Genetic structure of the INRA Capsicum spp. collection using SSR loci: refining the wild origin of cultivated C. annuum and impact of human selection on the structuration of genetic diversity in cultivar types

Anne-Marie Sage-Palloix, Maryse Nicolaï, Melissa Cantet, Véronique Lefebvre, Alain Palloix

Breakthroughs in the Genetics and Breeding of Capsicum and Eggplant

Torino, 2-4 September 2013

XV Meeting on Genetics and Breeding of Capsicum and Eggplant

EUCARPIA

EDITORS

Sergio Lanteri

Giuseppe Leonardo Rotino

Genetic structure of the INRA *Capsicum* spp. collection using SSR loci: refining the wild origin of cultivated *C. annuum* and impact of human selection on the structuration of genetic diversity in cultivar types.

Sage-Palloix A.M., Nicolaï M., Cantet M., Lefebvre V., Palloix A.

INRA PACA, UR1052, Unité de Génétique et Amélioration des Fruits et Légumes, CS 60094, F-84143 Montfavet Cedex, France.

alain.palloix@avignon.inra.fr

Abstract

Germplasm collections of cultivated plants constitute the source for further genetic progress and gained interest with approaches for tracking allelic variants associated to phenotypic variations within core collections. In order to explore the structure of genetic variation in pepper (*Capsicum* spp.) and to select core-collections maximizing the genetic and the phenotypic diversity, a pepper collection including 1352 non redundant accessions from 11 *Capsicum* species from 89 different countries was genotyped using 28 microsatellite (SSR) markers spanning the whole genome. Model-based analysis structured the collection into 6 clusters, with 3 clusters separating the main species complexes, including cultivated species and wild relatives, according to botanic classification (*C. frutescens/C. chinense, C. baccatum, C. pubescens*), and 3 additional clusters for *C. annuum*. The relationships between the cultivated *C. annuum* species and the wild relative (*C. annnuum* var *glabriusculum*) was refined. The 3 *C. annuum* clusters were significantly distinct for plant and fruit descriptors corresponding to cultivar types, showing that the genetic structuration of this cultivated species was strongly impacted by the long term human selection of landraces in primary as well as secondary diversification centres. We settled nested core-collections of 8, 16, 32, 64 and 128 *C. annuum* accessions capturing from 37% to 90% of the genetic diversity for further sequencing efforts and establishment of high through put genotyping assays. By compiling phenotypic and genotypic data, a larger core-collection of 332 accessions was established, capturing 97% of the *C. annuum* genetic and phenotypic diversity for further genetic association studies.

Keywords: *Capsicum*, germplasm collection, human selection, cultivar types, microsatellite, core collection

Introduction

In plant breeding, genetic diversity is the essential source of genetic progress which motivated the collection, maintenance and characterization of genetic resources for most cultivated plant species since the early 20th century. For most cultivated plants, the domestication process caused a loss of genetic variability compared to wild plant populations (Hammer et al. 2003;, Tang et al. 2010). Contrarily, thousands of years of human selection in multiple environments and cultural contexts, provided new mutants and allele combinations of agricultural interest which had poor probabilities to be retained under natural selection pressure. The thousands of local cultivars issued from field selection is a wide source of diversity for alleles of agricultural interest and local adaptations. Their contribution to further genetic progress and to the restoration of biodiversity in agrosystems is at least as promising as wild accessions, related species or exogenous gene sources.

Beyond the collection of these resources, their exploitation depends on our ability to characterize it. Association mapping or linkage disequilibrium (LD) mapping recently developed to track the allelic variants associated to phenotypic variations directly within core collections of plant genotypes, i.e. subsamples of genotypes which represent the genetic diversity of the crop with a minimal redundancy (Marita et al. 2000; Gupta et al. 2005; Zhu et al. 2008). The approach provides

access to multiple alleles and thus increases the efficiency of genetic resources exploitation. However, testing for statistical associations between genotypes and phenotypes in a population is directly affected by the presence of groups of related accessions with different allele frequencies (population structure) which may lead to false associations (Freedman et al. 2004). The core-collection also has to maximize the genetic diversity found in the whole collection and to span the full range of phenotypic variation (Ranc et al. 2010). Thus, analyses of the structure of the genetic diversity within the whole collection of accessions together with an evaluation of the range of phenotypic variation are prerequisite to the selection of core-collections for SNP mining and for association or LD mapping.

