Analysis of genetic diversity based on molecular markers (AFLP) and of heterosis in faba bean (*Vicia faba* L.)

Doctoral Dissertation

submitted for the degree of Doctor of Agricultural Sciences
of the Faculty of Agricultural Sciences
Georg-August-University Göttingen
Germany

by

Mahmoud Mohamed Zeid

from Alexandria, Egypt

Göttingen, February 2003

D7

Referent : Prof. Dr. Wolfgang Link

Korreferent : Prof. Dr. Reiner Finkeldey

Tag der mündlichen Prüfung : 6 Februar 2003

Table of Contents

The Roman numbers from I-III refer to the manuscripts presented in the pages to follow.

Appendix

Manuscripts I-III

The present thesis is based on the following manuscripts, which are referred to by their Roman numbers.

I. Zeid M, Schön CC, and Link W, (2001) Genetic diversity in a group of recent elite faba bean lines. Czech J Genet Plant Breed 37: 34-40

II. Zeid M, Schön CC, and Link W, (-) Assessing the genetic diversity in recent elite faba bean lines using AFLP markers (in preparation for Theoretical and Applied Genetics)

III. Zeid M, Schön CC, and Link W, (-) AFLP-based genetic similarity and hybrid performance in faba bean (in preparation for Crop Science)

1 Introduction

1.1 Importance and production of faba bean

Faba bean (*Vicia faba* L.) is a valuable protein-rich food that provides a large sector of the human populations in developing countries with a cheap protein source thus partly compensating for the large deficiency in animal protein sources. In developed countries faba bean provides an alternative to soybean meal for animal feed, this being particularly important in the more industrialized countries. The world area devoted to faba bean is continuously in decline, falling from 3.7 m ha in 1979-81 to 2.4 m ha in 2000-01. This reduction is mainly attributed to the unreliable yields and the poor returns from the crop.

After the harvest shortfall of soybean in the USA in 1973, the European Community realized the importance of domestic source of protein for animal feed. It is worth mentioning that the political decision to support protein crops production was set up 20 years after cereals and 12 years after oil seed regulations were made (Carrouee 1995). Under the current policy of the European Union it is expected that the area under protein crops will be at least maintained at around 1.3 m ha, thus ensuring a domestic production of 5-6 million tons of excellent quality protein (Hulot 1999). Today about 75% of protein-rich products (mainly soy meal and soybean) in Europe is imported, and is likely to increase with the ongoing ban on the use of meat and bone meal in the animal feed industry (Struan, 2002).

For decades, faba bean was the only grain legume crop which had been widely grown in Europe. During the eighties soybean and pea production suddenly rose and today faba bean shares sweet lupines in only about 19% of the 1.3 m ha planned for protein crops. Yield instability (Fig. 1.1) and low prices are behind the continuous drop in the area (Fig. 1.2) devoted to faba bean in Germany, France and the European Union as a whole. France has lately regained interest in the crop with a tremendous jump of 0.43 m ha in 2001 without any noticeable improvement in yield/ha (mean area being 0.13 m ha over the past 10 years). This could be explained by the fact that faba bean is replacing peas in areas infested with root rot *Aphanomyces euteiches* where reduction in pea yield amounts to 90% (Lacampagne 2001).

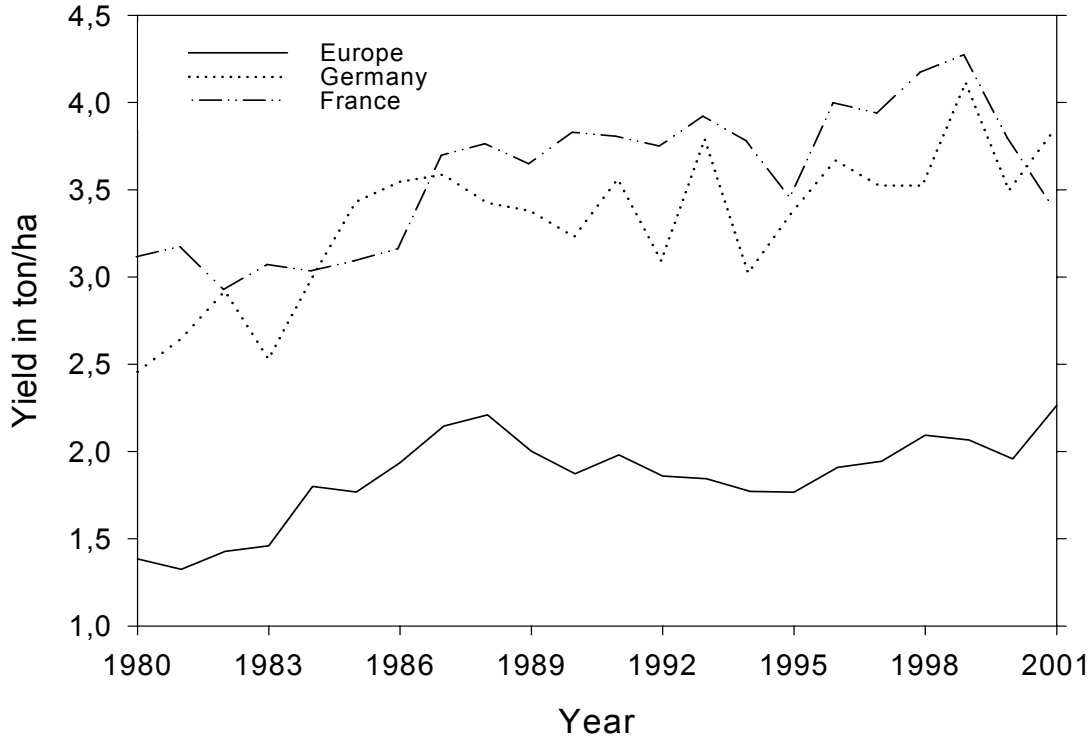


Fig. 1.1 Yiled of faba bean in the period 1980-2001
Source: FAO (2001)

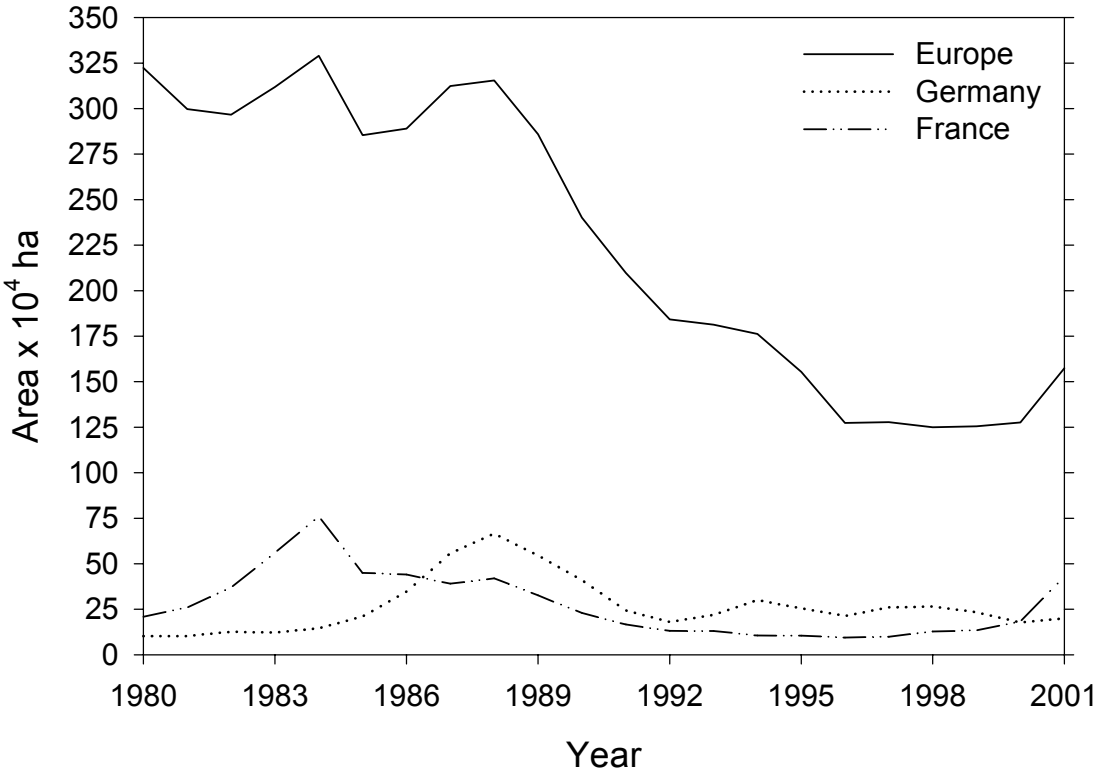


Fig. 1.2 Area devoted to faba bean in the period 1980-2001
Source: FAO (2001)

Faba bean plays a significant role in improving the productivity of the soil in the cereal-based rotations where it serves as a break crop; yields of cereal crops following faba bean are improved and needs for nitrogen fertilizer applications are reduced. Studies on the fixation of atmospheric nitrogen through symbiosis in organic farming (Schmidtke and Rauber 2000) have shown that faba bean surpasses peas in the amount of nitrogen fixed. In addition, an effective management of faba bean in crop rotations by reducing plant-available soil nitrogen before, during and after growing faba bean could achieve maximum symbiotic activity, low levels of nitrogen leaching and high yield of the succeeding non-legume crops. Schmidtke and Rauber (2000) reported that grain legumes play a more important role in organic farming than in conventional farming systems. In Germany, the area of arable land grown with faba bean and pea was 1.2% on conventional farms but 4.9% on organic farms. This demonstrates the need for supply of home grown legumes for animal feed and at the same time the ability of legumes to meet their own demand of nitrogen through symbiosis and the positive nitrogen effect of grain legumes on non-legume crops following in the crop rotation.

1.2 Distribution

Faba bean is widely believed to have originated in the Mediterranean-West Asia region probably in the Neolithic period (Cubero 1974). Throughout their long history as a cultivated crop, faba bean has been subjected to both natural selection and selection by farmers in the different environments where the crop has been grown. Although no successful crossing between faba bean and any of the other *Vicia* species has been reported, a wide range of genetic variation in the species still exists (Lawes et al., 1983). In spite of centuries of such selection, *Vicia faba* retains vestiges of its wild past and, in certain aspects, can be regarded as an incompletely domesticated species. The indeterminate nature of the growth habit and the existence of dehiscent pods in many populations can be cited as examples (Hanelt 1972). The mating system of the species, which stands between full autogamy and full allogamy, may be another (Hawtin 1982).

Distinct groups within faba bean based mainly on seed size, ranging from small seeded *minor* beans (0.2-0.5 g per seed) to medium seeded *equina* beans and the large seeded *major* beans (single seed weight of more than 2.0 g) have been

recognized by Muratova (1931). These groups can still be recognized in the major areas of production where the *minor* and *major* groups are grown in Central and Northwest Europe, both groups in addition to the *equina* group in South Europe, North Africa and up to West Asia and the *major* group in South America.

1.3 Breeding

Progress in breeding faba beans for resistance to biotic and abiotic stresses has been slow and with only few breakthroughs to be considered. Bond et al. (1994) reviewed the contributions of plant breeding facing these problems showing resistance to Chocolate spot (*Botrytis fabae*) has been identified in ICARDA lines coming from Ecuador, resistance to *Orobanche crenata* by breeding the cultivar 'Giza 402' in Egypt and breeding of the frost hardy cultivars Côte d'Or in France and Hiverna and Webo in Germany as the main breakthroughs. The problem of yield instability (mainly due to biotic and abiotic stresses) in faba bean however dominates over all achieved progress as indicated by the drastic reduction in the area devoted to the crop.

Although improvement in seed yield and yield stability are the primary objectives of most faba bean breeding programmes, in Germany, an annual increase of only 0.6% in yield could be achieved as compared to a yield improvement of 2% per year in case of wheat (Schön 1997). The number of approved faba bean cultivars is low and pedigree information regarding available cultivars is scarce and not well documented, especially since cultivars are not the outcome of a specific cross but from natural crosses or are populations improved through recurrent selection (Stelling et al., 1994).

The challenge in breeding faba bean resides mainly in its reproductive system being partial allogamous. Both cross and self-pollination may occur on the same plant as assisted by insects. Natural outcrossing also varies greatly, reported to be 35% (Bond and Poulsen 1983) and 60% on the average (Suso and Moreno 1999), depending on the genotype and environmental conditions. Results of Link et al. (1994b) showed that outcrossing in faba bean inbred lines ranged from 38 to 73% with an average of 54%. Ebmeyer (1988) has already pointed out that in open

pollinated populations of crossbred faba bean plants tend to self-pollinate, whereas inbred plants tend to cross-pollinate.

Heterosis has been realized by Bond (1966) in diallel crosses of winter bean types, with an average yield improvement of almost 23% above the best parent. Other studies showed 50% heterosis and more above the midparent (Von Kittlitz 1986; Link and Ruckenbauer 1988). Link et al. (1994a) stressed on the importance of utilizing the heterosis still available in faba bean, if yield stability is to be achieved. The genetic variability within open-pollinated faba bean varieties and the proportion of heterosis occurring under partial allogamy was studied by Ebmeyer and Stelling (1994). Their results showed that 70% heterosis for grain yield was achieved for crosses between inbred lines of different varieties and that only two thirds of this heterosis could be utilized in open pollination.

Two ways appear to be available to utilize heterosis in faba bean, namely: the production of hybrids and formulation of synthetics. Both methods appear to promise higher yield and yield stability than open-pollinated populations.

Results presented by Bond (1966) and Ebmeyer (1988) showed a large variation in the general combining ability and the low levels of specific combining ability in different faba bean crosses. These results are generally in favour of breeding synthetic varieties; however, after considering the high degree of self-fertilization occurring in the crop, hybrids appear to be more appropriate. Faba bean hybrids have also shown better adaptation to a wide range of abiotic conditions as compared to open pollinated or inbred cultivars, expressed in improved fertilization at high temperatures (Bond et al., 1994), and tolerance to lack of pollinating insects owing to their heterotic autofertility (Link, 1990), reduced winter damage (Bond et al., 1986) and better tolerance to drought stress (Abdelmula et al., 1999). Three–way hybrids, double hybrids, or single hybrids grown in blends were suggested as the key to yield improvement and stability in faba bean (Stelling et al., 1994).

Cytoplasmic male sterility (CMS) is a prerequisite for commercial hybrid variety production. More than 35 years after Bond et al. (1964) have discovered the first CMS system in faba bean, no hybrid cultivar is available. The latest work done by

Link et al. (1997) resulted in a more stable system (CMS199). Although studies regarding this system are still in progress (Martsch et al., 2001) hybrid varieties today are unfeasible mainly due to the instability of the available CMS systems.

Until hybrid varieties are ready for commercial production, synthetics offer a means of exploiting an intermediate level of heterosis from highly selected and characterised parents. A synthetic variety in faba bean was defined by Bond (1982) as any population, which has been constituted from a limited number of distinct and well-evaluated components (usually inbred lines). Four or five are common; this provides the best balance between the danger of too high sibbing and the inclusion of more lowly ranked components if the number is greater (Bond 1982).

A synthetic variety usually equals or exceeds the yield of the highest yielding component whichever one it is. Synthetics are more stable compared to their more inbred components due to the degree of hybridity they possess. The optimum number of components entering a synthetic (Link and Ederer 1993) and the suitable synthetic generation for maximising yield (Stelling et al., 1994) have already been defined and the amount of expected performance from the highest yielding synthetic is estimated to be at least 15% higher than the highest yielding inbred line.

Reviewing the pedigree information of some German cultivars, Stelling et al. (1994) reported that both the cultivar Minica (from the *major* group) and its offspring Alfred were frequently used as parents for many cultivars. Similarly, the cultivar Kleinberger Kleinkörniger was shared in the pedigree of almost all cultivars in Austria.

With the introduction of molecular marker techniques, new reliable tools that are neither affected by the surrounding environment, nor by growth stage of the plant (as in case of morphological characters) became available for the breeder. These can be applied for organizing germplasm, identification of cultivars, assisting in the selection of parents for hybridization and reducing the number of accessions needed to ensure sampling a broad range of genetic variability. Restriction fragment length polymorphism (RFLP) has been first employed by Van de Ven et al. (1990) in faba bean as a first step to create a linkage map. With the introduction of the PCR (polymerase chain reaction) more cheaper and less labour intensive marker

techniques than the RFLP became available. The random polymorphic DNA (RAPD) technique is one of those PCR-based methods that have become widely used in estimating the genetic relationships among genotypes. Link et al. (1995) have employed the RAPD technique to study the genetic diversity in European (*minor* and *major)* and Mediterranean faba bean germplasm. Their results were very promising, showing that within the European *minor* pool itself genetic diversity estimates were larger than that between the minor and major groups, implying the possibility of establishing genetically divergent heterotic groups even within the European minor germplasm.

Amplified fragment length polymorphism (AFLP), is another PCR based marker that exhibits several advantages over other markers available for this type of research, including: (i) the generation of a large number of markers in a single PCR reaction, (ii) a high level of polymorphism and (iii) a high reproducibility and reliability due to stringent PCR conditions (Vos et al. 1995).

The hypothesis of an association between molecular marker data and heterosis and/or hybrid performance is obvious. Experience taught breeders that there is a nearly linear relationship between pedigree distance of parents, i.e. heterozygosity and heterosis in the corresponding hybrid. Molecular markers allow inspection of an unselected sample of chromosomal loci for parental distance; hence, it is very tempting to forecast heterosis from this data. Prediction of heterosis has been investigated in different crops including maize (Lanza et al., 1997; Ajmone Marsan et al., 1998), soybean (Crena et al., 1997) and sunflower (Cheres et al., 2000). The aim of those studies were the prediction of the performance of single cross hybrids from marker information collected on their parental lines, thus predicting the performance of a large number of F1-hybrids beyond available capacity of any breeding program and at the same time saving the expenditure and time devoted for field evaluation. Contradicting results were reported ranging from suitability of the marker information for prediction purposes (Lanza et al., 1997) to lack of correlation between marker information and F1-hybrid performance or heterosis (Crena et al., 1997). The reasoning for the inconsistency was demonstrated by Frei et al. (1986), who showed that the usefulness of molecular markers for predicting hybrid performance depends

on whether crosses were produced including related lines or were a result of unrelated lines only.

2 The objectives of this work were to:

1- Study the structure of the genetic diversity within a selected sample of elite cultivars representing the actual faba bean material used for breeding in Asia, Europe and North Africa based on AFLP markers

2- Analyze both the reliability of the AFLP technique and consistency of the results obtained

3- Determine whether selection of genetically independent faba bean inbred lines within the elite European gene pool, (i.e. leading to non-inbred hybrids) is possible, if selection is solely based on available pedigree information. To achieve this objective:

(I) Available pedigree data on a sample of 18 inbred faba bean lines from the European gene pool were studied and independent parents were defined for hybrid production.

(II) The pattern of genetic diversity within this set of faba bean lines was investigated through high input assessment of AFLP markers.

(III) Associations between AFLP based genetic similarities of these inbred lines with agronomic performance and heterosis of their single cross hybrids were assessed.

3 Plant material

This investigation was based on a sample of 79 *Vicia faba* L. inbred lines derived from elite cultivars that were released in various world markets (Table 3.1). Inbred lines were developed in Germany at the Institute of Agronomy and Plant Breeding, Göttingen, and at the State Plant Breeding Institute, Stuttgart-Hohenheim, for one to more than 12 generations of selfing, employing the single-seed descent method.

Table 3.1 Country of origin, pedigree data and year of release of cultivars that were inbred and used in this study

Faba bean cultivar	Company/ Country of origin	Pedigree / Year of release	Inbreeding generation
European *minor*			
Alfred*	CEBECO / Netherlands	Minica x Horse bean/1983	F_3
Alpine	PBI/ UK	Victor x Toret/1997	F_3
Blaze	PBI/ UK	1976	F_3
Blitz	Redview Farms Ltd. / Canada	1994	F_2
Carola	Probstdorfer SZ / Austria	Minica (parent)/1986	F_8
Caspar	CEBECO / Netherlands	1992	F_7
Columbo	DLF Trifolium / Denmark	1996	F_3
Condor	NPZ Lembke / Germany	Kristall x Alfred/1990	F_8
Corvette	Twyford / Denmark	1996	F_3
Erfano	Gotha SZ/ Germany (GDR)	Minica x VF 1094-72/1988	F_5
Fatima	Rowland Uni. Sascat. / Canada	1993	F_2
Francks Ackerperle	Dr. Franck/ Germany	-	F_{10}
Franz	Wieselburger SZ / Austria	Selected from Wieselburger kleinkörnige/1990	F_4
Geo	Dr. Franck/ Germany	Herz Freya (parent) /1989	F_6
Gloria*	Winkler Gleisdorfer SZ / Austria	Kornb. Kl. Körnige x 757 /1994	F_8
Gobo*	Gotha SZ/ Germany (GDR)	1987	F_8
Hedin	Dr. Franck/ Germany	Herz Freya (relative)/ 1986	F_9
Herz Freya*	Herz / Germany	1935	F_{12}
Jantarnij	Russia	K-1408 x K-1663/1982	F_2

Table 3.1 continued

Karna*	Wieselburger SZ / Austria	Kornberger Kk. (parent)/1983	F_6
Kornberger-Kleinkörnige	Gleisdorfer SZ / Austria	1956	F_{10}
Kristall	Lochow Petkus / Germany	1973	F_{12}
Limbo	Lochow Petkus / Germany	Victor x Steuckart/1998	F_2
Luna	PBI/ UK	1991	F_2
Maris Bead*	PBI/ UK	1964	F_3
Mars*	Danisco / Denmark	Alfred x Maribo / 1993	F_3
Maya*	Serasem / France	Troy x ~Minica / 1995	F_3
Merkur*	Selgen / Czech Republic	Bolo x line x Fribo / 1997	F_3
Music*	Blondeau / Frankreich	1995	F_3
Nadwislanski	Poland	1955	F_7
Nixe	Germany	-	F_{12}
Orletzkij	Russia	Ackerperle x Tulunskie /1993	F_2
Pistache*	Joordens/ Netherlands	Felissia (parent)/1990	F_7
Scirocco*	NPZ Lembke / Germany	1992	F_7
Styria*	Gleisdorfer SZ / Austria	Carola x Mythos/1996	F_{12}
Topas	Lochow Petkus / Germany	Herz Freya x Kristall 1986	F_8
Troy*	NPZ Lembke / Germany	1985	F_{11}
Ukko	Boreal PB/ Finland	1984	F_7
Victor*	CEBECO / Netherlands	Minica x Cocksfieldsp./1988	F_{10}
Wieselb. Kk.	Wieselburger SZ / Austria	-	F_8

Table 3.1 continued

European *major*			
Canner Express	Nunhems Zaden / Netherlands	-	F_{12}
Con Amore	Nickerson Zwaan / Netherlands	1955	F_{12}
Dreifachweiß	UK	1905	F_7
Hylon	Gillet / UK	1980	F_7
Imperial Green W.	Zwaan / Netherlands	1930	F_7
Minica	Nickerson Zwaan / Netherlands	1973	F_{13}
L_1*	Germany	Minica x Canner Express	F_7
L_2*	Germany	Minica x Canner Express	F_7
Winter bean			
Clipper	PBI/ UK	Minica x PBI lines/1996	F_2
Côte d'Or	France	-	F_8
Hiverna	Littmann, Kiel / Germany	1986	F_8
Punch	PBI/ UK	1988	F_2
Striker	PBI/ UK	1994	F_2
Target	PBI/ UK	Minica x PBI lines/1996	F_2
Webo	Littman, Kiel / Germany	1997	F_7
South Europe			
Aquadulce	Spain	-	F_6
Enantia	Italy	-	F_3
Gemini	Italy	-	F_3
Peleponnes *	Greece / ICARDA	-	F_{12}

Table 3.1 continued

Pietranera	Univ. Palermo/ Italy	-	F_{10}
Polycarpe	FCPI Larisa / Greece	-	F_3
Sikania	Univ. Catania / Italy	-	F_3
Sikelia	Univ. Catania / Italy	-	F_3
North Africa			
F402/4	Egypt	-	F_4
Giza 3	Egypt	1979	F_{11}
Giza 402	Egypt	1979	F_4
Giza 461	Egypt	1990	F_4
Hudeiba	Sudan	Selection from Rebaya /1969	F_4
Morocco	Morocco / ICARDA	-	F_6
Rebaya	Egypt	-	F_8
Zeidab Local	Sudan	Selection from Zeidab population	F_4
Asia			
Cixi Dabeican	China	-	F_2
Kawachi-green	Sakatano-tane / Japan	-	F_2
Mairudo-green	Hokuetsu-nouji / Japan	-	F_2
Maya-Asia	Lang Li-juan Zhejiang / China	-	F_2
Otafuku	Sakatano-tane / Japan	-	F_2
Pinghu Zaojiazhong	Lang Li-juan Zhejiang / China	-	F_2
Pingyang Zhaodozi	Lang Li-juan Zhejiang / China	-	F_2
Shangya Tian	China	-	F_2

***** Lines were used to represent the European gene pool in the hybrid performance study

4 Results and discussion

4.1 Production and scoring of the AFLP markers (I, II, III)

AFLP markers were chosen because of their advantage over other markers available for this type of research. A successful AFLP assay of a genotype results in amplification products that are separated by gel electrophoreses, their presence is detected as a banding pattern seen on X-ray films as performed here in this study (Fig. 4.1). Scoring of the achieved banding pattern was performed visually, resulting in a 1/0 matrix indicating presence and absence of a band, respectively. Two phenomena were observed here, namely: the presence of monomorphic fragments and bands with markedly less intensity to be scored as present and at the same time not convincingly absent, these bands were then scored as (9) and were interpreted as missing data points in the data matrix analysis. Monomorphic fragments that amounted to 9.5% of the total number of fragments were excluded following the recommendation of Link et al. (1995). Preliminary studies performed here have shown that the inclusion of missing data points have no major effect on the reliability of the estimated genetic similarity values, even more, results have shown that fragments harboring one or more doubtful bands carry valuable information almost equivalent to the information provided by fragments free of doubtful bands.

