Euphytica **139:** 207–216, 2004.

Hybrid performance and AFLP- based genetic similarity in faba bean

Mahmoud Zeid[1], Chris-Carolin Schön[2] & Wolfgang Link[1,*]
[1]*Institute of Agronomy and Plant Breeding, Georg-August-University, D-37075 Göttingen, Germany;* [2]*State Plant Breeding Institute, University of Hohenheim, D - 70593 Stuttgart, Germany; (*author for correspondance: e-mail: wlink@gwdg.de)*

Received 14 April 2004; accepted 10 September 2004

Key words: AFLP, faba bean, genetic similarity, heterosis

Summary

Successful prediction of heterosis and performance of F_1-hybrids from the genetic similarity of their parents based on molecular markers has been reported in several crops and can be very helpful in hybrid breeding. The relationship 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 F_1-hybrids and their F_2-progenies were evaluated in field trials in four environments in Germany for seed yield, 1,000-seed weight and plant height. Results clearly demonstrated a stable superiority of the hybrids over their inbred parents and elite check cultivars, and showed a marked and varying amount of heterosis. Parental seed yield and F_2-hybrid yield were promising as predictors for F_1-hybrids. AFLP analysis of the 18 inbred lines using 26 *Eco*RI/*Mse*I primer combinations resulted in 1202 polymorphic fragments. Cluster analysis based on genetic similarity estimates unambiguously identified pedigree-related inbred lines. No clear separation of the 18 inbred lines into subgroups was detected. Correlation coefficients between genetic similarity estimates and either heterosis or F_1-hybrid performance were small and not useful. Also correlations between specific genetic similarity and specific combining ability were too small for all traits to be of predictive value. Results showed that AFLP-based genetic similarities are not useful to predict the performance of hybrids or heterosis within the elite European faba bean gene pool.

Introduction

Faba bean (*Vicia faba* L.) is one of many crops where hybrid breeding has been suggested as a solution for improving seed yield and yield stability (Stelling et al., 1994; Martsch et al., 2001). Its yield instability has been the main reason why farmers have been reluctant to grow the crop in spite of its many advantageous features. Yield improvement of 50–70% above the mid-parent values was achieved in crosses either between small-seeded *Vicia faba minor* and large-seeded *Vicia faba major* lines (von Kittlitz, 1986) or even within the *minor* group (Link & Ruckenbauer, 1988). Compared to open pollinated or inbred cultivars, faba bean hybrids have also shown better adaptation to a wide range of abiotic conditions (Bond et al., 1994; Abdelmula et al., 1999). Stelling et al. (1994) came to the conclusion that three-way hybrids, double-cross hybrids, or single-cross hybrids grown in blends are the key to yield improvement and stability in faba bean. A stable CMS system for hybrid seed production is not yet available in faba bean, but ongoing field research is conducted to develop such a system (Martsch et al., 2001). Presently, faba beans are bred as line cultivars, populations or synthetic varieties (Schnell, 1982).

Research aimed at studying the genetic diversity in faba bean (Link et al., 1995; Zeid et al., 2003), and identifying quantitative trait loci (QTLs) for agronomic traits (Roman et al., 2002) provide important clues for optimizing line and population breeding. However, the exploitation of heterosis in hybrid cultivars remains the best approach to maximize yield and yield stability. The success of hybrid breeding depends on the breeders' ability to identify parental lines which perform well in hybrid combinations, i.e., show good combining ability; evaluation of lines for combining ability is costly and time consuming and requires extensive field testing. In faba bean, F_1-hybrid seed production for field trials is prohibitively expensive. To produce hybrid seed, manual emasculation and fertilization of flowers is performed under pollen isolation conditions, with 30–50% pod set, one pod yielding two to four seeds. Only few authors (e.g., Ebmeyer, 1988; Link et al., 1996; Schill et al., 1998) published field data from large, multi-environmental trials of manually produced experimental hybrid seed. The use of F_2 instead of F_1 seed in exploratory trials would drastically lower the cost of hybrid testing, provided the performance of F_2-hybrids is a good predictor of the performance of the corresponding F_1-hybrid. In addition, the efficiency of hybrid breeding programs could be even more enhanced if parental line performance *per se* is proven to be a good predictor of hybrid performance. One feature of parental lines that eventually can be assessed as predictor for hybrid performance is the extent of genetic similarity between the parental lines of an experimental hybrid: heterosis and hybrid vigor are often greater when genetically dissimilar parents are used than in case of genetically similar ones. Breeders are since long highly interested in the relationship between genetic similarity measures (based on molecular markers or on coefficients of co-ancestry) and heterosis of their F_1-hybrids. DNA-fingerprints were successfully used to develop such measures of similarity (e.g., Smith & Smith, 1992). In few cases only, useful relationships between DNA-based similarities and heterosis were presented, e.g., for *Brassica napus* L. (Riaz et al., 2001) and for *Oryza sativa* L. (Sun et al., 2002). The detailed analysis of Charcosset & Essioux (1994) and of Melchinger (1999) separated the case (i) where F_1-hybrids are produced from parents that belong to the same gene pool (intra-pool hybrids) from the case where (ii) parents belong to different gene pools (inter-pool hybrids). According to Melchinger (1999), a significant correlation between parental genetic dissimilarity and yield of F_1 crosses is expected only in the first case. The intra-pool case was further subdivided into two groups, one group consisting of crosses from parents with varying levels of coancestry, and a second group where parents are not related by pedigree. The latter situation generally has to be faced by the breeder with no significant correlation between genetic dissimilarity and performance of F1 crosses to be expected *a priori*.