Since its domestication in pre-columbian times, peppers were subjected to successive migrations events through Atlantic and Pacific oceans to Europe, Africa and Asia. Trade routes between Europe, Middle-East and Asia also promoted reciprocal exchanges, so that complex introduction processes spread peppers throughout most tropical, mediterranean and temperate regions of the world (Somos 1984; Andrews 1984). In these secondary diversification centers, thousands of landraces have been selected for 4 to 5 centuries by growers to fit new environments and local consumption habits and trade, resulting in the phenotypic diversity of pepper cultivars (Nuez et al. 1996; Bosland and Votava 2000). The taxonomic structuration of the *Capsicum* genus was established from a multidisciplinary approach (Pickersgill et al. 1979; Pickersgil 1991) giving evidence for 5 distinct cultivated species originating from distinct domestication events, and grouped into 3 genetic pools. *C. annuum, C. frutescens* and *C. chinense* form the first genetic pool (the white flowered species) which was related to the wild progenitor *C. annuum* var *glabriusculum*. *C. baccatum* and the wild species *C. microcarpum* form the second genetic pool and *C. pubescens* together with the wild species *C. eximium* and *C. cardenasii* form the 3rd genetic pool. Since the nineties, the species nomenclature was consolidated (Barral and Bosland 2002) and analyses using isozymes, nuclear and chloroplastic DNA markers confirmed this structure and increased our knowledge of the relationships between wild and domesticated species (Walsh and Hoot 2001; Toquica et al. 2003; Ibiza et al. 2012). Many evaluations of genetic diversity were also published, showing that DNA polymorphism rate is rather constant within cultivated species whatever the markers used and generally higher than the polymorphism observed in other autogamous *Solanaceae* like tomato. Analyses of the structuration of genetic diversity reported the relationships between phylogenetic clusters and geographic distribution when species are considered in their primary diversification centers (Hernandez-Verdugo et al. 2001; Votava et al. 2002; Aguilar-Melèndez et al. 2009; Albrecht et al. 2012; Gonzalez-Jaral et al. 2012; Moses and Umaharan 2012; Pacheco-Olvera et al. 2012) and the narrow genetic basis of sweet and large fruited *C. annuum* cultivars compare to exotic landraces (Lefebvre et al. 2001; Tam et al. 2009).

These studies were always performed among restricted sets of accessions ($10<n<200$). Considering a larger panel of *Capsicum* genotypes, should provide a more complete view of the differentiation between pepper cultivars and landraces worldwide and enable us to establish core-collections for further studies of the impact of selection on genetic diversity (Nicolaï et al. submitted). With this aim, we genotyped the INRA *Capsicum* collection which includes 1352 non redundant accessions from 89 different countries, with a large majority of *C. annuum* landraces, but also representatives of 10 additional cultivated or wild species (Sage-Palloix et al. 2007), using 28 SSR loci. Model-based analysis structured this collection into 6 clusters, including 3 distinct clusters for *C. annuum*, which were related to large cultivar types differing in plant and fruit traits as a result of selection. These data were used to establish core collections with different sizes for further SNP mining or genetic association studies.

Material and Methods

Pepper germplasm collection and phenotypic trait measurements.

The pepper (*Capsicum* spp.) germplasm collection maintained at INRA *Unité de Génétique et Amélioration des Fruits et Légumes* includes 1352 non redundant accessions from 11 *Capsicum* species which were collected since 1959 from 89 distinct countries (Sage-Palloix et al. 2007). *Capsicum* accessions are mostly landraces from the cultivated species: *C. annuum* (1063 accessions, including 27 wild *C. annuum glabriusculum*), 92 *C. chinense*, 51 *C. frutescens*, 107 *C. baccatum*, 18 *C. pubescens* and representatives of 6 wild species: 13 *C. chacoense*, 3 *C. cardenasii*, 2 *C. eximium*, 1 *C. galapagoense*, 1 *C. microcarpum* and 1 *C. praetermissum*. These accessions are maintained and multiplied by controlled selfing. Accessions and passport data are registered in the European Solanaceae Network: http://www.ecpgr.cgiar.org/germplasm_databases/list_of_germplasm databases/crop_databases/crop_database_windows/pepper.html). The collection was phenotyped for 21 plant and fruit descriptors and resistance to several pathogens as in Sage-Palloix et al. (2007).

