Quantitative studies for investigating variation and its effect on heterosis of rice

A. Ahmadikhah[a,*], S. Nasrollanejad[b], O. Alishah[c]

[a]*Department of Agronomy and plant Breeding, Faculty of Agriculture, Gorgan University of Agricultural Sciences and Natural Resources, Beheshti St., P.O. Box 386, Gorgan, Iran.*
[b]*Department of Plant Pathology, Faculty of Agriculture, Gorgan University of Agricultural Sciences and Natural Resources, Beheshti St., P.O. Box 386, Gorgan, Iran.*
[c]*Cotton Research Institute of Iran, Gorgan, Iran.*
[*]Corresponding author. E-mail: ahmadikhah_a@gau.ac.ir

Abstract

An important step in plant breeding programs is identification of parents that are genetically distant from each other, to find crossing combinations with better prospects of hybrid vigor. The potential of identifying genetically distant parents depends on genetic diversity of population. 58 lines of rice from diverse origins have been employed to study genetic variation in order to identify the most effective components of grain yield. Using cluster analysis, the lines were classified into three distinct classes; some of classes containing lines from different origins. Factor analysis has been conducted to identify important components of variation and contributions of traits in each factor. Based on factor analysis, 6 components were identified, explaining ~82% of observed variation. Evaluation of heterosis of crosses, whose parents were selected based on genetic distances (GD), showed that maximum heterosis was obtained for crosses containing the furthest lines. All together, our results support the idea that the level of correlation between hybrid performance and genetic divergence is depends on the germplasm used.

Keywords: Rice; Variation; Quantitative; Heterosis

Introduction

Meeting the demands for rice as a staple food crop would be feasible by using of plant breeding methods and production of high-yielding varieties. A lack of diversity in germplasm limits long-term progress in plant breeding. The level of genetic diversity that characterizes commercially important crops is a matter of considerable concern, as it is generally agreed that access to genetic diversity has been and remains important in maintaining and increasing agricultural productivity (Ahmadikhah, 2006).

Development of high-yielding varieties is achieved via identification of genetic resources and characterization of genetic diversity present in plant populations and varieties

with desirable traits. Therefore, the genetic diversity is the basis of plant breeding. Plant breeders use the present diversity in plant populations and varieties to develop new varieties or to transmit the desirable traits into undesirable varieties (De et al., 1992). Correlation between traits has a specific importance because the correlations may aid the breeder in indirect selection for important traits via unimportant traits whose evaluation is easy (Bapu, 1992).

Plant breeding deals with high-yielding genotypes. However, distinguishing and choosing of the parents of these genotypes, remains a big challenge. Research on parent selection may be approached in two ways: *a priori* and *a posteriori* choice (Baenziger and Peterson, 1992). The former consists of selection methods based on *per se* parent performance, such as midparental value, divergence according to coefficient of parentage, character complementation, multivariate analysis and parental distances, least squares, parental complementation, and ideal genotype. In the latter, parents are evaluated on the basis of F1, F2 data and advanced generations. A long period of time is necessary to choose parents in this way, especially in perennial plants. Here, we focused on *a priori* choice based on parental distances determined from agronomic data.

To evaluate the genetic diversity of quantitative traits the breeders can use the multivariate statistical approaches such as cluster analysis. The method is used to investigate distance, similarity and relatedness of varieties or populations, so that similar varieties are classified into one group and dissimilar ones into distinct groups (Farshdfar, 1997). Therefore, the use of an alternative method rather than conducting a lot of crosses, one can identify superior individuals of all clusters in the given traits, and in the next step, cross between superior individuals from distant clusters (Nourmand-Moayed et al., 2000).

Factor analysis as a statistical model assumes that a small number of unobserved (i.e., latent) constructs are responsible for the correlation among a large number of observed variables (Bramel et al., 1984). The latent constructs, for example, academic ability, cannot be directly observed, but they affect observable variables, such as French, English, and Mathematics scores. Factor analysis assumes that the variance of each observed variable comes from two parts: a common part shared with other variables, which causes correlation among them, and a unique part that is different from other variables. The common parts are called factors, and these factors represent the latent constructs (Wang, 2001). Another using of factor analysis is the introduction of essential multivariate analysis (Guertin and Bailey, 1982).