The present study refers to the European faba bean gene pool, which has earlier been studied for its genetic diversity (Link et al., 1995) and for hybrid performance and heterosis (Ebmeyer, 1988; Link et al., 1996; Schill et al., 1998); heterotic groups within this pool have not been intentionally established, however, they may exist.

The objective of our study was to determine whether marker based genetic similarity and hybrid performance as well as heterosis of unrelated faba bean lines show a significant correlation. To achieve this objective, (I) available pedigree data were studied and unrelated lines were chosen to be mated, (II) the genetic diversity for AFLP markers and for hybrid performance and 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 inbred lines were 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 (Table 1). All lines, with the exception of lines L1, L2 and Peleponnes, belong to the European *minor* group and are considered a representative sample of the actual breeding material of the few European faba bean breeding programs. Four of these lines were used as testers (Table 1). Testers are unrelated and have outstanding agronomic potential. The testers were crossed manually to the remaining 14 experimental lines in a factorial design resulting in 56 single cross hybrids. In addition, the testers themselves were intermated in a diallel cross to generate six crosses. Thus a total of $56 + 6 = 62$ hybrids were studied. Except one case (Troy × Maya, cf. Table 1), testers were not pedigree-related with experimental lines. Seeds for lines, for F_1-progenies and F_2-progenies were produced under pollinator-free conditions, i.e., without access of honey bees (*Apis mellifera* L.) and bumble bees (*Bombus* sp.; Link, 1990). Seed for the F_2-progenies were bulk harvested from at least 40 selfed F_1-plants per hybrid.

For four of the crosses, having Gloria (Table 1) as a common maternal parent, the F_1-progenies could phenotypically be confirmed as being “true hybrids,” based on the monogenic recessive white flower color of Gloria (the other inbred lines are dominant, colored flowered).

Table 1. Pedigree and origin of the faba bean inbred lines used as parents of hybrids in the present study

Line[†]	Inbreeding generation	Source	Pedigree/Year of release
Testers			
1. Maya	3	Serasem/France	Troy × *equina*/1995
2. Merkur	3	Selgen/Czech Republic	Bolo × line × Fribo/1997
3. Scirocco	7	NPZ Lembke/Germany	[‡]/1992
4. Styria	12	Gleisdorfer SZ/Austria	Carola × Mythos/1996
Experimental lines			
1. Alfred	7	Cebeco/Netherlands	Minica × Horse bean/1983
2. Gloria	8	Wieselburger SZ/Austria	Kornb. Kl.körnige/1994
3. Gobo	8	Germany (GDR)	[‡]/1987
4. Herz Freya	12	Herz/Germany	[‡]/1935
5. Karna	6	Wieselburger SZ/Austria	Kornb. Kl.körnige/1983
6. L1[§]	7	Germany	Minica × Canner Express
7. L2[§]	7	Germany	Minica × Canner Express
8. Maris Bead	3	PBI/England	[‡]/1965
9. Mars	3	Vanisco/Denmark	Alfred × Maribo/1993
10. Music	3	Blondeau/Frankreich	[‡]/1995
11. Pistache	7	Joordens/Netherlands	[‡]/1990
12. Peleponnes	12	Greece/ICARDA	[‡]/[‡]
13. Troy	11	NPZ Lembke/Germany	[‡]/1985
14. Victor	10	Cebeco/Netherlands	Minica × Cocksfield/1988

[†] Names of inbred lines are identical with names of cultivars from which they were developped.
[‡] Pedigree information/year of release is not available.
[§] *Vicia faba major* (large-seeded, vegetable type).

No white flowering plants were found in the respective field and propagation plots. All 62 F_1-hybrids, their F_2-progenies, their 18 parental lines and two checks 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 6 m^2 (1.5 × 4.0 m) with a density of 35 plants/m^2 were sown within the first 10 days of April and harvested at the end of August. The experimental layout was a 5 × 4 lattice with four replicates for the parents and an 8 × 8 lattice with two replicates for the F_1-hybrids and the F_2-hybrids. The F_1- and F_2-hybrid trials shared a common field area. Faba bean hybrids are generally taller than their inbred parents. Thus, to avoid competition effects as reported by Stelling (1992) trials of parental lines were not grown in the same experiment as F_1- and F_2-hybrids but in a directly adjacent experiment.

Scirocco, a widely grown cultivar, and Limbo, a new cultivar on the market, were included as checks in each of these trials. Seed yield (Mg ha^{-1}), 1,000-seed weight (g), plant height (cm) and other morphological characters including flowering date, lodging and days to maturity were recorded on a plot basis. Yield and 1,000-seed weight were determined on 14% moisture content basis; three samples of 50 seeds per plot were counted and weighed to calculate 1,000-seed weight per plot. Plant height was measured from the ground to the top vegetative bud at physiological maturity on one to five typical plants in the center of each plot.