4.2 Consistency of the AFLP banding pattern (I, II)

Consistency of the banding pattern of the AFLP markers was studied to examine the amount of error and its different sources. This was conducted by examining differences in the banding patterns of duplicate assays (repeated lanes of one pre-amplification reaction on one gel) and replicate assays (lanes of two individuals of the same inbred line run on different gels). Results have shown that error for duplicate assays were negligible, whereas errors for replicate assays were markedly higher. It is thus concluded that scoring across gels should be backed by replicates and several appropriate check entries.

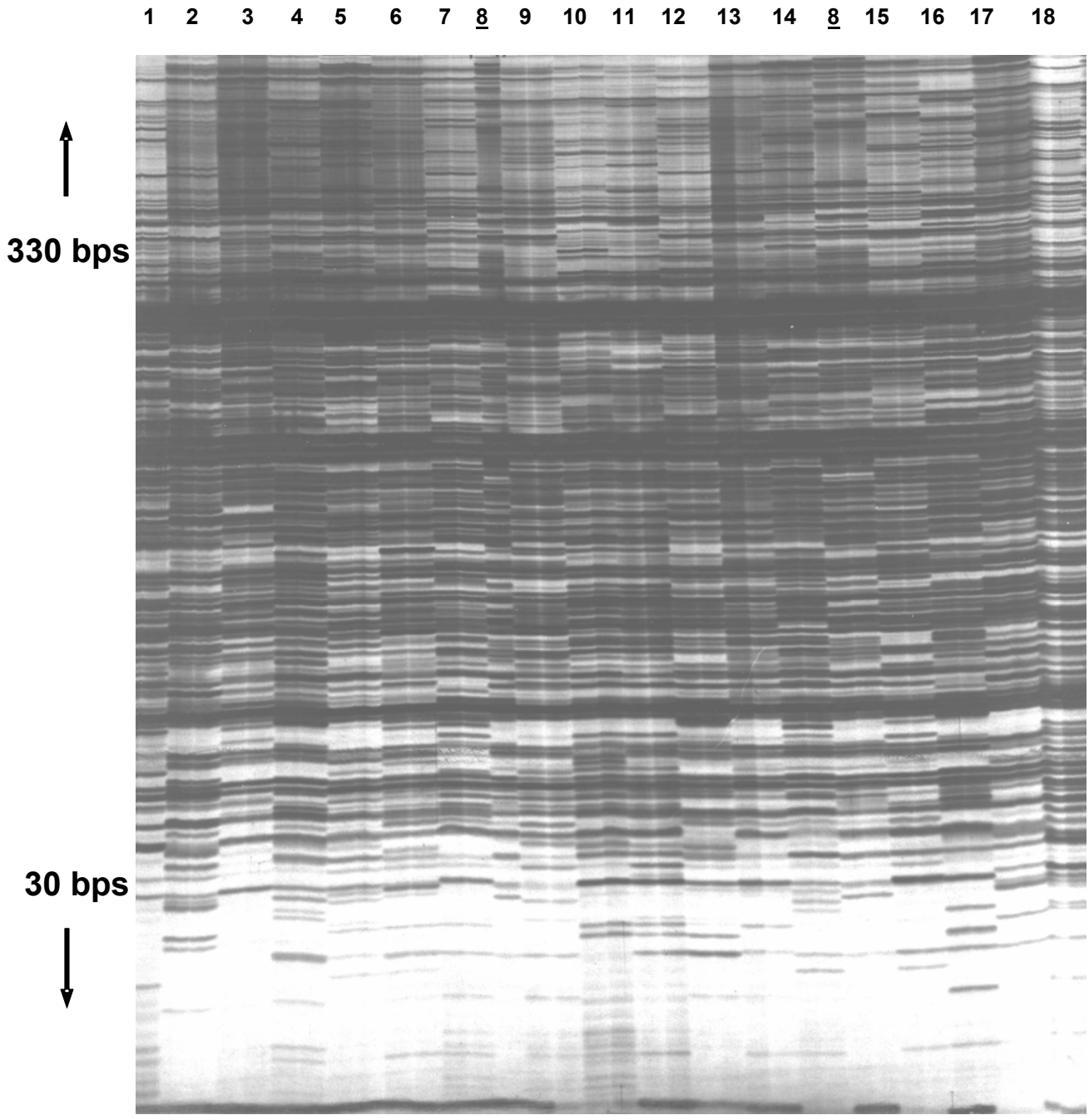


Fig. 4.1 The banding pattern of 18 faba bean lines using the primer combination E-ACA/M-CAA. Each line is represented by two lanes on the gel. The line number 8 (PeI) is an internal check repeated more than once on this gel to facilitate scoring

4.3 The appropriate genetic similarity coefficient (II)

The appropriate genetic similarity coefficient for marker data has often been debated (e.g. Piepho and Laidig 1997; Robinson and Harris 1999). Jaccard's coefficient of similarity (Jaccard 1908) was employed here after considering the type of markers (dominant markers) resulting from AFLP analysis, as explained by Engqvist and Becker (1994) and Link et al. (1995).

A Priori grouping of the 79 inbred lines according to their geographic origin and to their botanical group (cf. Table 1) resulted in six different groups namely: Asian, EU *minor*, EU *major*, South European, winter bean and North African germplasm groups. Results on the genetic similarity between and within groups (Fig. 4.2) showed that the EU *major*, EU *minor* and the Asian groups expressed a higher within group mean genetic similarity as compared to the other three groups. A higher genetic similarity among the EU *minor*, EU *major* and winter bean germplasm was also shown in comparison to genetic similarity of the other groups. These results are in agreement with the available information on the history and cultivation of faba bean in the studied regions.

4.4 Clustering and Principal Coordinate analysis (I, II, III)

Clustering of lines using the unweighted pair-group method with arithmetic means (UPGMA) method was applied to the genetic similarity matrix and the Principal Coordinate Analysis (PCoA) method for grouping of lines was performed. Applying both methods was recommended to extract the maximum amount of information from the matrix data (Messmer et al., 1992). Clustering was useful in detecting relationships among lines, while PCoA allowed a view on the relationships between groups. Failure of the PCoA to describe relationships between neighboring lines is attributed to the distortion of distances between lines due to the small proportion (<25%) of the total variation explained by the first three principal coordinates as indicated by Melchinger (1993). The amount of variation explained by the three principal coordinates shown in this study was below 25%. Reducing the number of entries tested preserving the proportion of entries per group (Fig. 4.3) showed a proportional increase in the amount of variation explained by the first three principal coordinates. This in turn showed the effect of the large number of entries and their lack of grouping on the amount of variation explained.

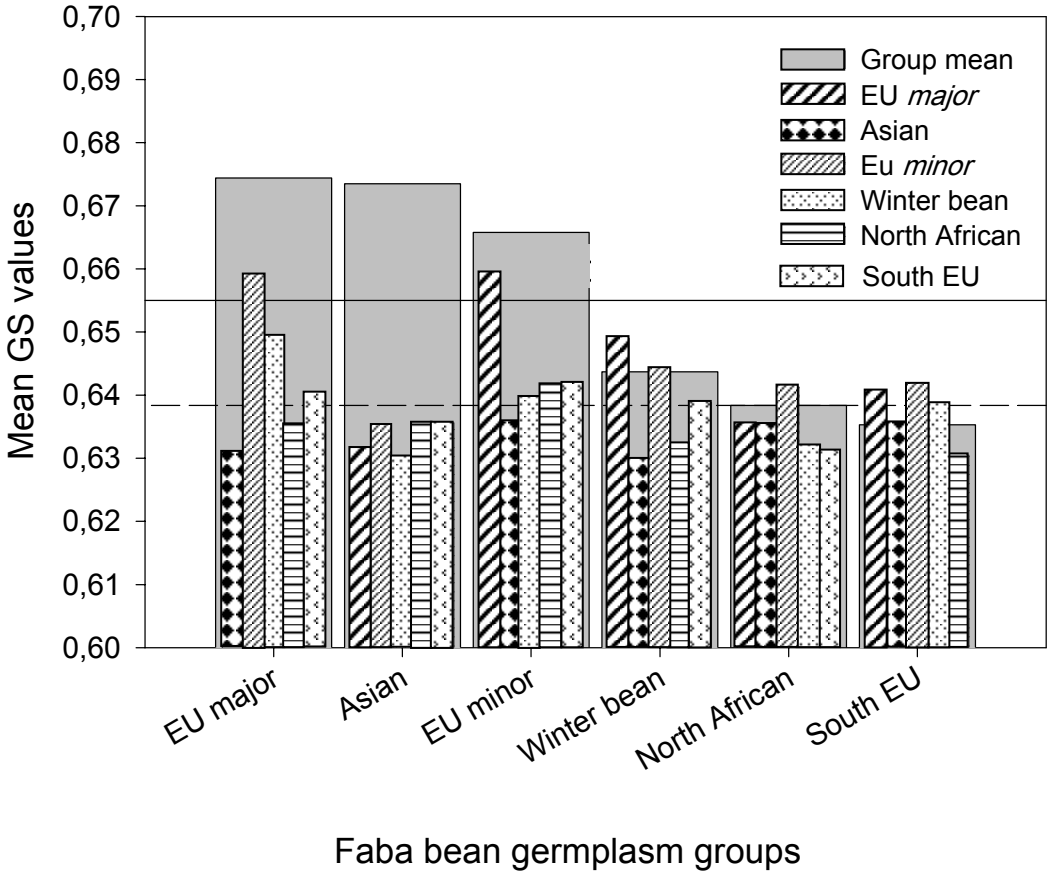


Fig. 4.2 Mean genetic similarity (Jaccard) within and between germplasm groups (The dotted line represents the mean of GS values between groups and the solid line represents that for within groups)

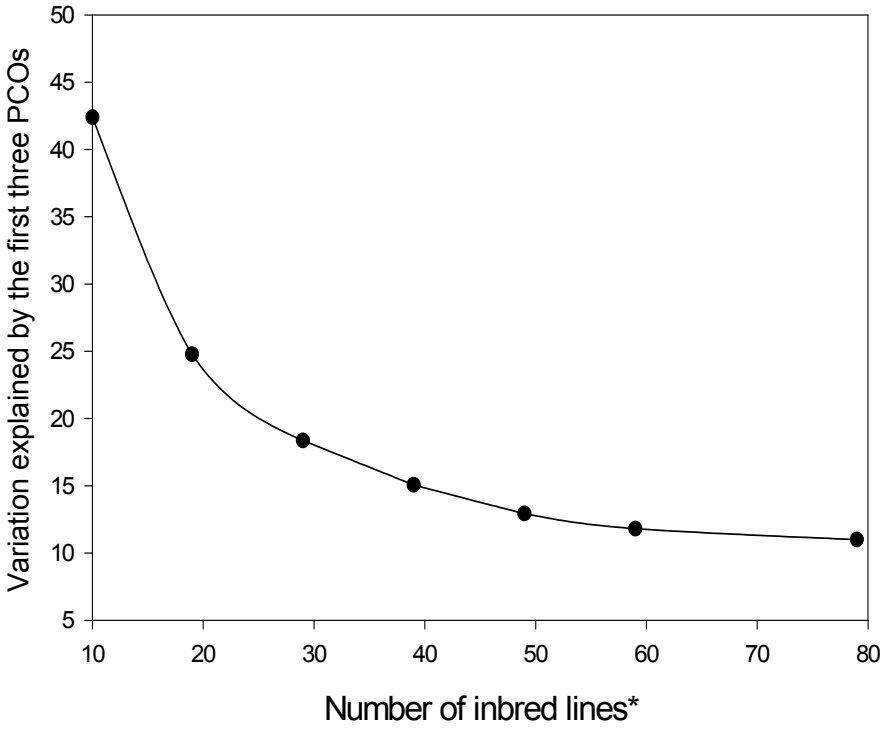


Fig. 4.3 Effect of reducing the number of inbred lines included in the PCoA on the amount of variation explained by the first three coordinates
* Reduction in number of lines was performed preserving the proportion of entries per group (c*eteris paribus*)

Tivang et al. (1994) illustrated that increasing the number of polymorphic bands provides more precise estimates of genetic relationships and reduces the variance caused by over- or under-sampling certain regions of the genome. Nevertheless, expenses in term of resources and time have to be considered. Correlation between the genetic similarity estimates for 18 inbred lines tested using 477 polymorphic fragments and those obtained from 1202 polymorphic fragments amounted to r = 0.705, (*P* = 0.01). Different resampling methods were used to investigate the stability of the classification of the inbred lines, in relation to the number of AFLP markers employed. Both bootstrapping (Felsenstein 1985) and disjoint subsampling have shown that about 200 markers (this is equivalent to four primer combinations with an average of 50 polymorphic markers / primer combination) were good enough for a stable classification.

4.5 Applications of AFLP markers in faba bean breeding

4.5.1 Genetic diversity studies (I, II, III)

The level of genetic diversity between inbred lines as indicated by the genetic similarities (GS) ranged between GS = 0.53 and GS = 0.88. Figure (4.4) shows the amount of variation in the genetic similarity when studying the pairwise comparisons of the 79 inbred lines, the amount sampled by studying only the 18 European lines and the further much smaller variation shown for pairwise comparisons of the parental lines of the 62 crosses tested in the field. Interesting is that the lowest GS value was shown between two inbred lines (Columbo and Blaze) belonging to the European *minor* gene pool. No clear grouping was indicated by the PCoA except for inbred lines from the Asian gene pool. Results of grouping of inbred lines according to their geographic origin and seed size were in agreement with available information on the history of spread and cultivation of faba bean in the studied regions.

4.5.2 Prediction of hybrid performance and heterosis (III)

The relation between genetic similarities based on AFLP markers and of 18 European faba bean lines and their hybrid performance and heterosis was investigated. Parental lines, 62 F1-hybrids and their F2-progenies were evaluated in field trials in four environments in Germany for their seed yield, 1000-seed weight and plant height. Correlation coefficients between the genetic similarity estimates or

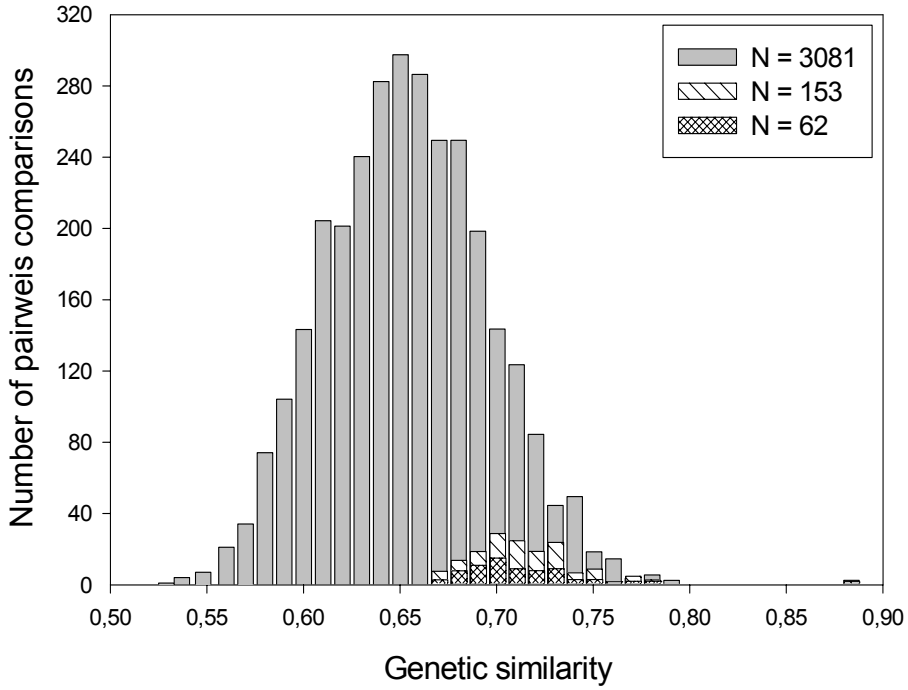


Fig. 4.4 Amount of variation between pairwise values of genetic similarity (Jaccard) when studying the 79 inbred lines (N = 3081 pairs), or only the 18 European lines (N = 153) or only those pairs of these 18 parents leading to the 62 crosses (N = 62) tested for heterosis

specific genetic similarity with either heterosis or F1-hybrid performance for all studied traits were too small to be of predictive value.

Results showed that: (i) the relationships among the faba bean inbred lines under study are in agreement with available pedigree information, indicating the usefulness of marker studies only in cases where pedigree information is doubtful or lacking. (ii) AFLP-based genetic similarities are not predictive of the performance of hybrids or heterosis within the elite European faba bean gene pool; at least as long as mating of parents which are related by pedigree is excluded.

5 Summary

Amplified fragment length polymorphism (AFLP) markers were used to assess the genetic diversity among 79 inbred lines of recent elite faba bean (*Vicia faba* L.) cultivars of Asian, European (Northern and Southern) and North African origin. Different sources of error arising along the steps of the analysis were investigated. Scoring across gels was found to be a major source of error, illustrating the need for sample replications and several appropriate check entries. Clustering of inbred lines based on Jaccard's similarity coefficient and the Principal Coordinate Analysis showed the Asian group of inbred lines to be the only visible group, other lines showed no marked further grouping. Nevertheless, pedigree relationships were verified. Pattern of genetic similarities between inbred lines, grouped according to their geographic origin corroborate available information on the history of spread and cultivation of faba bean in the studied regions. AFLP markers are thus of great value in allocation of genotypes where doubtful or no pedigree information is available and in organizing germplasm.

As further step, the relation between genetic similarities of carefully selected diverse 18 European faba bean lines and both performance and heterosis of their hybrids was investigated. Parental lines, 62 F1-hybrids and their F2-progenies were evaluated in field trials in four environments in Germany for their seed yield, 1000-seed weight and plant height. Correlation coefficients between the genetic similarity estimates and either heterosis or F1-hybrid performance and those between specific genetic similarity and specific combining ability for the three studied traits were too small to be of predictive value. Results showed that AFLP-based genetic similarities are not predictive of the performance of hybrids or heterosis within the elite European faba bean gene pool, at least as long as mating of parents that are related by pedigree is excluded.

6 Zusammenfassung

Analyse der mit molekularen Markern (AFLP) gemessenen genetischen Diversität und der Heterosis bei der Fababohne (*Vicia faba* L.)

Es wurden „Amplified fragment length polymorphism" (AFLP) Marker benutzt, um die genetische Diversität von 79 aktuellen Elite-Inzuchtlinien der Fababohne (*Vicia faba* L.) aus Asien, Europa (Norden und Süden) und Nordafrika zu studieren. Unterschiedliche Fehlerquellen, die entlang der Analyseschritte auftreten können, wurden untersucht. Das Scoren über Gele hinweg wurde als Hauptfehlerquelle erkannt, was die Notwendigkeit von Wiederholungen und mehreren geeigneten Kontroll-Prüfgliedern illustriert. Die auf Jaccards Ähnlichkeitskoeffizient beruhende Clusterung der Inzuchtlinien sowie die Hauptkoordinatenanalyse zeigte die asiatische Gruppe von Inzuchtlinien als einzige erkennbare Gruppe; die anderen Linien ließen keine weitere nennenswerte Gruppierung erkennen. Dennoch konnten Abstammungsdaten verifiziert werden. Das Muster der genetischen Ähnlichkeiten zwischen den Gruppen von Inzuchtlinien, gruppiert entsprechend ihrer geographischen Herkunft, stimmt mit den verfügbaren Erkenntnissen über die Geschichte der Verbreitung und Kultur der Fababohne in den betroffenen Regionen überein. Die AFLP Marker sind somit sehr wertvoll bei der Zuweisung von Genotypen zu Gruppen, wenn zweifelhafte oder keine Abstammungsdaten vorliegen und somit auch bei der Organisation von Zuchtmaterial. Als weiterer Schritt wurde die Beziehung zwischen umsichtig nach Diversität ausgesuchten 18 europäischen Fababohnen-Linien und der Leistung und Heterosis ihrer Hybriden untersucht. Elternlinien, 62 F1-Hybriden und ihre F2-Hybriden wurden in Feldversuchen in vier Umwelten in Deutschland auf Kornertrag, Tausendkorngewicht und Pflanzenhöhe evaluiert. Die Korrelationskoeffizienten zwischen der genetischen Ähnlichkeit und Heterosis bzw. F1-Hybrid-Leistung und auch zwischen der spezifischen genetischen Ähnlichkeit und der spezifischen Kombinationsfähigkeit der drei untersuchten Merkmale waren zu klein, um für eine Vorhersage tauglich zu sein. Die Ergebnisse zeigen, daß AFLP-basierte genetische Ähnlichkeiten innerhalb des europäischen Elite-Genpools der Fababohne keinen Vorhersagewert für die Leistung von Hybriden oder deren Heterosis haben, zumindest solange Paarungen von verwandten Eltern ausgeschlossen sind.

7 Arabic Summary

دراسة الأ ختلافات الوراثية فى الفول البلدى باستخدام المعلمات الوراثية AFLPs وعلاقتها بتفوق الهجين

- استخدمت تقنية ال AFLP لدراسة الأختلافات الوراثية على مستوى ال DNA بين 79 سلالة نقية من الفول البلدى (*Vicia faba* L.) مستنبطة من أصناف عالمية من أسيا، أوروبا (شمال و جنوب) و شمال أفريقيا.

- تم دراسة مصادر الخطأ و التى يمكن أن تؤثر على دقة النتائج المتحصل عليها و قد اوضحت النتائج ان المصدر الأساسى للا ختلافات يرجع الى صعوبة قراءة النتائج من جيل الى أخر: Scoring across gels

- بناء على نتائج ال Principal Coordinate Analysis and Cluster Analysis أظهرت المجموعة الأسيوية فقط أختلافات وراثية واضحة تفرقها عن باقى المجموعات تحت الدراسة.

- أتفقت النتائج المتحصل عليها فيما يخص العلاقة الوراثية بين المجموعات المختلفة مع تاريخ انتشار و زراعة الفول فى المناطق السابق ذكرها.

- اظهرت النتائج فائدة المعلمات الوراثية AFLPs فى تنظيم السلالات داخل المجموعات المختلفة بما يتطابق مع نسب كل سلالة و بالتالى فانه يمكن الاستغناء عن هذه التقنية فى حالة وجود معلومات كافية و موثوقة عن نسب السلالات تحت الدراسة.

- درست ايضا العلاقة بين التشابه الوراثى لعدد 18 من سلالات الفول ذات الاصل الاوروبى و التى تم انتخابها على اساس كونها ذات اصول وراثية متباعدة و بين قوة الهجن الناتجة (62 هجين) لكل من صفة المحصول، وزن الألف حبة و طول النبات.

- أظهرت النتائج عدم معنوية و ضعف العلاقة بين قوة الهجن و التشابه الوراثى للأباء. كذلك كانت العلاقة بين كل من:
Specific combining ability and Specific genetic similarity

مما يشير الى عدم امكانية التنبؤ بقوة الهجن من خلال معرفة درجة التشابه الوراثى للأباء

تحت الدراسة طالما ان الأباء لا ترتبط مع بعضها البعض بعلاقة نسب.

8 References

Abdelmula, A.A., W. Link, E. von Kitttlitz, and D. Stelling. 1999. Heterosis and inheritance of drought tolerance in faba bean, *Vicia faba* L. Plant Breeding 118:485-490.

Ajmone Marsan, P., P. Castiglioni, F. Fusari, M. Kuiper, and M. Motto. 1998. Genetic diversity and its relationship to hybrid performance in maize as revealed by RFLP and AFLP markers. Theor. Appl. Genet. 96:219-227.

Bond, D.A., J.M. Drayner, J.L. Fyfe, and G. Toynbee-Clarke. 1964. A male sterile bean (*Vicia faba* L.) inherited as a Mendellian recessive character. J. Agri. Sci. 63:229-234.

Bond, D.A. 1966. Yield and components of yield in diallel crosses between inbred lines of winter beans (*Vicia faba* L.). J. Agric. Sci. (Cambridge) 67:325-336.

Bond D.A. 1982. The development and performance of synthetic varieties of *Vicia faba* L. p. 41-51. *In* G. Hawtin and C. Webb (ed.) Faba bean improvement. The Netherlands.

Bond, D.A., S.J. Brown, G.J. Jellis, M. Pope, J.A. Hall and M.H.E. Clarke. 1986. Annual report of Plant Breeding Institute, pp. 42-46 Cambridge, UK.

Bond, D.A., and M.H. Poulsen. 1983. Classification, origin, breeding methods and objectives. p. 77-102 *In* P.D. Hebblethwaite (ed.) The faba bean (*Vicia faba* L.). A basis for improvement. Butterworths, London, England.

Bond, D.A., G.J. Jellis, G.G. Rowland, J. Le Guen, L.D. Robertson, S.A. Khalil, and L. Li Juan. 1994. Present status and future strategy in breeding faba beans (*Vicia faba* L.) for resistance to biotic and abiotic stresses. Euphytica. 73:151-166.

Carrouee, B. 1995. Grain legume in the European Union: Current status and prospects. p. 2. *In* AEP (ed) Second Conf. on Grain Legumes, AEP (ed.) Copenhagen, Denmark.

Cheres, M.T., J.F. Miller, J.M. Crane, and S.J. Knapp. 2000. Genetic distance as a predictor of heterosis and hybrid performance within and between heterotic groups in sunflower. Theor. Appl. Genet. 100:889-894.

Crena, F.J., S.R. Cianzio, A. Rafalski, S. Tingey, and D. Dyer. 1997. Relationship between seed yield heterosis and molecular marker heterozygosity in soybean. Theor. Appl. Genet. 95:460-467.

Cubero, J. 1974. On the evolution of *Vicia faba* L. Theor. Appl. Genet. 45:47-51.

Ebmeyer, E. 1988. Heterosis and genetic variances and their implications for breeding improved varieties of spring beans (*Vicia faba* L.). Plant Breeding 101:200-207.

Ebmeyer, E., D. Stelling. 1994. Genetic structure of three open pollinated faba bean varieties, *Vicia faba* L. Plant Breeding 112:17-23.

Engqvist, G., and H.C. Becker. 1994. Genetic diversity for allozymes, RFLPs and RAPDs in resynthezed rape. Biometrics in plant breeding: application of molecular markers. *In* The ninth meeting of the EUCARPIA section biometrics in plant breeding. 6-8 July 1994. Wageningen.

