Khoufi Sahari*, Pouilly Nicolas, Muños Stéphane, Bérard Aurélie, Ben Jeddi Fayçal, Vincourt Patrick and Brunel Dominique

Genetic Diversity and Core Collection Constitution for Subsequent Creation of New Sunflower Varieties in Tunisia

DOI 10.1515/helia-2016-0002
Received January 6, 2016; accepted March 21, 2016; previously published online April 12, 2016

Abstract: Sunflower (*Helianthus annuus* L.) is used as "seed snacks", so called "pipas", in the Mediterranean countries which may offer an interesting opportunity for agricultural diversification of this crop. The morpho-phenological variability already demonstrated in the Tunisian local populations can be a basis for the creation of new varieties well adapted to climate and soil conditions in that region. The molecular characterization of 59 accessions and reference lines generated 194 alleles from 30 SSR loci (3–10 alleles per locus) and 54 haplotypes, built from 117 SNP detected by NGS sequencing of 7 genes (4–16 haplotypes per gene). These data highlight some uniqueness of the Tunisian material compared to 7 control lines but a low genetic dispersion between accessions. However, a core collection of 8 populations, capturing 88% of the Tunisian genetic diversity, could be proposed for a future sunflower breeding program.

Keywords: core collection, genetic diversity, SNP, SSR, sunflower

***Corresponding author: Khoufi Sahari,** Laboratoire des Sciences Horticoles, Institut National Agronomique de Tunisie, 43, avenue Charles Nicolle 1082 Tunis-Mahrajène Tunisie, E-mail: sk111183@yahoo.fr
Pouilly Nicolas, Muños Stéphane, Laboratoire des Interactions Plantes Micro-organismes, INRA_CNRS, Chemin de Borde Rouge, B.P. 52627, 31326 Castanet Tolosan, France
Bérard Aurélie, US Etude du Polymorphisme des Génomes Végétaux, INRA, CEA/IG/Centre National de Génotypage, Université Paris Saclay, 2 rue Gaston Crémieux, 91057 Evry Cedex, France
Ben Jeddi Fayçal, Laboratoire des Sciences Horticoles, Institut National Agronomique de Tunisie, 43, avenue Charles Nicolle 1082 Tunis-Mahrajène Tunisie
Vincourt Patrick, Laboratoire des Interactions Plantes Micro-organismes, INRA_CNRS, Chemin de Borde Rouge, B.P. 52627, 31326 Castanet Tolosan, France
Brunel Dominique, US Etude du Polymorphisme des Génomes Végétaux, INRA, CEA/IG/Centre National de Génotypage, Université Paris Saclay, 2 rue Gaston Crémieux, 91057 Evry Cedex, France

Introduction

Sunflower (*Helianthus annuus* L.) is an important crop ranking 11th among world food crop in term of harvested areas (FAO, 2012). In Tunisia, its average cropping area ranges only from 10,000–12,000 ha per year (DGPA, 2013). It is used mainly for its seeds as "snacks". The Mediterranean climate and the nature of the vegetative cycle of sunflowers are among the factors that limit its expansion. Thus studying diversity aiming to select large seeds of sunflowers with high productivity in limited hydric conditions.

Moreover, supporting the sunflower production will be a weapon with two edges in Tunisia, developing mainly the confection/non-oil sunflower local market but also improving the oil content for prospective oil extraction, helping so, consumers to equilibrate its purchase power concerning vegetable oil (Benassi and Labonne, 2004). As the production of sunflower in Tunisia cover few hectares and still traditionally cultivated, it may guide breeders to focus on organic agriculture which is encouraged more and more in the Tunisia policy agriculture.

Morphological and phenological informations for Tunisian populations were already studied in a previous experiment (Khoufi *et al.*, 2013). Populations with interesting agronomic characteristics to start a breeding program were detected for short developmental cycle related to the least growth degree day, large seed size and 100 seed weight. To overcome the environmental effects and to better understand the genetic diversity of the Tunisian germplasm and the relationships between accessions, a study using molecular markers (SSR and SNP) was conducted.

Restriction fragment length polymorphism (RFLP), and random amplification of polymorphic DNA (RAPD) were the first markers used to assess sunflower diversity (Berry *et al.*, 1995, 1997, 2003; Gentzbittel *et al.*, 1995, 1999; Jan *et al.*, 1998; Rieseberg *et al.*, 1998), amplified fragment length polymorphism (AFLP) (Peerbolte and Peleman, 1996; Gedil *et al.*, 2001; Al-Chaarani *et al.*, 2004) and simple sequence repeats (SSR) (Brunel, 1994; Paniego *et al.*, 2002; Tang *et al.*, 2002, 2003; Yue *et al.*, 2002, 2003, 2008; Mokrani *et al.*, 2002; Zhang *et al.*, 2005) were the second type of markers, used to analyze the fine structure, dynamics of populations and genome mapping for sunflower crop (Hu *et al.*, 2010). Single nucleotide polymorphism (SNP) (Pashley *et al.*, 2006; Heesacker *et al.*, 2008; Lai *et al.*, 2005; Bachlava *et al.*, 2012; Hulke *et al.*, 2015; Pegadaraju *et al.*, 2013; Kolkman *et al.*, 2007; Livaja *et al.*, 2016) are largely used. In Tunisia, although the spread use of molecular markers for many crops, sunflower was not concerned.