AFLP analysis

DNA for AFLP analysis was isolated from young leaves of single plants of each of the 18 inbred lines according to Doyle & Doyle (1990). AFLP reactions were performed according to Zabeau & Vos (1993), using a commercially available kit (AFLP analysis System I, GIBCO BRL, Life Technologies, Inc.) and following the manufacturer's instructions. After pre-screening of

32 *Eco*RI/*Mse*I primer combinations, 26 proved to be useful and were further employed in this study. The *Eco*RI primers were radiolabeled with [γ-33 P]-dATP provided from Amersham Pharmacia Biotech (D-Freiburg). After selective amplification, the PCR-products were mixed with loading buffer, denatured and placed on ice. Seven μl of the mixture were loaded on a polyacrylamide gel. For each of the tested primer combinations, samples of all 18 inbred lines were run on the same gel. Two neighboring lanes together represented one inbred line. After electrophoresis, gels were fixed in 10% acetic acid and dried and 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. (2003).

Data analyses

The analysis of variance (ANOVA) for agronomic traits was performed on the data of each environment separately, followed by a combined analysis across environments based on the lattice adjusted entry means according to Cochran & Cox (1957) and following Boppenmaier et al. (1992), using PLABSTAT 2O (Utz, 2001). All effects in the statistical model were considered random. For one environment, only one replicate was available for yield and 1,000 seed-weight of the F_1- and F_2-hybrids. Here, weighted error terms were calculated for the combined analysis across environments as proposed by Cochran & Cox (1957) and implemented in PLABSTAT (Utz, 2001). Differences in checks' performance when grown in the parents trials *versus* when grown in the F_1- and F_2-hybrids trials were tested for significance for the four environments by ANOVA. Heritability (h) was calculated for the 18 parental lines and for the F_1-hybrids and the F_2-hybrids of the 62 crosses for the three studied traits as:

$$h^2 = \sigma_G^2/\sigma_P^2,$$

where σ_G^2: estimate of genotypic variance, σ_P^2: estimate of phenotypic variance; the expected composition of σ_P^2 was $\sigma_P^2 = (1/\mathrm{RE})\sigma^2 + (1/\mathrm{E})\sigma_{G\times E}^2 + \sigma_G^2$, where σ^2: estimate of error variance, $\sigma_{G\times E}^2$: estimate of genotype × environment interaction variance; R: number of replicates, and E: number of environments.

Midparent heterosis of each cross was calculated as [F_1-heterosis = F_1 – parental mean]; the estimation was based on genotype means in each environment and averaged across environments. Relative midparent heterosis was calculated referring to the parental mean as 100%.

A factorial analysis of general (GCA) and specific (SCA) combining ability was performed with the lattice-adjusted F_1 and F_2 data of the 56 factorial crosses, according to the North Carolina Design II (Comstock & Robinson, 1948), partitioning the mean performance of F_1-hybrids and F_2-hybrids (Y_{ij}) between inbreds *i* and *j* to be

$$\mathrm{Y}_{ij} = \mu + \mathrm{GCA}_i + \mathrm{GCA}_j + \mathrm{SCA}_{ij}$$

where μ is the overall mean performance of all possible hybrids, GCA_i and GCA_j are the GCA of the inbred lines *i* (testers) and *j* (experimental lines), respectively and SCA_{ij} is the SCA between inbred lines *i* and *j*. The testers, the experimental lines and the environments were considered as random factors.

AFLP-based genetic similarity estimates (GS) between all possible pairs of inbred lines were 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 F_1-hybrids could be partitioned into general (GGS) and specific (SGS) genetic similarity (Charcosset & Essioux, 1994), thus:

$$\mathrm{GS}_{ij} = \mu + \mathrm{GGS}_i + \mathrm{GGS}_j + \mathrm{SGS}_{ij}$$

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

Spearman's rank correlation coefficient was employed throughout, using PLABSTAT (Utz, 2001).

The similarity matrix based on the AFLP data was used to construct a dendrogram employing the unweighted pair-group method with arithmetic means (UPGMA). Bootstrap analysis to determine the confidence limits of the UPGMA based dendrogram was performed by drawing 1,000 random samples of the 1,202 fragments employed with replacement, using the software package "WinBoot" developed at IRRI (Yap & Nelson, 1996).

Results and discussion

Field trials

The average performance of the two checks (*cv.* Scirocco and *cv.* Limbo) was not significantly different between the trials of the parental lines and the F_1- and F_2-hybrids for seed yield, 1,000-seed weight and plant height ($P > 0.05$; data not shown). Genetic variances among parental lines as well as among the F_1- and F_2-hybrids were highly significant for all traits (for grain yield shown in Table 2).The largest genetic variance for grain yield of the three generations was found among the 62 F1-hybrids. Genotype × environment interaction effects were a significant and important source of phenotypic variation in all three generations. Heritability estimates for seed yield were $0.72 < h^2 < 0.79$; Table 2).