FAO, 2001. Food and Agriculture Organization. Production year book, Vol. 55.Rome, Italy.

Felsenstein, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39:783-791.

Frei, O.M., C.W. Stuber, and M.M. Goodman. 1986. Uses of allozymes as genetic markers for predicting performance in maize single cross hybrids. Crop Sci. 26:37-42.

Hanelt, P. 1972. Die infraspezifische Variabilität von *Vicia faba* L. und ihre Gliederung. Die Kulturpflanze 20:75-128.

Hawtin, G.C. 1982. The development and performance of synthetic varieties of *Vicia faba* L. p. 41-51. *In* G. Hawtin and C. Webb (ed.) Faba bean improvement. The Netherlands.

Hulot, J.F. 1999. Economic context of protein crops in the European Union. Grain Legumes 26:12-13.

Jaccard, P. 1908. Nouvelles researches sur la distribution florale. Bull. Soc. Vaud. Sci. Nat. 44:223-270.

Kittlitz, E.v. 1986. Some observations in reciprocal crosses between *Vicia faba* major. Biol. Zentrabl. 105:147-153.

Lacampagne J-P. 2001. Grain legume area recover in 2001. Grain legumes 34 (4): 24.

Lanza, L.L.B., Jr.C.L. de Souza, L.M.M. Ottoboni, M.L.C. Vieira, and A.P. de Souza. 1997. Genetic distance of inbred lines and prediction of maize single-cross performance using RAPD markers. Theor. Appl. Genet. 94:1023-1030.

Lawes D.A., D.A. Bond and M.H. Poulsen 1983. Classification, origin, breeding methods and objectives. p. 23-67. *In* P.D. Hebblethwaite (ed.) The faba bean (*Vicia faba* L.). A basis for improvement. Butterworths, London, England.

Link, W. 1990. Autofertility and rate of cross-fertilization: crucial characters for breeding synthetic varieties in faba beans (*Vicia faba* L.). Theor. Appl. Genet. 79: 713-717.

Link, W., and W. Ederer. 1993. The concept of varietal ability of partially allogamous crops. Plant Breeding 110:1-8.

Link, W., W. Ederer and E. von Kittlitz. 1994a. Zuchtmethodische Entwicklungen-Nutzung von Heterosis bei Fababohnen. Vortr. Pflanzenzüchtung 30:201-230.

Link, W., W. Ederer, P. Metz, H. Buiel and A.E. Melchinger. 1994b. Genotypic and environmental variation for degree of cross-fertilization in faba bean. Crop Sci. 34:960-964.

Link, W., C. Dixkens, M. Singh, M. Schwall, and A.E. Melchinger. 1995. Genetic diversity in European and Mediterranean faba bean germplasm revealed by RAPD markers. Theor. Appl. Genet. 90:27-32.

Link, W., W. Ederer, R.K. Gumber and A.E. Melchinger 1997. Detection and characterization of two new CMS systems in faba bean (*Vicia faba* L.) Plant breeding 116:158-162.

Martsch, R., J. Vaupel, C.C. Schön, and W. Link. 2001. Faba bean hybrid breeding: Current results with CMS199 p. 232-233. *In* AEP (ed.) Fourth Conf. on Grain Legumes, Cracow, Poland.

Melchinger, A.E. 1993. Use of RFLP markers for analysis of genetic relationships among breeding materials and prediction of hybrid performance. p. 621-628. *In* D.R. Buxton (ed.) First Int. Crop Sci. Congr. Crop Science society of America, Madison, Wis, USA.

Messmer, M., A.E. Melchinger, J. Boppenmaier, R.G. Herrmann, and E. Brunklaus-Jung. 1992. RFLP analyses of early-maturing European maize germplasm: I. Genetic diversity among flint and dent inbreds. Theor. Appl. Genet. 83:1003-1012.

Muratova, V. 1931. Common beans, *Vicia faba* L. Suppl. Bull. Appl. Bot. Genet. Pl. Br., Leningrad. 298 pp.

Piepho, H., and F. Laidig. 1997. A review of methods for cluster analysis of marker data. UPOV document BMT/47/7, 13 p. UPOV Genf.

Robinson, J., and S.A. Harris. 1999. Which DNA Marker for Which Purpose? *In* E.M. Gillet (ed.) Final compendium of the research project development, optimisation and validation of molecular tools for assessment of biodiversity in forest trees in the European Union DGXII Biotechnology FW IV Research Programme Molecular Tools for Biodiversity.

Schön, C.C. 1997. Züchtungsfortschritt bei Ackerbohnen. Vortr. Pflanzenzüchtung 39:122-125.

Schmidtke, K., R. Rauber. 2000. Grain legumes and nitrogen cycling in organic crop systems. Grain legumes 30:16-17.

Struan, S. 2002. EU shortfall plant protein. Grain Legumes 26:12-13.

Stelling, D., E. von Kittlitz, E. Ebmeyer, O. Sass, H. Jaiser, and W. Link. 1994. Erfolge und Perspektiven der züchterischen Verbesserung von Körnererbsen und Ackerbohnen in der EU. p131-147. *In* Ber. 45. Arbeitstagung AG Saatzuchtleiter 1994, Gumpenstein, Austria.

Suso, M.J., and M.T. Moreno. 1999. Variation in outcrossing rate and genetic structure on six cultivars of *Vicia faba* L. as affected by geographic location and year. Plant Breeding 118: 347-350.

Tivang, J., J. Nienhuis, and O.S. Smith. 1994. Estimation of sampling variance of molecular marker data using the bootstrap procedure. Theor. Appl. Genet. 89:259-264.

Van de Ven M., W. Powell, G. Ramsay, and R. Waugh. 1990. Restriction fragment length polymorphisms as genetic markers in Vicia. Heredity. 65:329-342.

Vos, P. R. Hogers, M. Bleeker, T. van de Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper, M. Zabeau. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic. Acids Res. 23:4407-4414.

Czech J. Genet. Plant Breed., 37, 2001 (1): 34–40

Genetic Diversity in a Group of Recent Elite Faba Bean Lines

MAHMOUD ZEID[1], CHRIS CAROLINE SCHÖN[2] and WOLFGANG LINK[1]

[1]*Institute of Agronomy and Plant Breeding, Georg-August-University, Göttingen,*
[2]*State Plant Breeding Institute, University of Hohenheim, Stuttgart, Germany*

Abstract: Genetic diversity among faba bean inbred lines descending from a group of elite cultivars belonging to both European and Mediterranean regions was measured based on amplified fragment length polymorphisms. Genetic similarities were calculated from these data using Jaccard's coefficient, a cluster was constructed and a Principal Coordinate Analysis was applied. The reliability of the genetic similarity estimates was investigated. Clustering of genotypes indicated consistency with relationships based on pedigree data, and as yet unknown relationships were detected. Several promising parental combinations for breeding purposes were tentatively proposed. Further studies on the precision of the genetic similarity estimates are in preparation.

Keywords: genetic diversity; AFLP; *Vicia faba*

Faba bean (*Vicia faba* L.) is an old world grain legume with high seed protein content. It is also characterized by a potential for high yield and low input requirements. The total area devoted to faba bean production has decreased over the last decades. A main reason for this decline has been insufficient breeding progress for yield and yield stability. Broadening the genetic base and systematic exploitation of heterosis in faba bean have been suggested as means to overcome these problems (KITTLITZ *et al.* 1993; BOND 1993). Information on the amount and structure of the genetic diversity in elite germplasm is far from being adequate for breeding purposes (LINK *et al.* 1995; POLIGNANO *et al.* 1993). Estimation of genetic relationships is useful in organizing germplasm, especially when selecting parents for line and hybrid breeding. Earlier, many studies devoted to faba bean and based on morphological and biochemical observations as well on geographical data have been dedicated to these objectives (e.g., AMET 1986; BOZZINI & CHIARETTI 1997; DE PACE *et al.* 1987; HANELT 1972; KÄSER & STEINER 1983; POLIGNANO *et al.* 1993; SERRADILLA *et al.* 1993; SABRAH & EL-METAINY 1985; SUSO *et al.* 1993). LINK *et al.* (1995) have employed the random amplified polymorphic DNA (RAPD) technique to study the genetic diversity within and between European minor and major germplasm and Mediterranean germplasm. They were able to group the European large seeded lines as an intermediate group between the Mediterranean and the European small seeded groups. Amplified Fragment Length Polymorphism (AFLP) has emerged lately as an important technique for genetic diversity studies (ZABEAU & VOS 1993). The advantage of the AFLP assay over the other DNA markers is the high number of polyporphisms amplified for a single PCR reaction, consequently increasing the speed of data generation.

The study presented here was conducted to (1) investigate the reliability of the AFLP technique for genetic studies in faba bean, and to (2) analyze genetic diversity among a group of inbred faba bean lines descending from elite cultivars originating from various geographical origins.

MATERIALS AND METHODS

Plant material: A total of 22 *Vicia faba* elite cultivars that were released in the European and the Mediterranean markets (18 European spring cultivars, one European winter bean and three Mediterranean cultivars; Table 1) were selected. All cultivars were inbred in Germany at the Institute of Agronomy and Plant Breeding, Göttingen, and at the State Plant Breeding Institute, Stuttgart-Hohenheim, for two to more than 12 generations. One F_1 hybrid resulting from crossing K25 and 34M was included, hence, the material comprised 23 genotypes.

DNA isolation and AFLP analysis: Approximately 0.2 g of young leaves harvested directly from 15 days old seedlings (one individual per genotype except for the line K25) was ground in liquid nitrogen to obtain a fine powder. DNA was extracted according to DOYLE and DOYLE (1990). AFLP reactions were performed using the GIBCOBRL AFLP System I (Cat. No. 10544) as described in the manufacture's manual.

Data scoring and analysis: A succesful AFLP assay of a genotype results in amplification products separat-

Table 1. Information on the faba bean lines used in the present study

Cultivar	Faba bean cultivar	Country of origin	Year of release	Inbreeding generation
European	Gloria	Austria	1983	6
	Karna	Austria	1994	4
	Styria	Austria	1996	10
	Merkur	Czech republic	1997	2
	Mars	Denmark	1993	2
	Maris Bead	England	1965	2
	Maya	France	1995	2
	Music	France	1995	2
	Herz Freya (H22)	Germany	1935	10
	Kristall (K25)	Germany	1973	>12
	Troy	Germany	1985	9
	Gobo	Germany	1987	6
	L1 M × CE	Germany	–	5
	L2 M × CE	Germany	–	5
	Scirocco	Germany	1992	5
	Hiverna (winter bean)	Germany	1986	7
	Alfred	Netherlands	1983	5
	Victor	Netherlands	1988	8
	Pistache	Netherlands	1990	5
Mediter-ranean	Giza 3	Egypt	~1979	9
	Morocco (34M)	Morocco	–	>8
	Peleponnes (Pel)	Greece	–	>11

ed by gel electrophoresis along the genotype's lane, located at positions on the gel that strictly correspond to the products' length. Any position on the gel which, when scored across genotypes, contains at least once a DNA amplification product will be called herein "fragment". Whenever a given fragment contains a scoreable amplification product, this will be termed "band". Hence, the number of fragments per genotype is a constant number for the experiment, whereas the number of bands per genotype varies across the genotypes. Monomorphic fragments contain an amplification product, a band, in each inbred line, polymorphic fragments don't. Bands were scored as (1), and absent bands were scores as (0). In cases where a band's intensity was markedly less strong than normal but not convincingly absent, a score of (9) was given. It was considered as missing data point and termed "doubtful band". Fragment scoring was performed visually from X-ray films and the resulting 1/0/9 matrix was then used to calculate similarity coefficients. Genetic similarity values (GS values) according to JACCARD (1908) were used throughout, applying the software NTSYS-pc version 1.8 (ROHLF 1993).

Monomorphic fragments were excluded from the analysis. The resulting similarity matrix was used to construct a cluster based on the unweighted pair-group method with arithmetic means (UPGMA); additionally, a Principal Coordinate Analysis (PCoA) was performed. The heterozygous F_1 genotype was excluded from these approaches. To test the precision of the generated GS values, the total data set of 526 polymorphic fragments was randomly subdivided into disjoint data subsamples. Thus five subsamples comprising 100 fragments each, i.e. each representing 19% of the data, three subsamples comprising 150 fragments each (28.5% of the data) and two subsamples, each harbouring 50% of the total data set were taken from the total data set. For all subsamples of AFLP fragments, estimates of all pairwise genetic similarities among the 22 inbred lines were determined, and for each subsample, the coefficient of variation among its GS values was calculated. Furthermore, the corresponding GS values of these subsamples were compared. Comparing only subsamples of equal size, ten comparisons among the five subsamples of 100 fragments were possible. Three were possible among the three subsamples of size 150, and one comparison between the two subsamples each comprising half of the fragments. Spearman rank correlations were calculated to quantify the similarity of GS values of pairs of subsamples.

An estimation of the error variance for the genetic similarity values was performed. For this purpose, for two genotypes two replicates were used. First, two individuals of the inbred line K25 entered into the AFLP assays and scoring procedure. Second, from the scored banding pattern of the genotypes K25 and 34M, the expected scores of the hybrid F_1(K25 × 34M) were generated by combining the parental scores, assuming dominant inheritance of the AFLP bands. Thus, a hypothetical genotype F_1(hyp) was constructed. Since an individual of the hybrid F_1(K25 × 34M) has been included in the material, a second case of two replicates was available: F_1(hyp) vs. F_1. The GS of the first K25 individual to all other lines was compared with corresponding GS values of the second K25 individual; the GS of the F_1 individual to all other lines was compared with corresponding GS values of the F_1(hyp) genotype. The GS values among the replicates of the same genotype ($0.968 < GS < 0.986$) were excluded from the analysis. An ANOVA was conducted to estimate the least significant differences when comparing GS values based on scores of non replicated individuals.

RESULTS

Screening 15 *Eco*R I/*Mse* I primer combinations indicated the usefulness of 12 combinations that resulted in a total of 662 fragments, 526 of which were polymorphic (Table 2). With the primer combinations E-ACC/M-CTC, /M-CTG and /M-CTT, the production of scorable fragments was not possible under the employed conditions. Primer combinations varied in the number of polymorphic fragments that were produced and feasible for scoring. Excluding the primer combination E-ACA/M-CAT that resulted in a very low number of scorable fragments, an average of 43.8 polymorphic fragments were produced per primer combination. AFLP fragment sizes ranged from 50–450 base pairs (bp). Scoreable fragments however, were detected between 60–250 bp.

Table 2. Primer combinations (*Eco*R I/*Mse* I) used for the AFLP analyses

Primer combination	Total number of fragments	% of polymorphic fragments
E-ACA/M-CAC	68	77.94
E-ACA/M-CAG	82	85.37
E-ACA/M-CAT	18	27.78
E-ACA/M-CTA	58	81.03
E-ACC/M-CAG	76	85.53
E-ACC/M-CAT	28	42.86
E-AGC/M-CAC	73	90.41
E-AGC/M-CAG	47	89.36
E-AGC/M-CTA	53	77.36
E-AGG/M-CAC	41	78.05
E-AGG/M-CTA	67	91.04
E-AGG/M-CTC	51	62.75

Different approaches to judge the reliability of the AFLP data produced for the studied faba bean lines were used. For the possible 253 GS values, the mean coefficient of variation (CV %) was calculated for a number of subsamples. The mean CV decreased from 13 to 9.48% when employing 19 and 50% of the data respectively (Fig. 1) and reached 8.2% upon considering the whole data set. The effect of the doubtful bands harboured by some fragments on the analysis was also investigated. Doubtful bands amounted to 7.5% of the data, and being

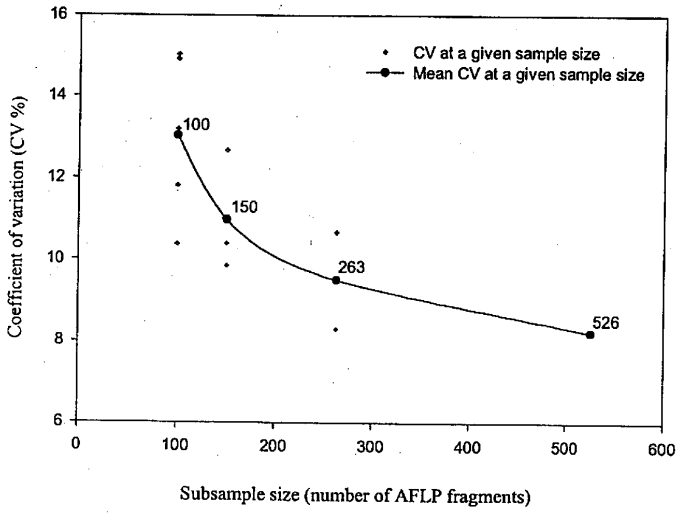


Fig. 1. Coefficient of variation of 253 GS values in disjoint, random subsamples of AFLP fragments

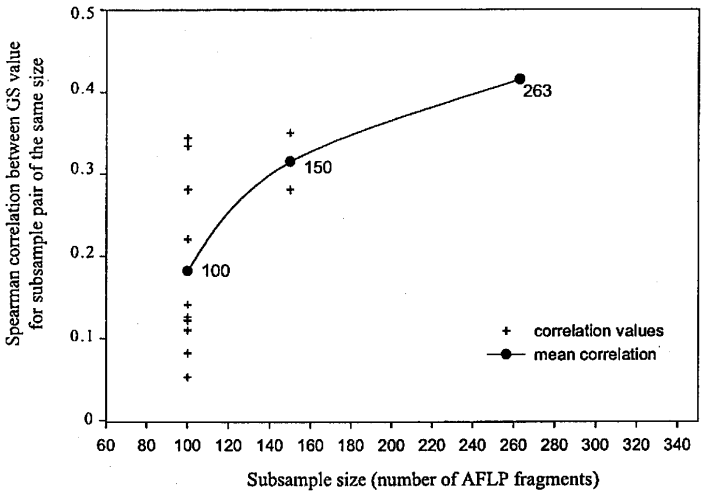


Fig. 2. Correlation of 253 pairwise comparisons (GS values) between any two data subsamples of the same size (for details refer to the text)

distributed across the 526 scored fragments, 297 fragments were scored with one to several doubtful bands. Those 297 fragments were analysed as a disjoint fragment sample and compared to the remaining 229 fragments. On correlating the 253 GS values of both, the 297 and the 229 sample, the Spearman rank correlation coefficient amounted to $r_s = 0.429$. Fig. 2 illustrates the relation between the number of fragments considered in a given subsample and the Spearman rank correlation coefficients for the 253 corresponding GS values in pairs of subsamples. A continuous increase in the mean correlation was shown as a result of increasing the number sampled fragments.

The number of differently scored fragments in the case of the two individuals of genotype K25 was two, the direct similarity between these two individuals was GS = 0.986; in the case of the F_1 a total of eleven fragments were scored to be different, and the direct comparison yielded GS = 0.968. The ANOVA (Table 3) revealed that the standard deviation among the regarded GS values (averages across two replicates) amounted to 10.20% ($[10.56 + 82.23 + 12.18]^{1/2}$) on a scale of 0% < GS < 100%. This variation is markedly inflated due to the inclusion of the F_1 into this consideration. The standard deviation in the total data

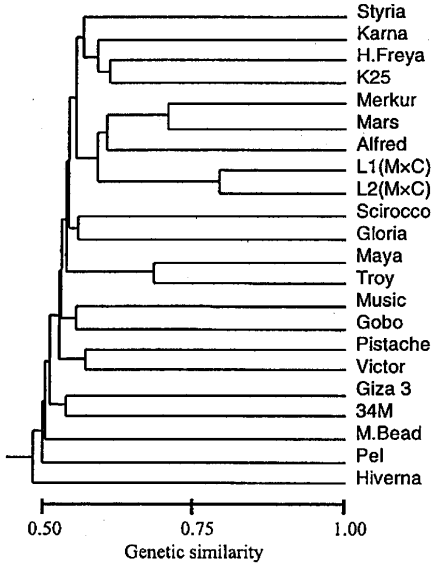


Fig. 3. Grouping of 22 faba bean lines based on 526 AFLP fragments using Jaccard's coefficient of similarity and UPGMA clustering

Table 3. Analysis of variance for two genotypes (G; K25 and F1) in two replicates (R), regarding their GS values to [21 + 2 × 2] = 25 other lines (L); GS values transformed to percent, i.e., 0% < GS < 100%

Source	df	Mean squares	Variance components	*F*
G	1	529.54	10.56	357**
L	24	330.42	82.23	223**
LG	20	25.85	12.18	17**
LRG	46	1.48	1.48	

**significant at 0.01 level of probability

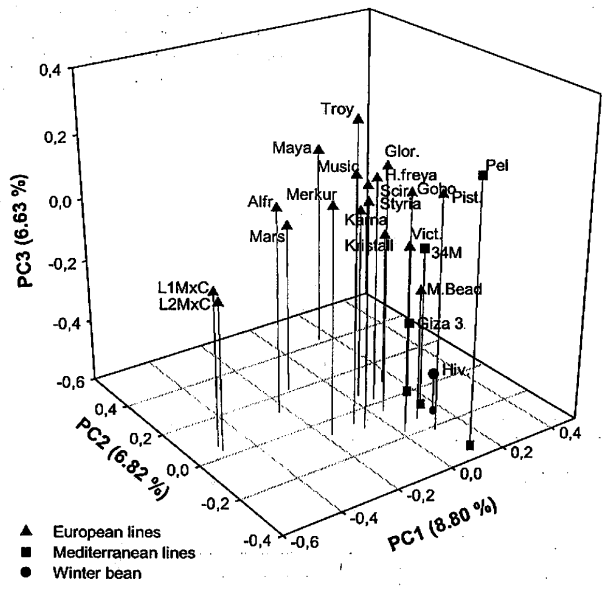


Fig. 4. Associations among 22 faba bean inbred lines revealed by Principal Coordinates Analysis performed on 526 polymorphic AFLP fragments

set, referring to 231 GS values each from one replicate only, excluding the heterozygous F_1, amounted to 4.40%. This may be compared with the error standard deviation of $(1.48)^{1/2}\% = 1.22\%$.

The GS values calculated from the whole data set were used to perform a cluster analysis using the UPGMA method. The resulting dendrogram (Fig. 3) did not show major "ball clusters". Some of the shown clusters were consistent with known pedigree data. The lines L1(M × C) and L2(M × C) showed the highest similarity of GS = 0.789 and clustered together, as expected for two lines originating from the same cross (Minica × Canner Express). The line Maya clustered with Troy (GS = 0.686), a relation that was explained when reviewing the pedigree of both lines: Troy was a parent of Maya. On the other hand, a new relationship was detected, shown by the clustering of Mars and Merkur with a GS = 0.704. Associations among lines revealed by the PcoA based on the GS values are presented in Fig. 4. The three principal coordinates encompassed 22% of the total variation. The winter bean line Hiverna was clearly separated from all other lines on the third principal coordinate (PC3) and a major clustering group comprised almost all of the European lines, while the Mediterranean lines were located along the border of the European cluster.

DISCUSSION

The AFLP technique has demonstrated its high potential for the generation of data in faba bean compared to the RAPD technique employed by LINK *et al.* (1995) who used similar plant material. While results indicated a very high similarity between the two K25 individuals and the hypothetical F_1 compared to the true F_1 individual, these similarities were not 100%. LINK *et al.* (1995) and DE RIEK *et al.* (1999) have pointed out the various sources of error that can influence the data along the different steps of the technique, starting from DNA extraction and up to band scoring. LAMBOY (1994) clearly demonstrated that both, consistent exclusion and consistent inclusion of error prone bands in a dominant marker system can cause marked bias when estimating genetic similarity coefficients. In this study we tried to overcome this problem by neither scoring doubtful bands as 1 nor as 0, but by taking them as missing data points; this feature is available in the NTSYS program. To study the effect of this approach on the GS values, we compared the results of 297 fragments comprising at least one doubtful band to those of the remaining 229 fragments (free of doubtful bands). The rank correlation between the corresponding GS values amounted to $r = 0.429$, which is comparable to the value of $r = 0.415$ (Fig. 2) obtained when comparing two disjoint, random subsamples of similar (~ 50% of the data set) size. This indicates that doubtful bands had no major effect on the reliability of the GS values and that fragments harbouring one or more doubtful bands carry valuable information almost equivalent to the information provided by those fragments free of doubtful bands. The ANOVA (Table 3) indicated that the least significant difference of a single GS value amounted to $LSD_{0.05} = 0.035$. This value is valid for comparisons among the orig-

inal GS values that entered this ANOVA; but it is considered to be valid for the whole data set as well. It considers the error that may be caused by minor genetic differences between individuals within the same inbred genotype, partly by non-dominant inheritance of AFLP bands, and by errors that evolve after DNA extraction, probably mainly when scoring.

SNEATH and SOKAL (1973) raised the important question of what is a sufficient number of markers to obtain a stable classification of genotypes. Here, this question was approached twice, by (1) considering the decrease in the variance between the 253 GS values when increasing the fragment subsample size, and by (2) considering the correlations between the GS values in all pairs of data subsamples of same size. As to the first approach, the variance between the GS values is expected to be inflated by error, the more, the smaller the subsample size. The asymptotic decline of the CV in Fig. 1 indicated that already with a number of 263 fragments (i.e. 50% of the total data set), the CV contains only a small residual error variance. As to the results given in Fig. 2 it seems doubtful whether with larger number of AFLP fragments, a correlation near to unity can be attained at all. Our full data set has shown to be very informative, nevertheless, increasing the number of AFLP fragments is advisable. Still this will not eliminate errors. Here some errors obviously occurred, leading to a standard error term of s.e. = 1.22% for the GS values.

Irrespective of these restrictions, the data clearly showed that (1) genotypes such as winter types and exotic genotypes were identified as being distant and that (2) there does not seem to be a very marked structure within the elite, European gene pool as represented with the given genotypes. It might be promising for breeders to establish such structures. A direct, though tentative result may be the identification of a promising parental combination for hybrid or line breeding, such as Scirocco × Gobo, that yielded a significantly lower similarity than, e.g., Scirocco × Mars (0.486 < 0.603). Further combinations of elite European lines with low GS values (0.477 < GS < 0.493) are Maris Bead × Victor, Troy × Victor, Troy × Merkur, and Styria × Victor. These proposals have of course to be confirmed by further studies, especially by running adequate field trials.

