B (2011) Genetic diversity studies in Cucurbits using molecular tools. In: Wang Y, Behera TK (eds) Cucurbits: Genetics, Genomics and Breeding in Crop plants, Science Publishers Inc, Enfield, New Hampshire, pp 25.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14: 2611-2620.

Ezura H, Fukino N (2009) Research tools for functional genomics in melon (Cucumis melo L.): Current status and prospects. Plant Biotechnol. 26:359-368.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using mulitlocus genotupe data: linked loci and correlated allele frequencies. Genetics 164:1567-1587.

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using mulitlocus genotype data: dominant markers and null alleles. Mol Ecol Notes 7:574-578.

Fernandez-Silva I, Eduardo I, Blanca J, Esteras C, Picó B, Nuez F, Arús P, Garcia-Mas J, Monforte AJ (2008) Bin mapping of genomic and EST-derived SSRs in melon (*Cucumis mel*o L.). Theor Appl Genet 118: 139-150.

Fergany M, Kaur B, Monforte AJ, Pitrat M, Rys C, Lecoq H, Dhillon NPS, Dhaliwal (2011) Variation in melon (*Cucumis melo*) landraces adapted to the humid tropics of southern India. Genetic Resources and Crop Evolution 58:225-243.

Fernandez-Silva I, Eduardo I, Blanca J, Esteras C, Pico B, Nuez F, Arus P, Garcia-Mas J, Monforte AJ (2008) Bin mapping of genomic and EST-derived SSRs in melon (Cucumis melo L.). Theor Appl Genet 118:139-150.

Ferriol M, Pico B, Nuez F (2003) Genetic diversity of a germplasm collection of *Cucurbita pepo* using SRAP and AFLP markers. Theor Appl Genet 107:271–282.

Ferriol M, Picó B (2008) Pumpkin and Winter Squash. In: Prohens J, Nuez F (eds) Handbook of Plant Breeding, Vegetables I, Springer, New York, pp 317-349.

Formisano G, Paris HS, Frusciante L. Ercolano MR (2010) Commercial Cucurbita pepo squash hybrids carrying disease resistance introgressed from Cucurbita moschata have high genetic similarity. Plant Genet. Res.8:198-203.

Gong L, Stift G, Kofler R, Pachner M, Lelley T (2008) Microsatellites for the genus *Cucurbita* and an SSR-based genetic linkage map of *Cucurbita pepo* L. Theor Appl Genet, 117:37-48.

Gonzalez VM, Rodríguez-Moreno L, Centeno E, Benjak A, Garcia-Mas J, Puigdomènech P, Aranda MA (2010) Genome-wide BAC-end sequencing of *Cucumis melo* using two BAC libraries. BMC Genomics 11:618.

Huang S, Li R, Zhang Z, Li L, Gu X, Fan W, Lucas WJ, Wang X, Xie B, Ni P, Ren Y, Zhu H, Li J, Lin K, Jin W, Fei Z, Li G, Staub J, Kilian A, van der Vossen EA, Wu Y, Guo J, He J, Jia Z, Ren Y, Tian G, Lu Y, Ruan J, Qian W, Wang M, Huang Q, Li B, Xuan Z, Cao J, Asan , Wu Z, Zhang J, Cai Q, Bai Y, Zhao B, Han Y, Li Y, Li X, Wang S, Shi Q, Liu S, Cho WK, Kim JY, Xu Y, Heller-Uszynska K, Miao H, Cheng Z, Zhang S, Wu J, Yang Y, Kang H, Li M, Liang H, Ren X, Shi Z, Wen M, Jian M, Yang H, Zhang G, Yang Z, Chen R, Liu S, Li J, Ma L, Liu H, Zhou Y, Zhao J, Fang X, Li G, Fang L, Li Y, Liu D, Zheng H, Zhang Y, Qin N, Li Z, Yang G, Yang S, Bolund L, Kristiansen K, Zheng H, Li S, Zhang X, Yang H, Wang J, Sun R, Zhang B, Jiang S, Wang J, Du Y, Li S (2009) The genome of the cucumber, Cucumis sativus L. Nat Genet 41:1275-1281.

Hubisz MJ, Falush D, Stephens M, Pritcard JK (2009) Inferring weak population structure with the assistance of sample group information. Mol Ecol Resources 9:1322-1332.