The highest yielding parental line in a single environment as well as on average across the four environments was the line Styria. Its mean yield across the four environments (4.51 Mg ha^{-1}) was 64% higher than the mean grain yield of all 18 lines. The check *cv.* Scirocco was slightly higher yielding (4.29 Mg ha^{-1} > 4.28 Mg ha^{-1}) than the check *cv.* Limbo at Göttingen in 2002, whereas Limbo outyielded Scirocco at the other three environments; Limbo's mean yield across the four environments was 3.72 Mg ha^{-1}, hence 35% higher than the mean of the 18 parental lines and 21% lower than the line Styria.

The superiority of the F_1-generation above the check *cv.* Limbo (4.81 Mg ha^{-1} > 3.72 Mg ha^{-1}; Table 3)was 29% for seed yield, proving the attractivity of F_1-hybrids in general when compared with most recent elite cultivars. The highest yielding F_1-hybrid was Styria × Alfred with 6.58 Mg ha^{-1} (Table 3), outyielding Limbo by 77% and the parental line Styria by 46%. When compared with Styria, a number of 21, 13, 13, and 4 F_1-hybrids were significantly ($P < 0.05$) higher yielding, respectively, in the four environments (details not shown). The three on average highest yielding F_1-hybrids (Styria × Alfred, Styria × Gloria, Styria × Scirocco; cf. Figures 1a–1d) outyielded their common parental line Styria by 41%, this superiority varied among the environments from 29% to 73% and was sifignicant for each hybrid in each single environment ($P < 0.05$, except two cases with $P < 0.10$), in spite of G × E being a marked and important source of variation (Table 3), demonstrating the dramatic improvement that could be achieved in faba bean with the introduction of hybrid varieties. Obviously, Styria inherits a high, stable and very promising heterosis and combining ability for seed yield to its hybrid offspring. From Figures 1a–1d it becomes obvious that the environment Hohenheim 2002 had a relatively large impact on the ultimate results. The three other environments showed lower correlations with means across environments ($r = 0.80$, $r = 0.77$, $r = 0.71$), hence, contributed more to the G × E interaction term (Figures 1a–1d; Table 3).

The highest yielding F_2-hybrid again traced back to Styria (Styria × Music), yielding 5.07 Mg ha^{-1}

Table 2. Variance component estimates, heritability and its confidence interval (CI) of 18 parental lines and of 62 factorial and diallel crosses in generations F_1 and F_2 for seed yield in four environments

Variance component	DF	Parents	DF	F_1	F_2
Yield (Mg ha^{-1})					
σ^2_G	17	0.35**	61	0.50**	0.35**
$\sigma^2_{G\times E}$	51	0.49**	183	0.37**	0.25**
h^2 (95% CI)		0.72 (0.32; 0.86)		0.79 (0.68; 0.86)	0.78 (0.66; 0.85)

**Significant at the 0.01 probability level.

Table 3. Mean, minimum and maximum performance of 62 F_1- and F_2-hybrids and their 62 corresponding parental mean values as well as midparent heterosis for seed yield, 1,000-seed weight and plant height evaluated in four environments

	Yield (Mg ha^{-1})	1,000 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
F_1 performance			
Mean	4.81	563.49	120.69
Min.	3.09	415.06	99.56
Max.	6.58	860.74	134.71
F_2 performance			
Mean	3.78	510.17	109.39
Min.	2.35	374.90	95.33
Max.	5.07	754.92	121.82
F_1-heterosis			
Mean	1.85	94.71	19.94
Min.	0.9	17.05	7.70
Max.	3.02	167.55	30.35
Check *cv.* Limbo	3.72	511.21	111.98

