

The genetic divergence of parents and variability of offspring quantitative traits. The reasons for discrepancies.

A. V. Smiryaev

Russian State Agrarian University—Timiryazev Agricultural Academy,

Moscow, 127550, Russia;

e-mail: genetics@timacad.ru

Abstract. We consider some indirect quantitative estimates of genetic divergence of the parental forms for the selection of pairs for mating. Verified by the simplistic notion of a positive correlation of overall allelic dissimilarity of parental forms and available for selection variability of any quantitative trait in the progeny. On data from field experiments using biometric-genetic analysis identified the possible causes and situations of multi-directional changes these correlations.

Genetic determination of quantitative traits variability in plant populations are usually unknown and is labile depend on growing conditions and genotype of parental forms [1]. Therefore, in recent decades for the selection of parental couple, aimed at increasing of available for selection variability of offspring quantitative traits, are using different estimates of genetic divergence (total dissimilarity of their alleles). In particular, for indirect estimating generalize a complex of quantitative traits of parental pairs in one value (number) - metric of parents dissimilarity, for example metric Mahalanobis, Euclidean distance, dissimilarity shape of parents reaction, etc. [2]. Others use metrics of dissimilarity based on the pedigrees of the parents [3], their biochemical [4], molecular [5, 6] and other markers. Simplistically assumed that the more common allelic dissimilarity of parental forms, the wider polymorphism of offspring population in genes that determines the variability of any quantitative trait. This should be manifested, for example, to increase the standard deviation of each valuable trait in a population of offspring. Purpose of investigation – to find the possible causes of disturbance such simplistic assumptions identified in the verification of the new

biometrics-genetic method of selection parental pairs based on a metric called dissimilarity shape of parents reaction [7].

Evaluation of genetic divergence of parents through the assessment of dissimilarity shape of their reactions

The notion of dissimilarity shape the reactions of genotypes and method of assessing dissimilarity have been proposed previously [8, 9] to determine the genetic divergence of parental genotypes. The basis of the experimental evaluation of the dissimilarity - multiple measurements of a growth trait - indicator of dissimilarity repeatedly measured in ontogenesis and under different environmental conditions of joint tests of the parents. For each parent in each test condition obtain the growth curve of this trait. The metric of dissimilarity between two parents is equal to the mean square of deviations values the repeatedly measured trait from the single square regression of relative variability these parents [10]. The metric of the dissimilarity of a pair of parents with respect to the shape of reaction is subdivided into two components. The first one is the “roughness” of the curve of the growth of the trait in one parent relative to that in the other parent during ontogeny for each variant of test conditions (year); the roughness is then summed up over years. The second component assesses the “divergence” of the curves of relative growth for the pair of parents in different years.

Figure 1 shows, as an example, experimental data for three years (1, 2, and 3). These are 28 measurements of plant height in centimeters (X), which were used as indicator for estimating this metric for two pairs of cultivars: (a) Mil'turum Pererod (X_7) and Ul'yanovka (X_9) and (b) Mil'turum Pererod and PPG-186 (X_{13}). For the latter pair, both the “roughness” and “divergence” of the curves was considerably greater compared to the former pair of cultivars. Calculations demonstrated that the dissimilarity metric of the latter pair was seven times higher than that of the former pair.