In determining the potential of genetically different lines and cultivars, breeders have to observe many different characters that influence yield. Accurate evaluation of these characters is more difficult due to the genotype by environment interaction (Tadesse and Bekele 2001).

The development of new promising rice varieties in Iran generally is based on the selection programs with or without crossing, mainly organized in Rice Research Institutes (Rasht and Amol, Northern Iran) and agricultural universities in Northern Iran. In such programs local varieties along with the entries from other countries (particularly from International Rice Research Institute, IRRI, Philippines) form the preliminary plant materials as a diverse source of variation and are evaluated for desirable traits such as plant height, biotic and abiotic resistances, yield and yield components etc. Then the entries are classified into some clusters to identify superior varieties. After identification of superior

varieties for the given traits, in the next step, superior individuals from distant clusters are crossed to make a new combination of genes, and the program is followed in an appropriate classical selection method.

The objectives of this study were the evaluation of genetic diversity, identification of main factors explaining observed variation, and the evaluation of the influence of genetic distance on heterosis of yield in several rice genotypes.

Material and methods

Fifty-eight inbred rice lines were chosen in order to display a large variation in quantitative traits and to represent a diversified sample of genotypes used both in Iran and International Rice Research Institute (IRRI) breeding programs (Table 1). The rice varieties were sown in a replicated field trial (with 4 replications) in the research field, Mazandaran University, Sari, Northen Iran, in May 1, 2000 and 2001. Seventy 20-day-old seedlings were transplanted in each plot of 2m × 2m dimensions, with 25 cm between seedlings. The NPK fertilizer (60-60-30 kg ha^{-1}) was supplied before transplanting in soil. The plants received a complementary urea fertilizer (100 kg ha^{-1}) at maximum tillering stage. At the appropriate times of growth, the following traits were measured: plant height (PH; cm), culm height (CH; cm), panicle height (PaH; cm), tillers per plant (Til), flag-leaf length (FLL; cm) and width (FLW; cm), flag-leaf length: width ratio (Ratio), leaves of main tiller (LMTil), total grains per panicle (TG), filled grains per panicle (FG), unfilled grains per panicle (SG), amount of panicle exertion from flag leaf (PaEx; cm), first date of flowering (FD; days), full maturity (FM; days), grain filling duration (GF), panicle fertility percentage (PaF), 1000-grain weight (1000-GW; g) and grain yield (Y; kg/ha).

Analysis of variance (ANOVA), multivariate regression, and factor- and cluster analysis were carried out to develop a quantitative understanding of genetic structure of the total population. SPSS10 and SAS software's were used to analyse data. For estimating genetic distance (GD) between varieties Euclidean distance was used as described (Kinnear and Colin, 2000). Cluster analysis was based on the unweighted paired-group method using arithmetic averages (UPGMA) of the SPSS10 software as described (Kinnear and Colin, 2000).

Statistical associations between different traits and yield were tested using stepwise multiple linear regressions (PROC REG in SAS software, SAS Institute, 1999) with P≤0.01 as the criterion for both entry and staying in the model.

Observations on 14 traits (from which ratios could be derived as additional traits) were recorded on 10 randomly selected plants in each plot (plot borders not considered for evaluation). Factor analysis calculations were performed using SPSS factor analysis programme. Estimates of factor loadings were based on data from all plants for all populations. The principal factor analysis method explained by Harman (1976) was followed in the extraction of the factor loadings. The array of communality, the amount of the variance of a variable accounted by the common factor together, was estimated by the highest correlation coefficient in each array as suggested by Seiller and Stafford (1985). The number of factors was estimated using the principal component method. The Varimax rotation method was used in order to make each factor uniquely defined as a distinct cluster of intercorrelated variables (Rao, 1952). The factor loadings of the rotated matrix, the

percentage variability explained by each factor and the communalities for each variable were determined.