References

AMET T.M. (1986): Geographical patterns of allozyme variation in a germplasm collection of faba bean (*Vicia faba* L.). Fabis Newsl., **16**: 5–12.

BOND D.A. (1993): Motor of *Vicia faba* evolution and breeding. Grain Legumes, **1**: 15–16.

BOZZINI A., CHIARETTI D. (1997): The genetic improvement of the Mediterranean faba bean (*Vicia faba* L.). I. Identification and characterization of parental lines. J. Genet. & Breed., **51**: 149–160.

DE PACE C., GENG S., FILIPPETTI A., SCARASCIA MUGNOZZA G.T. (1987) Phenotypic distances, full-sibs grouping, and *Vicia faba* breeding. Genet. Agr., **41**: 385–396.

DE RIEK J., DENDAUW J., MERTENS M., DE LOOSE M., HEURSEL J., VAN BLOCKSTAELE E. (1999): Theor. Appl. Genet., **99**: 1155–1165.

DOYLE J.J., DOYLE J.L. (1990): Isolation of plant DNA from fresh tissue. Focus, **12**: 13–15.

HANELT P. (1972): Die infraspezifische Variabilität von *Vicia faba* L. und ihre Gliederung. Kulturpflanze, **20**: 75–128.

JACCARD P. (1908): Nouvelles researches sur la distribution florale. Bull. Soc. Vaudoise Sci. Nat., **44**: 223–270.

KÄSER H.R., STEINER A.M. (1983): Subspecific classification of *Vicia* faba L. by protein and isozyme patterns. Fabis Newsl., **7**: 19–20.

KITTLITZ E. VON., IBRAHIM K.I., RUCKENBAUER P., ROBERTSON L.D. (1993): Analysis and use of interpool-crosses (Mediterranean × Central European) in faba beans (*Vicia faba* L.). 1. Performance of Mediterranean and Central European faba beans in Syria and Germany. Plant Breed., **110**: 307–314.

LAMBOY W. (1994): Computing genetic similarity coefficient from RAPD data: The effects of PCR artefacts. PCR Methods and Applications, **4**: 31–37.

LINK W., C. DIXKENS, M. SINGH, M. SCHWALL, MELCHINGER A.E. (1995): Genetic diversity in European and Mediterranean faba bean germplasm revealed by RAPD markers. Theor. Appl. Genet., **90**: 27–32.

POLIGNANO G.B., UGGENTI P., SCIPPA G. (1993): The pattern of genetic diversity in faba bean collections from Ethiopia and Afghanistan. Genet. Resour. Crop Evol., **40**: 71–75.

Rohlf F. J. (1993): NTSYS-PC. Numerical taxonomy and multivariate analysis system. Exeter Publ., Setauket. N.Y.

Sabrah N. S., El-Metainy A. Y. (1985): Genetic distances between local and exotic cultivars of *Vicia faba* L. based on esterase isozyme variation. Egypt. J. Genet. Cytol., **14**: 301–307.

SERRADILLA J.M., DE MORA T., MORENO M.T. (1993): Geographic dispersion and varietal diversity in *Vicia faba* L. Genet. Resour. Crop Evol., **40**: 143–151.

Sneath P. H., Sokal R. R. (1973): Numerical taxonomy. W. H. Freeman and Co. San Francisco.

SUSO, M.J., MORENO M.T., CUBERO J.I. (1993) Variation among Spanish faba beans cultivars: taxonomic and evolutionary implications. Genet. Resour. Crop Evol., **40**: 105–111.

ZABEAU M., VOS P. (1993): Selective restriction fragment amplification: a general method for DNA fingerprinting. Eur. Patent Office, Publication 0 534 858 A1.

Received for publication on May 30, 2000

Abstrakt

ZEID M., SCHÖN C.C., LINK W. (2001): **Genetická diverzita ve skupině současných elitních linií bobu**. Czech J. Genet. Plant Breed., **37**: 34–40.

Genetická diverzita mezi inbredními liniemi bobu pocházejícími ze skupiny současných odrůd z evropského i ze středozemního regionu byla zjišťována na základě AFLP (amplified fragment length polymorphism). Z těchto dat byly pomocí Jaccardova koeficientu vypočteny genetické podobnosti, byl vytvořen klaster a podroben analýze hlavních komponent. Zkoumala se spolehlivost odhadů genetické podobnosti. Shlukové seskupování genotypů odpovídalo poměrům založeným na rodokmenových údajích a byly objeveny i dosud neznámé vztahy. Bylo navrženo několik slibných rodičovských kombinací pro šlechtění. Další studie k přesnosti odhadů genetické podobnosti jsou připravovány.

Klíčová slova: genetická diverzita; AFLP; *Vicia faba*

Corresponding author:

M.Sc. MAHMOUD ZEID, Institute of Agronomy and Plant Breeding, Georg-August-University, Von Siebold Strasse 8, 37075 Göttingen, Germany, tel.: + 49 551 39 21 53, fax: + 49 551 39 46 01, e-mail: mkzeid@yahoo.com

Genetic diversity in recent elite faba bean lines using AFLP markers*

Key words Genetic diversity, *Vicia faba*, Bootstrap, Similarity coefficients, AFLP

Abstract Amplified fragment length polymorphism (AFLP) markers were used to study the genetic diversity among a large set (N = 79) of inbred lines of recent elite faba bean (*Vicia faba* L.) cultivars with Asian, European (Northern and Southern) and North African origin. The inbred lines were analyzed using eight selected AFLP primer combinations that produced 477 polymorphic fragments. Errors when scoring repeated lanes of one pre-amplification reaction on one gel were negligible, whereas errors when scoring lanes of two individuals of the same inbred line run on different gels were markedly higher. Scoring across gels should be backed by replicates and several appropriate check entries. Based on clustering with Jaccard's similarity coefficient and Principal Coordinate Analysis, only the Asian lines were distinct as group, the other lines showed no marked further grouping. Nevertheless, several known pedigree relationships were verified. *A priori* grouping of inbred lines (geographic origin and seed size) and AFLP data corroborate available information on the history of spread and cultivation of faba bean in the studied regions. Based on the diversity observed, studies especially concerning the relationship between genetic similarity based on AFLP markers and hybrid performance within the European elite germplasm have been launched.

---

*A modified, shortened version of this chapter is in preparation as manuscript for Theoretical and Applied Genetics with the following authors:

Mahmoud Zeid[1], Chris-Carolin Schön[2] and Wolfgang Link[1]

[1]Institute of Agronomy and Plant Breeding, Georg-August-University, 37075 Göttingen, Germany; [2]State Plant Breeding Institute, University of Hohenheim, 70593 Stuttgart, Germany

Introduction

Faba bean (*Vicia faba* L.) is a genetically isolated species with no successful crossing with any other *Vicia* species. Muratova as early as 1931 recognized distinct groups within faba bean based mainly on seed size, ranging from small seeded *minor* beans (0.2-0.8 g per seed) to medium seeded *equina* beans and the large seeded *major* beans (single seed weight of up to 2.6 g) that have become known later as botanical groups. Other taxonomic methods to study the evolution and genetic diversity in the species followed and were mainly based on morphology (Muratova 1931), geographic origin (Cubero 1974) and karyotype (Yamamoto 1973). The genetic variation currently available is not fully exploited and sterility barriers do not exist between subspecies or between botanical groups (Lawes et al. 1983).

Controlled hybridization between selected faba bean genotypes is the basis of improvement programs practiced by faba bean breeders. Crosses combining winter hardiness of *equina* winter types and absence of tannin originating from vegetable *major* spring beans (Bond 1976) is one of many examples of the intense recombinations between botanical groups that have been practiced in faba bean. Most breeding programs are based mainly on elite germplasm, including locally adapted and unadapted cultivars. Reliable information on the genetic relationships between such germplasm and the amount of genetic diversity the germplasm comprises is crucial in planning breeding programs. Such information is difficult to achieve if based only on morphological and agronomical characters that are affected by environment and growth stage of the plant.

Molecular markers have been repeatedly applied to study the genetic diversity in faba bean. Käser and Steiner (1983) employing protein and isozyme markers reported no clear grouping of 22 cultivars and 49 landraces from a world wide collection except for the 17 German cultivars studied. Link et al. (1995), applied the random amplified polymorphic DNA (RAPD) technique on 28 inbred lines to study the genetic diversity within and between European *minor* and *major* germplasm and Mediterranean germplasm. They grouped the European large seeded lines as an intermediate group between the Mediterranean and the European small seeded groups. The amplified fragment length polymorphism (AFLP) analysis was used to analyze in detail the genetic diversity among

22 faba bean inbred lines derived from elite cultivars (Zeid et al. 2001). Several approaches were used to analyze the marker-based genetic diversity of breeding material. If *a priori* defined groups exist, each entry may be characterized by its similarity to its own group compared to the other group(s) (e.g., Lübberstedt et al., 2000). Very often, clustering and distance-based ordination methods like Principal Coordinate Analysis (PCoA) were applied (e.g., Becker et al. 1995; Lübberstedt et al. 2000). A direct quantification of the degree of differentiation of a gene pool into sub-pools was proposed by Finkeldey (1994) and Excoffier (1994), focusing on codominant markers and allele frequencies.

Despite of irrefutable advantages, molecular data are not free of errors. The reproducibility of the AFLP profile in sugar beet was tested by Jones et al. (1997) across a network of six European laboratories. Using two *Mse*I/*Pst*I primer combinations that yielded 16 polymorphic fragments, only a single difference was observed (absence of a band) in the produced AFLP profile. Apart from the reliability of the marker profile, its consistency, i.e. whether an accession and a primer combination give the exact banding profile when the reaction is repeated under constant laboratory conditions is another important aspect. Running duplicates of the tested genotypes has been reported in some studies as an indication for reproducibility (Hansen et al. 1999; Winfield et al. 1998). However, it is rarely mentioned whether duplicates were re-runs of the same PCR products, re-amplifications or a replicate of a second individual within the same genotype. In most cases it was only a replicated lane in the same gel. Furthermore, the minimum number of AFLP markers required to achieve a given level of precision estimated by resampling of subsets of markers through the Bootstrap (Efron and Tibshirani 1986) and/or the Jacknife approach (Sokal and Rohlf 1981) is of great interest.

The objectives of this work were (1) to study the structure of the genetic diversity within a large sample of 79 selected, elite cultivars representing the actual faba bean material used for breeding in Asia, Europe and North Africa based on AFLP and (2) to analyze the precision of the results obtained.

Materials and methods

Plant material

This investigation was based on a large sample of (N = 79) of *Vicia faba* L. inbred lines derived from elite cultivars that were released in various world markets (eight Asian cultivars, 40 European spring *minor* cultivars, eight European spring *major* cultivars, eight South European cultivars, seven European winter bean cultivars and eight North African cultivars; Table 1). This collection comprises actual cultivars that are available on the market today, e.g. Giza 402 (Egypt), Scirocco (Germany) & Styria (Austria) and ancestral cultivars, e.g. Cixi Dabeican (China), Herz Freya (Germany) & Minica (Netherlands) that have dominated breeding and production for a considerable period of time and are now represented in many of the pedigrees of recent germplasm (Stelling et al. 1994). Inbred lines were developed in Germany at the Institute of Agronomy and Plant Breeding, Göttingen, and at the State Plant Breeding Institute, Stuttgart-Hohenheim, for one to more than 12 generations of selfing (Table 1), employing the single-seed descent method.

DNA isolation and AFLP analysis

One individual was used to represent one inbred line (exceptions are explained below). Approximately 0.20 g of young leaves harvested from 15 days old seedlings was directly ground in liquid nitrogen. DNA was isolated using CTAB extraction buffer according to Doyle and Doyle (1990). AFLP reactions were performed according to the procedure described by Vos et al. (1995), using a commercially available kit (AFLP analysis System I, GIBCO BRL, Life Technologies, Inc., Rockville, Md) and following the manufacture's instructions.

Eight out of 14 tested *Eco*RI/*Mse*I primer combinations were employed (Table 2). The *Eco*RI primers were radiolabeled with [γ-33 P]-dATP provided from Amersham Pharmacia Biotech. After selective amplification, the polymerase chain reaction (PCR) products were mixed with an equal volume of loading buffer [98% (v/v) formamide, 10 mM EDTA, 0.05% (v/v) xylene cyanol, and 0.05% (v/v) bromophenol blue] and denatured at 95ºC for 5 minutes and directly placed on ice. Seven µl of the mixture were loaded on an 8% (v/v) polyacrylamide gel (30x40 cm) containing 8.3 M urea that was pre-warmed for 20 minutes. Gels were run with 0.5x TBE electrophoresis buffer [50 mM Tris, 50 mM boric acid, 1mM

Table 1 Ten of these lines (‡) were used with two individuals (replicate assays), one (†) was used as replicate and as duplicate assay (one PCR product analyzed in two neighboring lanes); two were additionally run on every gel as internal checks (§)

Faba bean cultivar	Country of origin	Year of release	Inbreeding generation
<u>European *minor*</u>			
Alfred	Netherlands	1983	F_3
Alpine ‡	UK	1997	F_3
Blaze	UK	1976	F_{10}
Blitz ‡	Canada	1994	F_2
Carola	Austria	1986	F_8
Caspar	Netherlands	1992	F_7
Columbo	Denmark	1996	F_3
Condor	Germany	1990	F_8
Corvette	Denmark	1996	F_3
Erfano	Germany	1988	F_5
Fatima	Canada	1993	F_2
Francks Ackerperle	Germany	-	F_{10}
Franz	Austria	1990	F_4
Geo	Germany	1989	F_6
Gloria	Austria	1994	F_8
Gobo	Germany	1987	F_8

Table 1 (continued)

Faba bean cultivar	Country of origin	Year of release	Inbreeding generation
Hedin	Germany	1986	F_9
Herz Freya	Germany	1935	F_{12}
Jantarnij	Russia	-	F_2
Karna	Austria	1983	F_6
Kornberger Klein	Austria	1956	F_{10}
Kristall (K25) † §	Germany	1973	F_{12}
Limbo	Germany	1998	F_2
Luna	UK	1991	F_2
Maris Bead	UK	1964	F_3
Mars	Denmark	1993	F_3
Maya	France	1995	F_3
Merkur	Czech Republic	1997	F_3
Music	France	1995	F_3
Nadwislanski	Poland	1955	F_7
Nixe	Germany	-	F_{12}
Orletzkij	Russia	-	F_2
Pistache	Netherlands	1990	F_7
Scirocco	Germany	1992	F_7
Styria	Austria	1996	F_{12}
Topas	Germany	1986	F_8
Troy	Germany	1985	F_{11}

Table 1 (continued)

Faba bean cultivar	Country of origin	Year of release	Inbreeding generation
Ukko‡	Finland	1984	F_7
Victor	Netherlands	1988	F_{10}
Wieselb. KK.	Austria	-	F_8
<u>European *major*</u>			
Canner Express	Netherlands	-	F_{12}
Con Amore	Netherlands	1955	F_{12}
Dreifachweiß	UK	1905	F_7
Hylon	UK	1980	F_7
Imperial Green W. ‡	Netherlands	1930	F_7
Minica	Netherlands	1973	F_{13}
L_1*	Germany	-	F_7
L_2*	Germany	-	F_7
<u>Winter bean</u>			
Clipper	UK	1996	F_2
Côte d'Or ‡	France	-	F_8
Hiverna	Germany	1986	F_8
Punch	UK	1988	F_2
Striker	UK	1994	F_2
Target	UK	1996	F_2
Webo	Germany	1997	F_7

Table 1 (continued)

Faba bean cultivar	Country of origin	Year of release	Inbreeding generation
South Europe			
Aquadulce	Spain	-	F_6
Enantia ‡	Italy	-	F_3
Gemini	Italy	-	F_3
Peleponnes (Pel)§	Greece	-	F_{12}
Pietranera	Italy	-	F_{10}
Polycarpe	Greece	-	F_3
Sikania	Italy	-	F_3
Sikelia	Italy	-	F_3
North Africa			
F402/4	Egypt	-	F_4
Giza 3 ‡	Egypt	1979	F_{11}
Giza 402 ‡	Egypt	1979	F_4
Giza 461 ‡	Egypt	1990	F_4
Hudeiba	Sudan	1969	F_4
Morocco (34M)#	Morocco	-	F_6
Rebaya	Egypt	-	F_8
Zeidab Local ‡	Sudan	-	F_4

Table 1 (continued)

Faba bean cultivar	Country of origin	Year of release	Inbreeding generation
Asia			
Cixi Dabeican	China	-	F_2
Kawachi-green	Japan	-	F_2
Mairudo-green	Japan	-	F_2
Maya-Asia	China	-	F_2
Otafuku	Japan	-	F_2
Pinghu Zaojiazhong	China	-	F_2
Pingyang Zhaodozi	China	-	F_2
Shangya Tian	China	-	F_2

(#) A line obtained from ICARDA (BPL228)

(*) L1 and L2 are two inbred lines originating from the same cross Minica x Canner Express Pedigree information: Troy is parent of Maya. Alfred is parent of Mars. Minica is parent of Alfred and of Victor.

Table 2 Level of polymorphism revealed by the eight *Eco*RI/*Mse*I primer combinations with 79 faba bean inbred lines

Primer combination	Total number of fragments	Polymorphic fragments [%]
E-AAC/M-CAC	80	91.25
E-ACA/M-CTA	46	91.30
E-ACC/M-CAC	56	96.43
E-ACC/M-CAG	47	87.23
E-ACG/M-CTT	52	94.23
E-AGC/M-CTA	71	91.55
E-AGG/M-CAT	92	83.70
E-AGG/M-CTC	83	91.57

EDTA, pH 8.3] at 58-Watts constant power. After electrophoresis, gels were fixed in 10% acetic acid and dried. Dried gels were then exposed to X-ray films X-OMAT AR (Kodak) for 3-5 days depending on the intensity of the radiation signal.

Detection and scoring of AFLP fragments

AFLP amplification products, separated by gel electrophoresis along the inbred lines' lane, are located at positions on the gel that strictly correspond to the products' length. Any position on the gel which, when scored across genotypes, contains at least once a DNA amplification product will be called herein "fragment". Whenever a given fragment contains a scoreable amplification product, this will be termed "band". Hence, the total number of fragments is a constant number for the experiment, whereas the number of bands per genotype varies across genotypes. Monomorphic fragments contain an amplification product, a band, in each inbred line, polymorphic fragments do not. Fragments were scored as "1" when a band was present and were scored as "0" if not. In doubtful cases, mostly where a band's intensity was markedly less strong than that of checks but not convincingly absent, a score of "9" was given and it was considered as missing data point later in the analysis. Fragment scoring was performed visually from X-ray films. Throughout, two neighboring lanes on a gel were assigned to the amplified products of a single inbred line; a maximum of 24 inbred lines could be scored per gel. Thus, to test the 79 inbred lines with a single primer combination, at least four gels were run and scored. To facilitate scoring across gels, an AFLP DNA ladder (30-300bp; GIBCO BRL, Life Technologies, Inc., Rockville, Md) was used. Furthermore, for each of the tested primer combinations, amplification products of two inbred lines (K25 and Pel, Table 1) were included as additional checks in separate lanes on each gel.

Consistency of the AFLP profiles

The consistency of AFLP profiles was assessed by comparing the marker phenotypes from one duplicate assay (inbred line K25) and from replicate assays of 11 inbred lines (cf. Table 1) using seven primer combinations. The duplicate AFLP profiles were the outcome of two aliquots of one pre-amplification reaction of the entry K25 (Table 1), each of which was run in a separate gel. Replicate AFLP profiles were the product of replicate assays. A replicate assay comprised two individuals, termed replicate individual and reference individual. These two were prepared on separate occasions from seed harvested from the same selfed plant of a given inbred line. The replicate and reference

individual of inbred line K25 were run together in the same gel throughout. Further ten replicate assays as specified in Table 1 were run with their two individuals in separate gels, respectively.

Data analysis

The genetic similarity values (GS values) between the two individuals of each of the ten replicate assays were estimated employing three GS coefficients: GS_J (Jaccard 1908), GS_{SM} (Simple Matching; Sneath and Sokal 1973) and GS_{NL} (Nei and Li 1979). Analysis of data was performed for the primer combinations one at a time. It was analyzed whether the mean GS_J values across the ten replicate assays were significantly different as a result of the actual primer combination. Differences between the GS_J value of a replicate individual to any of the other 78 inbred lines and the GS_J value of its reference individual to the same 78 inbred lines were analyzed and used to estimate an error variance σ^2_e and a least significant difference (LSD) for these GS_J values as well as for the GS_J values of the main data set (see below).

For the 79 inbred lines, the 1/0/9 matrix that resulted from scoring of AFLP products of eight primer combinations (main data set) was used to calculate all possible pairwise similarity coefficients between inbred lines. After excluding monomorphic fragments, GS_J were used throughout, applying the software NTSYS-pc version 2.1q (Rohlf 2000). Jaccard's coefficient and exclusion of monomorphic fragments were implemented following the recommendation of Link et al. (1995). Bootstrapping (Felsenstein 1985) on the similarity matrix was performed. Marker samples of different sample sizes, starting from five fragments per sample and increasing in steps of 35 fragments per sample were randomly drawn from the total number of fragments (N = 477) with replacement. Two hundred marker samples for each sample size were established for two pairs of inbred lines; the most distant pair (Blaze & Columbo) and the most similar pair (L_1 & L_2). The mean coefficient of variation (CV %) for Jaccard's coefficient of similarity for each sample size was calculated using the software Gen Dist. 1.8 provided by A. Valentini, Universita` della Tuscia, upon request.

Based on the country of origin of the cultivars and on their seed size, it was obvious to divide the whole collection into six groups of germplasm namely: Asian, European (EU

major and EU *minor*), South European, winter beans and North African. For the large (N = 40) EU *minor* group and the other five smaller groups, the average intra-group GS_J-value was calculated [$(N^2-N)/2$ values per group]. Moreover, for each of the lines belonging to one of the five small germplasm groups, its average GS_J-value to the N = 40 EU *minor* lines was calculated. The difference between the mean inter-group similarity of e.g. the EU *major* lines to the EU *minor* lines, and the intra-group EU *minor* similarity was tested by T-test for significance of difference; the test was conducted for all five small groups. The GS_J mean values describing the similarity between groups of germplasm, taking the group as smallest unit of analyses, were used to construct a dendrogram based on the unweighted pair-group method with arithmetic means (UPGMA) using NTSYS-pc version 2.1q (Rohlf 2000).

The similarity matrix of the 79 individual inbred lines (main data set) was used to construct a further dendrogram using UPGMA. Associations among the 79 inbred lines using PCoA according to Gower (1966) were revealed based on the GS_J estimates employing NTSYS-pc version 2.1q (Rohlf 2000). Bootstrap analysis was performed by drawing 400 random samples of fragments, to determine the confidence limits of the UPGMA based dendrogram using the software package „WinBoot“ developed at IRRI (Yap and Nelson 1996).

Results

Analysis of 79 faba bean lines with eight *Eco*RI/*Mse*I primer combinations (Table 2) resulted in the amplification of 527 scoreable fragments. The number of scoreable fragments per primer combination ranged from 46 to 92 for the primer pairs E-ACA/M-CTA and E-AGG/M-CAT respectively (Table 2), with an average of 66 fragments. From those, 477 were polymorphic, thus monomorphic fragments amounted to 9.5% of the total number.

Consistency of the AFLP profiles

The inbred line K25 (highly inbred; F12), was used for a duplicate assay (two aliquots of one pre-amplification reaction run in different gels). Nine mismatches from a total of 477 polymorphic fragments (GS_J = 0.97) were found. A replicate assay of K25 (DNA from two individuals belonging to this line) run together in one gel was scored with only three

mismatches (GS_J = 0.99).

Examining the GS_J values between the ten replicate individuals and the 79 entries (including the corresponding ten reference individuals; Table 1) showed that each replicate individual expressed its maximum similarity with its corresponding reference individual. Results on the GS values for the replicate assays when estimated based on the three tested coefficients of similarity are summarized in Table 3. The mean genetic similarity based on Nei and Li between a replicated individual and its reference line amounted to GS_{NL} = 0.86 followed by mean GS_{SM} = 0.80 and mean GS_J = 0.76.

Consider two GS_J values comprising one common line, this common line being one of the ten replicate assays (e.g., GS_J [Alpine x Alfred] = 0.69 and GS_J [Alpine x Blaze] = 0.62, cf. Table 1). Hence, for the common line, data across two individuals are available. Comparing these two GS_J values, an error variance of $\sigma^2_e = 4.95 \times 10^{-4}$ was estimated, i.e., the error standard deviation amounts to 2.22×10^{-2} (cf. Table 4). The corresponding $LSD_{(0.05)}$ value amounts to 4.37×10^{-2}. Extending this to the main data set, where GS_J values are calculated from non-replicated data, the appropriate $LSD_{(0.05)}$ to compare GS_J values amounts to 6.17×10^{-2}. The error variance (4.95×10^{-4}) amounts to about 29% of the variance between all GS_J values in the main data set (17.32×10^{-4}).

Genetic similarities and PCoA

Genetic similarity values of pairs of inbred lines ranged from GS_J = 0.88 between the inbred lines L_1 & L_2 and GS_J = 0.53 between the inbred lines Blaze and Columbo. Clustering of inbred lines based on their GS_J values using the UPGMA, showed no major "ball clusters" (*sensu* Rohlf 1993). Six of the eight Asian inbred lines clustered together. Within this cluster (Fig. 1), the two Japanese inbred lines, Otafuku and Kawachi-green clustered together and were separated from the four Chinese inbred lines (Cixi Dabeican, Shangya Tian, Pinghu Zaojiazhong & Maya-Asia). None of the other groups formed a recognizable cluster. Highly similar lines were Mars & Alfred (GS_J = 0.79; Alfred is a parent of Mars), Condor & K25 (GS_J = 0.76; Kristall is a parent of Condor), Maya & Troy (GS_J = 0.76; Troy is a parent of Maya). Jantarnij & Orletzkij (GS_J = 0.81) as well as Mars& Merkur (GS_J = 0.79) are seemingly related but do not share common ancestors based on available pedigree data.