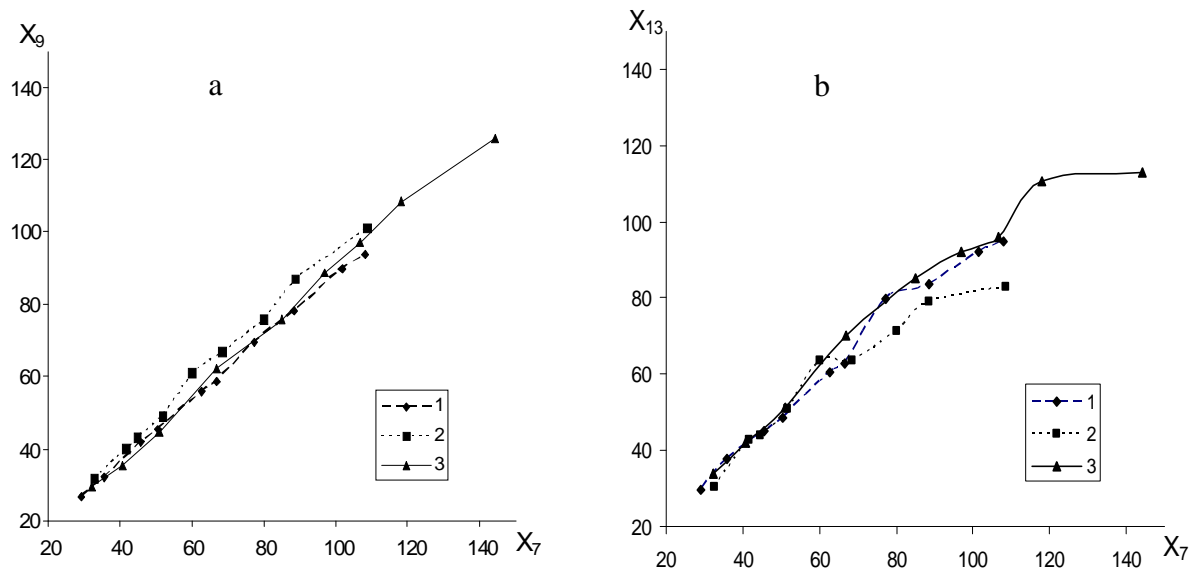


Figure 1. The curves of relative growth in the ontogeny for two pairs (a, b) of winter wheat varieties for three years field experiment [10]. Explanations in the text.

On the basis of this indicator ("plant height") similar results were obtained for all pairs of 22 winter wheat varieties and found positive correlation of allelic pair dissimilarity all pares with estimates of their dissimilarity in shape of its reaction [11]. This correlation can be explained as follows. Every substantial change in growth conditions in a field experiment causes deviations of the growth rates of many quantitative traits, the differences between these reactions of the two compared cultivars during ontogeny substantially depending on a large number of polymorphic loci. Later, the deviations of physiological processes, including growth processes, are partly compensated, but also in different ways depending on the allelic composition of the cultivars [12]. If repeated estimates of the growth trait are included into the metric of shape of reaction dissimilarity separately for each year of simultaneously testing the cultivars, additional information on the dissimilarity of their allelic compositions with respect to an increasing number of loci is obtained. This is a consequence of the so-called redefining of the genetic formula of a quantitative trait [1].

Thus, this metric for a specific growth trait (e.g., repeatedly measured height of the plants) is "accumulating" information about the dissimilarity of alleles,

manifested in the dissimilarities of deviations and compensations in two cultivars until the trait ceases to grow under different environmental conditions of simultaneous cultivar testing.

Check of the efficiency metrics for the selection of parental pairs

The possibility of using this metric pair dissimilarity of parents to predict the diversity available for selection on quantitative traits in populations of offspring, should check on the basis of experimental data. The material used – 6 cultivars and homozygous parental forms of spring wheat from the collection of the All-Russia Institute of Plant Breeding, including (1) k-58152, (2) Sibirskaya 3, (3) RG81220, (4) Planet, (5) St. Mercheisto, (6) SV Sonett and 15 hybrid populations of F_3 (2009) and F_4 (2010r.), obtained by diallel crossing of the parental forms.

Earlier in the four-year experiment was estimated pair dissimilarity shape of reaction (the dissimilarity metric is designated **H**) 6 these parents [10]. The obtained 15 values **H** used to predict quantitatively the diversity of 11 traits - the elements of the structure yields in the 15 hybrid populations F_3 and F_4 . Traits: plant height (the trait has the number 1), the length of the rod of the main ears of grain crops (2), productive tillering (3), the number of productive spikelets of the main ears (4), number of grains of the main ears (5), the mass of the main ears of grain (6), the number of grains of lateral ears (7), lateral mass of grain crops (8). In addition the three calculated traits was analyzed: the number of grains per plant (9), weight of grain per plant (10), 1000 grain weight (11). Each parent form and separately the hybrid population were represented as one plot in each of the 3 replications. On each plot values of the 11 traits in 30 randomly collected plants were measured.