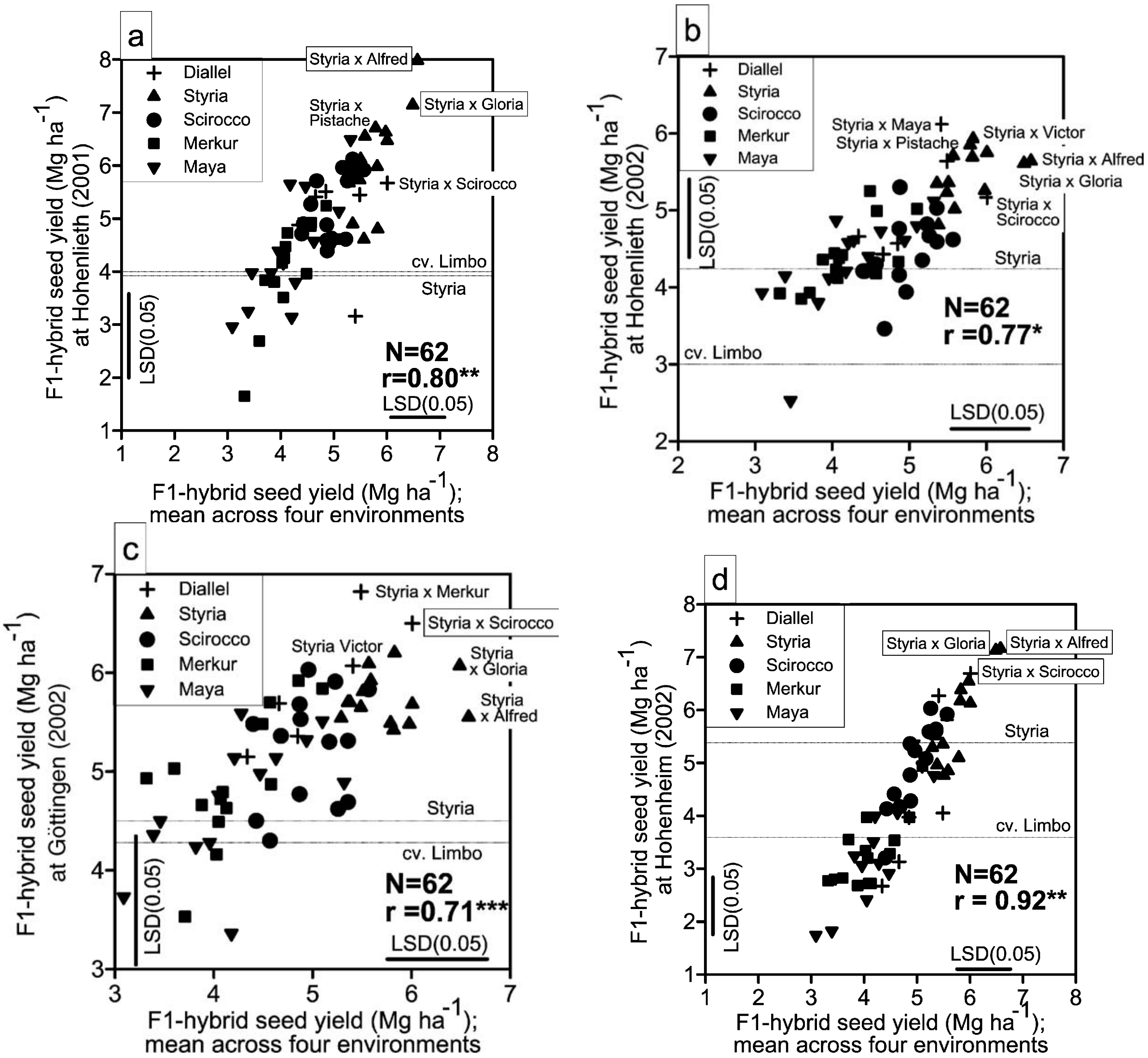

Figure 1. (a–d). F_1-hybrid grain yield at a single environment *vs.* F_1-hybrid grain yield averaged across four environments. If one of the three highest yielding hybrids across all four environments also belongs to the group of the locally three highest yielding ones it is named in frames.

(Figure 2a). Lowest seed yield in generations F_1 and F_2 was recorded for the cross Maya × Troy with 3.09 and 2.35 Mg ha^{-1}, respectively.

The F_1-hybrids realized a relative midparent heterosis of 70%, ranging from 40 to 119% among individual crosses. Compared to seed yield, midparent heterosis for 1,000-seed weight and plant height was less pronounced. Mean heterosis was 95 g for 1,000-seed weight and 20 cm for plant height, being about 20% for both traits (Table 3). Highly significant heterosis was detected for flowering date (2 days earlier) and lodging (23% increase), but not for maturity (data not shown). Similar results on hybrid superiority in faba bean were reported by Link & Ruckenbauer (1988) and Ebmeyer (1988). From basic quantitative genetic theory, the F_2-hybrids are expected to realize half of this midparent heterosis, i.e., the performance of an F_2-hybrid should be equal to the mean of its F_1-hybrid and its parental mean. Here, average seed yield in generation F_2, amounting to 3.78 Mg ha^{-1}, was very close to this expectation (0.5 [2.96+4.81 Mg ha^{-1}]) = 3.88 Mg ha^{-1}; cf. Table 3); for 1,000 seed-weight (516 g *vs.* 510 g) and plant height (111 *vs.* 109 cm), mean performance in F_2 was as well found to be very near to the

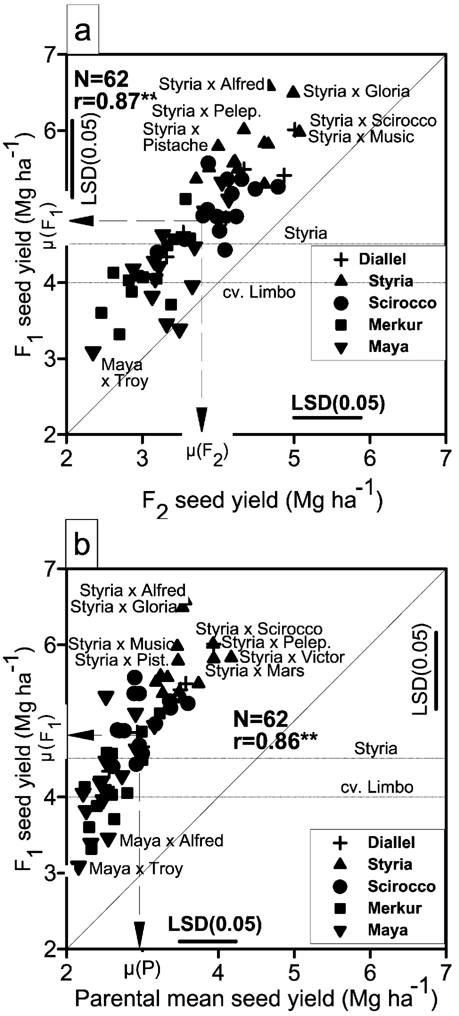

Figure 2. (a, b). F_1-hybrid *vs.* F_2-hybrid grain yield and F_1-hybrid *vs.* parental mean grain yield (means across four environments). Performance of check *cv.* Limbo and highest yielding line Styria are indicated.

expectation. The higher yielding F_1-hybrids (Figure 2a) did not tend to realize a smaller superiority compared to their corresponding F_2-hybrids than the lower yielding ones. Hence, in spite of surely less expensive F_2-seed it does not seem attractive for farmers to grow F_2-hybrids as substitutes for the more costly generation F_1.