Combining SSR and SNP data, we therefore aimed to generate a sunflower core collection, which presents the maximum diversity of the Tunisian genetic resources. The purpose of developing genetic core collections was to provide a restricted set of accessions with a limited loss of genetic variability of characters not yet studied as diseases resistances.

Materials and methods

Plant material

A total of 59 populations of sunflowers (*Helianthus annuus* L.) cultivated in the north of Tunisia (Figure 1 and Supplementary data 1) were collected in 2006 and was provided by National Institute of Agronomy of Tunisia (INAT) in form of decorticated five sunflower heads for each population. Seven reference genotypes with different origins, SF012 and SF193 (INRA, France); SF085 and SF332

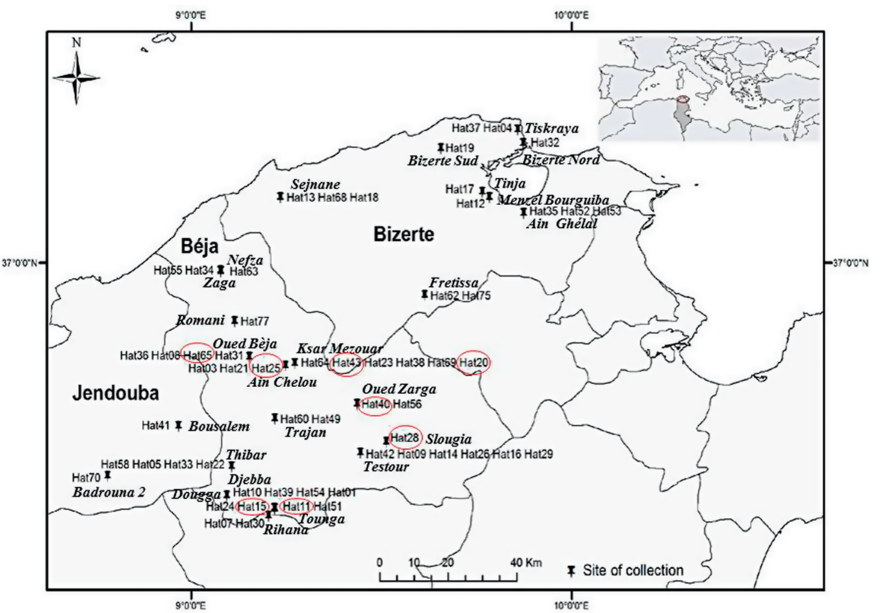

Figure 1: A map of the known distribution of *Helianthus annuus* in Tunisia and sites of sampled populations used in this study (Supplementary data 1 describes the different sampling location). Populations circled in red represent the core collection. (ESRI ® ArcMap 10.0).

(USDA); SF092 (Australia); SF109 (Maroc) and Var Turk (Turkey) were used for molecular diversity comparison.

DNA extraction

Seeds were put in Petri dishes for germination in an incubator at 26 °C for 2 days. The germinated grains were put in the growing chamber under 75 % humidity, 22 °C for 6 h of darkness and 22 °C for 18 h in light. After 15 days, leaf discs of the first pair of leaves were sampled from at least five plant of each population for DNA extraction. Total genomic DNA was extracted using the DNeasy 96 Plant Kit (*QIAGEN, Valencia, CA*), following the manufacturer's instructions and was suspended in $T_{10}E_1$ buffer (pH = 8).

DNA quality was evaluated by NanoDrop spectrophotometry and gel electrophoresis. DNA concentration was quantified using the *Quant-iT™ PicoGreen®* dsDNA Assay (*Invitrogen*) on an ABI7900 instrument (*Applied Biosystems, Foster City, CA, USA*). DNA concentrations were adjusted to 10 or 25 ng/µL for each sample, following the exigencies of the molecular protocols (SSR and SNP) used in this study.

SSR assay

SSR markers used in this study are published by Tang *et al.*, (2002) and were provided by the INRA group. 56 Primers are selected to cover the entire genome of 17 linkage groups (Supplementary data 2_a).

PCR reactions for SSR markers were carried out in 96 plates with a total volume of 10 µl. The reaction mix was: 20 ng DNA, 1X Taq buffer (*Promega*), 1.5 mM MgCl2, 75 µM dNTP (*Invitrogen*), 0.2 µM forward primer coupled to a 19-base M13 tail, 0.2 µM reverse primer, 0.2 µM of the 19-base M13 tail primer with 6-FAM, VIC, NED or PET fluorophore (*Applied Biosystems*) and 0.4 U Taq polymerase.