According to these data in each of the 15 populations-families F_3 and F_4 was evaluated m - mean value of plants for each of the 11 traits, as well as to assess the experimental variability of each trait was calculated \sqrt{D} - standard deviation of the trait. The latter estimates the diversity of the trait available for selection in the population. Finally calculate r - coefficients of pair correlation between the values

of the metric-predictor (**H**) for 15 pairs of parents and 15 experimental values of \sqrt{D} for each trait in their offspring (F_3 or F_4). The results are presented in Table. 1.

Table 1. The correlation coefficient ($r \times 100$) the values \sqrt{D} for 11 traits in the 15 hybrid populations F_3 , F_4 with the values **H** – 15 metrics of dissimilarity pairs for 6 parental forms (according to [7] with modifications)

Number of trait	1	2	3	4	5	6	7	8	9	10	11
F_3	-2	-32	-12	-1	-4	-17	-23	-28	-4	-64*	-25
F_4	-30	9	-59*	-1	-4	-52*	-53*	-52*	-57*	-63*	52*

Comparison values **H** with the observed variability (\sqrt{D}) in hybrid populations gave unexpected results. In F_3 the only significant correlation coefficient with the metric **H** manifested plant productivity (10-th trait), where coefficient r was negative (-0.64*). At F_4 it was also significantly negative (-0.63*) and the correlations with several other traits was also significantly negative. For all of them, the coefficients r had opposite sign than expected in the forecast. The exception was the 11-th trait (1000 grain weight), where the correlation \sqrt{D} with **H** was significantly positive (0.52*).

Possible causes of variability in the correlations

One of the main reasons for the variability of the correlation coefficient for two years is a strong genotype-environmental interaction for each trait. Fig. 2 clearly shows that over two years vary not only values, but ranks of \sqrt{D} .

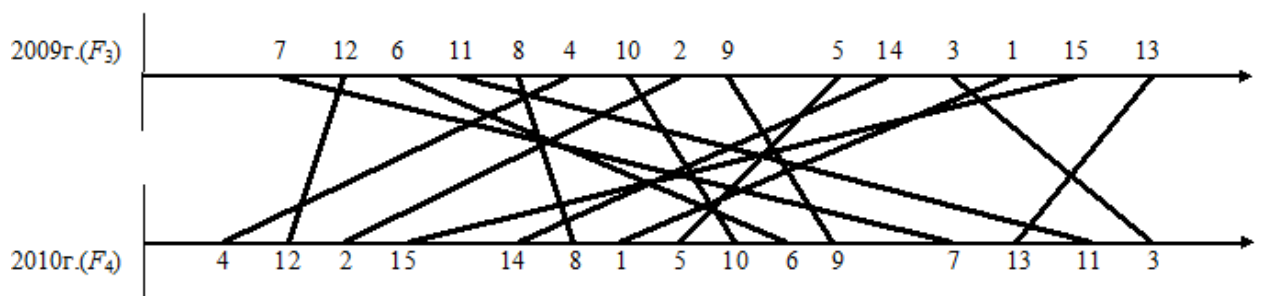


Figure 2. Variation of parameter \sqrt{D} in 15 hybrid populations on the example of 10-th trait (mass of grain per plant). Reduced to a single scale [7].

Recall that in contrast to the measure \sqrt{D} , values of **H** metric for pairs of parental forms have been evaluated previously in a 4-year data [10] and assumed to be constant as in generations, and on the traits. Therefore, the observed variability of the ranks for the indicator \sqrt{D} makes it difficult to carry out experimental estimation efficiency of the selection of pairs based on r for any metric that reflects the diversity of allelic composition of the parental forms.

The second reason for the lack of significant correlation coefficients, especially in generation F_3 (Table 1) may act as the destruction of successful genetic blocks of dissimilar parental genotypes by hybridization, caused by recombination. This results in reduced viability and rejection of many hybrid genotypes of offspring [13, 14].