The data collected on GDs of varieties was used for selecting the parents to be involved in crossing. Three types of crosses were conducted: the most distant lines (n=3) with a high GD, the intermediary distant lines (n=3) and the near lines (n=3) were crossed. The mid-parent and heterobeltiosis - based heterosis (H_{mp} and H_{hb} in percent, respectively) of obtained F_1s were calculated as following:

$H_{mp}=100[F_1-(P_1+P_2)/2]/[(P_1+P_2)/2]$

$H_{hb}=100[\times(F_1-Ps)/Ps]$; where Ps is superior parent involved in cross.

Table1. The varieties and lines from different regions evaluated in this study.

#	Variety	Origin	#	Variety	Origin	#	Variety	origin
1	Mohamadi-chaparsar	Iran	21	Khazar	Iran	41	IR67023-30-3-3-2-2	IRRI
2	Gerdeh	Iran	22	Sefid-Rood	Iran	42	IR67039-115-3-1	IRRI
3	Gharib	Iran	23	Dasht	Iran	43	IR67409-132-2-1-2-1	IRRI
4	Dom-Siah	Iran	24	Nemat	Iran	44	IR97410-143-1-9-1-1	IRRI
5	IR24	IRRI	25	Chz1	IRRI	45	IR65617-52-2-3-3-2-3	IRRI
6	IR28	IRRI	26	Bejar	Iran	46	Hari305	IRRI
7	IR36	IRRI	27	221	IRRI	47	338	IRRI
8	IR56	IRRI	28	IR65912-31-2-4-2-3-1	IRRI	48	IR67411-164-2-3-3-2	IRRI
9	IR50	IRRI	29	IR65610-205-2-5-2-2-2	IRRI	49	IR67413-64-1-1-1-1	IRRI
10	IR30	IRRI	30	7963	IRRI	50	IR67413-71-4-2-2	IRRI
11	IR64	IRRI	31	7879-IR92871-175-1-10	IRRI	51	IR67418-224-2-3-3-1	IRRI
12	BA-370	Pakistan	32	IR7924-72871-294-3-6	IRRI	52	IR64718-238-6-2-3-3	IRRI
13	PusaBa-834	Pakistan	33	7911	IRRI	53	IR67423-42-2-3-3	IRRI
14	JA-58	Pakistan	34	IR66233-234-2-1-2	IRRI	54	IR67423-53-2-3-3-2	IRRI
15	Fuji-Minori	Japan	35	IR66295-36-2	IRRI	55	MI18	IRRI
16	Amol3	Iran	36	IR66611-6-5-1-5-1(1036)	IRRI	56	Zhong Fan11	IRRI
17	Amol2	Iran	37	IR66611-6-5-1-5-1(1037)	IRRI	57	Usen	USA
18	Amol1	Iran	38	IR67017-124-2-4	IRRI	58	Hari304	IRRI
19	Champa	Iran	39	IR67017-13-3-3	IRRI			
20	Mehr	Iran	40	IR67017-180-2-1-2	IRRI			

IRRI- International Rice Research Institute.

Results

Analysis of variance (ANOVA)

ANOVA showed that the studied genotypes had significantly differences in all traits (data not shown); the between-group variation tended to exceed the within-group variation. Therefore, subsequent analyses were conducted.

Multivariate regression of yield on other traits

Stepwise regression was used to determine the contribution of studied traits in final grain yield. Ten traits remained in the model and as such are seen as significant determinants of final yield. The final regression of yield on other traits is:
Y= -19.8+ 34.1 (FM; days) + 44.3 (Til) + 1239.1 (PaF) + 53.8 (1000-GW; g) +14.3 (Ratio) -27.4 (PH; cm) + 1.06 (TG) + 8.60 (PaH; cm) -0.0936 (SG) -0.0953 (FG).

The coefficient of determination (R^2) of fitted model indicates that >99.9 % of the variation of the dependent variable (yield) is explained by the 10 traits in the equation.

Correlations of varietal means

Correlation analysis of varietal means showed that the highest correlation (0.959) was observed between plant height and culm height and the least (0.00025) between first date of flowering and flag-leaf length. The correlations with yield are shown in Table 2. As seen in the table, panicle height (in negative direction), first date of flowering and full maturity have the highest correlations with yield (-0.67, 0.62 and 0.62, respectively), followed by plant height (in negative direction), tiller number, leaves on main tiller, amount of panicle exertion from flag leaf (in negative direction) and 1000-grain weight (-0.56, 0.50, 0.46, -0.30 and 0.26, respectively).