PCR reactions were performed on Eppendorf Mastercycler, following a touch down program: initial 4 min denaturation step at 94 °C, was followed by 11 cycles. Each cycle consisted of 30 s, denaturation at 94 °C, 30 s at the annealing temperature which decreased by 1 °C per cycle from 60 °C to 50 °C, and 45 s elongation at 72 °C. The enrichment step was followed by 25 cycles, each consisting of 30 s denaturation at 94 °C, and 30 s at 50 °C and 45 s elongation at 72 °C. These cycles were followed by a final 5 min elongation step at 72 °C.

Capillary electrophoresis was performed on an ABI Prism 3730 DNA sequencer (*Applied Biosystems*) and the GeneScan internal size standard labelled with

LIZ-500 (*Applied Biosystems*) was used. Allele size was determined as base pair using the program GENEMAPPER (v 4.0) (*Applied Biosystems*), allelic data were used to calculate the number, range and distribution of amplified alleles to determine variation level in the studied collection of sunflower.

SNP assay

The selection of seven genes [*EFE* (GeneBank: 839345), *EXECUTER1* (GeneBank: 829504), *AVP1* (GeneBank: 838138), *CG068, P5CS2* (GeneBank: 824727), *LPT3a and LPT3b* (GeneBank: 836051)], was based on their role in direct or indirect regulation in metabolism process when the plant was subjected to abiotic stress (http://www.ncbi.nlm.nih.gov/genbank).

By using either the cDNA sequences or the corresponding genomic sequences identified on a local genome assembly of the *XRQ* line (*https://www.heliagene.org/HaT13*l), primers have been designed to produce PCR products that cover the whole gene sequence (Supplementary data _ 2b).

DNA sequencing

DNA amplifications were performed in Access Array™ System (*Fluidigm*) following the user guide instructions Access Array™ System for Illumina Sequencing Systems: http://www.fluidigm.com/home/fluidigm//Support/aa-illumina-ug-g1.pdf.

Sequencing was performed on a MiSeq sequencer (*Illumina*) using a 500 cycles MiSeq Reagent kit for a 2*250 bases paired-ends sequencing, following the user guide: http://supportres.illumina.com/documents/documentation/system_documentation/miseq/preparing-libraries-for-sequencing-on-miseq-15039740-d.pdf.

Bioinformatic analysis

A reference sequence of the seven candidate genes was built from the genome assembly of the *HA412* line (Kane *et al.*, 2011). Scaffolds larger than 1000 bps only were selected for the mapping of the paired-end read sequences (Supplementary data 3).

The MiSeq generated paired-end reads were analysed using CLC Genomics Workbench (v7) (http://www.clcbio.com/products/clc-genomics-workbench).

After a quality control analysis, the reads were trimmed for a Phred value higher than 30 and no ambiguity. The mapping on the reference genes sequences was done with the criteria of 98% of similarity on 98% of the read length. The variants detected by CLC tool were manually checked for each polymorphic site and then haplotypes were built.

Statistical analysis

To decipher the most distinctive SNP and SSR alleles, a Factorial Correspondence Analysis (FCA) was conducted using a matrix table of presence and absence of SNP and SSR alleles for each population. Analysis was performed with XLSTAT software (V1.06). Based on alleles and haplotypes detected in this study, M strategy (Schoen and Brown, 1983) was used for generating core collections by MSTRAT software version v4 (Gouesnard *et al.*, 2001). This strategy examines and singles out those that maximize the number of observed alleles at the "marker loci". These can then be chosen as final candidates for the core. The importance of the expected linkage disequilibrium between observed allelic richness at the marker loci ("active variable") and allelic richness on other loci such as morpho-phenological traits ("target variable") can be analyzed, as well as the comparison between a random sampling and the M method.

Results and discussion

Variability on molecular markers

From the 56 SSRs tested, 30 markers generated interpretable and polymorphic results (Supplementary data 4). A total of 194 alleles were noted; of which 162 alleles marked the Tunisian collection with an average of 5 alleles per marker. The SSR markers exhibiting more alleles were related to ORS078; ORS533; ORS733; SSU192 and ORS466. Almost all primers marked a difference in the total number of alleles between the Tunisian collection and reference varieties (Supplementary data 5_a). The total number of alleles in Tunisian collection ranged from 2 alleles revealed by ORS337, ORS316 and SSL412 to 9 alleles revealed by ORS078.

From 117 SNPs, a total of 58 haplotypes were built from 4 for *EXECUTER1, EFE, AVP1* genes to 16 for *LPT3b* gene (Supplementary data 5_b). Thirty-five

haplotypes are present in the reference collection and 50 in the Tunisian collection. The analysis of the different haplotype frequencies indicates that the Tunisian populations are marked by some original alleles and different percentage for some haplotypes in comparison with the reference genotypes.