DNA extraction and microsatellite genotyping.

DNA was extracted from pools of leaves of 6 young plantlets per accession as described by Fulton et al. (1995). A set of 28 published microsatellite markers was chosen on the basis of their distribution on the genetic map, spanning 11 of the 12 pepper chromosomes (available on demand). PCR amplifications were performed in a 10 µL reaction volume containing 25 ng of genomic DNA as template. Forward primers were 5'-end labelled with FAM, VIC, or NED for analysis on an Applied Biosystems 3730xI DNA Analyzer. GeneMapper 3.7 software (Applied Biosystems) was used to evaluate the size of the alleles.

Genetic diversity and structure analysis.

The number of allele, the number of genotypes, the Nei's unbiased gene diversity index (He), the observed heterozygosity (Ho), were calculated using the PowerMarker version 3.25 software (Liu and Muse 2005). To infer the population structure of the pepper collection, we used the model-based clustering algorithm implemented in the computer program Structure version 2.3.3 (Pritchard et al. 2000). From multilocus genotypes this algorithm identifies determined number (K) of clusters that have distinct allele frequencies and assigns portions of individual genomes to these clusters. We used the admixture model assuming correlation among allele frequencies. Ten runs were taken into account for each tested value of K, ranging from 1 to 10. In each run, we used a burn-period of 500,000 Markov Chain Monte Carlo iterations and then 250,000 iterations for estimating the parameters. The optimal K value (Kopt) was inferred according to Evanno et al. (2005). Individuals were assigned into a cluster when their proportion of membership into this cluster was higher than 50%. Genetic distance matrices between pairs of accessions were estimated from an index of dissimilarity based on the simple matching method for SSR alleles, and the standardized Euclidean distances for quantitative phenotypic traits. The graphical representation of the neighbour joining trees and principal coordinate analyses were performed with the DARwin 5.0.158 software (Perrier and Jacquemoud-Collet 2006).

Core collection sampling.

For sampling core collections, we used the Maximization (M) algorithm implemented in MSTRAT software version 4.1 (Gouesnard et al. 2001) which permits to maximize the number of alleles captured in the sample (allelic richness). Core collection's minimal size and accessions sampling were performed in 20 replicates with 30 iterations per replicate. The core collections were built using all SSR data alone (nested core collection) or together with phenotypic alleles for 3 plant phenotypic traits (flowering earliness, primary axis length and number of leaves), and 3 fruit traits (fruit length, fruit diameter and pericarp thickness). Phenotypic alleles were inferred from quantitative phenotypic data splitted into 10 classes of equal amplitudes.

Results and Discussion

Microsatellite diversity across the Capsicum species.

The diversity pattern of the 28 SSR loci across the 1352 *Capsicum* spp. accessions revealed 3 to 47 alleles per locus with an average of 18.2 alleles. The observed heterozygosity was very low (<0.085), as expected from accessions maintained and multiplied through selfing. Null alleles (no amplicons) were found for 2 to 3 SSRs in *C. pubescens, C. cardenasii, C. eximium, C. baccatum* and *C. chacoense*. Despite the 1352 *Capsicum* accessions were previously screened to remove redundant accessions according to passport and phenotypic descriptors, a few accessions displayed the same multilocus genotypes and were removed from further analyses, leading to a final panel of 1210 accessions. The allelic diversity (He index) was maximum in the *C. annuum* var *glabriusculum* sub-species (0.78) and lower but rather similar in the other species (0.47 to 0.59 for He).

Genetic structure of the collection.

The optimal number of genetic clusters in the complete collection (1210 accessions) was determined at 6 (Kopt = 6, Figure 1). The clusters 1, 2 and 3 included all the cultivated *C. annuum* accessions and displayed some admixture. The clusters 4, 5 and 6 displayed a clear cut structure with no or very few admixture. The 4th cluster included all the *C. frutescens* and *C. chinense* accessions, together with *C. galapagoense*. The 5th cluster included all the *C. baccatum* and *C. microcarpum* accessions. The 6th cluster included all the *C. pubescens, C. eximium, C. cardenasii, C. praetermissum* accessions and also the 10 *C. chacoense* accessions. This model based clustering closely corresponded to the known taxonomic groups of the *Capsicum* genus, except for the *C. annuum* var *glabriusculum* accessions which were distributed in several clusters.