Table 2. Correlations between different morphological traits and yield.

	Yield		Yield
PH	-0.56**	FM	0.62**
CH	0.11	GF	0.24
PaH	-0.67**	TG	0.05
Til	0.50**	SG	0.11
PaEx	-0.30*	FG	-0.00
FLL	-0.04	PaF	-0.12
FLW	-0.17	LMTil	0.46**
Ratio	0.09	1000-GW	0.26*
FD	0.62**		

* and ** represent significance at P≤0.05 and P≤0.01, respectively.

Cluster analysis

Cluster analysis of morphological data from samples allowed to discrimination of varieties based on individual plants. Genetic distances (GDs) among genotypes varied from

27 to 2643. According to the distance matrix, the least distance (27) belonged to variety numbers 7 and 29, and the highest distance (2643) belonged to variety numbers 1 and 7.

According to cluster analysis and cutting dendrogram in a single distance coefficient, studied rice varieties divided into three groups (Figure 1). Group A includes only variety number 1, an Iranian variety Mohamadi-chaparsar. Most rounded-grain varieties including Gerdeh, Gharib, Bejar (from northern Iran) and Fuji-Minori fell into group B. The remaining varieties, most of which are IRRI varieties, are classified in group C, which is subdivided into two subclasses of C1 and C2.
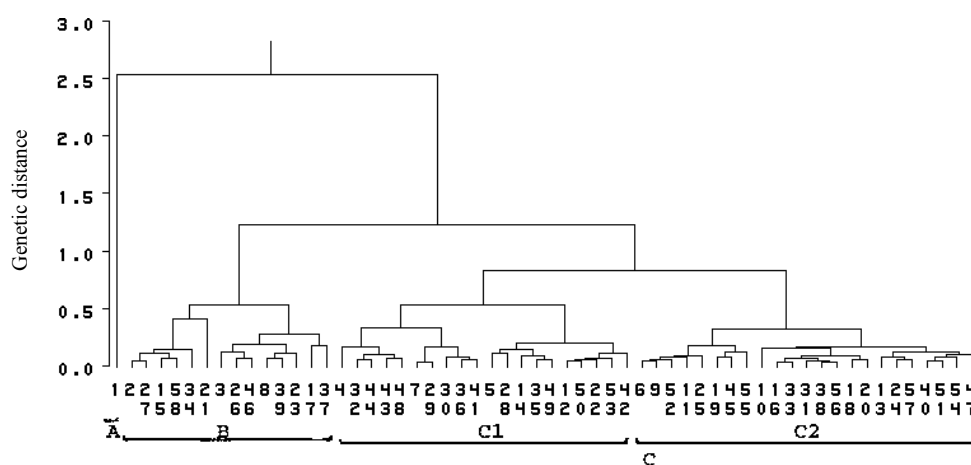

Figure 1. Dendrogram of 58 rice varieties and lines generated by morphological data using the UPGMA method.

Table 3. Total variance explained by different components. Extracted components with initial Eigenvalues >1 are in boldface.

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	**4.731**	**26.282**	**26.282**
2	**3.252**	**18.069**	**44.351**
3	**2.479**	**13.771**	**58.122**
4	**1.938**	**10.768**	**68.890**
5	**1.385**	**7.695**	**76.585**
6	**1.028**	**5.711**	**82.296**
7	0.871	4.838	87.134
8	0.776	4.311	91.445
9	0.700	3.891	95.336
10	0.461	2.559	97.895
11	0.191	1.061	98.956

Factor analysis

For identifying more important components contributing in total variation, factor analysis was conducted. The total variance and Eigenvalues explained by factors are

indicated in Table 3. The first six factors (with initial Eigenvalues >1) accounted for more than 82% of the total variance. The contributions of factors 1-6 to the total variance were 26.28, 18.07, 13.77, 10.77, 7.69 and 5.71 percent, respectively.