In summary, 162 alleles were noted within Tunisian populations and 32 within reference lines with an average of 6.4 alleles per SSR marker, however with SNP analysis we noted an average of 7.14 haplotypes per gene within Tunisian collection and 5 haplotypes per gene within references.

Factorial Correspondence Analysis (FCA) and grouping within reference varieties and Tunisian collection of sunflower

Diagrams generated by Factorial Correspondence Analysis (Figure 2 (a), and 2 (b)) distinguish three groups: the Tunisian collection, the reference genotypes and the Turkish variety. Another fact is the low dispersion of the Tunisian group. Factorial Correspondence Analysis (FCA) generated by SSR data explained 17 % of total variation in the local and reference collection. The first factor explained 11 %, the second factor explained 6 % (Supplementary data 6). Diagram generated by the haplotypes of the four genes explained 16 % of total variation in the local and reference collection. The first factor explained 9 % of total variability related mostly to haplotypes found in *LTP3* and *CG068*, the second factor explained 7 % characterized by haplotypes found in *EFE*, *EXECUTER*, *LTP3* and *CG068* (Supplementary data 6).

Factorial Correspondence Analysis (FCA) within Tunisian sunflower collection.

The diagrams obtained by considering only the Tunisian collection showed one major group that comprised most of the populations surrounded by other scattered ones (Figure 2 (c), and (d)).

Factor analysis generated by SSR explained 10 % of total variation in the local collection. The first factor explained 6 % of total variability. The second factor explained 4 % (Supplementary data 6). No particular SSR discriminates the populations.

FCA based on gene haplotypes explained 18 % of total variation in the local collection. The first factor explained 10 % of total variability related mostly to haplotypes found in *CG068* and *LTP3*, the second factor explained 8 % characterized by haplotypes found in *CG068*, *LTP3* and *EXECUTER* gene

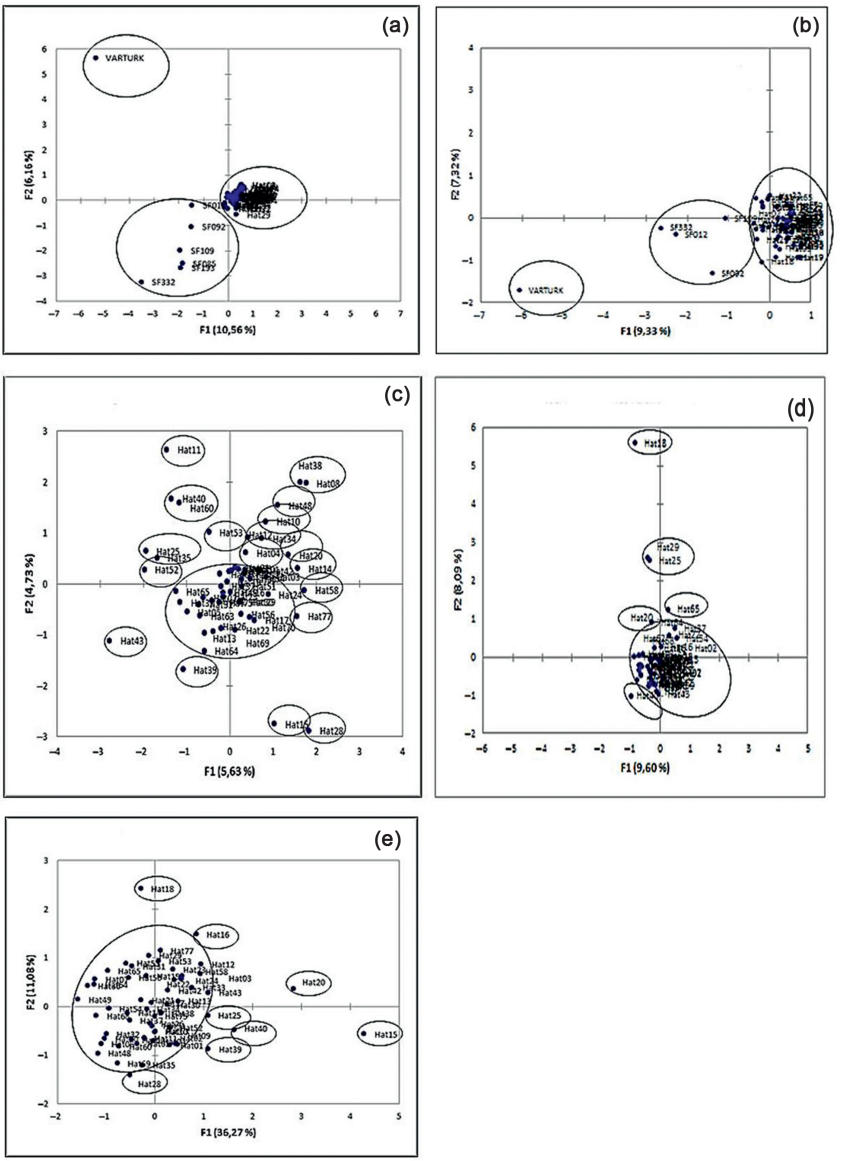

Figure 2: Distribution of reference varieties and Tunisian populations of sunflower according to factors 1 and 2 of FCA based on SSR (a) and SNP (b) markers. Distribution of Tunisian populations of sunflower according to factors 1 and 2 of FCA based on SSR (c) and SNP (d) markers. The main SSR/SNP markers and eigen values of alleles contributing to F1 and F2 are described in supplementary data 6. (e) Distribution of Tunisian populations of sunflower according to factors 1 and 2 of FCA based on morphological and phenological traits. The main morphological and phenological markers contributing to F1 and F2 are described in supplementary data 6.

(Supplementary data 6). *CG068* and *LTP3* are the major and redundant contributors.

Comparison between molecular and morpho-phenological data

Molecular results confirmed those obtained in a first study on 75 populations of sunflower and seven hybrids with the morpho-phenological traits (Khoufi *et al.*, 2013) that Tunisian sunflower populations formed a single group with some populations scattered around the main group (Figure 2 (e) and Supplementary data 6).

Comparison on SSR, SNP and morpho-phenology data (Table 1) exhibited few populations distinguished from other regardless of the analytical criterium. (ie. Hat15, Hat18, Hat20, Hat25, Hat28, Hat39, Hat40).

Table 1: The most polymorphic populations sorted out by each of SSR, SNP and morpho-phenological markers.

SSR	SNP	SSR et SNP	Morpho-phenological traits
Hat04			
Hat08			
Hat10			
Hat11		Hat11	
Hat14			
Hat15		Hat15	Hat15
			Hat16
	Hat18	Hat18	Hat18
Hat20	Hat20		Hat20
Hat25	Hat25	Hat25	Hat25
Hat28		Hat28	Hat28
	Hat29		
Hat34		Hat34	
Hat35			
Hat38			
Hat39		Hat39	Hat39
Hat40	Hat40	Hat40	Hat40
Hat43		Hat43	
Hat48			
Hat52		Hat52	
Hat53			
Hat58			
Hat60			
	Hat65	Hat65	
Hat77		Hat77	

No correlation between geographical and genetic distances

In order to find a geographical structuration of the population, a Mantel test (Miller, 1997) was done between geographical and Euclidian genetic distances. The lack of correlation ($r = 0.08$; $p = 0.18$) between these two estimates indicates no regional structuration of the diversity and suggests seed exchanges between farmers and/or agricultural institutes.

Toward a core collection based on MSTRAT analysis

Although not structured geographically, there is variability in Tunisian populations that will be helpful for future plant breeding. To ensure and determine the main populations with high diversity we build a core collection with the help of the MSTRAT tool (Gouesnard *et al*., 2001).

It appears that a core collection of 8 populations is enough to capture 90 % of molecular allelic (SSR and SNP) content of the 59 populations and a core of 16 for 98 % of the total allelic variability (Figure 3). The sampling efficiency of M strategy is superior to the random one to improve the choice of the more variable populations on molecular data (Figure 3_ graph of active variables). The morphological characters are entrained effectively (Figure 3_graph on target variables).

After 20 different runs and 50 iterations of MSTRAT, different core collections of 8 populations were built. From these data, 8 populations were chosen first because of their higher occurrence in the different runs and iterations. Second, it was checked their interesting morpho-phenological characteristics, mainly the shortness of the vegetative cycle. This core collection [Hat28 (Slougia), Hat15 (Dougga), Hat40 (Oued Zarga), Hat20 (Ksar Mezouar), Hat25 (Ain Chelou), Hat43 (Sejane), Hat11 (Tounga), and Hat65 (Oued Beja) (Figure 1)] contains 364 alleles, ie 88 % of the genetic variability of the Tunisian collection observed in our study. Comparing these populations with the other in Table 1, these 8 populations were already sorted out by either SNP or SSR markers. Five of them (Hat28, Hat15, Hat40, Hat20, Hat25) presented more distant morphological traits. Therefore, this core collection could be the basis to start a future selection for Tunisian sunflower.

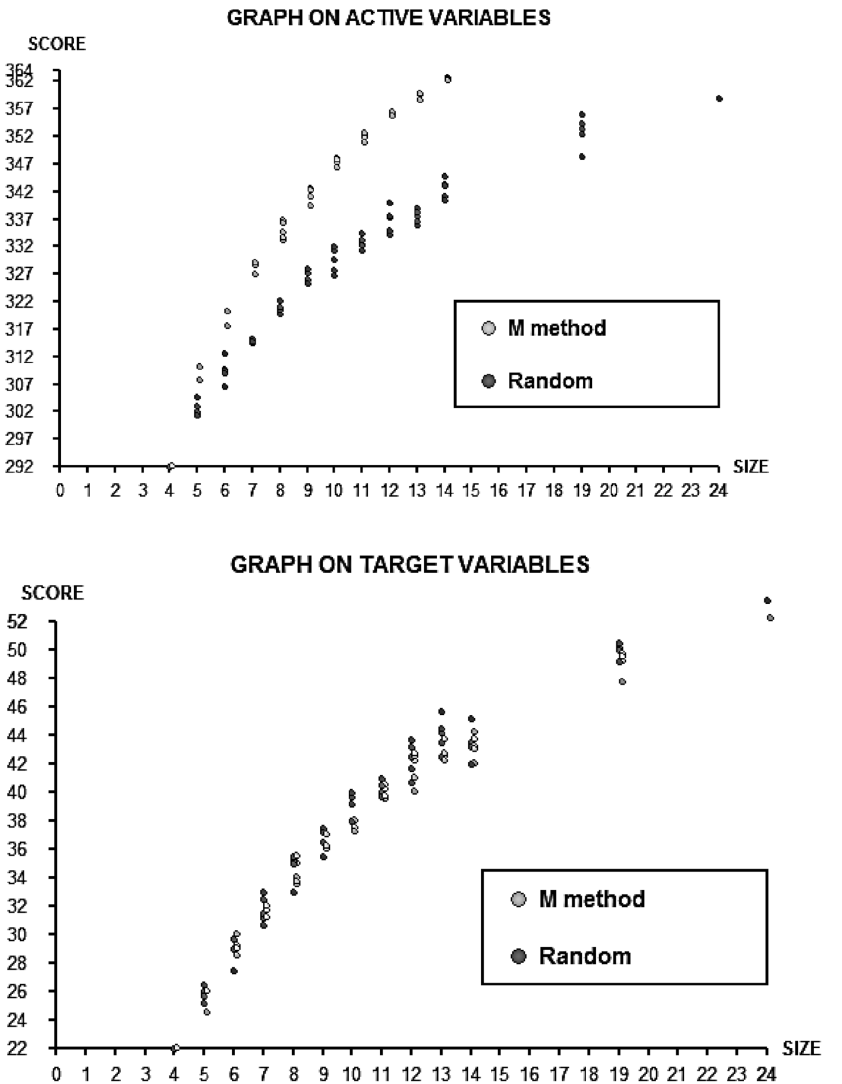

Figure 3: Comparison of the effectiveness in sampling genetic diversity between the M strategy (top curve) and the random strategy (bottom curve), according to the number of populations of the core collection. Molecular data (SSR and SNP) are the "active variables" and phenotype data, are the "target variables".

Conclusion

In conclusion, this is the first paper, to our knowledge, on reporting the genetic diversity of cultivated sunflower in Tunisia. Using different molecular markers, SSR and SNP, we have highlighted some genetic originality of the Tunisian material but without geographical structuration. However, these data were used to build a core collection of 8 populations that should provide a basis to expand the genetic base of the breeding material of sunflower in Tunisia.

Acknowledgements: This research was financially supported and performed by the laboratory of Plant-Microbe Interactions (LIPM) and US_EPGV (Etude du Polymorphisme des Génomes Végétaux) of the French National Institute for Agricultural Research (INRA). We are grateful also to Brigitte Gouesnard for the discussions about MSTRAT analysis and to Maha Ponnaiah for critical reading and help for English translation of the manuscript.

References

Al-Charaani, G.R., Gentzbittel, L., Huang, X., Sarrafi, A., 2004. Genotypic variation and identification of QTLs for agronomic traits using AFLP and SSR in recombinant inbred lines of sunflower (*Helianthus annuus* L.). Theoretical and Applied Genetics 109: 1353–1360.

Bachlava, E., Taylor, C.A., Tang, S., Bowers, J.E., Mandel, J.R., Burke, J.M., Knapp, S.J., 2012. SNP Discovery and development of a high-density genotyping array for sunflower. PLos one 7: 1–8.

Benassi, J.L., Labonne, M., 2004. Perspectives pour les oleagineux dans les pays du Maghreb: Algerie, Maroc et Tunisie 2000–2015. Oleagineux, Corps Gras, Lipides 11: 92–96.

Berry, S.T., Leon, A.J., Challis, P., Livin, C., Jones, R., Hanfrey, C.C., Griffiths, S., Roberts, A., 2003. Construction of a high density, composite RFLP linkage map for cultivated sunflower *Helianthus annuus*. *In*: Proc. Int. Sunflower Conference. Beijing, China, p. 14, 12–20th June, 2003.

Berry, S.T., Leon, A.J., Hanfrey, C.C., Challis, P., Burkholz, A., Barnes, S.R., Rufener, G.K., Lee, M., Caligari, P.D.S., 1995. Molecular marker analysis of *Helianthus annuus* L. and construction of an RFLP map for cultivated sunflower. Theoretical and Applied Genetics 91: 195–199.

Berry, S.T., León, A.J., Peerbolte, R., Challis, P., Livini, C., Jones, R., Feingold, S. 1997. Presentation of the Advanta sunflower RFLP linkage map for public research. *In*: Proc. 19th Sunflower Res., Fargo, USA.

Brunel, D., 1994. A microsatellite marker in *Helianthus annuus* L. Plant Molecular Biology 24: 397–400.

DGPA, 2013. Rapport annuel de suivi des emblavures, direction des grandes cultures, direction générale de la production agricole. Ministère de l'agriculture, de l'environnement et des ressources hydrauliques.

FAO., 2012. FAOSTAT Database from the Food and Agriculture Organization. www.fao.org.

Gedil, M.A., Wye, C., Berry, S., Segers, B., Peleman, J., Jones, R., Leon, A., Slabaugh, M.B., Knapp, S.J., 2001. An integrated restriction fragment length polymorphismamplified fragment length polymorphism linkage map for cultivated sunflower. Genome 44: 213–221.

Gentzbittel, L., Mestries, E., Mouzeyar, S., Mazyrat, F., Badaoul, S., Vear, F., Tourvieille de Labrouche, D., Nicolas, P., 1999. A composite map of expressed sequences and phenotypic traits of sunflower (*Helianthus annuus* L.) genome. Theoretical and Applied Genetics 99: 218–234.

Gentzbittel, L., Vear, F., Zhang, Y.-X., Bervillé, A., Nicolas, P., 1995. Development of a consensus linakge RFLP map of cultivated sunflower (*Helianthus annuus* L.). Theoretical and Applied Genetics 90: 1079–1086.

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. Journal of heredity 92: 93–94.

Heesacker, A., Kishore, V.K., Gao, W., Tang, S., Kolkman, J.M., Gingle, A., Matvienko, M., Kozik, A., Michelmore, R., Lai, Z., Rieseberg, L.H., Knapp, S.J., 2008. SSRs and INDELs mined from the sunflower EST database: Abundance, polymorphisms and cross-taxa utility. Theoretical and Applied Genetics 117: 1021–1029. http://www.ncbi.nlm.nih.gov/genbank.

Hu, J., Seiler, G., Kole, C., 2010. Genetics, Genomics and Breeding of Sunflower, Science publishers, British Channel Islands, p. 326.

Hulke, B.S., Grassa, C.J., Bowers, J.E., Burke, J.M., Qi, L., Talukder, Z.I., Rieseberg, L.H., 2015. A unified SNP map of sunflower (*Helianthus annuus* L.) derived from the Sunflower Genome Project resources. Crop Science 55: 1696–1702.

Jan, C.C., Vick, B.A., Miller, J.F., Kahler, A.L., Butler, E.T., 1998. Construction of an RFLP linkage map for cultivated sunflower. Theoretical and Applied Genetics 96: 15–22.

Kane, N.C., Gill, N., King, M.G., Bowers, J.E., Berges, H., Gouzy, J., Bachlava, E., Langlade, N.B., Lai, Z., Stewart, M., Vincourt, P., Knapp, S.J., Rieseberg, L.H., 2011. Progress towards a reference genome for sunflower. Botany 89: 429–437.

Khoufi, S., Khamassi, K., Teixeira da Silva, J.A., Aoun, N., Rezgui, S., Ben Jeddi, F., 2013. Assessment of diversity of phenologically and morphologically related traits among adapted populations of sunflower (*Helianthus annuus* L.). Helia 36: 29–40.

Kolkman, J.M., Berry, S.T., Leon, A.J., Slabaugh, M.B., Tang, S., Gao, W., Shintani, D.K., Burke, J.M., Knapp, S.J., 2007. Single nucleotide polymorphisms and linkage disequilibrium in sunflower. Genetics 177: 457–468.

Lai, Z., Livingstone, K., Zou, Y., Church, S.A., Knapp, S.J., Andrews, J., Rieseberg, L.H., 2005. Identification and mapping of SNPs from ESTs in sunflower. Theoretical and Applied Genetics 111: 1532–1544.

Livaja, M., Unterseer, S., Erath, W., Lehermeier, C., Wieseke, R., Plieske, J., Polley, A., Luerßen, H., Wieckhorst, S., Mascher, M., Hahn, V., Ouzunova, M., Schön, C.C., Ganal, M.W., 2016. Diversity analysis and genomic prediction of Sclerotinia resistance in sunflower using a new 25 K SNP genotyping array. Theoretical Applied Genetics 129: 317–329.

Miller, M.P., 1997. Tools for Population Genetic Analysis. *Version 1.3*, Department of Biological Sciences, Northern Arizona University, Flagstaff.

Mokrani, L., Gentzbittel, L., Azanza, F., Al-Chaarani, G., Sarrafi, A., 2002. Mapping and analysis of quantitative trait loci for grain oil content and agronomic traits using AFLP and SSR in sunflower (*Helianthus annuus* L.). Theoretical Applied Genetics 106: 149–156.

Paniego, N., Echaide, M., Muñoz, M., Fernández, L., Torales, S., Faccio, P., Fuxan, I., Carrera, M., Zandomeni, R., Suárez, I., Hopp, E.H., 2002. Microsatellite isolation and characterisation in sunflower (*Helianthus annuus* L.). Genome 45: 34–43.

Pashley, C.H., Ellis, J.R., Mccauley, D.E., Burke, J.M., 2006. EST databases as a source for molecular markers: Lessons from *Helianthus*. Journal of Heredity 4: 381–388.

Peerbolte, R.P. and Peleman, J., 1996. The Cartisol sunflower RFLP map (146 loci) extended with 291 AFLP markers. *In*: Proc. 18th Sunflower Res. Forum. Fargo, N.D., USA, 18, 1–12th January, 1996.

Pegadaraju, V., Nipper, R., Hulke, B.S., Qi, L., Schultz, Q., 2013. *De novo* sequencing of sunflower genome for SNP discovery using RAD (Restriction site Associated DNA) approach. BMC Genomics 14: 556.

Rieseberg, L., Baird, S.J.E., Desrochers, A.M., 1998. Patterns of mating in wild sunflower hybrid zones. Evolution 52: 713–726.

Schoen, D.E., Brown, A.H.D., 1983. Conservation of allelic richness in wild crop relatives is aided by assessment of genetic markers. Proceedings of the National Academy of Sciences 90: 10623–10627.

Tang, S., Heesacker, A., Kishore, V.K., Fernandez, A., Sadik, E.S., Cole, G., Knapp, S.J., 2003. Genetic mapping of the Or5 gene for resistance to Orobanche race E in sunflower. Crop Science 43: 1021–1028.

Tang, S., Slabaugh, M.B., Shintani, D.K., Knapp, S.J., 2002. Simple sequence repeat map of the sunflower genome. Theoretical and Applied Genetics 105: 1124–1136.

Yu, J.K., Mangor, J., Thompson, L., Edwards, K.J., Slabaugh, M.B., Knapp, S.J., 2002. Allelic diversity of simple sequence repeat markers among elite inbred lines in cultivated sunflower. Genome 45: 652–660.

Yu, J.K., Tang, S., Slabaugh, M.B., Heesacker, A., Cole, G., Herring, M., Soper, J., Han, F., Chu, W.C., Webb, D.M., Thompson, L., Edwards, K.J., Berry, S., Leon, A.J., Grondona, M., Olungu, C., Maes, N., Knapp, S.J., 2003. Towards a saturated molecular genetic linkage map for cultivated sunflower. Crop Science 43: 367–387.

Yue, B., Radi, S.A., Vick, B.A., Cai, X., Tang, S., Knapp, S.J., Gulya, T.J., Miller, J.F., Hu, J., 2008. Identifying quantitative trait loci for resistance to Sclerotinia head rot in two USDA sunflower germplasms. Phytopathology 98: 926–931.

Zhang, L.S., Le Clerc, V., Li, S., Zhang, D., 2005. Establishment of an effective set of simple sequence repeat markers for sunflower variety identification and diversity assessment. Canadian Journal of Botany 83: 66–72.

Supplemental Material: The online version of this article (DOI: 10.1515/helia-2016-0002) offers supplementary material, available to authorized users.

Analyse des ressources génétiques et constitution d'une core-collection pour la création de nouvelles variétés de tournesol en Tunisie.

Résumé

Le Tournesol (*Helianthus annuus* L.) est utilisé dans les pays méditerranéens en tant que fruit sec (pipas) pourrait offrir un débouché intéressant de diversification

de cette culture. La variabilité morpho-phénologique déjà mise en évidence dans les populations locales de Tunisie peut être une base à la création de nouvelles variétés bien adaptées aux conditions climatiques et pédologiques de cette région. La caractérisation moléculaire des lignées références et des 59 accessions a permis d'identifier 194 allèles pour trente locus microsatellites (soit 3 à 10 allèles par locus) et 54 haplotypes à partir de 117 SNPs détectés par le séquençage NGS de 7 gènes (soit 4 à 16 haplotypes par gène). Ces résultats confirme l'originalité génétique du matériel tunisien par rapport à des lignées témoin ainsi qu'une faible distance génétique entre les accessions. Néanmoins, une core collection de 8 populations, capturant 88 % de la diversité génétique, peut être proposée comme base pour un futur programme d'amélioration du Tournesol.

Mots clefs: Core collection, diversité génétique, SNP, SSR, Tournesol