Lebeda A, Widrlechner MP, Staub J, Ezura H, Zalapa J, Kristkova E (2007) Cucurbits (Cucurbitaceae; Cucumis spp., Cucurbita spp., Citrullus spp.). In: Singh RJ (ed) Genetic resources, chromosome engineering, and crop improvement, vol. 3. CRC Press, Boca Raton, pp 271–376

Li Z, Huang S, Liu S, Pan J, Zhang Z, Tao Q, Shi Q, Jia Z, Zhang W, Chen H, Si L, Zhu L, Cai R (2009) Molecular isolation of the M gene suggests that a conserved-residue conversion induces the formation of bisexual flowers in cucumber plants. Genetics, 182:1381-1385.

Lira R, Montes S (1994) Cucurbits (*Cucurbita* spp.). In: Hernandez JE, Leon J (eds) Neglected crops, 1492 from a different perspective. F.A.O., Rome, pp 63–77

Liu K, Muse SV (2005) Powermarker: integrated analysis enviroment for genetic marker data. Bioinformatics 21:2128-2129.

Mantel N (1967) The detection of disease clustering and a generalized regression approach. Cancer Research 27: 209-220.

Metzker ML (2010). Applications of Next-Generation Sequencing. Sequencing technologies the next generation. Nature Reviews Genetics 11:31-46.

Nei M Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76:5269-5273.

Nesom GL (2011) Toward consistency of taxonomic rank in wild/domesticated Cucurbitaceae. Phytoneuron 13:1-33.

Paris HS (1986) A proposed subspecific classification for *Cucurbita pepo*. Phytologia 61:133-138.

Paris HS (2000) History of the cultivar-groups of *Cucurbita pepo*. In: Janick J (ed) Hort Revs, Vol 25 pp., John Wiley & Sons, Inc, pp 71–170.

Paris HS (2008) Summer squash. In: Prohens J, Nuez F (eds) Handbook of Plant Breeding, Vegetables I. Springer, New York, pp 351–379.

Paris HS, Brown RN (2005) The genes of pumpkin and squash. HortScience 40:1620–1630

Paris HS, Yonash N, Portnoy V, Mozes-Daube N, Tzuri G, Katzir N (2003) Assessment of genetic relationships in 4 *Cucurbita pepo* (Cucurbitaceae) using DNA markers. Theor Appl Genet 106:971–978.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945-959.

Rohlf JF (1998) NTSYS: Numerical taxonomy and multivariate analysis system, version 2.02. Exeter Software, Setauket, NY.

Sanjur OI, Piperno DR, Andres TC, Wessel-Beaver L (2002) Phylogenetic relationships among domesticated and wild species of *Cucurbita* (Cucurbitaceae) inferred from a mitochondrial gene: Implications for crop plant evolution and areas of origin. Proc Natl Acad Sci USA 99:535-540.

Smith BD (1997) The initial domestication of Cucurbita pepo in the Americas 10,000 years ago. Science 276:932–934.

Table 1. Summer squash (*C. pepo* ssp. *pepo*) Spanish landraces, other landraces, commercial cultivars and hybrids used in this study. *C. pepo* ssp. *texana* included as controls