A principal factor matrix after Varimax rotation with Kaiser Normalization for these six factors is given in Table 4. The values in the table for factor loadings indicate the contribution of each variable to the factors. To interpret the result, only those factor loadings having greater values which are in boldface in Table 4 are considered. Factor 1, which accounted for about 26% of the variation, was strongly associated with 1000-grain weight (1000-GW), leaves on main tiller (LMTil), first date of flowering (FD) and full maturity (FM).This factor was regarded as a phenological factor, since it consisted mostly of date-dependent traits. All important variables in factor 1 had positive loadings. The sign of the loading in Table 4 indicates the direction of the relationship between the factor and the variable. Factor 2, which accounts for about 18% of the variation, was named growth factor since it consisted of height components including plant height (PH) and culm height (CH). In this factor both variables, also had positive loadings. Factor 3, was named fertility factor, which is positively associated with sterile grains per panicle (SG) and negatively with panicle fertility (PaF). This factor accounted for about 14% of the total variation. The forth factor which is named sink factor, accounts for about 11% of the variation. The variables number of filled grains/panicle (FG) and number of total grains/panicle (TG) had high positive loadings in this factor. The fifth factor which is named source factor, accounts for about 8% of the variation. Variables flag-leaf length (FLL) and flag-leaf width (FLW) and their ratio (Ratio) had high value of loadings in this factor, along with panicle height (PaH). Finally, the sixth factor, which is named yield factor, accounts for about 6% of the variation. The variables number of tillers/plant (Til) (negative), grain filling duration (GF) and grain yield (Y) had high values in this factor.

Table 4. Component score coefficient matrix after Varimax rotation with Kaiser Normalization. The coefficient values with a high weight (higher than 0.20) are shown in boldface.

Trait	Component					
	1	2	3	4	5	6
	Phonological	Growth	Fertility	Sink	Source	yield
PH	0.064	**0.321**	0.006	-0.008	0.052	-0.031
PaH	0.137	0.112	0.005	0.075	**0.224**	0.011
CH	0.027	**0.323**	0.005	-0.034	-0.017	-0.036
Til	0.060	-0.051	-0.045	0.099	0.048	**-0.650**
PaEx	-0.041	0.146	0.026	-0.146	0.034	-0.029
FLL	-0.009	0.090	0.079	0.082	**0.408**	-0.165
FLW	0.045	0.165	0.021	0.152	**-0.222**	-0.211
Ratio	-0.032	-0.056	0.043	-0.081	**0.436**	0.029
FD	**0.220**	0.009	0.087	-0.051	0.017	0.031
FM	**0.213**	-0.008	0.076	-0.048	0.006	0.101
GF	0.060	-0.115	-0.076	-0.027	-0.011	**0.531**
TG	-0.068	-0.028	0.051	**0.401**	0.003	-0.052
SG	-0.057	0.018	**0.426**	-0.016	0.054	-0.016
FG	-0.050	-0.046	-0.202	**0.521**	-0.031	-0.055
PaF	0.022	-0.035	**-0.480**	0.184	-0.049	-0.007
LMTil	**0.247**	0.008	-0.051	-0.027	-0.073	0.017
1000-GW	**0.300**	0.149	-0.191	-0.131	-0.049	-0.110
Y	0.197	-0.184	-.105	0.050	0.055	**-0.235**

Heterosis study

Evaluation of heterosis (H_{mp} and H_{hb}) indicated in table 5. As seen in the table, maximum heterosis was obtained for crosses containing the furthest lines; that is, crossing of lines number 1 and 29 (H_{hb}=5.09%) with GD=2633.7, and lines number 1 and 7 (H_{hb}=4.0%) with GD=2642.57, which slightly did not differ significantly from each other, and only slightly differed from a cross containing the lines number 7 and 34 (H_{hb}=1.32%) with an intermediate GD (GD=1572.1). Minimum heterosis was obtained for crosses containing the genetically near lines [lines number 7 and 29 with GD=26.9 (H_{hb}= -1.69%); lines number 38 and 56 with GD= 28.1 (H_{hb}= -1.14%)]. The cross between lines number 1 and 17 with GD=1423.4, also gave a low heterosis (H_{hb}= -1.01%), which did not differ significantly from formers.