The third reason - changes in the level of residual heterozygosity inside 15 populations of progeny families, in the transition from F_3 to F_4 .

The above factors that may cause variability or reduction of the correlation coefficients r , are known for a long time from the literature. However, they do not explain the causes of multi-directional correlations, manifested in the F_4 (Table 1).

Reasons for multi-directional correlations the forecast with the actual variability of traits

Consider the effect on the coefficients r the effects of genes that cause correlated variability of quantitative traits themselves and are reflected in the coefficients of their correlation. There are two basic level of pair correlations between the studied traits (Fig. 3): intra- and interfamilial

The first correlations are explained by the pleiotropic effects of genes that determine the variability of traits, and may be by plot allocompetition of plants in each offspring population. Inside each of the 15 families for any pair of traits the correlation coefficient were estimated from measurements of 30 plants in three replications, i.e., of 90 plant. The second (interfamily) - estimate based on the generalized parameters obtained in each of the 15 families in three repetitions: m -

mean value of trait or \sqrt{D} - standard deviation of the trait. Pleiotropic effects in this case, affect the interfamily correlation m or \sqrt{D} indirectly.

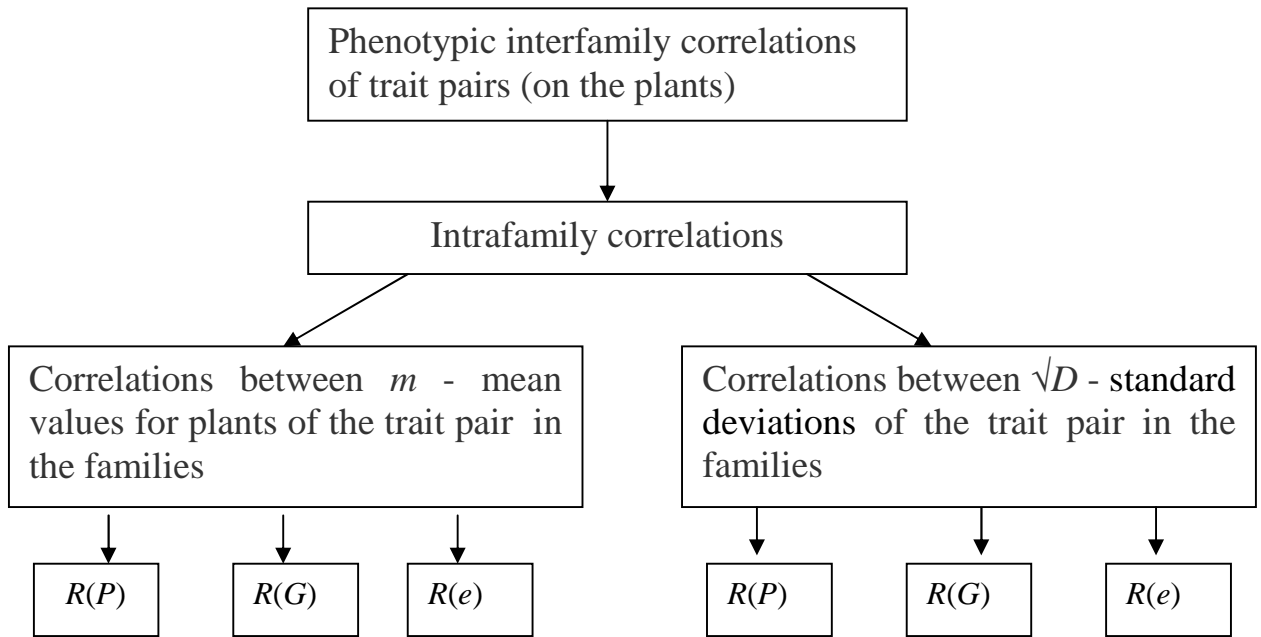


Figure 3. Block diagram of the traits pair correlation characteristics. $R(P)$, $R(G)$, $R(e)$ - interfamily phenotypic, genotypic, environmental correlation.