Table 3 Range and mean of the three genetic similarity estimators for the ten replicate assays based on results of 7 primer combinations

Estimator	Min.	Max.	Mean
Simple Matching	0.75	0.83	0.80
Nei & Li	0.83	0.88	0.86
Jaccard	0.71	0.79	0.76[#]

[#]cf. Fig. 1

Table 4 Analysis of variance for GS_J values calculated between ten replicate assays (I) with two individuals each (R), and N = 79 inbred lines (L)

Source	DF	Mean squares ($x10^{-4}$)	Variance components ($x10^{-4}$)	LSD (0.05) ($x10^{-2}$)
I	9	238.78	1.48	0.49
L	78	100.33	4.77	1.38
LI	692	4.34	-0.31	4.37
LRI	780	4.95	4.95	-

Bootstrap values indicated on the dendrogram showed that the two inbred lines L_1 & L_2 were grouped together in 100% of the cases. Similarly, a bootstrap value *P* = 97.8% was shown for the inbred lines Jantarnij & Orletzkij (Fig. 1). Other inbred lines grouping together with a relatively high confidence level, although not reaching the 95% bootstrap *P* value (Felsenstein 1985) were e.g., Maya & Troy (*P* = 69.3) and Pinghu Zaoajiazhong & Maya-Asia (*P* = 70.8).

The mean intra-group similarity within the EU *minor* lines amounted to $GS_J = 0.67$; Mars, Herz Freya, Kristall, Condor and Victor showed the highest values ($GS_J \approx 0.70$), whereas Blaze, Nadwislanski, Nixe, Erfano and Fatima showed the smallest ($0.61 < GS_J < 0.67$). The same mean value ($GS_J = 0.67$) was found for the intra-group similarity within the EU *major* lines and within the Asian bean lines. The mean intra-group similarity within the winter beans amounted to $GS_J = 0.64$; the same mean value was found for the intra-group similarity within the South European and within the North African inbred lines. The mean genetic similarity between the EU *major* and EU *minor* lines amounted to $GS_J = 0.66$. This was not significantly different (P = 0.05) from the intra-group EU *minor* value ($GS_J = 0.67$, cf. Fig. 2). The similarities of the other small germplasm groups to the EU *minor* group lines were significantly different (P = 0.01) from the intra-group EU *minor* value (Fig. 2), the Asian beans showing the smallest value. Upon clustering of the six groups using the UPGMA based on their mean GS_J values (Fig. 3), the Asian group was convincingly separated from all other groups. These could be regarded as two further groups: EU *minor*, EU *major* and winter bean vs. South European and North African. Yet, this further grouping is less convincing.

The first two principal coordinates (Fig. 4) explained 8.16% of the total variability, with the Asian group partly separated from the others. Both, the EU *major* and *minor* groups were spread across the principal coordinates and no grouping was seen. The pattern shown was not much clearer when the other three groups were regarded (Fig. 4).

Bootstrapping of AFLP fragments revealed that the coefficient of variation (CV%) of Jaccard's similarity coefficient between the inbred lines Blaze & Columbo (the most distant inbred lines) decreased from 50% for samples of size five fragments to 5.1% for samples of N = 477 fragments (Fig. 5). A coefficient of variation of 2.1% for samples of N = 477

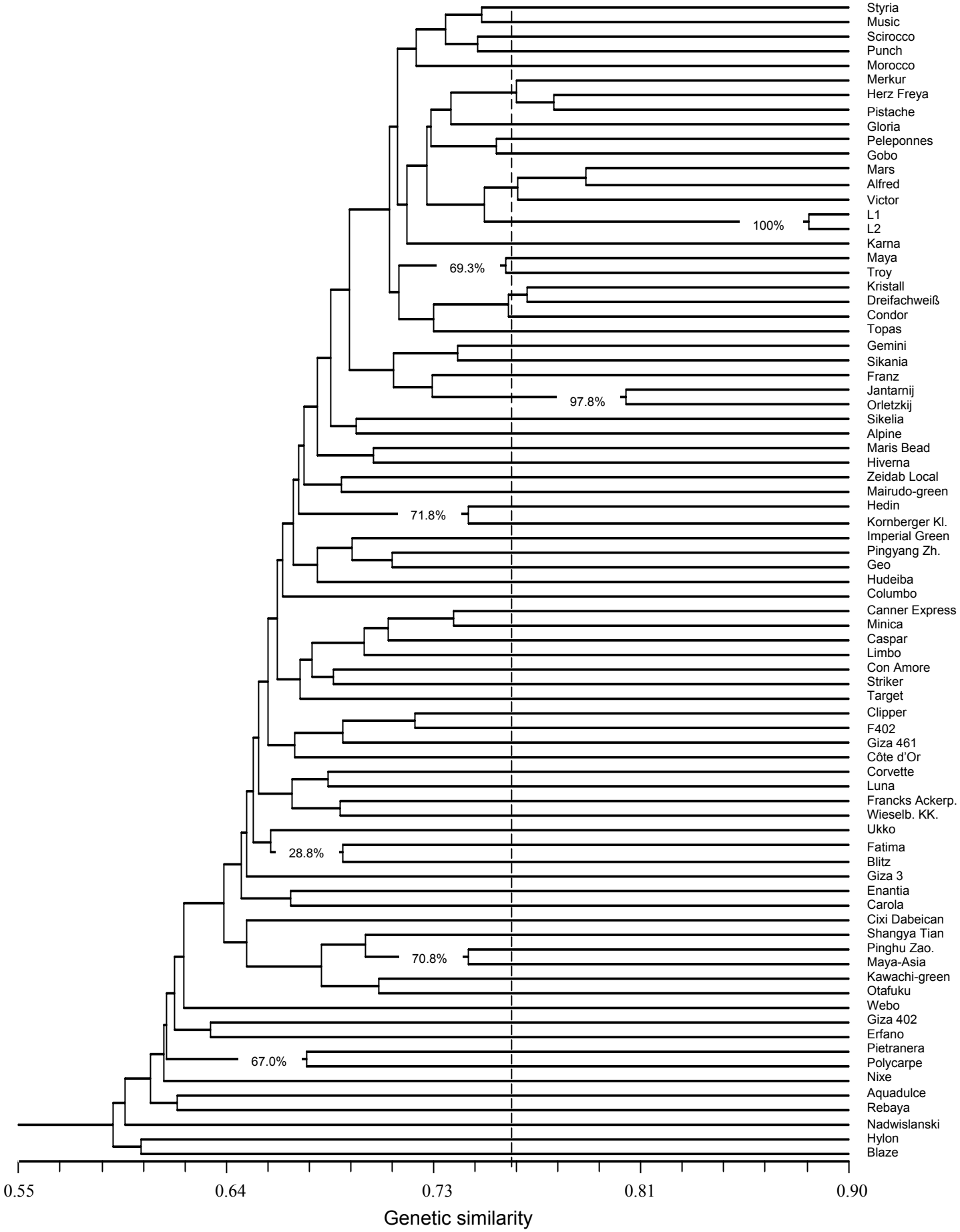


Fig. 1 Grouping of 79 faba bean lines based on 477 polymorphic fragments using Jaccard's coefficient of similarity and UPGMA clustering. The dotted line indicates the threshold value of 0.76 (cf. Table 3), nodes to the right of this line present tightly related inbred lines. Numbers shown at different nodes represent percentage confidence limits of the bootstrap analysis

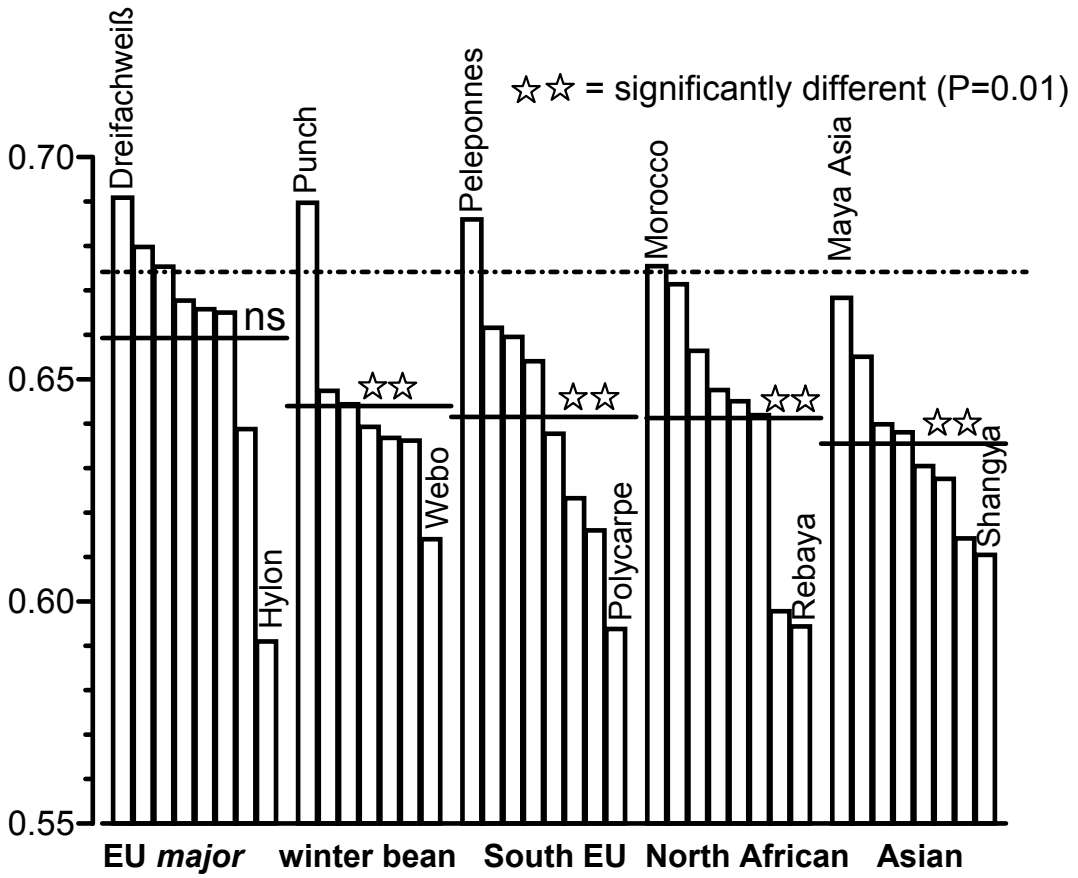


Fig. 2 Mean GS_J values between the individual non EU *minor* inbred lines and the 40 EU *minor* lines. The solid lines show group mean similarities, the broken line shows the average intra-EU *minor* similarity for comparison. Intra-EU *minor* similarity and the mean similarities of the groups were tested by T-Test for significance of difference

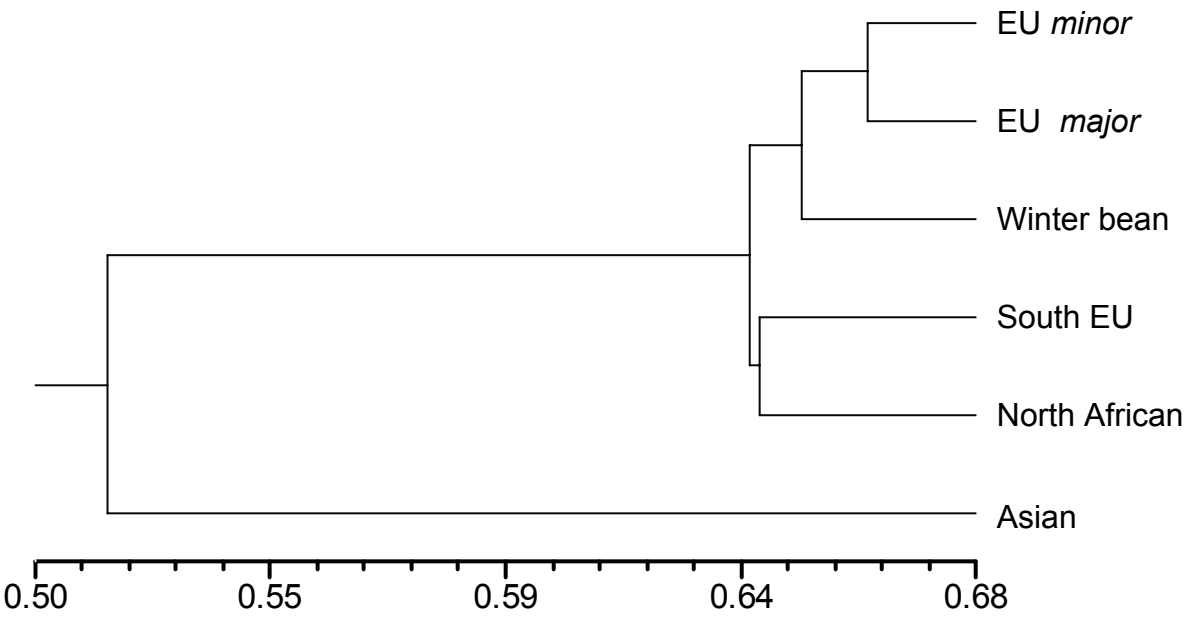


Fig. 3 Dendrogram generated by UPGMA for mean GS values of germplasm groups

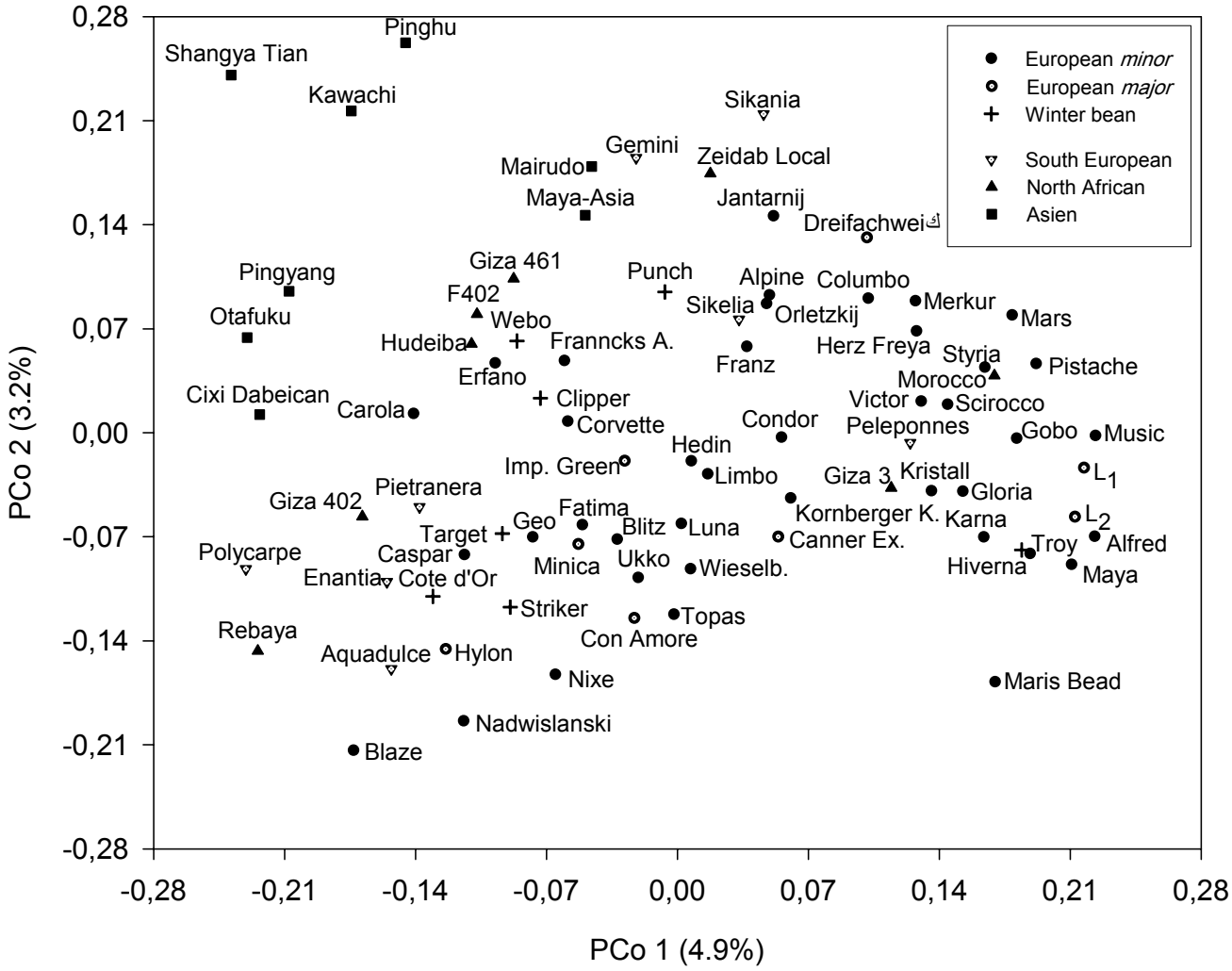


Fig. 4 Associations among 79 faba bean inbred lines revealed by Principal Coordinate Analysis performed on 477 polymorphic fragments

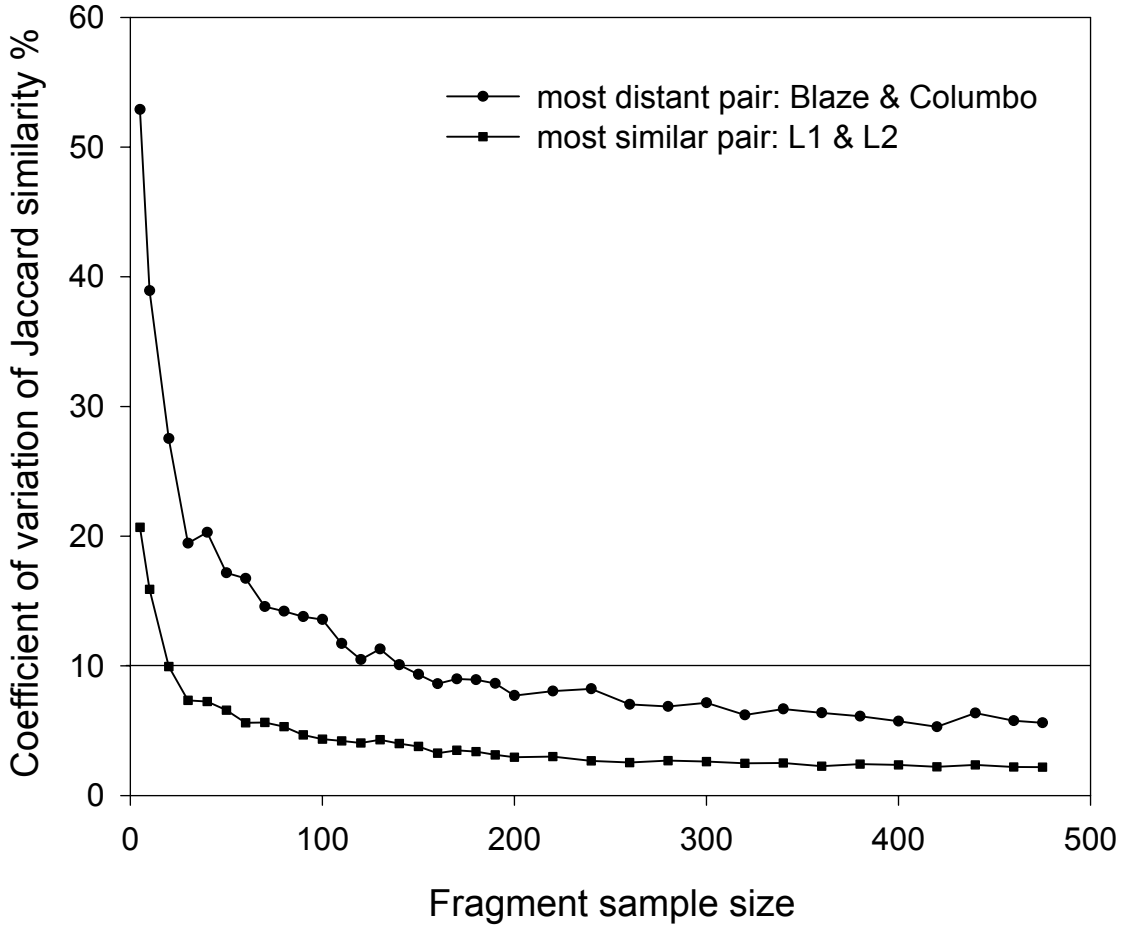


Fig. 5 Coefficient of variation of Jaccard's genetic similarity in relation to marker sample size (N = 200 bootstraps). Fragment samples of different size were sampled with replacement from two pairs of individuals; the most similar and most distant

fragments, representing the whole data set, was achieved for the closest inbred lines: L_1 & L_2.

Discussion

Error estimation

Not only the amount of diversity (fixed when sampling the inbred lines), but the amount of error in the scored data as well are decisive for the outcome and usefulness of a diversity analysis. DNA quality differences, affecting the AFLP banding pattern, may result in errors (Vos et al. 1995). Hansen et al. (1999) evaluated the usefulness of the AFLP technique for genetic analysis of sugar beet. Gel resolution was among the sources of error reported in their study. In the present study replicate and duplicate assays were employed to allow for inspection and quantification of error in the results. The replicate (GS_J = 0.99) and duplicate assay (GS_J = 0.97) with K25 demonstrated the high consistency of the AFLP profiles obtained here. This is in agreement with Winfield et al. (1998) where estimated GS values for duplicates of *Populus nigra* ranged between 0.96 < GS < 1.00; this seems to reflect the amount of technique-born errors for AFLPs. As to the amount of error estimated, replicate assays of ten inbred lines run in different gels showed an average of 82 mismatches with an average genetic similarity GS_J = 0.76. Theoretically, this value should be near to unity as shown for the replicate assay of K25 (GS_J = 0.99). However, K25 itself was one of the two internal checks, which resulted in a better scoring, since all of its bands were traced precisely across gels. Bands of the ten replicate assays scored across gels on the other hand, had the bands of the internal checks (K25 and Pel) as a reference, this obviously attributed to high error rates during scoring. These results are supported by observations of Schwarz et al. (2000), showing greater inter-gel differences in band size of identical samples than intra-gel differences for fluorescence-based semi-automated AFLP analysis in barley and wheat.

Sneath and Sokal (1973) noted the usefulness of testing the significance of the difference between similarity coefficients. The least significant difference in the main data set was $LSD_{0.05}$ = 0.062. This value focuses on differences between individuals within the same inbred line (replicate assays) caused by residual segregation within the lines and by errors after DNA extraction, especially at scoring across gels. Applying this LSD to GS_J values of both, L1 & L2, to any other inbred line showed no significant differences between these

values; the same was true for the inbred line pair Jantarnij & Orletzkij. Opposed to this, 96.4% of the GS_J values of any EU *major* line to any other EU *major* lines was surpassing this difference when compared with the corresponding GS_J values to any Asian line (cf. Fig. 2), reflecting the larger AFLP similarities among EU *major* lines compared to those between EU *major* and Asian lines. These findings substantiate the usefulness of the LSD-value.

Primer combinations to employ

Ellis et al. (1997) observed that 80% of the provided genetic information could already be achieved by selecting six "good" primer combinations. Selection was based on the number of polymorphic fragments revealed and the number of genotypes uniquely identified by each primer combination. We compared the GS_J values of each of ten replicate assays for one primer combination at a time. All primer combinations were equally useful (Table 5), no significant differences were observed between them. Consequently all primer combinations were employed.

The appropriate genetic similarity coefficient

Appropriateness of similarity coefficients has been often debated (Piepho & Laidig 1997 and Robinson & Harris 1999). In the present case (Table 3), where two replicate, sister individuals taken from the same, selfed parental plant were studied, we suggest employing GS_{SM} that takes (0-0) negative matches into account. Negative matches here uncover genetic similarity, since the cause for absence of a band in those two individuals should be identical and inherited from the same ancestor (identity by descent, Smith and Smith 1992). This however is not the case when comparing non-related inbred lines, where the absence of a given band is not necessarily based on the same genetic cause: it may or may not be caused by the same DNA variant. Thus, in such cases a similarity coefficient that ignores negative matches is more appropriate. Both Jaccard and Nei & Li disregard negative matches. Piepho and Laidig (1997) demonstrated the monotonic, non-linear relationship between both namely: $GS_{NIxy} = 2\ GS_{Jxy} / (1 + G_{Jxy})$; hence, the rank correlation between GS_{NL} values and GS_J is r = 1.00, which could be corroborated by our findings.

Lamboy (1994) studied the bias in an estimated GS value, defined as difference between estimated and true GS value, caused by false positive bands or false negative bands. He

Table 5 Analysis of variance for GS_J values calculated for ten replicate assays (I), with two individuals each, when tested using seven different primer combinations (P)

Source	DF	Mean squares ($\times 10^{-4}$)	Variance components ($\times 10^{-4}$)	F
I	9	47.1	-1.51	0.82
P	6	73.3	1.56	1.27
IP	53	58.0	58.0	-

regarded dominant marker data, and used both Jaccard's and Nei & Li's coefficients of similarity. Simulations demonstrated smaller bias by Nei & Li than by Jaccard. A disadvantage of the Nei & Li coefficient in case of dominant markers is, it is not a linear function of the coancestry coefficient of the lines under study, whereas the Jaccard coefficient indeed is, provided the entries are homozygous lines (Link et al. 1995). Decision on which of these two coefficients should be used will depend on the type of molecular marker employed as explained by Engqvist and Becker (1994) and Link et al. (1995), i.e. Jaccard for dominant markers, Nei & Li for codominant markers.

Here, not all lines were highly homozygous (Table 1), which could blur the results. Less inbred individuals tend to have more bands; they do share more common bands with others; for a heterozygous AFLP locus the recessive AFLP allele ("no band") is ignored. On average, N = 311.6 bands were amplified per highly inbred line (> F9) in comparison to N = 317.1 bands for lines that were selfed for only three generations or less. The average number of bands expected from crosses (F1) among the highly inbred lines, following a dominant inheritance of AFLPs, amounted to N = 384.2 bands. Compared with this result, obviously a high inbreeding level ($311.6 \approx 317.1 << 384.2$) was achieved even with our less inbred entries. This is expected; the inbred lines were obtained by selfing individuals from elite cultivars that are already rather narrow populations or even lines.