Table 5. Genetic distances and yield of parents selected for crossing, yield of resultant F_1s and the respected mid parent (H_{mp}) and heterobeltiosis-based (H_{hb}) heterosis.

Cross	Genetic distance	Parental Yield		F_1 yield (kg/ha)	H_{mp} (%)	H_{hb} (%)
		P1	P2			
1 x 7	2642.5	3823.8	6464.4	6723	30.69	4.00
1 x 29	2633.7	3823.8	6455.9	6710	32.01	5.09
1 x 36	2538.8	3823.8	6360.9	6575	24.99	0.06
1 x 8	1328.0	3823.8	5149.6	5105	13.78	-0.87
1 x 17	1423.4	3823.8	5242.8	5190	14.49	-1.01
7 x 34	1572.1	6464.4	4894	6550	15.33	1.32
7 x 29	26.9	6464.4	6455.9	6355	- 1.63	- 1.69
16 x 33	27.9	5450.5	5436.8	5460	3.00	0.17
38 x 56	28.1	5450	5462	5400	-1.03	-1.14

Discussion

Based on the analysis of variance, multivariate regression, and factor- and cluster analysis we could establish a quantitative understanding of genetic structure of the studied population of rice varieties. Results showed that all traits had adequate genetic variation, the among-group variation tended to exceed the within-group variation, indicating the total observed variation mainly was formed from the genetic variation. Analysis of correlations with yield revealed that the phonological traits including panicle height (in negative direction), first date of flowering and full maturity had the highest correlations with yield; that is, the genetic loci increasing the panicle height give rise to the decrease in yield, and inversely, the loci increasing the period of flowering time and ripening stage, decrease the yield. In this line, He et al (2002) suggested that effect-increasing and effect-decreasing loci can greatly improve correlation coefficient and may be used to predict the yield and yield components. The use of these criteria may be useful in the improvement of rice varieties in breeding programs.

A multivariate regression model has been established to identify the most effective components of grain yield. We applied our suggested model for the extrapolated prediction

of yield on some rice varieties (n=20) other than that used in this study (data not shown). Results of this evaluation showed that >98.0% of the variation of the dependent variable ($\hat{Y}$) is explained by the 10 traits in the equation.

Using cluster analysis, different cultivars were classified into three distinct classes; some of classes containing lines from different origins. This finding is consistent with other reports (Jelodar et al., 1999; Aghazadeh, 2000; Shefaodin et al., 1998). Most IRRI varieties were classified in group C, which is subdivided into two subclasses of C1 and C2. Their difference in plant height may explain this division. Mean plant height of C1 subclass (consisted of 20 varieties) was ~92 cm and that of C2 subclass (consisted of 23 varieties) was 104.5 cm.

Factor analysis made it possible to identify important components of variation and contributions of traits in each factor. We recognized six factors explaining about 82% of total observed variation in the population.

The most important factor (explaining 26% of the variation) was factor 1 (phonological factor), which was strongly associated with 1000-grain weight (1000-GW), leaves on main tiller (LMTil), first date of flowering (FD) and full maturity (FM). Factor 2 (growth factor) was consisted of height components including plant height (PH) and culm height (CH). Factor 3 (fertility factor) was positively associated with sterile grains per panicle (SG) and negatively with panicle fertility (PaF). The forth factor (sink factor) consisted of the number of filled grains/panicle (FG) and number of total grains/panicle (TG). The fifth factor (source factor) consisted of flag-leaf length (FLL) and flag-leaf width (FLW) and their ratio (Ratio), and panicle height (PaH). Finally, the sixth factor (yield factor) was formed from number of tillers/plant (Til), grain filling duration (GF) and grain yield (Y). Such type of dividing the total variation into most important components may be useful to conduct the indirect selection for improving rice yield. These results well in coordination with the results obtained by other researchers and confirm them (e. g. Aghazadeh, 2000).