With replications data, each $R(P)$ – pair interfamily phenotypic correlation, as between the 15-th values m , and \sqrt{D} , can be divided into $R(G)$ – genotypic and $R(e)$ – environmental.

Analysis interfamily correlations of \sqrt{D} in 2010 [7] revealed a constellation of four closely correlated traits with the numbers 7, 8, 9, 10: estimates of $R(P)$ - phenotypic correlations ranged from 0.85 to 0.98. $R(G)$ - genotypic correlations were even higher. The same applies to the correlation of these 4 traits on m - the average values of families (from 0.73 to 0.98). In fact, these groups of 4 traits evident as a trait, i.e. their variability on the 15 families is mainly determined by one set of polymorphic genes. Therefore, in particular, each of these 4 traits in 2010 showed a similar magnitude negative correlation \sqrt{D} with \mathbf{H} - dissimilarity metric of parent pairs (Table 1).

Correlation \mathbf{H} with \sqrt{D} for the 11-th trait (the mass of 1000 grains), in contrast, was significantly positive (Table 1). And the 11-th trait is not correlated with any

of the other 10. Recall that this trait is calculated – Its value (x_{11}) for each plant within the family is:

$$x_{11} = x_{10}/x_9. \quad (1)$$

Intrafamily correlations of 10-th trait (x_{10} - mass of grain per plant) with 9-th (x_9 - number of seeds per plant) within all of the 15 families are also very high: from 0.83 to 0.96. Consequently, we can reliably enough to express one trait (x_{10}) through another (x_9) in any family:

$$x_{10} = a + b \cdot x_9, \quad (2)$$

where a, b - coefficients of linear regression.

Substituting equation (2) in (1) and dividing the numerator on the right side by the denominator, we obtain:

$$x_{11} = a/x_9 + b. \quad (3)$$

From Handbook of Lloyd E. and W. Lederman [15] taking into account equation (3) we obtain connection parameters \sqrt{D} for 9th and 11th traits within each family:

$$\sqrt{D_{11}} \approx a \cdot \sqrt{D_9}/m_9^2, \quad (4)$$

where m_9 - expectation of the 9-th trait in the family.

Considering (4) for 15 families, under the assumption of weak fluctuations of the coefficients a for different families (which is valid for a close interfamily correlation between $\sqrt{D_9}$ and $\sqrt{D_{10}}$), we can draw the following important conclusion. If under a variation of values $\sqrt{D_9}$ in families, denominator (m_9^2) in equation (4) changes correlated with $\sqrt{D_9}$, but m_9^2 grows faster than $\sqrt{D_9}$, then decrease $\sqrt{D_9}$ by families accompanied by an increase $\sqrt{D_{11}}$ in them. These conditions are sufficient for the occurrence of a positive correlation **H** with $\sqrt{D_{11}}$ at a negative correlation **H** with $\sqrt{D_9}$. If the denominator in (4) grows roughly in sync with the numerator, then, despite the correlation **H** with $\sqrt{D_9}$, a significant correlation **H** with $\sqrt{D_{11}}$ will not, i.e. connection $\sqrt{D_{11}}$ by families with a change in genetic divergence of parental varieties will not occur: $r(\mathbf{H}, \sqrt{D_{11}}) \approx 0$. This conclusion is valid not only for the metric **H**, but also for any other measure of genetic divergence of parents.

Note that from equation (1) that the variability of the 11-th trait is determined by the same genes that the 9-th and 10-th. Recall that the last two traits relate to one trait constellation: they are highly correlated with each other as in every family, and by 15 families. That is, the variability (\sqrt{D}) of all three traits (9, 10, 11), although depends almost from one set of polymorphic genes, but expressions of this gene set differ in traits. The same applies to the parameter m_9 - the value of trait, averaged by plants of the family. Within each family m_9 is calculated on the same experimental data (measurements x_9) as $\sqrt{D_9}$ - the standard deviation of the trait. Therefore, variability in both parameters depends on some polymorphic genes, but these genes manifest in different ways in m_9 and $\sqrt{D_9}$.