Code[a]	**Name**	**Morphotype**	**source**	**Resistances**[b]
		C.pepo* ssp. *pepo		
		Commercial hybrids		
ZU-GIO	Giove F1	Zucchini	Petoseed	WMV, ZYMV
ZU-ZU1805	ZU 1805 F1	Zucchini	Petoseed	PM, CMV, WMV, ZYMV
ZU-MIK	Mikonos F1	Zucchini	Syngenta	PM, CMV, WMV, ZYMV
ZU-QUI	Quine F1	Zucchini	Syngenta	PM, CMV, WMV, ZYMV
VM-TON	Tonya F1	V.Marrow	Syngenta	ZYMV
CO-GS2386	GS2386 F1	Cocozelle	Syngenta	PM, ZYMV
Code	**Name**	**Morphotype**		**source**
		Commercial varieties		
ZU-NVM	Nano Verde di Milano	Zucchini		La Semiorto Sementi
ZU-TRF	True French	Zucchini		Thompson & Morgan
ZU-BB	Black Beauty	Zucchini		Semillas Battle
CO-ROM	Romanesco	Cocozelle		Semitalia
CO-LUF	Lungo Fiorentino	Cocozelle		La Semiorto Sementi
CO-ODF	Ortolana di Faenza	Cocozelle		La Semiorto Sementi
CO-LBS	Lungo Bianco	Cocozelle		La Semiorto Sementi
CO-BDT	Bianca di Trieste	Cocozelle		La Semiorto Sementi
CO-SPQ	San Pasquale	Cocozelle		La Semiorto Sementi
CO-VAL	Alberello Sel. Valery	Cocozelle		La Semiorto Sementi
PU-TON	Tondo chiaro di Nizza	Pumpkin		La Semiorto Sementi
PU-TOP	Tondo di Piacenza	Pumpkin		La Semiorto Sementi
Code	**Name**	**Morphotype**		**origin**
		Spanish landraces		
ZU-A13	Calabacín Blanco	Zucchini		Huesca
ZU-E10	Calabacín	Zucchini		Cáceres
ZU-E27	Calabacín	Zucchini		Cáceres
ZU-MU16	Calabacín	Zucchini		Murcia
ZU-MU20	Calabacín verde	Zucchini		Murcia
VM-A12	Calabacín oscuro	V.Marrow		Huesca
VM-AN23	Calabacín de mesa	V.Marrow		Málaga
VM-AN27	Calabaza	V.Marrow		Cádiz
VM-AN113	Calabacín de freir	V.Marrow		Almeria
VM-CL19	Calabacín	V.Marrow		Segovia
VM-CL21	Calabacín	V.Marrow		Valladolid
VM-CM32	Calabacín	V.Marrow		Cuenca
VM-CM47	Calabacín	V.Marrow		Ciudad Real
VM-CM949	Calabacín blanco	V.Marrow		Guadalajara
VM-V10	Calabacín	V.Marrow		Valencia
VM-V21	Calabacín	V.Marrow		Valencia

CO-A2	Calabacín francés	Cocozelle	Teruel
CO-AN75	Rastrero blanco	Cocozelle	Córdoba
CO-C9	Calabazón	Cocozelle	Barcelona
CO-V74	Calabaza de freir	Cocozelle	Valencia
CO-V116	Calabaza parda	Cocozelle	Castellón
CO-V185	Calabacín	Cocozelle	Alicante
CO-VPAS	Calabacín	Cocozelle	Castellón
		Other landraces	
ZU-ECU227	Calabacín Zaguin	Zucchini	Ecuador
VM-AFR12	Calabacín	V.Marrow	Morroco
PU-PI 171678[c]	pumpkin	Pumpkin	Turkey
PU-Styrian	Oil-seed pumpkin	Pumpkin	Austria
		C. pepo* ssp. *texana	
CR-NSL5227[c]		Crookneck	USA
AC-PI615111[c]		Acorn	USA
SC-UPV196		Scallop	Valencia

[a]The accession code indicates the morphotype/group followed by the cultivar name (commercial cultivars) or COMAV´s Genebank code (Spanish landraces).

[b]Resistance to PM= powdery mildew, ZYMV= zucchini yellow mosaic virus, CMV= cucumber mosaic virus, WMV= watermelon mosaic virus, PRSV = papaya ringspot virus.

[c]These accessions were kindly provided by NPGS of the USDA.