The correlation coefficient between the parental means and their F1-hybrids for seed yield was $r = 0.86$ ($P < 0.01$; Figure 2b), very similar to the correlation between F2 and F1 ($r = 0.87$; $P < 0.01$; Figure 2a). These correlations were remarkably high and originated mainly from the on average lower yielding crosses with the two tester lines Merkur and Maya (F2 *vs.* F1 and parental means *vs.* F1: $0.69 < r < 0.77$, $P < 0.01$). For crosses with the higher performing tester lines Scirocco and Styria these correlations were smaller: not significantly different from zero, $r = 0.41$ and $r = 0.27$ ($P > 0.1$), respectively, in case of the correlation between parental means and F_1-hybrids; and of medium strength, $r = 0.57$ and 0.62, ($P < 0.05$), in case of F2- *vs.* F1-hybrids. The magnitude of the correlation is encouraging for hybrid breeding and can be used to predict hybrid performance, at least based on F2-hybrids (cf. Lonnquist & Lindsey, 1964). However, it is questionable whether these correlations will remain high once heterotic pools have been established.

Estimates of the variance components of GCA of the four testers and of the 14 experimental lines and of SCA as based on the 56 factorial crosses were highly significant for all three traits (data not shown). As expected from quantitative genetic theory for unselected material, SCA was an important source of variation; the SCA variance for grain yield was nearly as large as the GCA variance. These results show that the creation of heterotic gene pools would facilitate hybrid breeding because GCA effects become more important once heterotic pools have been established and the efficiency of early testing of inbred lines could be increased (Melchinger, 1999).

AFLP analysis

Testing 32 *Eco*RI/*Mse*I primer combinations showed the usefulness of 26 combinations that amplified 1,662 fragments, 1,202 of which were polymorphic, with an average of 46 polymorphic fragments per primer combination. Among all possible pairwise combinations the highest GS value amounted to $GS = 0.833$ for L1 *versus* L2 (not crossed, hence cross not field-tested; cf. Table 1), followed by $GS = 0.736$ for Mars × Merkur and by $GS = 0.705$ for Maya × Troy; the smallest value ($GS = 0.561$) was found for Styria × Maris Bead. The variance for the GS values of all parental combinations was $\sigma^2_{GS} = 1.36 \times 10^{-3}$, whereas the variance of the GS values from the field-tested hybrids was only $\sigma^2_{GS} = 0.91 \times 10^{-3}$ (with a mean value of $\mu(GS) = 0.615$); the reduction in variance presumably reflects the exclusion of crosses from pedigree-related parents. The variance became even smaller ($\sigma^2_{GS} = 0.79 \times 10^{-3}$ and $\mu(\text{GS}) = 0.614$) when excluding (the field-tested) cross Maya × Troy (*cv.* Troy was parent of *cv.* Maya).

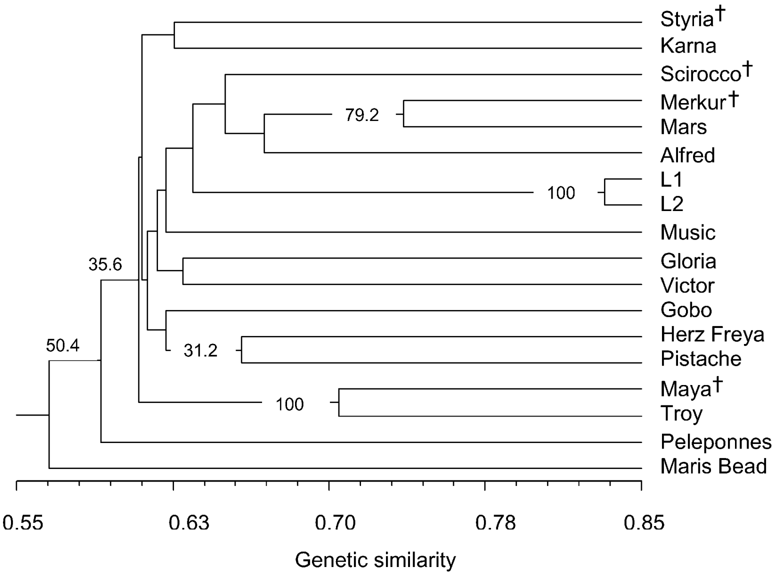

Figure 3. Grouping of 18 faba bean lines based on c using Jaccard's coefficient of similarity and UPGMA clustering (numbers shown at different nodes represent percentage confidence limits of the bootstrap analysis). †Lines were used as testers.