Resampling

A stable classification of genotypes is not obtained based on a small number of polymorphic fragments (Sneath and Sokal 1973). Bootstrap analysis was applied to the genetic similarity matrix of the 79 inbred lines to estimate the sampling variance of the genetic similarities. Bootstrapping of 477 polymorphic fragments for all GS_J values resulted in a mean CV of less than 4% (details not shown), indicating that a sufficient number of markers were applied here. To achieve a CV < 10%, only 110 AFLP fragments were required for the most distant pair of inbred lines, closer pairs of inbred lines required even fewer fragments. A further increase in the number of fragments did not result in a proportional decrease in the CV value. These results are in accordance with those obtained on other marker types and species (Tivang et al. 1994; Ajmone-Marsan 2001) indicating only marginal gains in precision upon increasing the number of fragments beyond a certain limit. A further point, not regarded here (because mapping data for the

fragments scored here is not available) is the localization and distribution of the marker loci across the faba bean genome (Satovic et al. 1996); an even distribution should be favorable (Castiglioni et al. 1999).

Pedigree and historical aspects

Considering the long history of cultivation of faba bean in Central and Western Europe, the European *minor* group is represented by a dominating, large sample of selected cultivars in this study (N = 40). Yet narrowing down this sample to eight lines (eight is the number in other groups) would not alter the final conclusion deduced from this experiment (details not shown). Central and South America were not represented here, though faba bean was introduced there in the sixteenth century by the Spaniard and the Portuguese and surely underwent an independent development there (Lawes et al. 1983).

For the 79 inbred lines, the cluster analysis indicated that no clear grouping was detected neither based on seed size (EU *minor*, EU *major*) nor on geographic origin. In one case only, for the Asian group, six of the respective eight inbred lines clustered together (Fig. 1). Yet, cases of known pedigree relationships like Alfred & Mars were convincingly corroborated (cf. Table 1); “non related pedigrees” were uncovered as doubtful because of temptingly high genetic similarity (e.g., Jantarnij & Orletzkij). Nadwislanski, Blaze and Hylon were uncovered as very loosely related to any other line; hence, it is very tempting to consider these for crosses with recent elite lines not being that isolated, like Mars, Condor and Limbo. The PCoA did not show a more marked pattern in the AFLP-based genetic diversity either. The first two principal coordinates of the PCoA (Fig. 5) explained 8.16% of the variation. This is, in spite of the first impression, not a small quantity: a high number of entries (N = 79) drastically restrict this quantity. In maize, Messmer et al. (1992) employing RFLP markers and Lübberstedt et al. (2000) employing AFLP markers found quantities of 14.6-17.5% with smaller numbers (N = 57 and N = 51), even though lines of two clearly distinct maize gene pools were analyzed in their studies.

A rather simple approach was to *a priori* group the inbred lines according to their geographic origin and their botanical group (Fig. 2; Fig. 3). Hanelt (1972) reported that in Europe, faba bean *minor* was the first group to be cultivated and surprisingly late, after about the year 500 the *major* group did emerge. Only after the seventeenth century the first true winter hardy stocks were developed, those were most probably selected from the

European *minor* group (Hanelt 1972). Our results showed a higher genetic similarity among the EU *minor*, EU *major* and the winter bean germplasm in comparison to genetic similarity to the other groups. With the Near or Middle east suggested to be the center of diversity, the spread of the crop took place towards Central and Northwest Europe through North Africa: hence, some similarity is expected between European and North African germplasm. In the Far East, faba bean was probably introduced already in about 100 BC (Tao 1981). Regarding the very far geographic distance of the European and Mediterranean area to China and Japan, the results of the molecular data (Fig. 3) are in accord with the expectations.

Results from the marker data presented in this study revealed the broad genetic base of the faba bean germplasm investigated. Although crossing between faba bean germplasm groups has occurred in the past and is still an ongoing process, a great potential for further cultivar improvement and hybrid production is still available. Wide crosses between materials from the European germplasm and that from Asia are of special importance, e.g. if traits like winter hardiness of the Asian material are of interest. The European germplasm, although showing an elevated intra-group similarity, comprises quite diverse lines that provide a good base for utilizing readily adapted elite material for future hybrid breeding programs. Studies concerning the relationship between genetic similarity based on AFLP markers and hybrid performance within the European elite germplasm were launched and results are currently under evaluation.

Acknowledgements

Thanks are due to all who provided us with the seeds used in this study. The financial support from the Gemeinschaft zur Förderung der privaten deutschen Pflanzenzüchtung e. V. (GFP) is gratefully acknowledged

References

Ajmone Marsan P, Negrini R, Crepalde P, Milanesi E, Gorni C, Valentini A, Cicogna M (2001) Assessing genetic diversity in Italian goat populations using AFLP markers. Intr Soc Animal Genet 32: 281-288

Becker HC, Engquvist GM, Karlsson B (1995) Comparison of rapeseed cultivars and resynthesized lines based on allozyme and RFLP markers. Theor Appl Genet 91: 62 -67

Bond DA (1976) Field bean, *Vicia faba*. In: Simmonds NW (ed) Evolution of crop plants, Longman, London, pp 179-182

Castiglioni P, Ajmone Marsan P, van Wijk R (1999) AFLP markers in a molecular linkage map of maize: Codominant scoring and linkage group distribution. Theor Appl Genet 99: 425-431

Cubero JI (1974) On the evolution of *Vicia faba* L. Theor Appl Genet 45: 47-51

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

Efron B, Tibishirani R (1986) Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. Stat Sci 1: 54-77

Ellis RP, Mc Nicol JW, Baird E, Booth A, Lawrence P, Thomas B, Powell W (1997) The use of AFLPs to examine genetic relatedness in barley. Mol Breed 3: 359-369

Engqvist GM, Becker HC (1994) Genetic diversity for allozymes, RFLPs and RAPDs in resynthezed rape. In: van Ooijen JW, Jansen J (eds) Biometrics in plant breeding: application of molecular markers. Proc of the ninth meeting of the EUCARPIA section biometrics in plant breeding. 6-8 July 1994. Wageningen, pp 85-90

Excoffier L (1994) The statistical analysis of molecular data for inferring population genetic structure: The AMOVA framework. In: Perrin N, Keller L, Goudet J (eds) Evolution of structured populations, Bull Soc Vaud Sc Nat 83:159-160

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783-791

Finkeldey R (1994) A simple derivation of the partitioning of genetic differentiation within subdivided populations. Theor Appl Genet 89: 198-200

Gower JC (1966) Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53: 325-338

Hanelt, P (1972) Die infraspezifische Variabilität von *Vicia faba* L. und ihre Gliederung. Die Kulturpflanze 20: 75-128

Hansen M, Kraft T, Christiansson M, Nilsson NO (1999) Evaluation of AFLP in Beta. Theor Appl Genet 98: 845-852

Jaccard, P (1908) Nouvelles researches sur la distribution florale. Bull Soc Vaud Sci Nat 44: 223-270

Jones CJ, Edwards KJ, Castaglione S, Winfield MO, Sala F, van de Wiel C, Bredemeijer G, Vosman B, Matthes M, Daly A, Brettschneider R, Bettini P, Buiatti M, Maestri E, Malcevschi A, Marmiroli N, Aert R, Volckaert G, Rueda J, Linacero R, Vazquez A, Karp A (1997) Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. Mol Breed 3: 381-390

Käser HR, Steiner AM (1983) Subspecific classification of *Vicia* faba L. by protein and isozyme patterns. Fabis Newsl 7: 19-20

Lamboy W (1994) Computing genetic similarity coefficients from RAPD data: The effects of PCR artefacts. PCR Methods and Applications 4: 31-37

Lawes DA, Bond DA, Poulsen MH (1983) Classification, origin, breeding methods and objectives. In: Hebblethwaite PD (ed) The faba bean (*Vicia faba* L.) A basis for improvement. Butterworths, London, England, pp 23-67

Link W, Dixkens C, Singh M, Schwall M, Melchinger AE (1995) Genetic diversity in European and Mediterranean faba bean germplasm revealed by RAPD markers. Theor Appl Genet 90: 27-32

Lübberstedt T, Melchinger AE, Dußle C, Vuylsteke M, Kuiper M (2000) Relationships among early European maize inbreds: IV Genetic diversity revealed with AFLP markers and comparison with RFLP, RAPD and pedigree data. Crop Sci 40: 783-791

Messmer MM, Melchinger AE, Boppenmaier J, Brunklaus-Jung E, Herrmann RG (1992) Relationships among early European maize (*Zea mays* L.) inbreds: I. Genetic diversity among flint and dent lines revealed by RFLPs. Crop Sci 32: 1301-1309

Muratova V (1931) Common beans, *Vicia faba* L. Suppl Bull Appl Bot Genet Pl Br, Leningrad, pp 298

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

Piepho HP, Laidig F (1997) A review of methods for cluster analysis of marker data. UPOV document BMT/47/7. UPOV, Genf, p13

Robinson J, Harris SA (1999) Which DNA Marker for Which Purpose? In: Gillet EM (ed) Final compendium of the research project development, optimisation and validation of molecular tools for assessment of biodiversity in forest trees in the European Union. DGXII Biotechnology FW IV Research Programme Molecular Tools for Biodiversity.

Rohlf FJ (1993) NTSYS-PC. Numerical taxonomy and multivariate analysis system. Exeter Software, New York

Rohlf FJ (2000) NTSYS-pc. Numerical taxonomy and multivariate analysis system Exeter Software, New York

Satovic Z, Torres AM, Cubero JI (1996) Genetic mapping of new morphological, isozyme and RAPD markers in *Vicia faba* L. using trisomics. Theor Appl Genet 93: 1130-1138

Schwarz G, Herz M, Huang XQ, Michalek W, Jahoor A, Wenzel G, Mohler V (2000) Application of fluorescence-based semi-automated AFLP analysis in barley and wheat. Theor Appl Genet 100: 545-551

Smith JSC, Smith OS (1992) Fingerprinting crop varieties. Adv Agron 47: 85-140

Sneath PH, Sokal RR (1973) Numerical taxonomy. WH Freeman & Co. San Francisco. pp 573

Sokal RR, Rohlf J (1981) Biometry. WH Freeman and Company, New York.

Stelling D, von Kittlitz E, Ebmeyer E, Sass O, Jaiser H, Link W (1994) Erfolge und Perspektiven der züchterischen Verbesserung von Körnererbsen und Ackerbohnen in der EU. Bericht über die 45. Arbeitstagung der Arbeitsgemeinschaft der Saatzuchtleiter im Rahmen der „Vereinigung österreischer Pflanzenzüchter“, BAL Gumpenstein, Nov 1994, pp 22-24

Tao ZX (1981) Faba bean production and research in China. FABIS 3: 24-25

Tivang JG, Nienhuis J, Smith OS (1994) Estimation of sampling variance of molecular marker data using the bootstrap procedure. Theor Appl Genet 89: 259-264

Vos P, Hogers R, Bleeker M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23: 4407-4414

Winfield MO, Arnold GM, Cooper F, Le Ray M, White J, Karp A, and Edwards KJ (1998) A study of genetic diversity in *Populus nigra* subsp *betulifolia* in the Upper Severn area of the UK using AFLP markers. Mol Ecol 7: 3-10

Yamamoto K (1973) Karyotaxonomical studies on *Vicia.* 1. On the karyotype and character of some annual species of *Vicia*. Jpn J Genet 48: 315-327

Yap IV, Nelson RJ (1996) WINBOOT: a program for performing bootstrap analysis of binary data to determine the confidence limits of UPGMA-based dendrograms. IRRI Disc Pap Ser 14

Zeid M, Schön CC, and Link W, (2001) Genetic diversity in a group of recent elite faba bean lines. Czech J Genet Plant Breed 37: 34-40

AFLP- based genetic similarity and hybrid performance in faba bean*

ABSTRACT

Successful prediction of heterosis and performance of F1-hybrids from the genetic similarity of their parents based on molecular markers has been reported in various crops. The relation between genetic similarities based on amplified fragment length polymorphism (AFLP) of 18 European faba bean lines and their hybrid performance and heterosis was investigated. Parental lines, 62 F1-hybrids and their F2-progenies were evaluated in field trials in four environments in Germany for their seed yield, 1000-seed weight and plant height. AFLP analysis of the 18 inbred lines using 26 *Eco*RI/*Mse*I primer combinations resulted in 1202 polymorphic fragments. Cluster analysis and Principal Coordinate Analysis based on genetic similarity estimates were in agreement with available pedigree information. Correlation coefficients between the genetic similarity estimates and either heterosis or F1-hybrid performance, even those between specific genetic similarity and specific combining ability for the three studied traits were too small to be of predictive value. Results showed that AFLP-based genetic similarities are not predictive of the performance of hybrids or heterosis within the elite European faba bean gene pool, at least as long as mating of parents that are related by pedigree is excluded. Furthermore, data clearly demonstrated a marked amount of genetic diversity and heterosis available even within the European *minor* gene pool.

---

* A modified, shortened version of this chapter is in preparation as manuscript for Crop Science with the following authors:

Mahmoud Zeid[1], Chris-Carolin Schön[2] and Wolfgang Link[1]

[1]Institute of Agronomy and Plant Breeding, Georg-August-University, 37075 Göttingen, Germany;

[2]State Plant Breeding Institute, University of Hohenheim, 70593 Stuttgart, Germany

INTRODUCTION

Faba bean is one of many crops where hybrid breeding has been suggested as a solution for improving both seed yield and yield stability (Stelling et al., 1994; Martsch et al., 2001). Its yield instability has been the main reason why farmers, world wide and especially in Europe were reluctant to grow faba bean in spite of many advantageous features of this crop. Yield improvement of 50-70% above the midparents was achieved in crosses either between *minor* and *major* lines (von Kittlitz, 1986) or even within the *minor* group (Link and Ruckenbauer, 1988).

Faba bean hybrids have also shown better adaptation to a wide range of abiotic conditions as compared to open pollinated or inbred cultivars, expressed in improved fertilization at high temperatures (Bond et al., 1994), and tolerance to lack of pollinating insects owing to their heterotic autofertility (Link, 1990), reduced winter damage (Bond et al., 1986), and better tolerance to drought stress (Abdelmula et al., 1999). Based on a detailed work studying the effects of heterozygosity and heterogeneity on yield stability in faba bean, Stelling et al. (1994) came to the conclusion that three–way hybrids, double hybrids, or single hybrids grown in blends are the key to yield improvement and stability in faba bean.

For commercial hybrid seed production in faba bean, a reliable cytoplasmic male sterility (CMS) system must be established (Duc, 1980). CMS in faba bean has been first reported by Bond et al. (1964), however, this first system (CMS447) failed due to its instability. Other CMS systems followed namely: CMS350 (Berthelem, 1970) and most recently the system (CMS199) developed by Link et al. (1997). Reports on this latest system (Martsch et al., 2001) still showed insufficient stability due to complex inheritance mechanisms, indicating that the system is not yet ready for commercial use. Presently, faba beans are bred as line cultivars, populations or synthetic varieties.

Research aimed at studying the genetic diversity in faba bean (Link et al., 1995; Zeid et al., 2001), and identifying quantitative trait loci (QTLs) for agronomic traits (Vaz Patto et al., 1999) provide important clues for optimizing crossing plans and selection strategies for line and population breeding. Hybrids, however, remain the method of choice where maximum gain from heterosis could be achieved (Stelling et al., 1994). Identification of parental lines leading to superior hybrids is a crucial factor in hybrid breeding programs.

As in case of maize, where hybrid breeding has been practiced for many decades, evaluation of combining ability of new lines is a costly and time consuming process that requires extensive field tests. Moreover, the large number of possible hybrid combinations to be produced from a small number of inbred lines, renders the practical evaluation of all possible combinations infeasible (Bernardo, 1994). Enhancing the efficiency of hybrid breeding programs could be achieved if the inbred lines *per se* could be screened and the superior crosses predicted before field evaluation (Melchinger et al., 1990a).

Breeders have since long expressed their desire to determine whether measures of similarity based on markers are correlated with measures of diversity, i.e. pedigree, F1 performance and heterosis. DNA-fingerprints were indeed used to develop measures of similarity between genotypes based on their marker profile (Smith and Smith., 1992). Furthermore, correlations of genetic dissimilarities based on RFLP data and yield performance of hybrids mainly in maize were not consistent among studies, e.g., Smith et al. (1990), Melchinger et al. (1990a,b) and Boppenmaier et al. (1992). Genetic dissimilarities based on RAPD markers were successfully employed for predicting grain yield in maize Lanza et al. (1997). Ajmone Marsan et al. (1998) pointed out the usefulness of AFLP markers over those of RFLP in predicting hybrid performance in maize.

A key study regarding the association between hybrid performance and genetic distance based on isozymes in maize was that of Frei et al. (1986). The authors related the usefulness of markers for predicting hybrid performance to whether crosses are produced including related lines or pairwise unrelated lines only. Their results demonstrated that a high correlation between genetic dissimilarity and yield of F1 crosses is only to be expected for lines with common pedigree background. The highly significant correlations between RFLP-based genetic dissimilarity and both heterosis (r_S = 0.87) and F1 performance (r_S = 0.93) for grain yield shown by Smith et al. (1990) for a large number of maize inbred lines including a large proportion of crosses between related lines is in general agreement to the observations of Frei et al. (1986). This was further elaborated by Melchinger (1999): a tight association between genetic dissimilarities and mid parent heterosis can be expected if there is marked variation for the relatedness of parental lines. It was also shown that in this case the coefficient of ancestry (f) and marker based

genetic dissimilarity are linearly interdependent; both should equally correlate with heterosis. In case of intra-group crosses, crosses between markedly related lines occur in addition to crosses between unrelated lines, at least as long as this is not intentionally avoided. The lack of association between heterosis and genetic dissimilarities for inter-group hybrids is explained by absence of crosses between related parents, i.e. by the absence of variation for parental relatedness: all crosses have unrelated parents.

The European faba bean gene pool has been studied for its genetic diversity (Link et al., 1995) and for hybrid performance and heterosis of intra- and inter-pool crosses (Link et al., 1996; Schill et al., 1998); heterotic groups within that pool however, are not yet established.

The primary objective of this study was to determine whether selection of genetically independent faba bean inbred lines within the elite European gene pool, useful for hybrid production is possible, if selection is solely based on available pedigree information. To achieve this objective (I) available pedigree data were studied and independent pairs of parents were defined, (II) the pattern of genetic diversity for AFLP markers and for heterosis within this set of faba bean lines was investigated and (III) associations between AFLP based genetic similarities of these inbred lines with agronomic performance and heterosis of their single cross hybrids were assessed.

MATERIALS AND METHODS

Plant materials and field trials

Eighteen faba bean (*Vicia faba* L.) lines derived from modern European (e.g., Scirocco, Styria) and ancestral cultivars (e.g., Herz Freya, Alfred), that are represented in the pedigree of recent germplasm were inbred for several generations (Table 1). Inbred lines were grouped into four testers, belonging to the *minor* group carefully selected to be of independent genetic background and to have outstanding agronomic potential and 14 experimental lines (also belonging to the *minor* group except for the lines L1, L2 and

Table 1 Pedigree, source and inbreeding stage of 18 faba bean lines

Line		Inbreeding	Source	Pedigree / Year of release
''Testers''				
1	Maya	I3	Serasem / France	Troy x ~Minica / 1995
2	Merkur	I3	Selgen / Czech Republic	Bolo x line x Fribo / 1997
3	Scirocco	I7	NPZ Lembke / Germany	# 1992
4	Styria	I12	Gleisdorfer SZ / Austria	Carola x Mythos / 1996
''Experimental lines''				
1	Alfred	I7	CEBECO/ Netherlands	Minica x Horse bean / 1983
2	Gloria	I8	Wieselburger SZ / Austria	Kornb. Kl. körnige / 1994
3	Gobo	I8	Germany (GDR)	# / 1987
4	Herz Freya	I12	Herz / Germany	# / 1935
5	Karna	I6	Wieselburger SZ / Austria	Kornb. Kl. Körnige / 1983
6	L1	I7	Germany	Minica x Canner Express
7	L2	I7	Germany	Minica x Canner Express
8	Maris Bead	I3	PBI / England	# /1965
9	Mars	I3	Danisco / Denmark	Alfred x Maribo / 1993
10	Music	I3	Blondeau / Frankreich	# / 1995
11	Pistache	I7	Joordens / Netherlands	# / 1990
12	Peleponnes	I12	Greece / ICARDA	# / #
13	Troy	I11	NPZ Lembke / Germany	# / 1985
14	Victor	I10	CEBECO / Netherlands	Minica x Cocksfieldsp./ 1988

Pedigree information or year of release is not available.

Peleponnes). The inbred lines were used to produce 62 single cross hybrids by controlled manual crossing. Testers were crossed with the experimental lines in a factorial design resulting in 56 hybrids and the former were intermated in a diallel cross with no reciprocals, to generate six further hybrids. Seeds for lines and for F2-progenies were produced by tripping-assisted, controlled selfing (Link, 1990).

All 62 single cross hybrids, their F2-progenies and their 18 parental lines were evaluated at four environments representing a broad range of agro-ecological conditions in Germany. Evaluation trials were conducted in the North (Hohenlieth in 2001 and in 2002), Middle (Göttingen in 2002) and South (Hohenheim in 2002) of Germany. Plots of $6m^2$ with plant density of 35 plants/m^2 were sown within the first 10 days of April and harvested at the end of August, with the varieties Scirocco and Limbo used as additional entries. The experimental layout was a 5x4 lattice with 4 replicates for the parents and an 8x8 lattice with 2 replicates for the F1-hybrids and for the F2-hybrids. The trials of parental lines were grown directly adjacent to the hybrids; the F1 and F2 progenies´ experiments shared a common field. Seed yield (ton ha^{-1}), 1000-seed weight (g), plant height (cm) and other morphological characters including flowering date, lodging and days to maturity were recorded on a plot basis. Differences in seed yield for the two checks (Scirocco and Limbo) in the trials of the parental lines and the hybrid trials were tested for significance. Midparent heterosis of each cross was calculated for F1-hybrids [F1-heterosis = F1 - parental mean] and their corresponding F2-hybrids [F2-heterosis = 2 x (F2 - parental mean)]. Heterosis and mid parent results were calculated at the level of genotypic means of each of the four environments.

AFLP analysis DNA was isolated from young leaves of single plants of each of the 18 inbred lines using CTAB extraction buffer according to Doyle and Doyle (1990). AFLP reactions were performed according to the procedure described by Zabeau and Vos (1993), using a commercially available kit (AFLP analysis System I, GIBCO BRL, Life Technologies, Inc., Rockville, Md) and following the manufacturer's instructions. After pre-screening of 32 *Eco*RI/*Mse*I primer combinations, 26 proved to be useful (Table 2) and were further employed in this study. The *Eco*RI primers were radiolabeled with [γ-33 P]-dATP provided from Amersham Pharmacia Biotech. After selective amplification, the

Table 2 Primer combinations (*Eco*R I/ *Mse* I) used for the AFLP analysis

Primer combination	Total number of fragments	% of polymorphic fragments
E-AAC/M-CAA	90	58.9
E-AAC/M-CAC	80	75.0
E-AAC/M-CAG	64	65.6
E-AAC/M-CAT	36	44.4
E-AAC/M-CTA	42	54.8
E-AAC/M-CTC	59	59.3
E-AAC/M-CTT	89	68.5
E-ACA/M-CAA	105	69.5
E-ACA/M-CAC	91	82.4
E-ACA/M-CAG	79	86.1
E-ACA/M-CAT	18	27.8
E-ACA/M-CTA	46	50.0
E-ACC/M-CAA	76	84.2
E-ACC/M-CAC	56	80.4
E-ACC/M-CAG	47	68.1
E-ACC/M-CAT	28	42.9
E-ACG/M-CTT	52	76.9
E-ACT/M-CTC	84	72.6
E-AGC/M-CAC	73	90.4
E-AGC/M-CAG	47	89.4
E-AGC/M-CTA	71	70.4
E-AGC/M-CTG	46	76.1
E-AGG/M-CAC	41	78.1
E-AGG/M-CAT	92	67.4
E-AGG/M-CTA	67	89.6
E-AGG/M-CTC	83	80.7

polymerase chain reaction (PCR) products were mixed with an equal volume of loading buffer [98% (v/v) formamide, 10 mM EDTA, 0.05% (v/v) xylene cyanol, and 0.05% (v/v) bromophenol blue] and denatured at 95ºC for 5 minutes and directly placed on ice. Seven µl of the mixture were loaded on an 8% (v/v) polyacrylamide gel containing 8.3 M urea that was pre-warmed for 20 minutes. For each of the tested primer combination, samples of all 18 inbred lines were run on the same gel. Two neighboring lanes represented each inbred line. Gels were run with 0.5x TBE electrophoresis buffer [50 mM Tris, 50 mM boric acid, 1mM EDTA, pH 8.3] at 58-Watts constant power. After electrophoresis, gels were fixed in 10% acetic acid and dried. Dried gels were then exposed to X-ray films X-OMAT AR (Kodak) for 3-5 days depending on the intensity of the radiation signal. Fragment scoring was performed visually from X-ray films, where bands were scored as present (1), absent (0) or doubtful (9), as previously described by Zeid et al. (2001)

Data analyses

Analysis of variance for agronomic traits was performed on the data of each environment separately, followed by a combined analysis across environments based on the lattice adjusted means of F1 and F2 and on the crosswise parental means according to Cochran and Cox (1957) using Plabstat 2N (Utz, 1997). A factorial analysis of general (GCA) and specific (SCA) combining ability was performed with F1 and F2 data of the 56 crosses only, according to the North Carolina Design II (Comstock and Robinson, 1948), partitioning the performance of F1-hybrids and F2-hybrids (Y_{ij}) between inbreds *i* and *j* to be:

$$Y_{ij} = \mu + gca_i + gca_j + sca_{ij} \qquad (1)$$

where μ is the overall mean performance of all possible hybrids, gca_i and gca_j are the general combining ability effects (GCA) of the inbred lines *i* (testers) and *j* (experimental lines) respectively and sca_{ij} is the specific combining ability (SCA) between inbred lines *i* and *j*. The four testers were considered as fixed factor.