Genetic distances (GDs) among genotypes calculated based on morphological data largely ranged. The least distance belonged to variety numbers 7 and 29 (GD=26.9), and the highest distance belonged to variety numbers 1 and 7 (GD=2642.5), followed by varieties number 7 and 29 (GD=2633.7). Therefore, it can be expected that crossing between these morphologically distant varieties will be resulted to a high degree of heterosis (Jain et al., 1994; Tsaftaris, 1995; Xiao et al., 1993; Syed and Chen, 2005) although for each trait this must be evaluated. For evaluating the GD-based selection of parental lines in crossing and subsequent evaluating of heterosis, we have formed a combination of the lines suitable for our purpose to be crossed. However, in selecting the parents for crossing, we taken into account the possible F_1 sterility emerging between *indica* and *japonica* varieties. Therefore, our crosses only contained *indica × indica* combinations. Our evaluations showed that maximum heterosis is obtained for crosses containing the furthest lines, although this is always not a case (for example, a cross containing the lines number 7 and 34 with an intermediate GD, had a positive heterosis (H_{hb}=1.32%)). Also we observed an intermediate significant correlation (0.615) between GD and heterobeltiosis- based heterosis (Figure 2). Using molecular markers, Kwon et al (2002) revealed that correlation values of GDs with F_1 performance of 'Tongil'-type rice [in particular, a *japonica × indica* hybrid cultivar, named Tongil, developed and released in the early 1970s by Korean breeders (Park et al., 2003)] were significant for yield, culm length

and spikelets per panicle. However, despite the lack of direct correlation between the genetic distance and the degree of heterosis, genetic diversity forms a very useful guide not only for investigating the relationships among genotypes but also in the selection of parents for heterotic hybrid combinations (Jain et al., 1993). These results support the idea that the level of correlation between hybrid performance and genetic divergence is dependent on the germplasm used.
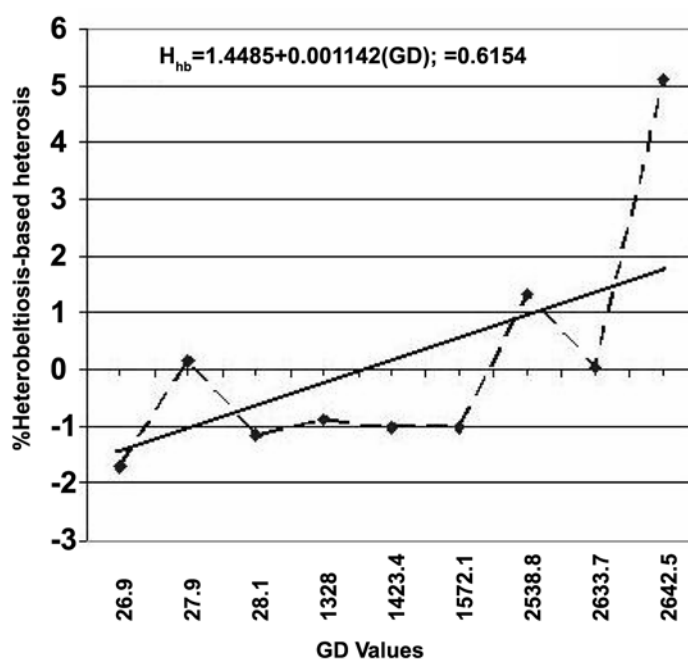

Figure 2. The relationship between genetic distance (GD values) and heterobeltiosis-based heterosis (H_{hb}).

References

Aghazadeh, R., 2000. Determination of genetic variation of rice using morphological and molecular RAPD markers. MS.c. Thesis. Mazandaran University, pp 130.

Ahmadikhah, A., 2006. Molecular tagging of CMS and fertility restoration genes in rice. Ph.D thesis: Moscow, pp. 162.

Baenziger, P.S., Peterson, C.J., 1992. Genetic variation: its origin and use for breeding self-pollinated species. In: Plant Breeding in the 1990s. Proceedings of the Symposium on Plant Breeding in the 1999s, Raleigh, NC, USA (Stalker, H.T. and Murphy, J.P., eds.). CAB International, Wallingford, UK. pp. 69-100.