The dendrogram (Figure 3) confirmed previously known pedigree information of experimental lines (Table 1): the lines L1 and L2, originating from the same cross, clustered tightly together, as well as the related lines Troy and Maya. Yet, the presumably unrelated lines Mars and Merkur clustered together with a relatively high similarity ($GS = 0.736$). A high confidence level from bootstrapping with $P = 100\%$ resulted only for the lines Troy and Maya and for L1 and L2. For the genetic material tested, AFLP markers provided additional information compared to pedigree data. However, the dendrogram did not uncover a convincing separation of the 18 lines into possible heterotic groups.

Association of genetic similarity with hybrid performance

Correlations between the GS of the 62 pairs of parents and the performance of their 62 F_1-hybrids were not significant for seed yield ($r = -0.05$), 1,000-seed weight ($r = -0.17$), and plant height ($r = 0.16$). Correlations between F_1-heterosis and GS were also not significant for all three traits being $r = -0.08$ for seed yield (Figure 4), $r = -0.23$ for 1,000-seed weight, and $r = -0.09$ for plant height. Partitioning the GS estimates into their components resulted in a slightly higher correlation coefficient between SGS and SCA (not shown): correlations between SGS and SCA for seed yield amounted to $r = -0.21$ ($P > 0.05$); for 1,000-seed weight it was $r = -0.23$ ($P > 0.05$), and for plant height $r = -0.18$ ($P > 0.05$). Focusing on the $N = 56$ factorial crosses, the tester Merkur showing the highest mean GS-value with the experimental lines ($GS = 0.622$) was found to give the smallest mean heterosis for seed yield (1.55 Mg ha^{-1}), whereas the

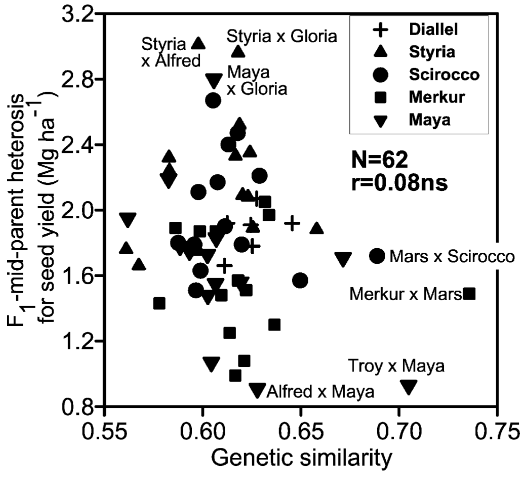

Figure 4. Mid-parent heterosis of 62 F1-hybrids for grain yield (assessed across four environments) *versus* AFLP-based genetic similarity.

least similar tester, Styria ($GS = 0.607$) exhibited the largest mean heterosis (2.20 Mg ha^{-1}). However, these relationships were still too weak to be useful in predicting the best single crosses in faba bean and were much weaker than the correlation between parental mean and F$_1$-performance (for seed yield $r = 0.86$; $P < 0.01$; Figure 2b).

Poor correlations between GS and heterosis can be attributed to various reasons. Markers used here are (i) not mapped, hence, probably not evenly distributed across the genome. Genetic similarities might (ii) be based on an inadequate number of marker loci. With the 1202 polymorphic markers employed in this study, we reached a high level of precision of the estimated genetic similarity (*cf.* Zeid et al., 2003). The conclusion of Melchinger et al. (1990), comprising marker distribution and number, was that better marker coverage of the genome does not by itself provide a key to increase the predictive power for hybrid performance. Here, we have (iii) not used AFLP fragments with known linkage to QTL. Bernardo (1992) has set conditions for effective prediction of hybrid performance based on molecular similarity of parents: At least 30–50% of the QTL have to be in linkage disequilibrium (LD) with molecular markers and not more than 20–30% of the markers should be randomly dispersed or non-predictive for QTLs. Hence, genetic similarities estimated from random, unmapped markers are not promising for the prediction of performance via the LD approach (Jordan et al., 2003).

In the present study, the intention was to predict the degree of heterozygosity at QTLs from the degree of heterozygosity at marker loci. It is generally understood that heterozygosity at the trait-determining loci would be strongly correlated with heterosis in the traits. Since both, markers and QTLs, are widely scattered across the genome, few will show significant linkage. Hence, any correlation between the heterozygosity of the two classes of loci in the hybrids depends upon two conditions: (i) the extent to which both classes of loci – regarding to their degree of heterozygosity – represent the heterozygosity of the whole genome and (ii), the extent to which there is variation among the hybrids in heterozygosity. For yield, the first condition is understood to be met, because yield integrates the effect of many different genes; the AFLP method sampled a large number of markers anyway. The second condition depends on the degree and pattern of relatedness of the parents used (Charcosset & Essioux, 1994; Melchinger, 1999; Jordan et al., 2003). Pedigree relationships between the four tester lines and the 14 experimental lines were deliberately excluded. The presence of just one exceptional cross with related parents obviously already created an association between the SGS and SCA as indicated by a significant Pearson's correlation coefficient of $r = -0.33$ ($P < 0.05$) for seed yield and $r = -0.41$ ($P < 0.01$) for 1,000-seed weight. Neutralizing its major effect on the correlation by using Spearman's rank correlation resulted in very weak, non-significant associations, $r = -0.21$ for seed yield and $r = -0.23$ for 1,000-seed weight as mentioned above.