AFLP-based genetic similarity (GS) between all possible pairs of inbred lines was calculated according to Jaccard (1908), applying the software NTSYS-pc version 2.1q (Rohlf, 2000). As outlined for the agronomic traits, the parental GS values of the 56 F1-hybrids could be partitioned into general (GGS) and specific (SGS) genetic similarity

(Charcosset and Essioux, 1994), thus:

$$GS_{ij} = \mu + ggs_i + ggs_j + sgs_{ij} \qquad (2)$$

where μ is the overall mean genetic similarity (GS) of all possible line combinations, ggs_i and ggs_j are the general genetic similarities (GGS) of the inbred lines *i* and *j* respectively and sgs_{ij} is the specific genetic similarity (SGS) between the inbred lines *i* and *j*.

Both (r) simple and (r_S) Spearman's rank (Spearman 1904) correlation coefficients were employed. Heritability ($h^2 = \sigma_g^2 / \sigma_p^2$, where σ_g^2 = estimate of genotypic variance, σ_p^2 = estimate of phenotypic variance) was calculated taking the 62 crosses as entries for parental mean, F1-hybrids and F2-hybrids for the three studied traits.

The similarity matrix based on the AFLP data was used to construct a dendrogram employing the unweighted pair-group method with arithmetic means (UPGMA). Associations among the inbred lines using Principal Coordinate Analysis (PCoA) according to Gower (1966) were analyzed using NTSYS-pc version 2.1q (Rohlf, 2000). Bootstrap analysis was performed by drawing 1000 random samples of the 1202 fragments employed with replacement, to determine the confidence limits of the UPGMA based dendrogram using the software package „WinBoot" developed at IRRI (Yap and Nelson, 1996).

RESULTS

Field trials

Differences in yield, 1000-seed weight and plant height of the checks (Scirocco and Limbo) in the trials of the parental lines and the hybrid trials were not significant (details not shown). Mean, minimum and maximum values for performance and heterosis of seed yield, 1000-seed weight and plant height across the four environments of the 62 F1-hybrids, their parental lines and F2-hybrids are presented in Table 3. F1-hybrids yielded 4.8 t ha^{-1} on the average, with a surplus in seed yield ranging from 40 to 119% and a mean of 70% above their corresponding parental mean. The highest yielding F1-hybrid was Styria x Alfred with 6.58 t ha^{-1} and the lowest seed yield of 3.09 t ha^{-1} was recorded for the cross Maya x Troy. A smaller increase in 1000-seed weight and plant height due

to heterosis was detected: a mean heterosis of 94 g for 1000-seed weight and 20 cm for plant height, amounting to about 20% heterosis in both traits. The highest heritability estimate for F1-hybrid performance (h^2 = 0.97) occurred for 1000-seed weight followed by h^2 = 0.79 for seed yield and h^2 = 0.68 for plant height. Heritability estimates for a given trait were similar in the three generations except for plant height, where it was somewhat lower in F2 (Table 3; Table 4).

Estimates of the variance components of GCA and SCA for the F1-hybrids and F2-hybrids as shown in Table 5 were highly significant for all three traits. For seed yield, the *per se* performance of the 14 experimental lines (details not shown) was significantly correlated with r_S = 0.417 (P<0.01) to their GCA-effects estimated in F1-hybrids and with r_S = 0.127 (P>0.05) to their GCA-effects estimated from F2-hybrids. For 1000-seed weight, these relationships were similar, (r_S = 0.466, P<0.01 and r_S = 0.122, P>0.05, respectively), but much weaker for plant height (r_S = 0.231, P>0.05 and r_S = 0.002, P>0.05, respectively). The variances due to GCA of the experimental lines were generally larger than that due to SCA of the F1-hybrids. Variance of SCA was very small compared to GCA in 1000-seed weight (~2%) and small in plant height (~22%), but rather large in seed yield (~63%) when compared to the GCA variance estimated in F1-hybrids. Comparable results were shown for F2-hybrids with the SCA variance being only 1% for 100-seed weight, 24% for plant height and 42% for seed yield of the GCA. The reduction of variance components from F1 to F2 was larger for SCA than for GCA in seed yield and 1000-seed weight and similar in plant height.

AFLP analysis

Testing 32 *Eco*RI/*Mse*I primer combinations showed the usefulness of 26 combinations that amplified 1662 fragments, 1202 of which were polymorphic (Table 2), with an average of 46 polymorphic fragments per primer combination. Genetic similarity values of those 62 pairs of inbred lines tested as hybrids in the field ranged from GS = 0.736 for the inbred lines Merkur and Mars to GS = 0.561 between the inbred lines Styria and Maris Bead (cf. Table1). The overall highest value was GS = 0.833 for L1 vs. L2 (not tested as hybrids), followed by GS = 0.736 for Mars x Merkur and GS = 0.705 for Maya x Troy (both hybrids were tested). The GS values among the four testers were rather similar, the smallest GS value was GS = 0.611 for Merkur vs. Scirocco and the largest value was GS = 0.646 for Merkur vs. Styria, hence, none of the four testers was markedly similar or

Table 3 Heritability (h^2) estimates of parental mean, F1- and F2-hybrid performance in addition to mean, minimum and maximum values (also for F1-heterosis) across environments for seed yield (t ha^{-1}), 1000-seed weight (g) and plant height (cm)

	Yield (t ha-1)	1000 seed- weight (g)	Plant height (cm)
Parental mean			
Mean	2.96	468.78	100.75
Min.	2.16	360.35	87.49
Max.	4.17	722.84	113.15
h^2	0.76	0.98	0.74
F1 performance			
Mean	4.81	563.49	120.69
Min.	3.09	415.06	99.56
Max.	6.58	860.74	134.71
h^2	0.79	0.97	0.68
F2 performance			
Mean	3.78	510.17	109.39
Min.	2.35	374.9	95.33
Max.	5.07	754.92	121.82
h^2	0.78	0.97	0.58
F1- heterosis			
Mean	1.85	94.71	19.94
Min.	0.9	17.05	7.7
Max.	3.02	167.55	30.35

Table 4 Variance component estimates of misprint, F1- and F2-hybrid performance for seed yield (t ha^{-1}), 1000-seed weight (g) and plant height (cm) across environments

Variance component	DF	Midparent	F1	F2
Yield				
σ^2_G	61	0.188**	0.504**	0.345**
σ^2_{GxE}	183	0.243§	0.528	0.395
1000-seed weight				
σ^2_G	61	9259**	12042**	7933**
σ^2_{GxE}	183	894	1302	1114
Plant height				
σ^2_G	61	26.63**	39.06**	18.59**
σ^2_{GxE}	183	37.43	72.94	54.87

** Significant at the 0.01 level of probability

§ Significance of σ^2_{GxE} was not tested

Table 5 Variance component estimates of F1- and F2-hybrids for the GCA of the14 experimental lines (σ_L^2) and their SCA (σ_{TxL}^2) of seed yield (t ha^{-1}), 1000-seed weight (g) and plant height (cm) across environments

Variance Component	DF	F1	F2
Yield			
σ_L^2	13	0.099**	0.084**
σ_{TxL}^2	39	0.062**	0.035*
1000-seed weight			
σ_L^2	13	13101**	8772**
σ_{TxL}^2	39	224**	85*
Plant height			
σ_L^2	13	32.00**	19.71**
σ_{TxL}^2	39	7.15**	4.81*

*,** Significant at the 0.05 and 0.01 levels of probability, respectively
ns not significant

dissimilar to any other tester. Similarly, the mean GS of the four testers to the 14 experimental lines was GS = 0.614. The dendrogram (Fig. 1) confirmed previously known pedigrees of experimental lines (Table 1): The lines L1 & L2 clustered tightly together (GS = 0.833), also the lines Maya & Troy clustered together (GS = 0.705). Yet, the presumably unrelated lines Mars & Merkur clustered together with a relatively high similarity (GS = 0.736). A high confidence level resulted from bootstrapping, with *P* = 100% for the lines Maya & Troy and L1 & L2 (Fig. 1). Although not reaching the 95% bootstrap *P* value (Felsenstein, 1985) the lines Mars and Merkur showed a *P* = 79.2%.

The first three principal co-ordinates of the PCoA (Fig. 2) explained about 24% of the total variability. Associations among lines were presented after joining the lines L1 and L2 into one entry (L1\L2), thus avoiding bias in the ordination result due to their extremely high similarity (GS = 0.883). No clearly separated groups were observed except for the entries Alfred, L1\L2 and Victor. The line Maris Bead was shown to be rather distant from all other lines, as shown from the third principal coordinate (Fig. 2).

Association of genetic similarity with hybrid performance

Correlations between the GS of the 62 pairs of parents and the performance of their 62 F1-hybrids were not significant for seed yield (r_S = -0.045) and 1000-seed weight (r_S = -0.166). For plant height, the correlation was even positive (r_S = 0.157). Correlation between F1-heterosis and GS was also not significant for all three traits being r_S = -0.226 for 1000-seed weight, r_S = -0.079 for seed yield (Fig. 3) and r_S = -0.093 for plant height. As shown in Fig. 4, significant correlations were found between SGS and SCA only for seed yield (r = -0.332; P<0.05) and 1000-seed weight (r = -0.412; P<0.01); for plant height the correlation was not significant (r = -0.251). These correlations are obviously inflated by the extreme behavior of the cross Troy x Maya. Hence, when applying Spearman's rank correlation, associations were not significant as shown in Figure 4a (r_S = -0.207), 4b (r_S = -0.234) and 4c (r_S = -0.175), respectively.

A positive highly significant correlation of intermediate strength (r_S = 0.558) was shown between F1-heterosis and F2-heterosis for seed yield (Fig. 5). For the other two traits, also highly significant positive correlations of r_S = 0.482 for 1000-seed weight) and r_S = 0.316 for plant height were detected. The SCA-effects calculated for the 56 F1-

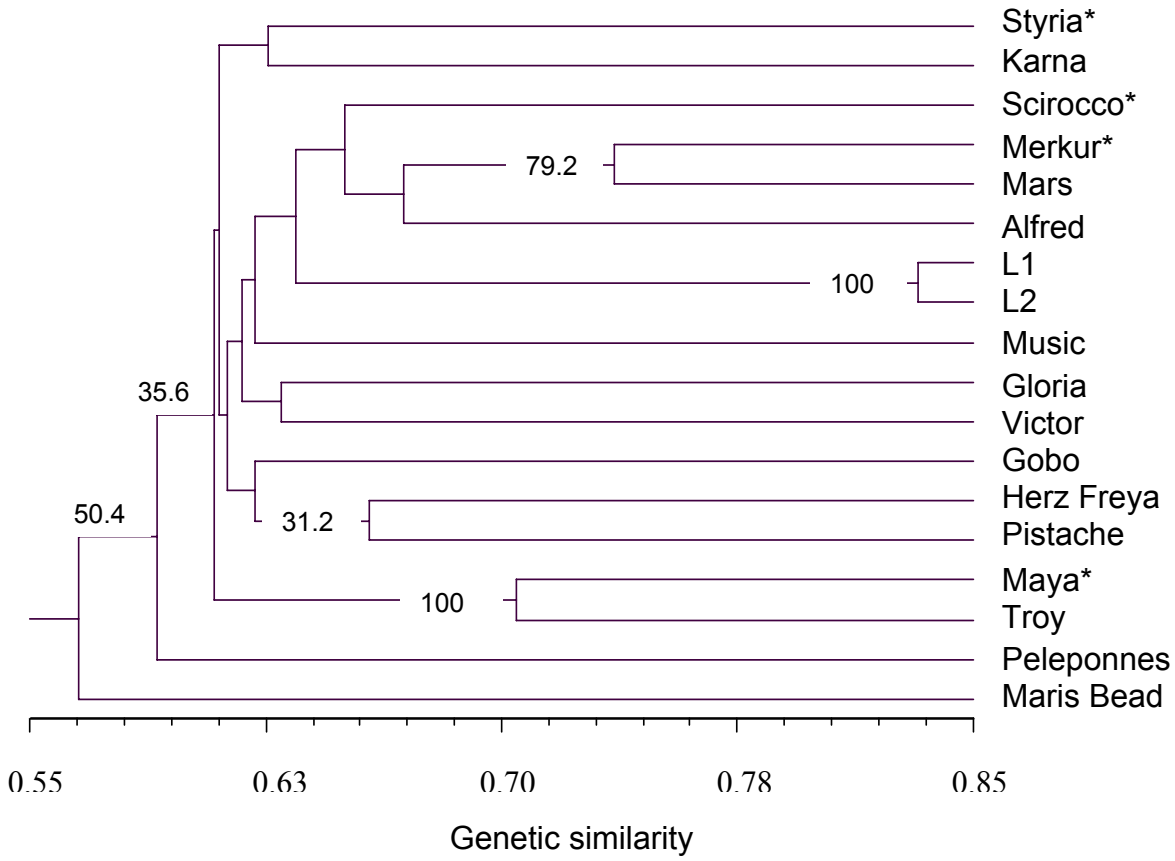


Fig. 1 Grouping of 18 faba bean lines based on 1202 polymorphic fragments using Jaccard's coefficient of similarity and UPGMA clustering (numbers shown at different nodes represent percentage confidence limits of the bootstrap analysis). * Lines were used as testers in the factorial analyses

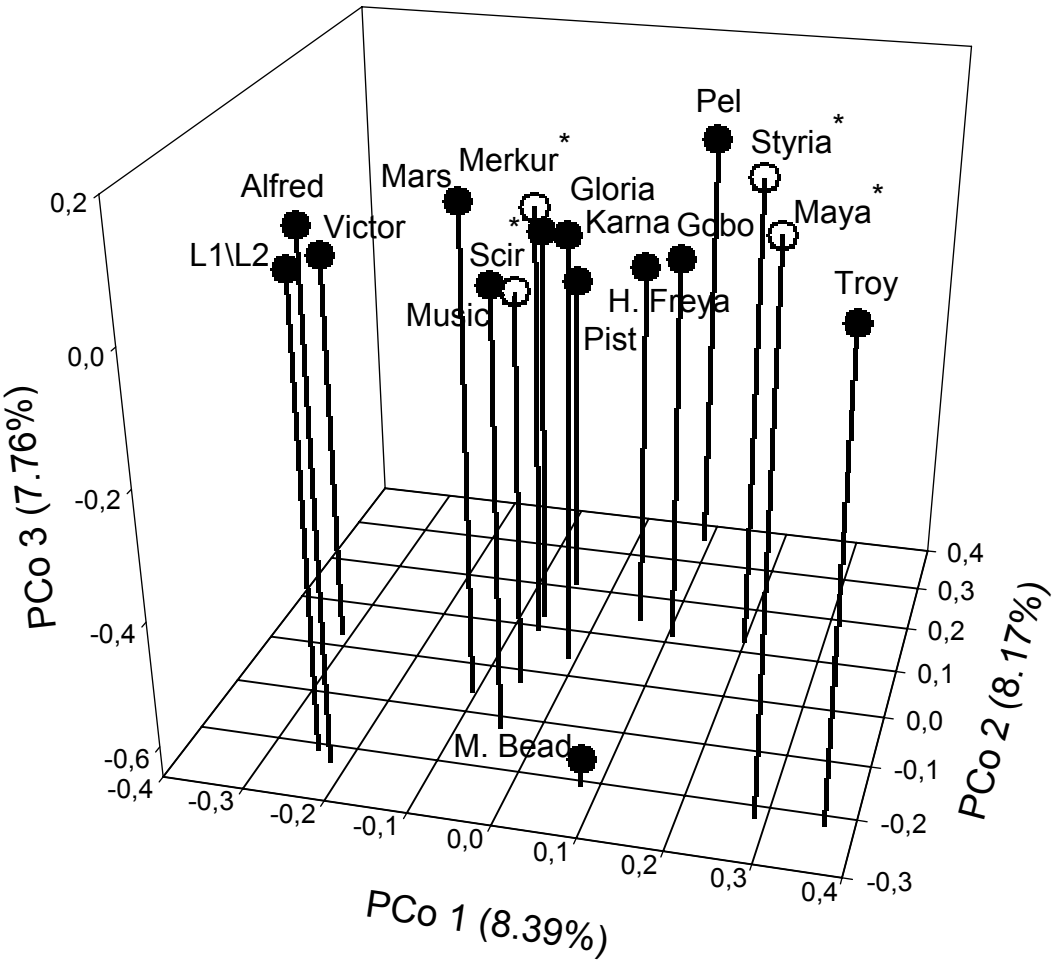


Fig. 2 Associations among 18 faba bean inbred lines (see text for details) revealed by Principal Coordinates Analysis performed on 1202 polymorphic AFLP fragments. Solid circles represent the experimental lines and empty circles denote the four testers

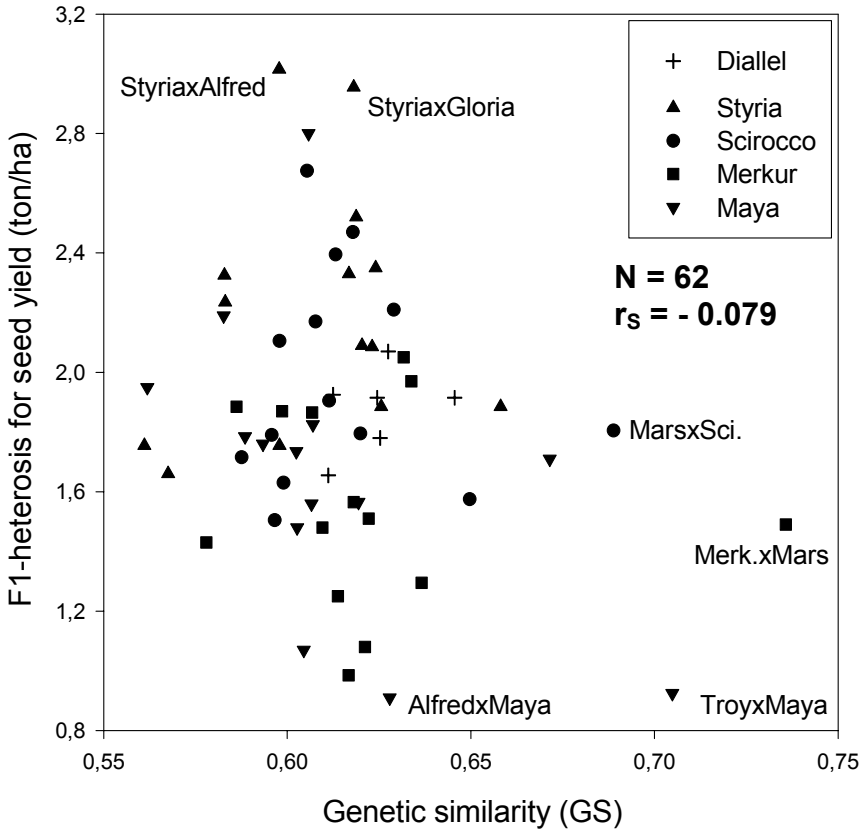
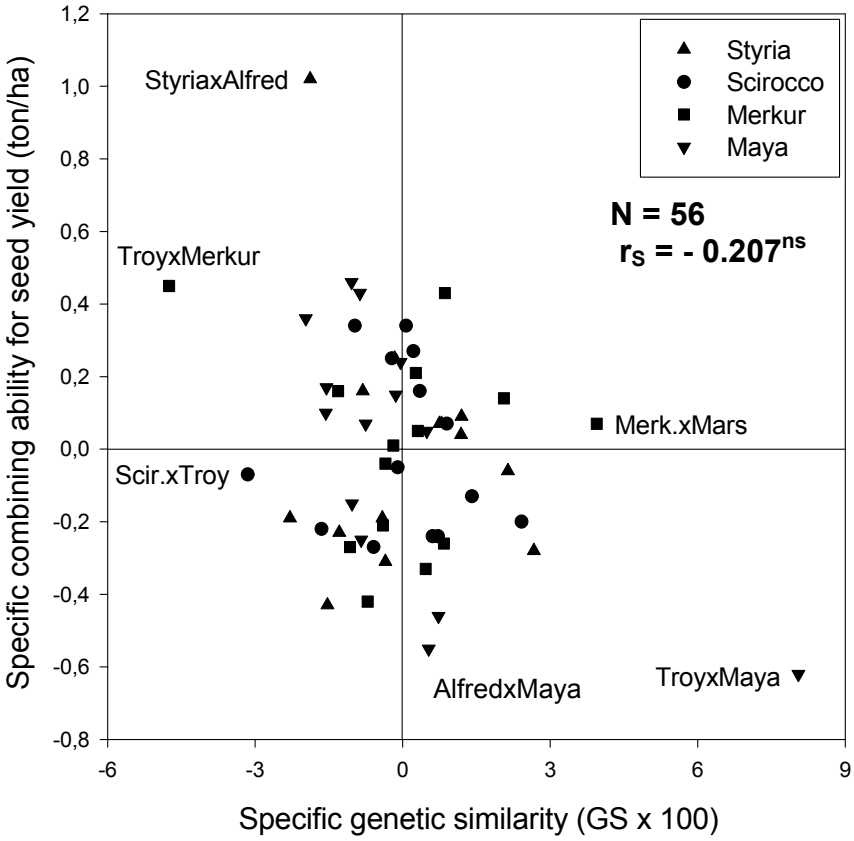
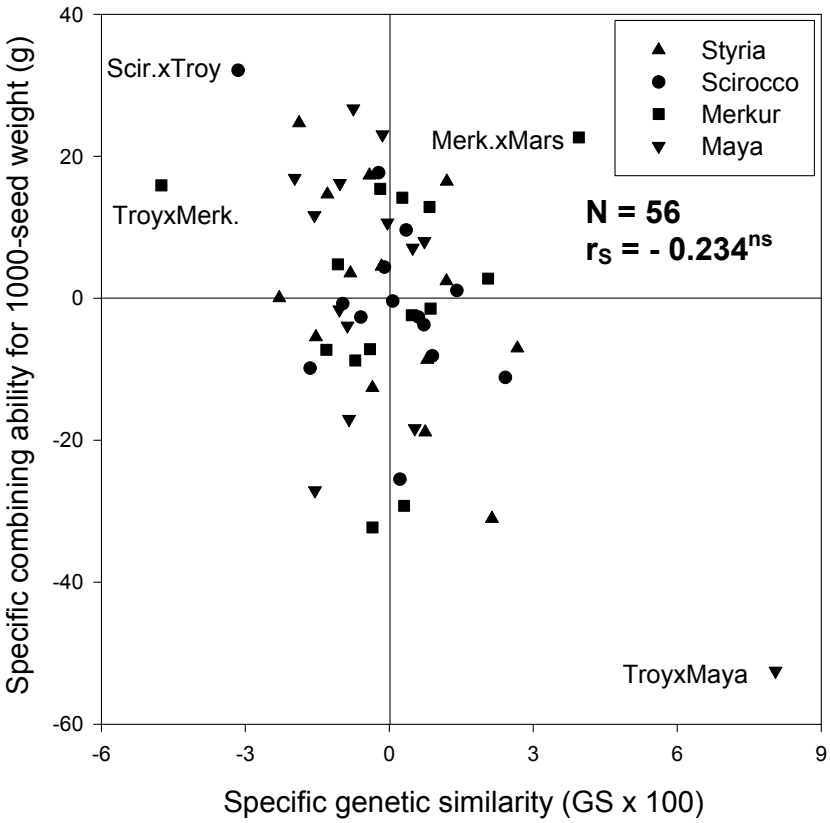


Fig. 3 Genetic similarity (GS) vs. heterosis of 62 F1-hybrids for seed yield

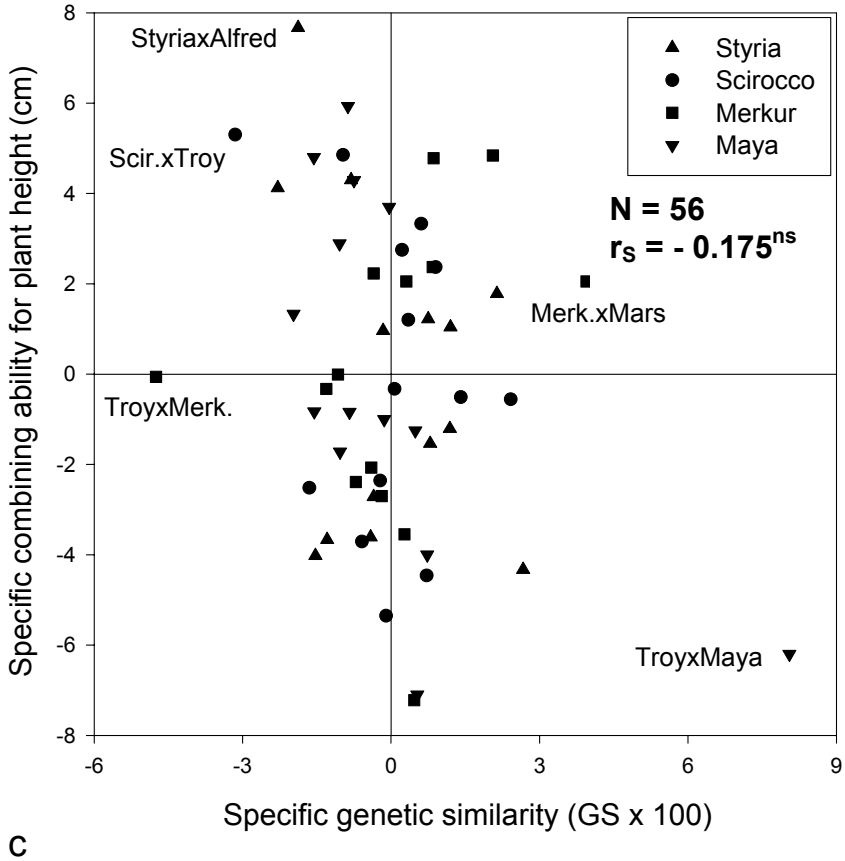


a

Fig. 4 Specific genetic similarity (GSx100) vs. specific combining ability for (a) seed yield (b) 1000-seed weight (c) plant height in a factorial design of four testers (Styria, Scirocco, Merkur, Maya) and 14 experimental lines


Specific combining ability for 1000-seed weight (g)
Specific genetic similarity (GS x 100)
Styria
Scirocco
Merkur
Maya
Scir.xTroy
Merk.xMars
TroyxMerk.
TroyxMaya
N = 56
r_S = - 0.234^{ns}
40
20
0
-20
-40
-60
-6
-3
0
3
6
9

b


Specific combining ability for plant height (cm)
Specific genetic similarity (GS x 100)
Styria
Scirocco
Merkur
Maya
StyriaxAlfred
Scir.xTroy
Merk.xMars
TroyxMerk.
TroyxMaya
N = 56
r_S = - 0.175^{ns}
8
6
4
2
0
-2
-4
-6
-8
-6
-3
0
3
6
9

c

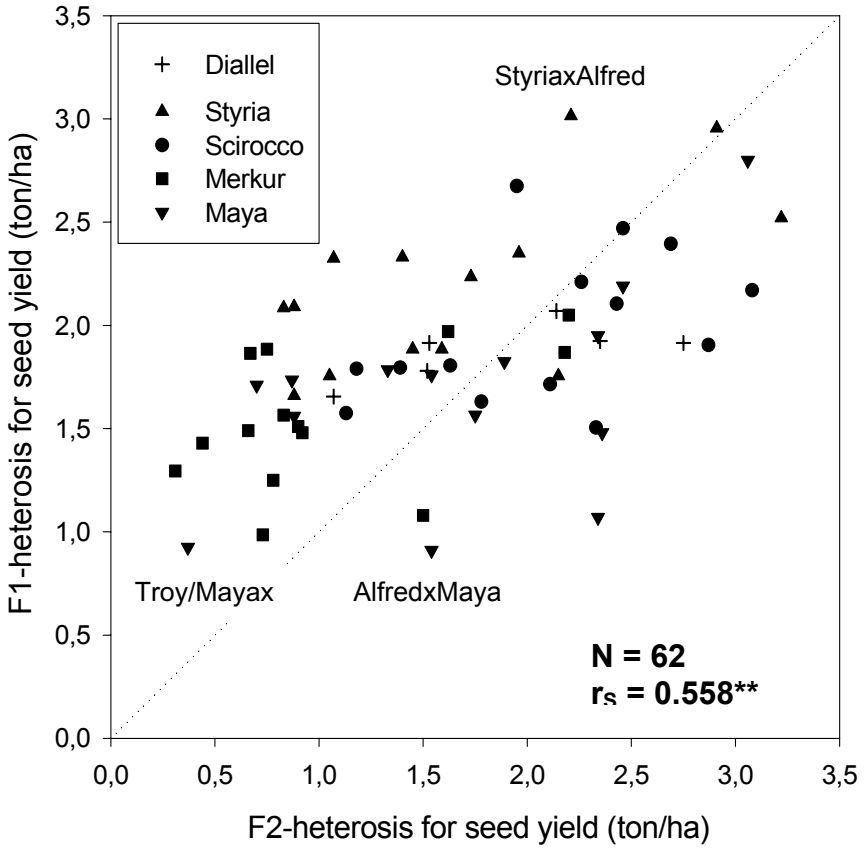


Fig. 5 F2-heterosis vs. F1-heterosis for 62 crosses seed yield

hybrids and F2-hybrids were correlated with r_S = 0.440 in case of seed yield, r_S = 0.400 for 1000-seed weight and r_S = 0.359 for plant height all being highly significant.