Table 2. Morphological characterization of the Summer squash accessions

Code	Growth habit[a]	Branching[a]	Spines[a]	DMF[a]	NMF[a]	DFF[a]	NFF[a]	NºMF[a]	NºFF[a]	NºFF/MF[a]
Commercial hybrids										
ZU-GIO	B	0	0	30.7	2	41.6	7	12	7.3	0.6
ZU-ZU1805	B	0	0	32.7	2.3	45.7	10	17.7	4.7	0.3
ZU-MIK	B	0	0	35.6	1.7	42	8.7	11.3	5.3	0.5
ZU-QUI	B	0	0	30.7	1.7	43.7	7.3	18	6.7	0.4
VM-TON	B	0.5	0.5	34.3	1	52.6	8.7	23	7.7	0.3
CO-GS2386	B	0	0	34.3	1.7	46	11	16.7	7.3	0.4
Commercial varieties										
ZU-NVM	B	0	0	34	3	52.6	12	20.3	5.3	0.3
ZU-TRF	B	0	0	30,6	4,3	48,3	11	18.7	8.7	0.5
ZU-BB	B	0	0	33	3	45,2	9	22.2	8,1	0.4
CO-ROM	I	0	1	32	2	43.3	10.3	14	8.7	0.5
CO-LUF	B	0	1	39,3	2.6	47,3	13	14.6	5.3	0.4
CO-ODF	B	1	0	37.7	3	56.5	11	31	3.5	0.1
CO-LBS	I	0.5	0.5	44	1	53.3	12.7	29	5	0.2
CO-BDT	I	0	0.5	44.3	2.3	49.3	14	65	9.3	0.1
CO-SPQ	B	0.5	0.5	44	3.7	61.3	11.7	37.7	7.7	0.2
CO-VAL	V	1	0	45.3	1.7	53.3	16	48.3	4.3	0.1
PU-TON	I	0.5	1	32	1.3	48.3	12.7	25	6	0.2
PU-TOP	B	0	0.5	33.7	1	55.3	9.7	38	8	0.2
Spanish landraces										
ZU-A13	I	0.5	0	32	2.5	52	16.5	40	4	0.1
ZU-E10	B	1	0.5	49	2	68	12	32	2	0.1
ZU-E27	B	0	0	35.5	1	45.5	8.5	10.5	6	0.6
ZU-MU16	B	0	0	32.5	1.5	41	7	11	7	0.6
ZU-MU20	I	1	0.5	37	2	57	15	55	4	0.1
VM-A12	I	1	0.5	39	3	56.5	15	60	10	0.2
VM-AN23	V	1	1	62.5	11.5	65	24	60	5.5	0.1
VM-AN27	I	1	0.5	47	1	58.5	12	37.5	2.5	0.1
VM-AN113	V	1	1	39.5	1.5	50.5	13	45	5	0.1
VM-CL19	V	1	0.5	33	3	50	21.5	60	3.5	0.1
VM-CM32	V	1	1	39.5	2	54	15.5	45	4.5	0.1
VM-CM47	I	1	1	32.5	2	45	15	25	4	0.2
VM-CM949	V	1	1	47.5	3	61	18.5	60	5.5	0.1
VM-V10	I	1	1	48.5	2	56.5	18.5	60	3.5	0.1
VM-V21	I	1	1	46	2	58	17	55	4	0.1
CO-A2	I	1	1	47	4	62.5	16	72.5	2.5	0.05
CO-AN75	V	1	1	43	1.5	58	19.5	60	7	0.1
CO-C9	I	1	0.5	34	2	54	22.5	60	2.5	0.05
CO-CL21	I	1	0.5	41	6	77	21	60	7	0.1
CO-V74	V	1	1	42	3.5	53	18	60	4.5	0.1
CO-V116	V	1	1	40.5	3.5	57	24	60	5.5	0.1

CO-V185	B	0	1	30	1	58	12	14	6	0.05
CO-VPAS	I	1	1	50.5	3.5	58	13	38.5	6	0.2
				Other						
ZU-ECU227	B	1	0.5	49	2	60	6	36	8	0.2
VM-AFR12	I	1	1	34.5	1	2	11.5	60	4.5	0.1

[a]growth habit: B=bushy, V= viney, I=intermediate; lateral branch development: 0=no-branching, 0.5=moderately branched, 1=highly branched; presence of spines: 0=spineless, 0.5=sparse and small spicules, 1= dense, large and sharp spicules; DMF, DFF: days to male and female flowering from transplant; NMF, NFF: node in which the first male/female flower appear; NºMF and NºFF: number of male/female flowers 7 days after the opening of the first female flower; NºFF/ NºMF: femaleness ratio

Table 3. Characteristics of the Genomic SSRs selected for genotyping the Summer squash collection.