A phylogenetic tree for the whole collection, based on Nei's genetic pairwise distances was constructed using UPGMA procedure. This tree generally confirms the previous model based clustering, but brings more precisions in agreement with the taxonomic classification of *Capsicum* species (Figure 2). Indeed, the *C. chacoense* accessions are clearly separated from the *C. pubescens* accessions, similarly, the *C. chinense* accessions are clearly in a distinct branching than the *C. frutescens* accessions. The wild *C. annuum* var *glabriusculum* were distributed in different branches: at the root of and within the *C. annuum* branches (14 accessions originating from Mexico and North-America) or at the root of the *C. frutescens and C. chinense* branches (11 accessions originating from Central America), or close to the *C. chacoense* group (2 accessions from Columbia). Most of these accessions were given by B. Pickersgill in 1976 (former var. *aviculare*) who suggested this subspecies to relate to the 3 cultivated white flowered species (Pickergill et al. 1979). Further analyses (Pickersgill 1991; Moscone et al. 2007) revealed distinct karyotypes in those accessions, with one or two pairs of (sub)telomeric chromosomes which are specific of *C. frutescens* and *C. chinense* or of the domesticated *C. annuum* respectively. Interestingly, their position in the tree validates and refines their distribution in the white flowered group and suggests distinguishing the wild relatives of *C. annuum* which originated from Mexico and North America from the wild relatives of *C. frutescens* and/or *C. chinense* originating from Central and South Americas.

Finally, the cultivated *C. annuum* accessions are distributed in several branches in the lowest half of the tree with a large group corresponding to the previous cluster 1, but displayed a slightly more complex pattern for the previous clusters 2 and 3 which are subdivided into several subgroups.

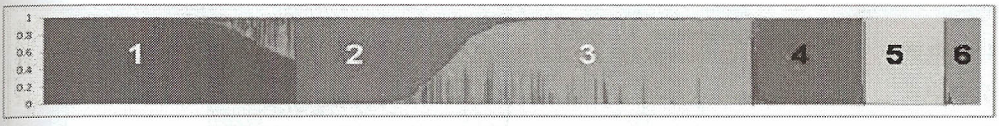

Figure 1: Model-based structuration in the whole *Capsicum* collection (1210 accessions) based on allelic variants at the 28 SSR loci. Six clusters were defined following the method of Evanno et al. (2005). Clusters 1, 2 and 3 included 908 cultivated *C. annuum* and 9 var *glabriusculum*, cluster 4 included 136 *C. frutescens C. chinense* , *C. galapagoense* plus 4 *var glabriusculum* accessions, cluster 5 included 105 *C. baccatum* and *C. microcarpum* accessions, cluster 6 included 34 *C. chacoense, C. pubescens, eximium, cardenasii, praetermissum* and 14 var *glabriusculum* accessions.

Genetic and phenotypic diversity of the 3 clusters of C. annuum.

The model based analysis with the Structure software delivered 3 distinct clusters within the *C. annuum* accessions (Figure 1) with admixted accessions. These 3 *C. annuum* clusters differ in their genetic diversity, with a higher He value for the cluster 1 (He 0.64) than for clusters 2 and 3 (He 0.44 and 0.40 respectively). In a first attempt to explore the diversity in phenotypes and origin of these 3 clusters, their means for plant and fruit parameters were compared and revealed significantly different average values for most plant and fruit descriptors (Figure 3).

The cluster 1 was characterized by late flowering plants (+ 3 days), with a long primary axis (28 cm) developing at least 14 leaves before flowering. Fruits from these accessions were small in length and particularly in diameter (1.9 cm in average) resulting in an elongated shape (4.7 times longer than large), with a pointed blossom end, and a thin pericarp (1.5 mm). That is characteristic for most small and elongated fruited peppers which represent 82% of the accessions from this cluster. This cluster was highly diversified in the geographic origins of the accessions, including traditional Mexican cultivars ('Pasilla', 'Anaheim', 'Serrano' types), from Asia ('Perennial' from India, 'Nanjing early' from China) and from Africa ('H3' from Ethiopia, 'Chatah' from Sudan). These late flowering and small-fruited cultivars mostly originate from subtropical areas but the cluster also includes many cultivars that became traditional in temperate and mediterranean countries like 'Espelette' pepper from France which appeared close to the Mexican 'Pasilla Apaseo'.