With the help of oligogene biometric-genetic models can easily show that in such situations, the interfamilial schemes of inheritance of three traits (9, 10, 11) vary. For example, if type of inheritance for the 9-th and the 10-th traits corresponds to additive-dominant scheme, then in the scheme of the 11-th feature is almost certainly manifest epistasis. A similar situation holds for interfamilial variability parameters m and \sqrt{D} of any trait.

In the more general case without the assumption of a close correlation $\sqrt{D_{10}}$, $\sqrt{D_9}$ and m_9 between families, using estimates from the handbook [15] we can show:

$$\sqrt{D_{11}} \approx |V_{10} - V_9| \cdot m_{10} / m_9,$$

where $|V_{10} - V_9|$ - modulus of the difference coefficients of variation 10 and 9th traits.

In certain situations of co-varying the parameters of the 10th and 9th traits, positive correlation $\sqrt{D_{11}}$ in families of offspring with any metric of genetic divergence of parents will lead to negative correlation of this metric with parameters $\sqrt{D_9}$, $\sqrt{D_{10}}$ or a lack of correlation with them. It all depends on differences in gene expression that determine the variability of the parameters m , \sqrt{D} , V of the 9-th and 10-th traits.

There are other forms of the connection between the original and calculated signs. For example, the total mass of plants (x_0) can be calculated as the sum of the masses (x_1) useful parts such as ears, and the rest (x_2):

$$x_0 = x_1 + x_2.$$

If the intrafamily correlation coefficients between x_1 and x_2 are close to -1 (for example, due to intrafamily polymorphism of attraction plastic substances from the straw in the ear), the parameter $\sqrt{D_0}$ in any family will be associated with the variances D_1 and D_2 as follows:

$$\sqrt{D_0} \approx |\sqrt{D_1} - \sqrt{D_2}|.$$

If, moreover, a close positive relationship interfamily regression $\sqrt{D_1}$ and $\sqrt{D_2}$ will provide an increase in their reduction for the difference, it will lead to a decrease $\sqrt{D_0}$ in families. The correlation coefficients of any genetic divergence metric of parents with $\sqrt{D_1}$ and $\sqrt{D_2}$ in the families of the offspring will again have the opposite sign compared to correlation coefficients with $\sqrt{D_0}$.

These examples show that biometric-genetic analysis allows us to formulate a variety of conditions under which useful for selection variability of calculated quantitative trait in the offspring would not be correlated with the genetic dissimilarity of the parents or will have the opposite sign of correlation compared to the initial traits. Perhaps these very factors rather than defects of a specific metric of parents divergence (**H**), explain the causes of multi-directional correlations of this metric with variability (\sqrt{D}) of some traits in the families and the absence of significant correlations **H** with \sqrt{D} for others. In addition, the concepts of "calculated" and "original" traits are conventional. We can consider the 10-th trait as calculated, and the 9th and 11th as original. Note also that for almost any quantitative trait we can pick up a two new "original" traits and this trait will be "calculated". We must conclude that the increase in genetic divergence of parents, i.e. total dissimilarity of their allelic composition, does not guarantee increasing of the useful diversity in the offspring for all the studied quantitative traits. Some traits may manifest lack of response variability or even a regular decrease in response variability in the progeny.

References

1. Драгавцев В.А. К проблеме генетического анализа полигенных количественных признаков растений. С.-Петербург: ВИР, 2003. 35 с.
2. Смиряев А.В., Кильчевский А. В., Генетика популяций и количественных признаков // М: Изд-во «КолосС», 2007, 270 с.
3. Мартынов С.П. Генеалогический подход к анализу генетического разнообразия зерновых культур с помощью информационно-аналитической системы генетических ресурсов: Методические указания СПб.: ВИР, 2006. – 87с.
4. Reif J.C., Melchinger A.E., Frisch M. Genetic and mathematical properties of similarity and dissimilarity coefficients applied in plant breeding and seed bank management // *Crop Sci.* – 2005. – Vol.45. – pp. 1-7.
5. Burkhamer R.L., Lanning S.P., Martens R.J. et al. