

## DIVERSITY AMONG CULTIVATED SUNFLOWER RESOURCES AND USE IN BREEDING

---

Vear, F.<sup>1</sup>, Cadic, E.<sup>2</sup>, Vincourt, P.<sup>3</sup>

---

<sup>1</sup>INRA, UMR 1095, Domaine de Crouelle, 234, Ave du Brezet,  
63000 Clermont-Ferrand, France

<sup>2</sup>INRA, LIPM, UMR 2594, Auzeville, BP 52627, 31326 Castanet-Tolosan cedex,  
France Biogemma, Domaine de Sandreau, 31700 Mondonville, France

<sup>3</sup>INRA, LIPM, UMR 2594, Auzeville, BP 52627,  
31326 Castanet-Tolosan cedex, France

*Received: November 15, 2011*

*Accepted: December 10, 2011*

### SUMMARY

The first significant breeding to high oil content in sunflowers, which took place in Russia from 1930, concerned open pollinated varieties. These were the basis of modern breeding programmes, which started in 1950-60, mainly in Europe, Canada, the USA and Argentina, but work in that period in countries such as Morocco, Australia and Chile also produced genotypes of direct agronomic interest, or which widened the genetic variability available. Studies of the structure of this variability have become possible following genotyping of large numbers of inbred lines by microsatellites (2 or 3 per linkage group) and the definition of core collections which cover a large part of the diversity in cultivated sunflower. Although some geographic effects on combining ability have been reported, the main groups which appear are characterised by *cms* maintenance or male fertility restoration. The collections defined should be useful in similar studies for the determination of genome sequences controlling important phenotypic characters, although for characters determined by parts of the genome poorly represented by the markers, enriched or specific collections may be necessary. Such knowledge should help to optimize the use of sunflower genetic resources in breeding.

**Key words:** sunflower, genetic resources, breeding

### INTRODUCTION

Genetic resources for cultivated sunflower exist either as populations or as inbred lines. Their maintenance is expensive and efforts are required to determine how to obtain maximum efficiency, maintaining the maximum diversity with the least possible means. Two separate questions are involved: firstly, the determination of genetic diversity and secondly, the maintenance techniques. This paper will deal mainly with the first question, but we will deal shortly with the second as an introduction to a subject that should be dealt with in the discussion period of this symposium.

Populations

ПОДСОЛНЕЧНИК	
КРУГЛИК 7-15-163	КРАСНОДАРЕЦ
КРУГЛИК 631	ЛУЧ
КРУГЛИК А-41	ВЫМПЕЛ
КРУГЛИК 1846	САЛЮТ
КРУГЛИК 1975	ПЕРЕДОВИК улучшенный
АРМАВИРСКИЙ 611	СМЕНА улучшенный
АРМАВИРСКИЙ 768	ВНИИМК 8540 улучшенный
АРМАВИРСКИЙ 762	ВНИИМК 8931 улучшенный
ВНИИМК 1646	ВНИИМК 8883 улучшенный
ВНИИМК 1813	ВНИИМК 1646 улучшенный
ВНИИМК 3519	КРАСНОДАРСКИЙ
ВНИИМК 4036	КРУГЛИК
ВНИИМК 4966	ПРОГРЕСС
ВНИИМК 6540	НОВИНКА
ВНИИМК 8931	РЕЗЕРВ
ВНИИМК 8932	НАХОДКА
ВНИИМК 8883	ПОДАРОК
ПЕРЕДОВИК	ОКТЯБРЬ
СМЕНА	ЮБИЛЕЙНЫЙ 60
ВОСТОК	НАДЕЖНЫЙ
КУБАНЕЦ	ВНИИМК 43
ДРУГИЕ КУЛЬТУРЫ	

Figure 1: List of open pollinated varieties bred by V.V. Pustovoit at VNIIMK, Krasnodar, Russia

The first widely grown sunflowers with high content of oil were open-pollinated varieties, bred especially by V.V. Pustovoit at Krasnodar in Russia, the first, in about 1930, was “Kruglik”, with several different forms, followed by many other varieties (Figure 1, the poster of Pustovoit varieties at Krasnodar).

Table 1 gives some details of the relations between these varieties, found in INRA field notebooks dating from 1964 and in the pedigree map published by Korell *et al.* in 1992. However, this data is incomplete. It would be useful if somebody (Krasnodar?) could complete the Table. Additional forms were produced, either intentionally, with improvement in the yield or oil content at Krasnodar, or by

genetic drift in different environmental conditions in other countries.

Table 1: Origins of Russian open pollinated varieties

Variety	Date	Origin
Kruglik A-41	1930	?
VNIIMK 1813	?	?
VNIIMK1646	1938	?
VNIIMK 6540	1950	V 1813 × V1646
ARMAVIR 3497	1953	V 1646
VNIIMK 8931	1953	V 3452 × V 6540
VNIIMK 8883	1955	V 6540
Peredovik	1960	V 6540 × V8931
Armavir 9345		V 8883 × A 3497
Cernianka 66		
Progress	?	?
Novinka	?	?
Smena	?	?
Vostoc	?	?
Kubanec	?	?
Saliut	?	?
Luc	?	?
others	?	?

For example, at INRA we have had 6 samples of Armavir 3497, Armavir 9345 and VNIIMK 1646, 7 of VNIIMK 6540, 8 of VNIIMK 8931, 10 of VNIIMK 8883 and 16 of Peredovik, from different countries at different dates, starting with 1962. In addition, it is probable that some other populations with different names, intro-

duced from other countries, are in fact also forms of these Russian varieties. Some can be identified, like "NSP...", as selections from Novi Sad based on Peredovik, but others, with just the name of the place, or the person who provided the sample cannot be identified.

Table 2: Populations/open pollinated varieties for which INRA has several samples at different dates or from different counties which appear phenotypically identical

Variety	Identical accessions	Origins	Total accessions
Admiral	2	Canada France	3
Armavir 3497	2	Russia Croatia Italy	4
Beresanski	2	Serbia Russia	2
Comangir INTA	2	Argentina (INTA Cargill)	2
Donski 695	2	Russia Bulgaria	3
Fajba / GK70	2	Hungary	2
Gris Strie	2	Franc	4
Guyacan INTA	3	Argentina INTA (Codisol)	3
Impira INTA	2	Argentina	2
Issanka	3	France Tunisia	3
Jogovstokny	2	Russia (VIR) (Codisol)	2
Kisvardai	4	Hungary	6
Lovaspatonai	2	Hungary (Budapest Szeged)	2
Mennonite	2	Canada USA	3
Nagykorosi	2	Hungary (Budapest Szeged)	2
Nain Noir	2	France (Montpellier Cargill)	3
Novosadka 4	2	Serbia Croatia	4
Oro 9	3	Morocco	3
Peredovik	2+3	Russia VNIIMK (Codisol)	16
Population Branesti	3	Romania	3
Record	4	Romania Morocco	4
Saratov 169	2	France (Montpellier Versailles)	2
Saratov Ranni / Ranni	2+2	Russia France (Montpellier Versailles)	4
Smena/Puntano Smena	5	Russia Croatia Chile	11
Sunrise	3	Canada (Morden Ottawa)	3
VNIIMK 8883	2+4	Russia (VIR VNIIMK) Serbia	10
VNIIMK 8931	2	Romania Serbia	8
Yenissei	4	Croatia Romania France	5

There are other rather different open-pollinated varieties, for example from Canada (very early), Argentine (mostly late, although with Russian varieties as ancestors (Romano and Vazquez, 2003)), North and tropical Africa, as well as confectionary sunflowers from Spain, China and central Asia, including Amerindian selections from the USA. Only a few French open-pollinated varieties still exist, either confectionary/bird seed (Nain Noir and Gris Strie) or Peredovik adapted to France and Issanka, an early selection from VNIIMK 1646. There are also popula-

tions from many countries whose status have not been determined; they may have been grown or may have only been developed as sources of variability for breeding.

At INRA, including "pools" developed by INRA or by USDA, we had a total of more than 650 accessions, which posed an important maintenance problem. With the help of French sunflower breeders, some are regenerated each year under netting cages, but many have had to be maintained by sib crosses between 10 to 20 plants, probably with a consequent loss in variability. As the first step in optimisation of maintenance, from 2009 to 2010, we made phenotypic observations of these population samples, grouped according to name or possible relation and have found a significant number strictly identical and others with only a small number of phenotypic differences. Table 2 lists the most important origins with different introductions to France which appear strictly identical, and which we have started to combine as regenerations.

These phenotypic observations have also showed that some entries, considered as "populations", are quite homogeneous and can be considered as inbred lines, while some are certainly homozygous. On the other hand, others may result from the loss in variability during maintenance. Their homozygosity was checked by selfing and comparison of the following generation with a sample of the origin, to be sure that there was no effect of inbreeding. These results enabled 120 accessions to be transferred to the inbred line collection, selfing under bags being much cheaper than inter-pollination in netting cages. However, there is the problem with some self-sterile lines, but these should be maintained by only a few sib crosses.

Overall, these observations allowed us to reduce from 650 to 400 population accessions, including 60 pools of known origin, the number of accessions to be maintained separately (Table S1 in Klane *et al.*, 2011, gives the list and places and dates of origin, excluding the INRA and USDA breeding pools). How to go further? The evident reply is molecular analyses, but this has still to be put into practice. We have identified 66 pairs or groups of accessions with very similar phenotypes, and classed them in groups which will have to be combined in case molecular studies do not show any significant difference between them. It remains to be seen whether molecular analyses can go further than phenotypic observations in the case of polymorphic populations. Some molecular analyses with AFLP or SSR were made in the past, but generally the markers were not adapted (because their adaptation would differentiate alleles from wild species or because there was a single sample for each population, making it difficult to show intra population variability).

Quantitative PCR and SNP studies could make it possible to improve this situation, to determine whether similar samples have the same genotype, perhaps whether populations with different names are the same. Going further, it would be most useful to determine whether different Russian populations, developed one from another, really provide different alleles that need to be kept by all breeding and research programmes. For example, if VNIIMK 8883 is confirmed as an earlier, shorter version of VNIIMK 6540, which it was derived from, if it has no new alleles,

it will not be necessary for research organisations, other perhaps than VNIIMK, to maintain both on a large scale. This may also be applicable for the series of Argentinean populations whose ancestors are also known (Romano and Vazquez, 2003).

It would be very useful if further phenotypic and molecular comparisons could be made with international collaboration.

### **Inbred lines**

The INRA collection of inbred lines includes about 2200 species, both those bred by INRA and those introduced from other countries (600 maintainers + *cms*, 1000 maintainers and 600 restorers). The origins of some are quite clear, whereas other just have a number and the country of origin (BU 1507, BU1740, MO502, MO60, P1010, P1015, MH.1, MH.2). All details in introduction lists may be important. As an example, a line denoted F 200 was provided by Freeman-Johnson, Cargill USA to P. Leclercq in 1969. We thought, for a long time, that "F" was for Freeman-Johnson, and when this line was presented for registration, we realised that it might be the same as a parent of a Cargill hybrid. However, in other old papers, we found a detailed list and realised that "F" stood for France, a selection from Peredovik that Leclercq had started and Freeman-Johnson continued!

In general, phenotypic studies of inbred lines are easier than those of populations and more knowledge has been obtained over the years, which should be kept as valid if the lines are highly, if not strictly, homozygous. Thus, there should be less possibility of occurrence of several identical lines. But, what is their structure and differentiation and which are the most important? The first "core" collections were made up before molecular studies, where the series of lines which were thought to be representative served to search for resistance when a new disease appeared. As a consequence, the lines related to the best ones were tested. This happened in particular in the case of resistance to *Phomopsis*, broomrape and the quantitative resistance to downy mildew. The choices for these studies have been useful in starting core collections for molecular analyses.

Molecular studies started about 20 years ago and have developed considerably. Their general interest is to provide information on the structure and variability of cultivated sunflower, but there have already been several applications:

Firstly, in the variety of registration, they serve to checked homozygosity and that hybrid seed is derived from the parents announced.

Secondly, the data obtained have made it possible to construct core collections which truly represent allele diversity. This is directly used in association studies to identify agronomically important genes. Perhaps in the future, genomic data will help to identify lines that maximise heterosis.

Concerning genetic resources, the question is to decide which lines are most important to retain, among old lines whose origin is not well known and among recent breeding lines, which present different combinations of alleles, but not necessarily any original ones.

The first studies were made with RFLP, using only a few lines but with quite wide genome coverage. For example, Gentzbittel *et al.* (1994) analysed 17 inbred lines by RFLP, with a total of 331 probe-enzyme combinations, 101 of which showed polymorphism giving 614 banding patterns.

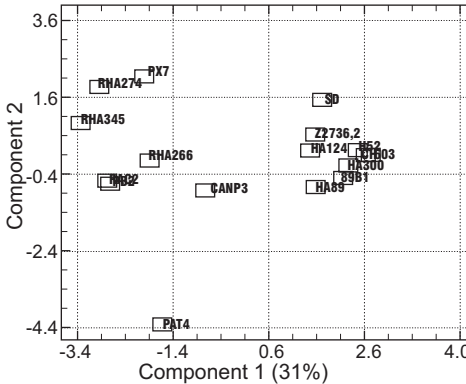


Figure 2: PCA of Gentzbittel *et al.*, 1994.

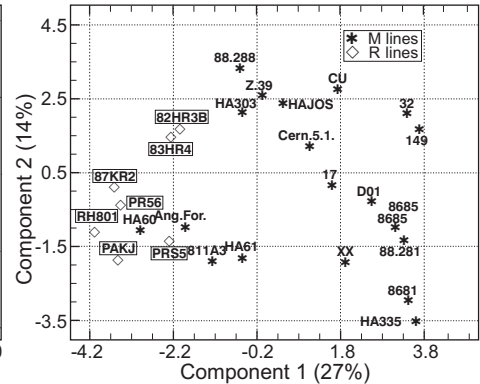


Figure 3: PCA of Zhang *et al.*, 1995.

Figure 2 presents their principal component analysis, dividing the lines very clearly into maintainers and restorers, except for one maintainer line CANP3 (derived from Armavir 9345). The wide difference of the line PAT4 was explained by the fact that it probably contained some genes from *Helianthus tuberosus*, but this has not been proven. Zhang *et al.* (1995) made a similar study with 26 inbred lines, 81 probe-enzyme combinations and 345 polymorphic fragments. Again (Figure 3) restorers and maintainers were distinguished. More recently, mapped SSR have made it possible to cover equally all linkage groups, but with a rather small number of markers, but on larger numbers of lines... Coque *et al.* (2008) used 51 SSR (3 per linkage group) on more than 500 lines, to make up a core collection of 384 lines, including a “public “core of 96 lines (INRA+USDA, available for research) which accounts for 60% of total diversity (some private lines were from interspecific introgression). With a few exceptions, restorers and maintainers were again separated, but the restorers also appeared to form 2 groups (Figure 4). Twelve lines were defined as the basic core “kernel”, because they are the parents of widely used RIL families (-XRQ PSC8, RHA266 PAC2, FU PAZ2) or because of their historical importance internationally or in France (HA89, RHA274, 2603, 83HR4, PST5, 92A6). It is interesting to compare their positions in the different PCA.

Using two SSR per linkage group, Mandel *et al.* (2011), studied 433 accessions including the 86 lines from the INRA core and lines from the cores of Brothers and Miller (1999) and also some confectionary lines. They distinguished the oil restorers from the other lines (Figure 5), which suggests that the branching gene is important, since many confectionary restorers are unbranched, like maintainer lines. Some of the INRA restorers appear to form the second restorer group.

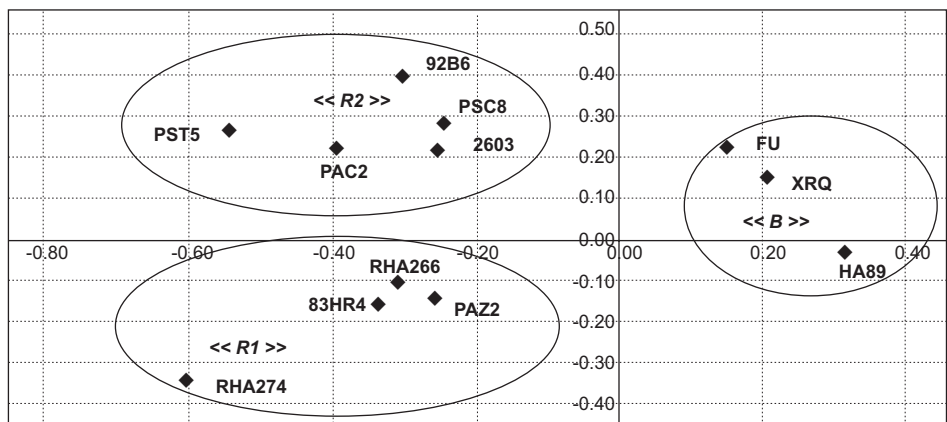


Figure 4: PCA of Coque et al., 2008, showing 12 central core lines in 3 groups defined by 384 lines

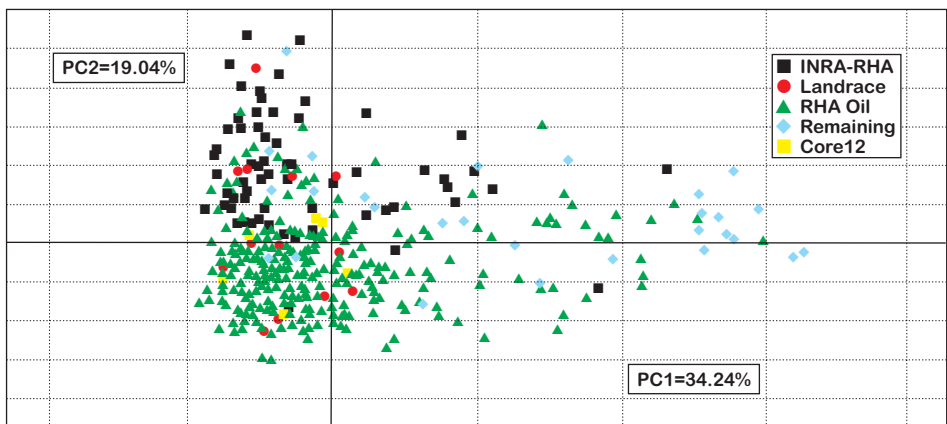


Figure 5: PCA of Mandel et al., 2011 on USDA and INRA lines

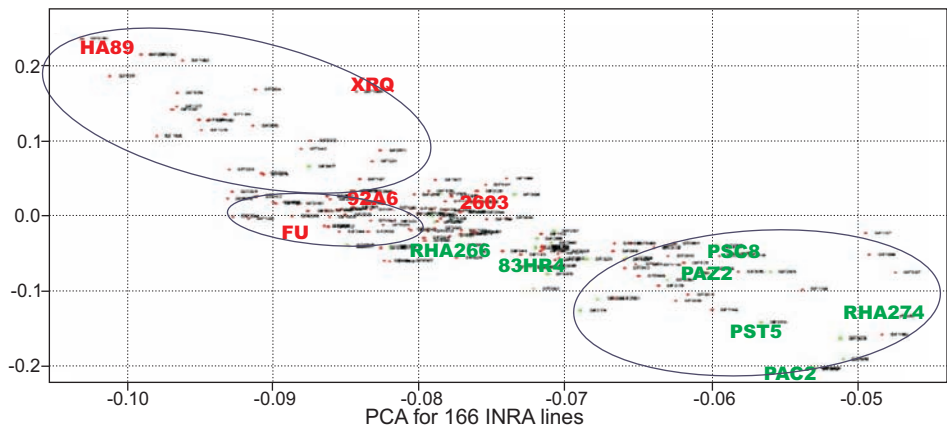


Figure 6: PCA for 166 INRA lines, using 4154 SNP, PC1-PC2

Our most recent results, with 4154 SNP on the INRA 96 core together with 70 other INRA lines show a quite similar picture, although with a general gradation from maintainers to restorers (Figure 6). In this case, certain alleles from HA89, and also the related downy mildew resistant forms HA335 and HA338 and RHA340, and the Spanish oleic version HA-OL-09, appear to have considerable weight and these origins should be separated from the restorer group. The top left of the figure is entirely composed of such lines. Lines of Bulgarian origin, which do not appear phenotypically very polymorphic but which were chosen in the 96 core to represent variability, also appear grouped, to the left below the HA89 lines. The restorers do not show 2 groups but are all close together, except for RHA266, unbranched, and 83HR4, which is placed in the middle of diverse maintainer lines. However, the 12 kernel lines of the Coque *et al.* (2008) collection still appear to cover the variability of all the material.

Comparison of results with SSR and SNP for the 96 core shows highly significant correlations:

PC1-PC2:  $r=0.413$   $p=0.0001$  R: 17%

PC2-PC3:  $r=0.656$   $p=0.0000$  R: 43%

Since so many more markers have been used in the SNP study, this can be considered as the best representation of genetic variability, but the SSR data are in general agreement. As more markers, or genes, become available, PCA should be increasingly similar and lines should be quite stable in their positioning. It will be of great interest to add additional lines that have not been studied so far, to determine whether they fill blank spaces on the graph.

It should also be possible to study groups of lines in detail to determine their similarities and differences. For example, the 96 core was phenotyped for reaction to *Sclerotinia capitulum* attack, together with an additional series of highly resistant lines. It was found that the core did not represent differences in reaction in a way which is satisfactory enough, with a large excess of susceptible lines (Vear and Grezes-Besset, 2010). It may be that only a relatively few genes are involved so that differences were not made evident by the 51 SSR. It would be most useful to be able to place all lines on a PCA, to identify not only identical positions, but also groups of lines with similar or contrasting characteristics.

### Maintenance

Sunflower seed is generally dried at 30-35°C to 3-6% RH and then kept in cold chambers at about 4°C and low RH. Experience shows that its lifespan is very variable, from 5 years to about 30 years, but with no evident explanation, the lines with low oil contents probably remain viable longer than those with high oil contents, but the relation is not strict. In 2001 and 2002, the Montpellier team (Y. Griveau) dried seed of a number of samples of populations down to 1-2% RH and then stocked them in airtight aluminium foil/plastic bags. These have been kept in a commercial frozen chamber at -18°C. We tested 9 samples in March 2011, allowing them to



warm up slowly, 24h in an ice box and then 2 days at 19°C (room temperature) with the bag open to rehumidify slowly, before soaking and germination at 23°C. We found the germination as good, or even better, as the same samples kept in the normal cold chamber. Table 3 indicates final results after 7 days. It should be noted that most samples were not frozen during the year of multiplication, so less than 100% germination may be from harvest or during first stocking in the cold chamber.

Table 3: Percentage germination of seed samples of sunflower populations kept in a cold chamber for 11 to 15 years or in a freezer for the last 10 years

Population	Production year	% germination	
		cold chamber	freezer
Assad 10/83.68	1999	82	86
Argentine.35	1998	93	100
Cirmis.1893	1996	82	76
Ella.124	1998	73	79
Peredovik.1219	1998	84	85
Progress.1815	2000	89	89
Sintetic 2.437	1997	91	90
Timisoara.2.RA.172	1999	100	75
Z 2736.S.6.399	1997	97	99

The conclusion of this study is that freezing is a quite possible solution to genetic resource maintenance. The questions for the future will be how often to test samples and how long they will last, as well as if freezing without special drying would give the same results and are the bags used of any importance.

Supposing a satisfactory reply is obtained to these questions, the other that remains is that if scientists, breeders and technicians only see lines once or twice in their careers (every 20 years?), how will they know what this material is like, what might be useful in research or breeding programmes? That is the reason why the long-term genetic resource management needs to be carefully organised so that observations are more frequent than regenerations.

## CONCLUSIONS

Genetic resource management has so far been based on phenotypic observations and the knowledge of the origin of accessions. If the former is probably quite satisfactory, there are many gaps concerning origins and ancestry and it is important that these should be filled and become available internationally. With automation of molecular analyses, determination of genetic variability is becoming a reality. The data available at present appear to correspond quite closely to the known ancestry and breeding methods. What remains to be compared is the phenotypic and genotypic differentiation between accessions in order to decide the best combination of the two in management of resources and in breeding.

## ACKNOWLEDGEMENTS

*We would like to thank our colleagues at INRA and in breeding firms for their help in maintenance of resources. Core collection analyses were supported by a French "Agence Nationale pour la Recherche" programme "Sunyfuel" and by «Oleosol», a programme with contributions from Midi-Pyrenees region, Aveyron department, Rodez, Biogemma, Syngenta Seeds, Soltis and RAGT. E.Cadic receives CIFRE thesis support from Biogemma.*

## REFERENCES

- Brothers, M.E., Miller, J.F., 1999. Core subset for the cultivated sunflower collection. Proc. 21<sup>st</sup> Sunflower Workshop Fargo, ND, USA. Pp. 124-127.
- Coque, M., Mesnildrey, S., Romestant, M., Grezes-Besset, B., Vear, F., Langlade, N., Vincourt, P., 2008. Sunflower line core collections for association studies and phenomics. Proc. 17<sup>th</sup> Int. Sunflower Conf., 8-12/6/2008, Cordoba, Spain. Pp. 725-728.
- Gentzbittel, L., Zhang, Y-X., Vear, F., Griveau, Y. & Nicolas, P., 1994. Nuclear RFLP to assess genetic relationships among inbred lines of cultivated sunflower *Helianthus annuus* L.: evidence for distinct restorer and maintainer germplasm pools. Theor. Appl. Genet. 89: 419-425.
- Kane, N.C., Burke, J.M., Marek, L., Knapp, S.J., Seiler, G., Vear, F., Vincourt, P., Rieseberg, L.H., 2011. Sunflower genetic, genomic, and ecological resources. Molecular Ecology Resources (proposed).
- Korell, M., Mosges, G., Friedt, W., 1992. Construction of a sunflower pedigree map. Helia 15: 7-16.
- Mandel, J.R., Dechaine, J.M., Marek, L.F., Burke, J.M., 2011. Genetic diversity and population structure in cultivated sunflower and a comparison to its wild progenitor *Helianthus annuus* L. Theor. Appl. Genet. DOI 10.1007/s00122-011-1619-3.
- Romano, A.B., and Vazquez, A.N., 2003. Origin of the argentine sunflower varieties. Helia 26: 127-136.
- Vear, F., Grezes-Besset, B., 2010. Progress in breeding sunflowers for resistance to *Sclerotinia*. Proc. I.S.A. Symposium "Sunflower breeding for resistance to diseases", Krasnodar, Russia, 19-21/6/2010: 30-35. Plenary lecture.
- Zhang, Y-X., Gentzbittel, L., Vear, F. & Nicolas, P., 1995. Assessment of inter and intra-inbred line variability in sunflower (*Helianthus annuus*) by RFLPs. Genome 38: 1040-1048.