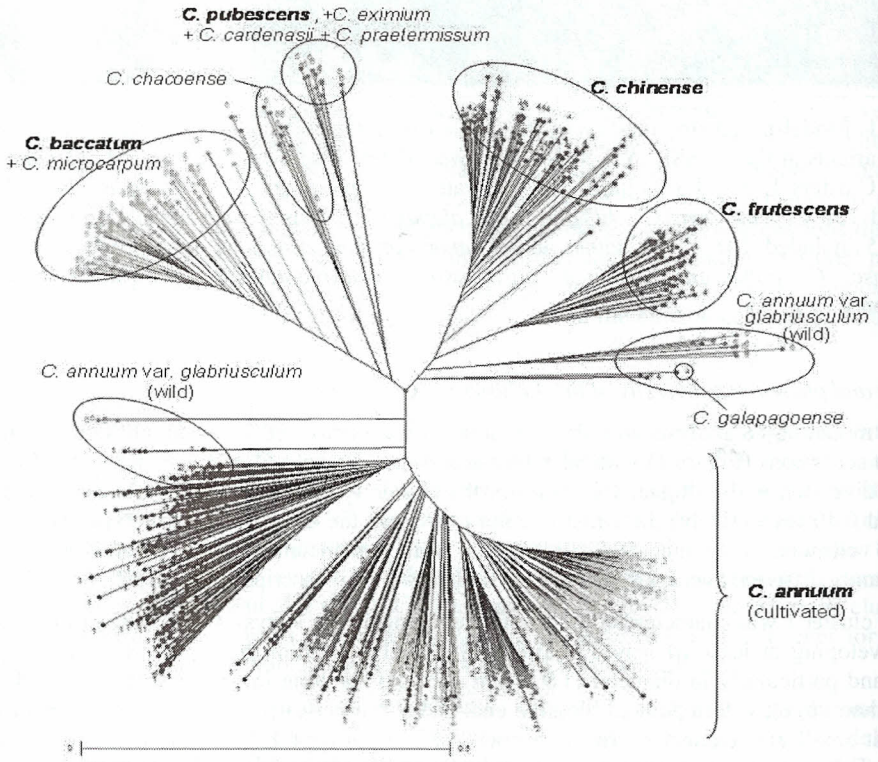

Figure 2: Phylogenetic tree showing the genetic diversity of the pepper germplasm collection (1210 accessions). The tree was produced using the neighbor-joining UPGMA method based on the 28 SSR markers.

The cluster 2 is characterized by early flowering plants (-3.5 days in average) with shorter primary axis (22 cm) bearing a lower number of leaves (9.5). The fruits are longer and larger than the previous cluster, resulting in a 2.8 length/width ratio with an obtuse apical end and a much thicker pericarp (4.2 mm). This corresponds to the triangular and horn shaped peppers but also to some elongated peppers, which represent 40% and 35% of the accessions respectively. Considering the geographic origins, this cluster displayed a clear predominance of central European origin. Indeed, 147 of the 201 accessions from cluster 2 (73%) are local cultivars originating from central Europe (Hungary, former Yugoslavia and Czechoslovakia, Romania, Poland, South Russia, Bulgaria) whereas these countries represent only 18% of the origins of the whole *C. annuum* collection. This cluster can be characterized by the traditional cultivars with elongated fruits like 'Hatvani', conical fruits with light green or ivory immature color like 'Podarok Moldavia', 'Feherozon', 'Cecei', but also a few blocky fruits with ivory color like 'Bela Krupna' or 'Paradicsom'. Another characteristic of this cluster is the presence of the traditional Turkish cultivars with horn or conical shaped fruits (10 landraces from the 'Sivri' and 'Carliston' types).