The present study clearly demonstrated a marked amount of genetic diversity and heterosis available even within the European *minor* gene pool. However, AFLP markers were not useful to predict this heterosis, whereas parental mean and F$_2$-hybrid performance indeed proved useful to predict F$_1$-hybrid performance. Establishing mutually heterotic gene pools is a task of eminent importance in hybrid breeding. The lines Alfred, Victor, L1 and L2 and further genotypes sharing Minica as common parent (Table 1) appear to be a nucleus for a perceptible group within the European elite gene pool. Exploitation of this potential should be facilitated by further progress in CMS research.

Acknowledgments

We are very grateful for the cooperation of Dr. Olaf Sass, NPZ Lembke, Germany. Fred Stoddard's (Wolverhampton, UK) contributions to make a complicated part of the text understandable are highly acknowledged. Funding for this study was provided by the Federal Minister of Consumer Protection, Food and Agriculture, Berlin (GFP ÖE 112 + 113) and in part by DAAD (German Academic Exchange Service), Bonn. This paper was dedicated to Professor 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.

References

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

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

Bond, D.A., G.J. Jellis, G.G. Rowland, J. Le Guen, L.D. Robertson, S.A. Khalil & 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 & R.G. Herrmann, 1992. Genetic diversity for RFLPs in European mais inbreds: I. Relation of performance of Flint × Dent crosses for forage traits. Crop Sci 32: 895–902.

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

Cochran, W.G. & G.M. Cox, 1957. Experimental Designs, 2nd ed. Wiley, New York.

Comstock, R.E. & 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. & J.L. Doyle, 1990. Isolation of plant DNA from fresh tissue. Focus 12: 13–15.

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.

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

Jordan, D.R., Y. Tao, I.D. Godwin, R.G. Henzell, M. Cooper & C.L. McIntyre, 2003. Prediction of hybrid performance in grain sorghum using RFLP markers. Theor Appl Genet 106: 559–567.

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

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. & P. Ruckenbauer, 1988. Aspekte der Nutzung von Heterosis bei der Pferdbohne (*Vicia faba* L.). In: Ber. 38. Arbeitstagung der Vereinigung österreichischer Pflanzenzüchter, pp. 147–162. Gumpenstein, Austria.

Link, W., C. Dixkens, M. Singh, M. Schwall & 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 & 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.

Lonnquist, J.H. & M.F. Lindsey, 1964. Topcross *versus* S_1 line performance in corn. Crop Sci 4: 580–584.

Martsch, R., J. Vaupel, C.C. Schön & W. Link, 2001. Faba bean hybrid breeding: Current results with CMS199. In: Proceedings, 4th European Conference on Grain Legumes, Cracow, Poland, pp. 232–233. Association Européenne de Recherche sur les Protéagineaux (AEP), Paris.

Melchinger, A.E., 1999. Genetic diversity and heterosis. In: J.G. Coors & S. Pandey (Eds.), Genetics and exploitation of heterosis in crops, pp. 99–118. Crop Sci Soc of America, Madison, Wisconsin, USA.

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

Riaz, A., G. Li, Z. Quresh, M.S. Swati & C.F. Quiros, 2001. Genetic diversity of oilseed *Brassica napus* inbred lines based on sequence-related amplified polymorphism and its relation to hybrid performance. Plant Breeding 120: 411–415.

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

Roman, B., A.M. Torres, D. Rubiales, J.I. Cubero & Z. Sativic, 2002. Mapping of quantitative trait loci controlling broomrape (*Orobanche crenata* Forsk.) resistance in faba bean (*Vicia faba* L.). Genome 45: 1057–1063.

Schill, B., A.E. Melchinger, R.K. Gumber & 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, J.S.C. & O.S. Smith, 1992. Fingerprinting crop varieties. Advances in Agronomy 47: 85–140.

Stelling, D., 1992. Core harvesting – essential for the measurement of yield and yield related characters in field trials. In: 8th International Conference on Mechanization of Field Experiments, pp. 30–43, IAMFE, Paderborn, Germany.

Stelling, D., E. von Kittlitz, E. Ebmeyer, O. Sass, H. Jaiser & W. Link, 1994. Erfolge und Perspektiven der züchterischen Verbesserung von Körnererbsen und Ackerbohnen in der EU. In: Ber. 45. Arbeitstagung der Vereinigung österreichischer Pflanzenzüchter, pp. 131–147. Gumpenstein, Austria.

Sun, C.Q., T.B. Jiang, Y.C. Fu & X.K. Wang, 2002. Indica-japonica differentiation of paddy rice and its relationship with heterosis. Plant Breeding 121: 330–337.

Utz, H.F., 2001. A computer program for statistical analysis of plant breeding experiments. Version 2O. Institute of Plant Breeding, Seed Science and Population Genetics, University of Hohenheim, Germany.

Yap, I.V. & 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. Los Baños, Philippines.

Zabeau, M. & 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 & W. Link, 2003. Genetic diversity in recent elite faba bean lines using AFLP markers. Theor Appl Genet 107: 1304–1314.