DISCUSSION

The dendrogram of the 18 entries (Fig. 1) clustered the lines L1 and L2 together as well as Maya and Troy. These results are in agreement with the available pedigree data (L1 and L2 originated from the cross Minica x Canner Express, and Troy was a parent of Maya; cf. Table 1). The high confidence level estimated by the bootstrap analysis for those relations supports this clustering. A marked relationship was detected between the lines Mars and Merkur (GS = 0.735). Available pedigree data regarding those two lines do not explain this close relationship. The lines Peleponnes and Maris Bead joined the cluster at its end (Fig.1), showing their small genetic similarity to the other lines. The PCoA (Fig. 2) illustrates that the four testers are of different genetic background. The entries Alfred, Victor and L1\L2 fall together in one group, which is in agreement with the pedigree data showing a common parental cultivar (Minica) for all of them.

Results on performance of the 62 single crosses showed superiority of the hybrids above the check varieties (Limbo and Scirocco) with on average 43% for seed yield. The highest yielding hybrid (Styria x Alfred) outyielded the checks by 96%. Similar results on hybrid superiority were reported by Link and Ruckenbauer (1988) and Ebmeyer (1988) in faba bean. In this study highly significant heterosis were detected for flowering date (2 days earlier) and lodging (23% increase), while maturity was not much affected.

In faba bean, F1-hybrid seed production for scientific field trials is very demanding and expensive. To produce hybrids, hand emasculation and fertilization of single flowers is performed under pollen isolation conditions. Use of F2 seeds instead of F1 would cut back the costs of hybrid trials dramatically. Provided that the F2-hybrids are good predictors for heterosis in F1, F2-hybrids would be taken as a substitute for F1-hybrids in pertinent field trials. The correlation between F1-heterosis and F2-heterosis (Fig. 5) was highly significant (r_S = 0.558), however not adequate to predict the heterosis of F1-hybrids from that of their F2-progeny. Most probably the heterogeneity in the F2-plots did cause the difference between these two estimators of heterosis (cf. Link and Ruckenbauer, 1988).

The variance due to GCA of the 14 experimental lines for yield in F1-generation was only little larger than in F2 (Table 5). A reduction in variance from F1-generation to F2-generation is expected, since in F2 the dominance effects of the loci coding for yield contribute only half as much to the GCA-effects than in F1, whereas the contribution of the additive effects to GCA remain unchanged. With unknown allele effects and allele frequencies, no definite expectation as to the relation between these variances in F1-generations and F2-generations exists. The variance of SCA effects in F2 was markedly smaller than in F1. Here, the expectation is that in F2 the variance is ¼ of that in F1 (Weir and Cockerham, 1977; Duc et al., 1990). When taking the standard error of these variances into account (0.03 in F1 and 0.02 in F2), the results achieved here are in accordance with this expectation. Melchinger (1999) demonstrated the major increase of SCA variance if intra-pool crosses are included rather than excluded. Already from the findings of Schnell (1982) with a basic quantitative genetic model, it becomes clear that the dominance variance (i.e., SCA variance, Falconer and Mackay 1996) between intra-pool hybrids in larger than between inter-pool hybrids. The difference increases as the allele frequency of the gene-pools under consideration are increasingly different and as their mean frequency increasingly deviates from p = 0.5. Here, at least for yield SCA variance was too important to be neglected.

With predominance of the GCA variance over that of SCA, early testing of inbred lines would become more effective: promising hybrids could be identified based on GCA effects. Predominance of GCA makes hybrid breeding more efficient (Melchinger 1999).

This study demonstrated a non significant correlation between genetic similarity of faba bean parental lines from the elite European gene pool and heterosis of their hybrids expressed in seed yield, 1000-seed weight and plant height. Partitioning the GS estimates into their components (equation 2) resulted in a higher correlation between SGS and SCA, however, the values were still too low to be useful in predicting the best single crosses in faba bean.

Various reasons could attribute to these poor correlations. Markers used here are not mapped hence, their distribution across the genome is not known. Virk et al. (2000) compared the pattern of genetic diversity based on mapped and unmapped AFLP markers in rice; they showed that there was no advantage in using evenly distributed markers for

assessing diversity. Furthermore, wide distribution of AFLP markers across the genome has been reported in soybean (Keim et al., 1997) and in barley (Becker et al., 1995). Saliba-Colombani et al. (2000), on the other hand have shown that AFLP markers were grouped in clusters around putative centromeric regions in tomato. A further reason for the poor correlations might be an inadequate number of markers. Tivang et al. (1994) showed that precision of the estimated genetic similarity is improved as more marker loci are screened. Zeid et al. (2001) estimated the coefficient of variation (CV) for genetic similarities in faba bean; a CV of less than 4 % was achieved by employing 500 AFLP markers. With the 1202 polymorphic markers employed in this study, we assume quite a high level of precision of the estimated genetic similarity.

Melchinger et al. (1990b) explained that better marker coverage of the genome does not by itself provide a key to increase the predictive power for hybrid performance. Bernardo (1992) has set conditions for effective prediction of hybrid performance based on molecular similarity of parents: At least 30-50% of the QTLs have to be in linkage disequilibrium with molecular markers and not more than 20-30% of the markers have to be randomly dispersed or non-predictive for QTLs. Furthermore, Charcosset et al. (1991) explained that the QTLs which are not marked by marker loci and marker loci which do not mark any QTL play symmetrical roles and the consequence is a decrease in the correlation between the genetic similarities and heterosis. Here, we have no information on the percentage of AFLPs marking QTLs.

The approach to mark individual QTLs with markers and hence predict heterosis and performance is probably promising; still, it is fundamentally different from the idea to predict heterosis from parental fingerprints. For the first approach, linkage disequilibria between tightly linked markers and QTL alleles are needed. This was shown even to exist e.g. in a set of cultivated US maize germplasm (Ching et al., 2002). Within gene fragments of 300-500 bp size, 36% of SNPs (single nucleotide polymorphisms) showed highly significant disequilibrium. For the latter approach, to meet the conditions mentioned above, linkage disequilibria have to exist between QTL and less tightly linked and unlinked marker loci. In the data set of Ching et al. (2002), only 0.3% of the unlinked SNPs were in disequilibrium, hence, this material would not meet the conditions.

The basic task with our approach here is to predict degree of heterozygosity of F1-hybrids at the trait loci (QTLs) using the scored degree of heterozygosity at the marker loci (from fingerprints of parents). It should be made clear that the translation of heterozygosity at QTLs into trait heterosis anyway relies on the trait- and population-specific genetic causes of heterosis.

Given that the marker loci and the QTLs are unlinked, they should be regarded as two (independent) samples of the genome‘s loci. Hence, correlation of their degrees of heterozygosity depend on (i) whether they both are really representative for the genome and on (ii) whether there is marked variation between the hybrids for the genome-wide degree of heterozygosity. With monogenic and oligogenic traits, the QTLs will not represent the genome, still, for yield this assumption may hold. The variation for heterozygosity of the genomes of F1-hybrids depends on the structure of the population (Pritchard et al., 2000). There is a main cause for a marked variation: the population is composed of groups. If so, some parental lines of crosses are related, they originate from the same group, some are not related. Structured population show linkage disequilibrium and may meet the conditions mentioned above. Still, if mating of related parents is successfully and intentionally avoided, mating gametes are not in linkage disequilibrium, the conditions are not met. This was obviously the case in the material presented here: parental lines were carefully selected to exclude relationships between tester lines and experimental lines and experimental lines based on pedigree data. It is worth mentioning at this point that the cross Troy x Maya was already under field evaluation as pedigree information (Troy being a parent of Maya) became available. Presence of this cross obviously inflated the association between the SGS and SCA as indicated by a significant correlation of $r = -0.332$ ($P<0.05$) for seed yield and $r = -0.412$ ($P<0.01$) for 1000-seed weight. Neutralizing its major effect on the correlation coefficient by using Spearman’s rank correlation (Fig. 4 a, b & c) resulted in a poor association of $r_S = -0.207$ for seed yield and $r_S = -0.234$ for 1000-seed weight and $r_S = -0.175$ for plant height.

Our finding, a low correlation between marker similarity and heterosis, is not in contradiction to a negative relationship between genetic similarity and heterosis, but it reflects the fact that the crosses here did not include related parents (one exception). The variation of marker similarity observed here should be understood merely as sampling variance, barely reflecting a true (if any) variation of genome-wide heterozygosity of the

hybrids used here. Indeed, variance for the GS values of all parental combinations, i.e. supposedly including "true" variance (all possible intra-pool combinations) was GS = 1.36 x 10^{-3}, whereas the variance of those GS values that indeed represented the field-tested hybrids was only GS = 0.91 x 10^{-3}, and was even smaller when excluding the cross Maya x Troy (GS = 0.79 x 10^{-3}).

Results from this study illustrated the usefulness of AFLP markers as substitute for missing pedigree data. Yet, AFLP markers were not needed to avoid single crosses with related parents within the elite European faba bean gene pool: knowledge on pedigree data has been adequate and could be used efficiently.

Stelling et al. (1994) showed that both Minica and its offspring Alfred have been frequently used as parents for many cultivars in Germany. The four experimental lines tested here, sharing Minica as a parent (Alfred, Victor, L1 and L2) appear to be a promising nucleus for one perceptible group within the European elite gene pool. Even within the *minor* pool selection of parental lines has demonstrated promising results. The mean heterosis of the 50 intra-pool *minor* x *minor* crosses studied here amounted to 1.87 t ha^{-1}. Link et al. (1996) recommended the *minor* x *major* European inter-pool crosses for being superior to intra-pool crosses. On the other hand, Schill et al. (1998) did not favor the *minor* x *major* compared to the *minor* x *minor* crosses. The present data clearly demonstrated a marked amount of genetic diversity and heterosis available even within the European *minor* gene pool.

ACKNOWLEDGMENTS

This paper was dedicated to Professor Dr. Dr. h.c. F.W. Schnell on the occasion of his 90th birthday. His lectures and publications on heterosis are indispensable for researchers in this field. Financial support from the Gemeinschaft zur Förderung der privaten deutschen Pflanzenzüchtung e.V. (GFP) and BLE/BMVEL is gratefully acknowledged. We are very grateful for the cooperation of Dr. Loaf Sass, NPZ, Lembke KG.

REFERENCES

Abdelmula, A.A., W. Link, E. von Kittlitz, and D. Stelling. 1999. Heterosis and inheritance of drought tolerance in faba bean, *Vicia faba* L. Plant Breeding 118:485-490.

Ajmone Marsan, P., P. Castiglioni, F. Fusari, M. Kuiper, and M. Motto. 1998. Genetic diversity and its relationship to hybrid performance in maize as revealed by RFLP and AFLP markers. Theor. Appl. Genet. 96:219-227.

Becker, J., P. Vos, M. Kuiper, F. Salamini, and M. Heun. 1995. Combined mapping of AFLP and RFLP markers in Barley. Mol. Gen. Genet. 249:65-73.

Berthelem, P. 1970. Rapport d'activite des la Station d'Ameleoration des Plantes de Rennes (INRA), Rennes, France.

Bernardo, R. 1992. Relationship between single-cross performance and molecular marker heterozygosity. Theor. Appl. Genet. 83:628–634.

Bernardo, R. 1994. Prediction of single-cross performance using RFLPs and information from related hybrids. Crop Sci. 34:20-25.

Bond, D.A., J.M. Drayner, J.L. Fyfe, and G. Toynbee-Clarke. 1964. A male sterile bean (*Vicia faba* L.) inherited as a Mendellian recessive character. J. Agri. Sci. 63:229-234.

Bond, D.A., S.J. Brown, G.J. Jellis, M. Pope, J.A. Hall, and M.H.E. Clarke. 1986. Annual report of Plant Breeding Institute, pp. 42-46 Cambridge, UK.

Bond, D.A., G.J. Jellis, G.G. Rowland, J. Le Guen, L.D. Robertson, S.A. Khalil, and L. Li Juan. 1994. Present status and future strategy in breeding faba beans (*Vicia faba* L.) for resistance to biotic and abiotic stresses. Euphytica 73:151-166.

Boppenmaier, J., A.E. Melchinger, E. Brunklaus-Jung, H.H. Geiger, and R.G. Herrmann. 1992. Genetic diversity for RFLPs in European maize inbreds. I. Relation to performance of Flint x dent crosses for forage traits. Crop Sci. 32:895-902.

Charcosset, A., M. Lefort-Buson, and A. Gallais. 1991. Relationship between heterosis and heterozygosity at marker loci: A theoretical computation. Theor. Appl. Genet. 81:571-575.

Charcosset, A., and L. Essioux. 1994. The effect of population structure on the relationship between heterosis and heterozygosity at marker location. Theor. Appl. Genet. 89:336-343.

Ching, A., K.S. Caldwell, M. Jung, M. Dolan, O.S. Smith, S. Tingey, M. Morgante, and A.J. Rafalski. 2002. SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. BMC Genetics 3:19.

Cochran, W.G., and G.M. Cox. 1957. Experimental designs, 2nd ed. Wiley, New York.

Comstock, R.E., and H.F. Robinson. 1948. The components of genetic variance in populations of biparental progenies and their use in estimating the average degree of dominance. Biometrics 4:254-266.

Doyle, J.J., and J.L. Doyle. 1990. Isolation of plant DNA from fresh tissue. Focus 12:13-15.

Duc, G., 1980. Effect of the environment on the instability of two sources of cytoplasmic male sterility in faba bean FABIS 2:29.

Duc, G., N. Berthaut, R. Pelletier, and A. Carteron 1990. Study on the inheritance and relationship of some morphological and fertility characters in faba bean (*Vicia faba* L.). Plant Breeding 105:126-136.

Ebmeyer, E. 1988. Heterosis and genetic variances and their implications for breeding improved varieties of spring beans (*Vicia faba* L.). Plant Breeding 101:200-207.

Falconer, D.S., and T.F.C. Mackay. 1996. Introduction to Quantitative Genetics. 4th ed. Longman Group Ltd. Essex, England.

Felsenstein, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39:783-791.

Frei, O.M., C.W. Stuber, and M.M. Goodman. 1986. Use of allozymes as genetic markers for predicting performance in maize single cross hybrids. Crop Sci. 26:37-42.

Gower, J.C. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53:325-338.

Jaccard, P. 1908. Nouvelles researches sur la distribution florale. Bull. Soc. Vaud. Sci. Nat. 44:223-270.

Kittlitz, E.v. 1986. Some observations in reciprocal crosses between *Vicia faba* major. Biol. Zentrabl. 105:147-153.

Keim, P., J.M. Schupp, S.E. Travis, K. Clayton, T. Zhu, L. Shi, A. Ferreira, and D.M. Webb. 1997. A high-density soybean genetic map based on AFLP markers. Crop Sci. 37:537-543.

Lanza, L.L.B., Jr.C.L. de Souza, L.M.M. Ottoboni, M.L.C. Vieira, and A.P. de Souza. 1997. Genetic distance of inbred lines and prediction of maize single-cross performance using RAPD markers. Theor. Appl. Genet. 94:1023-1030.

Link, W. 1990. Autofertility and rate of cross-fertilization: crucial characters for breeding synthetic varieties in faba beans (*Vicia faba* L.). Theor. Appl. Genet. 79:713-717.

Link, W., and P. Ruckenbauer. 1988. Aspekte der Nutzung von Heterosis bei der Pferdbohne (*Vicia faba* L.). p.147-162. *In* Ber. 38. Arbeitstagung der Vereinigung österreichischer Pflanzenzüchter,1987, Gumpenstein, Austria.

Link, W., C. Dixkens, M. Singh, M. Schwall, and A.E. Melchinger. 1995. Genetic diversity in European and Mediterranean faba bean germplasm revealed by RAPD markers. Theor. Appl. Genet. 90:27-32.

Link, W., B. Schill, A.C. Barbera, J.I. Cubero, A. Filippetti, L. Stringi, E. von Kittlitz and A.E. Melchinger 1996. Comparison of intra- and inter-pool crosses in faba bean (*Vicia faba* L.). I. Hybrid performance and heterosis in Mediterranean and German environments. Plant breeding 115:352-360.

Link, W., W. Ederer, R.K. Gumber, and A.E. Melchinger 1997. Detection and characterization of two new CMS systems in faba bean (*Vicia faba* L.) Plant breeding 116:158-162.

Martsch, R., J. Vaupel, C.C. Schön, and W. Link. 2001. Faba bean hybrid breeding: Current results with CMS199 p. 232-233. *In* Fourth Conf. on Grain Legumes, AEP (ed.) Cracow, Poland.

Melchinger, A.E., M. Lee, K.R. Lamkey, and W.L. Woodman.1990a. Genetic diversity for restriction fragment length polymorphisms: Relation to estimate genetic effects in maize inbreds. Crop Sci. 31:1033–1040.

Melchinger, A.E., M. Lee, K.R. Lamkey, A.R. Hallauer, and W.L. Woodman. 1990b. Genetic diversity for restriction fragment length polymorphisms and heterosis for two diallel sets of maize inbreds. Theor. Appl. Genet. 80:488–496.

Melchinger, A.E. 1999. Genetic diversity and heterosis p. 99-118. *In* J.G. Coors, and S. Pandey (ed.) Genetics and exploitation of heterosis in crops. Crop Sci. Soc. of America, Madison, Wisconsin.

Pritchard, J.K., M. Stephens, N. Rosenberg, and P. Donnelly. 2000. Association mapping in structured populations. Am. J. Hum. Genet. 67:170-181.

Rohlf, F.J. 2000. NTSYS-pc. Numerical taxonomy and multivariate analysis system. Exeter Software, New York.

Saliba-Colombani V., M. Causse, L. Gervais, and J. Philouze. 2000. Efficiency of RFLP, RAPD, and AFLP markers for the construction of an intraspecific map of the tomato genome Genome 43:29-40.

Schill, B., A.E. Melchinger, R.K. Gumber, and W. Link. 1998. Comparison of intra- and inter-pool crosses in faba bean (*Vicia faba* L.). II. Genetic effects estimated from generation means in Mediterranean and German environments. Plant breeding 117:351-359.

Schnell, F.W. 1982. A synoptic study of the methods and categories of plant breeding. Z. Pflanzenzüchtg 89:1-18.

Smith, O.S., J.S.C. Smith, S.L. Bowen, R.A. Tenborg, and S.J. Wall. 1990. Similarities among a group of elite maize inbreds as measured by pedigree, F1 grain yield, grain yield, heterosis, and RFLPs. Theor. Appl. Genet. 80:833–840.

Smith, J.S.C., and O.S Smith. 1992. Fingerprinting crop varieties. Advances in agronomy, 47:85-140.

Spearman, C. 1904. The proof and measurement of association between two things. American Journal of Psychology 15:72-101.

Stelling, D., E. von Kittlitz, E. Ebmeyer, O. Sass, H. Jaiser, and W. Link. 1994. Erfolge und Perspektiven der züchterischen Verbesserung von Körnererbsen und Ackerbohnen in der EU. p131-147. *In* Ber. 45. Arbeitstagung AG Saatzuchtleiter 1994, Gumpenstein, Austria.

Tivang, J.G., J. Nienhuis, and O.S. Smith. 1994. Estimation of sampling variance of molecular marker data using the bootstrap procedure. Theor. Appl. Genet. 89:259-264.

Utz, H.F. 1997. A computer program for statistical analysis of plant breeding experiments. Version 2N. Institute of Plant Breeding, Seed Science and Population Genetics, University of Hohenheim.

Virk, P.S., H.J. Newbury, M.T. Jackson, and B.V. Frod-Lloyd. 2000. Are mapped markers more useful for assessing genetic diversity? Theor. Appl. Genet. 100:607-613.

Vaz Patto, M.C., A.M. Torres, A. Koblizkova, J. Macas, and J.I. Cubero. 1999. Development of a genetic composite map of *Vicia faba* using F2 populations derived from trisomic plants. Theor. Appl. Genet. 98:736-743.

Weir, B.S., and C.C. Cockerham. 1977. Two-locus theory in quantitative genetics. p. 347-369. *In* E. Pollack, O. Kempthorene and T.B. Bailey (ed). Proc. Intern. Conference on quantitative genetics, 1976, Iowa, Ames, USA.

Yap, I.V., and R.J. Nelson. 1996. WINBOOT: a program for performing bootstrap analysis of binary data to determine the confidence limits of UPGMA-based dendrograms. IRRI Disc. Pap. Ser. 14.

Zabeau, M., and P. Vos. 1993. Selective restriction fragment amplification: a general method for DNA fingerprinting. European Patent Office, publication 0 534 858 A1.

Zeid M., C.C. Schön, and W. Link. 2001. Genetic diversity in a group of recent elite faba bean lines. Czech J. Genet. Plant Breed 1:34-40.

Acknowledgments

Praise be to ALLAH for granting me the health, strength and time to finish this work.

I am truly in debt to Prof. Dr. Wolfgang Link for his excellent supervision, extreme patience and scientific and moral support.

The generous support of the DAAD during the language course and the years of my study in Germany is much appreciated. My personal thanks goes to Frau Margret Leopold for her kindness and understanding.

I am grateful to Prof. Dr. Heiko Becker for his kindness and never ending support.

My special thanks goes to Frau Chris Carolin Schön in Hohenheim, Stuttgart for her extremely valuable advices and her professional suggestions regarding the manuscript.

My appreciation to Prof. Dr. Finkeldey from the Department of Forest Genetics for his valuable suggestions and ideas regarding the manuscript.

Thanks are due to Regina Martsch (Göttingen) and Martina Russ (Hohenheim) for the excellent experiments in the field and for the countless crosses made in the greenhouse.

Thanks are also due to Bettina Onken for her indispensable technical support.

I am grateful to Dr. Wolfgang Ecke for his useful suggestions and his willingness to help, also to Sabine, Ulli, and Christian for the good and well organized team work.

I am thankful to Stefan who kindly revised the text, for Rubens, Karla, Seif, Andreas and Tesfaye for the moral support and my warmest thanks to all friends at the institute: Antja, Jens, Anja, Uwe, Rosi, Nicole, Sonja, Kemal, Adeferis, Tefera, Memmen, Andi, Agung, Serafin, Birgitte, Bernd, Hilke, Hom, Joe, Alexandra, John, Thomas, Volker; without you all around me, I could not have done it.

To all who shared their good times with me, I enjoyed every minute, thank you all.

CURRICULUM VITAE

Personal Data

Name	Mahmoud Mohamed Zeid
Date of birth	26 December 1970
Place of birth	Alexandria, Egypt

Education

1976-1985	Attended an elementary and junior high school at Victory College, Alexandria, Egypt
1986-1988	Attended a secondary school at Victory College, Alexandria, Egypt
1989-1992	Studied at the Faculty of Agriculture, Alexandria University, Alexandria, Egypt and graduated with a Degree of Bachelor of Agricultural Sciences
1992-1997	Studied at the Faculty of Agriculture, Alexandria University, Alexandria, Egypt and was awarded the Degree of Master Science in Crop Science
1999-2002	Studied at the Institute of Agronomy and Plant Breeding, Georg-August University, Göttingen, Germany and awarded a Ph. D. in Plant Breeding

Professional Career

1992-present time	Serving as teaching assistant at the Department of Crop Science, Faculty of Agriculture, Alexandria University, Egypt