SSR *locus*[a]	**Linkage group[a]**	**Allele number reported[a] /observed**	**Motif[a]**	**N° Rep[a]**	**Expected size[a]**	**Observed size**	**PIC**
CMTp193	1	5/2	ga	18	186	177-200	0.37
CMTp98	2	3/6	aag	9	213	224-255	0.55
CMTp131	3a	3/3	ccg	7	117	109-135	0.38
CMTp187	3b	3/2	cag+caa	6+4	189	200-300	0.37
CMTp63	4	3/4	ttc	10	152	145-175	0.49
CMTp88	5	5/3	tc	12	167	178-197	0.55
CMTp235	5	6/9	gtt	12	148	145-200	0.74
CMTp256	6	2/3	atc	5	154	175-200	0.10
CMTp224	6	5/3	caa	7	151	169-175	0.41
CMTp248	7	4/1	gga	5	154	145-200	0
CMTp142	8	7/6	tc	12	158	145-240	0.64
CMTp257	9	3/5	cgt	11	138	100-200	0.60
CMTp58	9	2/1	ga+t	6+10	102	100	0
CMTp145	10a	4/4	cat	7	100	117-128	0.42
CMTp66	10a	3/4	gaa	9	128	120-175	0.10
CMTp260	11	4/4	cat	7	155	140-174	0.63
CMTp245	11	5/3	gcg	9	134	100-200	0.43
CMTp36	12	3/4	aac	5	151	164-185	0.47
CMTp69	13	3/4	tatt	4	70	77-92	0.19
CMTp176	14	4/4	tc	12	111	100-145	0.24
CMTp33	14	4/3	gaa	7	171	189-210	0.27
CMTp86	15	3/3	cca	9	133	151-192	0.23
CMTp169	15	2/4	gaa	11	158	175-200	0.42
CMTp231	16	9/-	ag	38	170	--	-
CMTp208	17	2/4	gtt	5	117	133-136	0.25
CMTp209	17	3/3	gtt	6	116	100-145	0.10
CMTp183	18	3/5	cat	6	196	216-303	0.56
CMTp188	18	2/4	cat	7	147	145-350	0.69
CMTp132	19	4/5	gat	12	151	145-200	0.46
CMTp47	20	3/3	ag	9	154	165-173	0.44

**locus* name, linkage group, motif, number of repetitions and allele number reported, according to Gong et al., 2008

Table 4. Characteristics of the EST-SSRs selected for genotyping the Summer squash collection.

SSR *locus*[a]	Position within EST[a]	Allele number reported[a]/observed	Motif[a]	N° rep[a]	Expected size[a]	Observed size	PIC	Arabidopsis Ortholog [b]	Melón Ortholog[b]	Gene description[b]	Functional categories[b]
CUTC001906	UTR	2/3	acgg	7	198	200-255	0.45	AT2G36320	-	zinc finger (AN1-like) family protein	response to stress
CUTC002749	ORF	4/4	tgc	10	183	175-200	0.53	AT1G60030	MU21505	nucleobase-ascorbate transporter 7 (NAT7)	transport
CUTC004158	ORF	2/2	aag	7	188	200-204	0.35	AT4G19100	MU19453	unknown	-
CUTC004307	ORF	4/5	aag	8	100	120-145	0.54	AT5G20190	MU27127	tetratricopeptide repeat (TPR)-like superfamily protein	binding function
CUTC004399	ORF	3/3	acg	7	165	175-200	0.19	AT1G12830	MU38445	unknown protein	-
CUTC004782	UTR	3/2	agg	7	164	175-200	0.30	-	-	unknown protein	-
CUTC004991	UTR	3/4	agc	7	173	175-200	0.32	-	-	hydratase,cytoplasmatic, (cucurbita maxima)	metabolism
CUTC005739	ORF	2/2	aag	8	194	200-204	0.32	AT5G11270	MU24155	overexpressor of cationic peroxidase 3. Transcription factor	transcription regulator, defense response
CUTC005800	ORF	5/5	aag	11	189	200-240	0.66	AT2G47460	-	domain protein 12 MYB12	transcription regulator
CUTC006209	ORF	3/3	aac	9	200	200-300	0.42	AT3G04930	MU21252	transcription regulator	transcription regulator
CUTC006703	ORF	2/2	acc	7	123	145,00	0.27	AT2G42260,	MU23778	UV-B-insensitive 4 UVI4	Cell cycle
CUTC006891	ORF	3/3	ag	12	154	145-175	0.26	AT2G44620	-	mitochondrial acyl carrier protein 1	transport
CUTC007942	ORF	3/5	aac	8	152	175-200	0.16	-	MU37594	BEL1-LIKE homeodomain 1, BLH1	transcription regulator