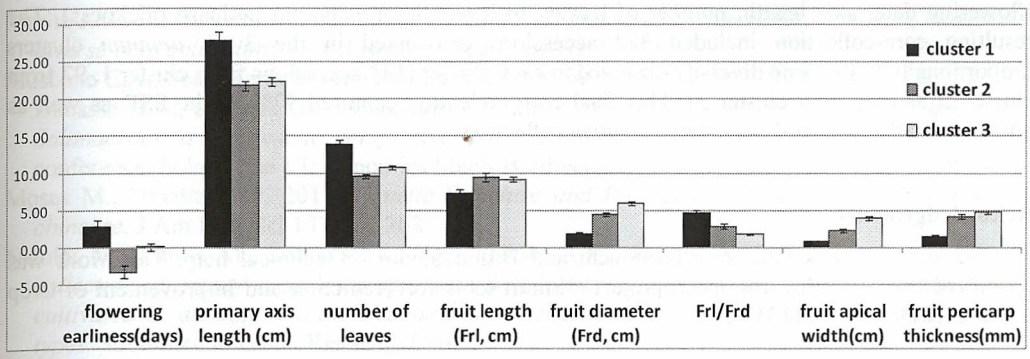

Figure 3 : Average values of the 3 cultivated *C. annuum* clusters defined by the model-based analysis for 8 plant and fruit traits. Vertical bars represent the 95% confidence interval.

The cluster 3 included plants with intermediate earliness but axis growth and development close to the cluster 2. Fruits are close to the cluster 2 in length but significantly larger (6 cm) resulting in an average length/width ratio of 1.8. The mean apical end is much more large and lobate and the pericarp is thick (4.8 mm). This clearly corresponds to the large fruited peppers with blocky or rectangular shape which contribute to 61% of the accessions of this cluster. Geographic origins in this cluster are diversified, including the traditional cultivars with very large (up to 600g) and rectangular fruits from Mediterranean Europe ('Largo de Reus' and 'Largo Valenciano' from Spain, 'Lagnes' from France), the large blocky fruits from Italy ('Quadrato Asti'), smaller blocky fruits from USA ('Yolo Wonder', 'California Wonder'), from Netherlands ('Mavras'), Poland ('Oda'), China ('Zao Feng', 'Ben Xi'). In this cluster were also located several accessions with thick pericarp but triangular, tomato, cherry or heart shaped fruits like 'Fresno' or 'Cherry bomb' from USA, 'Morron Conserva' or 'Ñora' from Spain. These cultivars present an admixted genome between the cluster 3 and 2 or 1.

Except for landraces from cluster 2 which originated from Central Europe and Turkey, the structuration by geographic origins of the clusters 1 and 3 was weakly visible. A more detailed analysis within each cultivar type reveals clusters of accessions with common geographic origins. However, these clusters also include cultivars collected from exotic countries. Genetic differentiation between cultivar types interferes or tends to dominate differentiation between geographic origins. It also attests the numerous and complex migration events of pepper genotypes, their adoption in a new country resulting from human migration and local selection.

Construction of core collections.

Sub-samples of 8, 16, 32, 64, and 128 accessions of *C. annuum* were selected, based on their genotypes at the 28 SSR loci. In this strategy, the accessions from the smaller samples were included in the successive larger samples (nested core-collections) These successive core collections captured 37 %, 55 %, 71 %, 85 %, and 89 % of the alleles from the whole *C. annuum* collection. The M strategy algorithm, using alleles at SSR loci, permits to select the smallest samples while maximizing the genetic diversity which maybe favourable for sequence diversity analyses and SNP mining. However the deficit in accessions from the cluster 3 (large fruited cultivars) and cluster 2 (early flowering plants and conical or long fruited cultivars) in these small core collections affects their representativeness for horticultural traits, which may not be favourable to association analyses with these traits. Thus, a larger core collection of *Capsicum annuum* was built with the objective to optimize the contribution of the *C. annuum* clusters, and to maximize both the genetic and phenotypic diversity. This was achieved, using the same M strategy, but based on their alleles at the 28 SSR markers and on their 'phenotypic alleles' for 6 primary traits

(flowering date, axis length, number of leaves, fruit length, fruit width, pericarp thickness). The resulting core-collection included 332 accessions distributed in the 3 *C. annuum* clusters proportionally to the gene diversity observed in each cluster (142 accessions from cluster 1, 97 from cluster 2, and 93 from cluster 3). This final core collection captured 97% of the SSR as well as phenotypic alleles from the whole *C. annuum* collection.