Bapu, J.R.K., 1992. Genotypic association and path analysis in F3 generation of rice crosses. Madras Agriculture Jour. 76, 619–623.

Bramel, P.L., Hinz, P.N., Green, D.E., Shibler, R.M., 1984. Use of principal factor analysis in the study of three stems termination types of soybean. Euphytica 33, 388-400.

De, R.N., Rreddy, J.N., Suriara, A.V., Mohanty, K.K., 1992. Genetic divergence in early rice under two situations. Indian Jour. Genet. 52, 225 - 229.

Farshdfar, E., 1997 'The Use of Quantitative Genetics in Plants.' University of Razi Press: Kermanshah, Iran.

Guertin, W.H., Bailey, J.P., 1982. 'Introduction to Modern Factor Analysis.' Edvards Brothers Inc.: Michigan.

Harman, H.H., 1976. Modern Factor Analysis. University of Chicago Press: Chicago.

He, G.H., Hou, L., Li, D.M., Luo, X.Y., Niu, G.Q., Tang, M., Pei, Y., 2002. Prediction of yield and yield components in hybrid rice by using molecular markers. Acata Genetica Sinicia 29, 438-444.

Kinnear, P.R., Colin, D.G., 2000. SPSS for Windows made simple: Release 10. Hove, UK: Psychology Press.

Kwon, S.J., Ha, W.G., Hwang, H.G., Yang, S.J., Choi, H.C., Moon, H.P., Ahn, S.N., 2002. Relationship between heterosis and genetic divergence in 'Tongil'-type rice. Plant Breeding 121, 487-492.

Jain, A., Bhatia, S., Banga, S.S., Prakash, S., Lakshmikumaran, M., 1994. Potential use of random amplified polymorphic DNA (RAPD) technique to study the genetic diversity in Indian mustard (*Brassica juncea*) and its relationship to heterosis. Theor Aappl Genet. 88, 116-122.

Jelodar, B.N., Nemazadeh, Gh., Mirlouhi, E., Rezaei, A., 1999. The study of genetic variation of Iranaian rice germplasm based on morphologic traits. 6th Agronomy and Plant Breeding Congress, Babolsar, Iran.

Noormand-Moayed, F., 2000. The study of genetic diversity of barely lines from ICARDA using cluster analysis for selection of crossing parents. 6th Agronomy and Plant Breeding Congress, Babolsar, Iran.

Park, S.-Y., Milgroom, M.G., Han, S.-S., Kang, S., Lee, Y.-H., 2003. Diversity of Pathotypes and DNA Fingerprint Haplotypes in Populations of *Magnaporthe grisea* in Korea over Two Decades. Phytopathology, 93, 1378-1385.

Rao, C.R., 1952. Advanced Statistical Method in Biometric Research. John Wiley and Sons: NewYork.

SAS Institute Inc., 1999. *SAS Procedures Guide, Version 8*, Cary, NC: SAS Institute Inc.

Shefaodin, S., Mohammadi, M.S., Vojdani, P., 1998. The study of genetic diversity of Iranian landraces rice varieties of national plant gene bank. 5th Agronomy and Plant Breeding Congress, Karj, Iran, p. 434.

Seiler, G.J., Stafford, R.E., 1979. Factor analysis of components of yield in guar. Crop Sci. 25, 905-908.

Syed, N.H., Chen, Z.J., 2005. Molecular marker genotypes, heterozygosity and genetic interactions explain heterosis in *Arabidopsis thaliana*. Heredity 94, 295-304.

Tadesse, W., Bekele, E., 2001. Factor analysis of components of yield in grasspea (*Lathyrus sativus* L.). Lathyrus Lathyrism Newsl. 2, 91-93.

Tsaftaris, S.A., 1995. Molecular aspects of heterosis in plants. Physiologia Plantarum 94, 362-370.

Wang, Y., 2001. What is factor analysis? StatNews, 48.

Xiao, J., Li, J., Yuan, L., McCouch, S.R., 1993. Genetic diversity and its relationship to hybrid performance and heterosis in rice as revealed by PCR-based markers. Theor. Appl. Genet. 92, 637-643.