CUTC008357	ORF	3/3	acc	8	136	156,00	0.41	AT1G23860	-	RS containing zinc finger protein 21	DNA splicing
CUTC008409	ORF	3/3	tgc	8	107	120-145	0.43	AT3G12920	MU22243	protein binding	defense response
CUTC008659	ORF	1/1	aag	7	221	241,00	0.00	AT1G64770,	MU34026	NDH-dependent cyclic electron flow 1, NDF2	catabolism
CUTC046645	ORF	4/4	atc	9	194	200-240	0.28	-	MU26160	Acyl-CoA binding protein	binding function
CUTC009316	ORF	3/2	gca	7	300	320,00	0.15	-	-	RNA binding	binding function
CUTC009607	UTR	4/4	ag	11	146	145-200	0.52	AT3G13510	MU35247	unknown protein	-
CUTC009760	ORF	3/2	aac	7	159	179,00	0.26	-	-	EIN3- like protein (Cucumis melo)	transcription regulator
CUTC011336*	ORF	5/5	agg	9	149	145-175	0.53	AT3G15070	-	zinc finger (C3HC4 Type RING FINGER) family protein	binding function
CUTC012342	ORF	3/2	aag	7	182	200-204	0.18	AT4G34630	-	unknown protein	-
CUTC017708	ORF	4/4	agc	8	223	250-300	0.47	-	-	trasporter (Arabidopsis thaliana) NAT6	transport
CUTC018879	ORF	2/2	aag	7	164	175-200	0.37	-	-	unknown protein	-
CUTC020992	ORF	4/3	aag	10	156	175,00	0.42	-	MU23978	Similar to factor bHLH147	transcription regulator
CUTC022867	ORF	2/5	agc	8	140	145-175	0.48	-	-	Similar to AT1G73230.1, nascent polypeptide-associated complex (NAC) domain-containing protein	response to salt stress
CUTC023363	ORF	4/2	agc	7	244	264,00	0.02	-	-	similar to predicted protein (Populus trichocarpa)	-

[a]*locus* name, location in the EST (ORF: open reading frame; UTR, untranscribed region), motif, , number of repetitions, expected size and allele number reported, according to Blanca et al. (2011).

[b]Annotation results reported in Blanca et al. (2011) are also included, *Arabidopsis* and melon orthologs were identified by searching in *arabidopsis* (http://www.arabidopsis.org) and ICUGI, International Cucurbit Genomics Initiative melon, (http://www.icugi.org database).

Table 5. Average Polymorphism Information Content (PIC) among cultivar groups relative to the SSR origin (EST-SSR or gSSR) and taking both types together (All SSRs)

Variety origin	**EST-SSRs**	**g-SSRs**	**All SSRs**
		Cocozelle	
Commercial	0,32	0,37	0,36
Spanish landraces	0,29	0,33	0,31
		Vegetable marrow	
Commercial	0.19	0.32	0.26
Spanish landraces	0.24	0.33	0.31
		Zucchini	
Commercial	0.17	0.27	0.23
Spanish landraces	0.24	0.31	0.31
		All morphotypes	
Commercial varieties	0.32	0.34	0.34
Commercial hybrids	0.22	0.32	0.28
Spanish landraces	0.29	0.34	0.33

Figure captions

Fig.1 Phenotipic variation of mature fruits of the Summer squash Spanish landraces used in this study. Accessions are located in the map according with their origin.

Fig.2 NJ tree showing relationships among the 48 accessions of Summer Squash using genomic SSRs and EST-SSRs markers based on DICE distance. Bootstrap values over 500 are indicated and are based on 1,000 re-samplings of the data set. Spanish landraces are underlined. Cluster analysis clearly separate the control accessions of *C.pepo* ssp. *texana* (yellow). The population structure estimated by STRUCTURE analysis is depicted in the right lower quadrant, revealing two major populations (green or red), only accession CO-BDT (blue) could be separated clearly for the two previous populations. The populations inferred by STRUCTURE are connected with the corresponding accessions in the NJ tree by colored arrows. Branches of the NJ tree are colored according the population that was assigned from STRUCTURE analysis.

Fig.3 Diagram showing the relationships among the 48 accessions of Summer squash based on principal coordinates analysis using both genomic and EST-SSRs. Two-dimensional scatter plot using the first and second principal coordinates is shown. Spanish landraces are underlined. Cocozelle, Vegetable marrow, Zucchini and Pumpkin Groups are indicated in red, yellow, green and blue, respectively.