Ackowledgements:

The authors thank Ghislaine Nemouchi and Bruno Savio for technical help. This work was supported by the SPICY European project ("Smart tools for Prediction and Improvement of Crop Yield", KBBE-2008-211347).

References

Aguilar-Melèndez A., Morrell P.L., Roose M.L., Kim S.C. 2009. *Genetic diversity and structure in semiwild and domesticated chili peppers (Capsicum annuum; Solanaceae) from Mexico.* Am J Bot 96:1190–1202

Albrecht E., Zhang D., Saftner R.A., Stommel J.R. 2012. *Genetic diversity and population structure of Capsicum baccatum genetic resources.* Genet Res Crop Evol 59:517–538

Andrews J. 1984. Pepper, The domesticated *Capsicum.* , University of Texas Press, Austin, Texas.

Baral J.B., Bosland P. 2002. *An updated synthesis of the Capsicum genus.* Capsicum and Eggplant Newsl 21:11-21

Bosland P.W., Votava E.J. 2000. *Peppers: vegetable and spice Capsicum.* CABI Publishing, Oxon, UK and New York

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

Freedman M.L., Reich D., Penney K.L. et al. 2004. *Assessing the impact of population stratification on genetic association studies.* Nat Genet 36:388-393

Fulton T.M., Chunwongse J., Tanksley S.D. 1995. *Microprep Protocol for Extraction of DNA from Tomato and other Herbaceous Plants.* Plant Mol Biol Rep 13: 207-209

Gouesnard B., Bataillon T.M., Decoux G., Rozale C., Schoen D.J., David J.L. 2001. *MSTRAT: An Algorithm for Building GermPlasm Core Collections by Maximizing Allelic or Phenotypic Richness.* J Hered 92:93-94

Gonzalez-Jaral P., Moreno-Letelier A., Fraile A., Pinero D., Garcia-Arenal F. 2012. *Impact of Human Management on the Genetic Variation of Wild Pepper, Capsicum annuum var. glabriusculum.* PLoS ONE 6(12): e28715. doi:10.1371/journal.pone.0028715

Gupta P.K., Rustgi S., Kulwal P.L. 2005. *Linkage disequilibrium and association studies in higher plants: present status and future prospects.* Plant Mol Biol 57: 461–485

Hammer K., Arrowsmith N., Gladis T. 2003. *Agrobiodiversity with emphasis on plant genetic resources.* Naturwissenschaften, 90(6): 241–250. doi:10.1007/s00114-003-0433-4

Hernandez-Verdugo S., Luna-Reyes R., Oyama K. 2001. *Genetic structure and differentiation of wild and domesticated populations of Capsicum annuum (Solanaceae) from Mexico.* Plant Syst Evol 226:129–142

Ibiza V.P., Blanca J., Canizares J., Nuez F. 2012. *Taxonomy and genetic diversity of domesticated Capsicum species in the Andean region.* Genet Resour Crop Evol 59:1077–1088

Lefebvre V., Goffinet B., Chauvet J.C., Caromel B., Signoret P., Brand R., Palloix A. 2001. *Evaluation of genetic distances between pepper inbred lines for cultivar protection purposes: comparison of AFLP, RAPD and phenotypic data.* Theor Appl Genet 102:741–750

Liu K., Muse S.V. 2005. *PowerMarker: an integrated analysis environment for genetic marker analysis.* Bioinforma 21:2128-2129

Marita J.M., Rodriguez J.M., Nienhuis J. 2000. *Development of an algorithm identifying maximally diverse core collections.* Genet Resour Crop Evol 47:515-526

Moscone E.A., Scaldaferro M.A., Grabiele M., Cecchini N.M., Garcia Y.S., Jarret R., Davina J.R., Ducasse D.A., Barboza G.E., Ehrendorfer F. 2007. *The evolution of chili peppers (Capsicum—Solanaceae): a cytogenetic perspective.* In: Proceedings of the VIth international Solanaceae conference. Solanaceae VI: Genomics Meets Biodiversity, Acta Horticulturae 745:137-169

Moses M., Umaharan P. 2012. *Genetic Structure and Phylogenetic Relationships of Capsicum chinense.* J Am Hort Sci 137:250-262

Nicolaï M., Cantet M., Lefebvre V., Sage-Palloix A.M., Palloix A. 2013. *Genotyping a large collection of pepper (Capsicum spp) with SSR loci brings new evidence for the wild origin of cultivated C. annuum and the structuration of genetic diversity by human selection of cultivar types.* Submitted to Genet Res Crop Evol.

Nuez F., Gil R., Costa J. 1996. *El cultivo de pimientos, chiles y ajies.* Mundi_Prensa, Madrid

Pacheco-Olvera A., Hernandez-Verdugo S., Rocha-Ramirez V., Gonzalez-Rodriguez A., Oyama K. 2012. *Genetic Diversity and Structure of Pepper (Capsicum Annuum L.) from Northwestern Mexico Analyzed by Microsatellite Markers.* Crop Sci 52:231-241

Perrier X., Jacquemoud-Collet J.P. 2006. *DARwin software.* http://darwin.cirad.fr/

Pickersgil B. 1991. *Cytogenetics and evolution of Capsicum L.* In: Tsuchiya T, Gupta PK (eds) Chromosome engineering in plants:genetics, breeding, evolution, Elsevier, Amsterdam, 139-160

Pickersgill B., Heiser C.B., McNeill J. 1979. *Numerical taxonomic studies of variation and domestication in some species of Capsicum.* In: JG Hawkes (ed) The biology and taxonomy of the *Solanaceae*, Academic Press, New York, pp 679-699

Pritchard J.K., Stephens M., Rosenberg N.A., Donnelly P. 2000. *Association Mapping in Structured Populations.* Am J Hum Genet 67:170-181

Ranc N., Muños S., Xu J., Le Paslier M.C., Chauveau A., Bounon R., Rolland S., Bouchet J.P., Brunel D., Causse M. 2010. *Genome-wide association mapping in tomato(Solanum lycopersicum)is possible using genome admixture of Solanum lycopersicum var. cerasiforme.* Genes Genomes Genet 2: 853-864

Sage-Palloix A.M., Jourdan F., Phaly T., Nemouchi G., Lefebvre V., Palloix A. 2007. *Structuring genetic diversity in pepper genetic resources: distribution of horticultural and resistance traits in the INRA pepper germplasm.* In Niemirowicz-Szczytt K (Ed), Progress in research on *Capsicum* & Eggplant. Warsaw University of Life Sciences Press, Warsaw, Poland, pp.33:42

Somos A. 1984. *The paprika.* Akademiai Kiado, Budapest, ISBN 963 05 3299 9, 302 p

Tam S.M., Lefebvre V., Palloix A., Sage-Palloix A.M., Grandbastien M.A. 2009. *LTR-retrotransposons Tnt1 and T135 markers reveal genetic diversity and evolutionary relationships of domesticated peppers.* Theor Appl Genet 119:973-989

Tang H., Sezen U., Paterson A.H. 2010. *Domestication and plant genomes.* Curr Opin Plant Biol 13:160-166

Toquica S.P., Rodriguez F., Martinez E., Duque M.C., Tohme J. 2003. *Molecular Characterization by AFLPs of Capsicum Germplasm from the Amazon Department in Colombia.* Genet Res Crop Evol 50: 639–647

Votava E.J., Nabham G.P., Bosland P.W. 2002. *Genetic diversity and similarity revealed via molecular analysis among and within an in situ population and ex situ accessions of chiltepin (Capsicum annuum var. glabriusculum).* Conserv Genet 3:123–129

Walsh B.M., Hoot S.B. 2001. *The phylogenetic relationships of Capsicum (Solanaceae) using DNA sequences from two non-coding regions: the chloroplast atpB-rbcL spacer region and nuclear waxy introns.* Int J Plant Sci 162:1409-1418

Zhu C., Gore M., Buckler E.S., Yu J. 2008. *Status and prospects of association mapping in plants.* The Plant Genome 1: 5–20

Zwendie Y., Tong N., Bosland P. 2004. *Establishing a core collection of Capsicum using a cluster analysis with enlightened selection of accessions.* Genet Res Crop Evol 51: 147-151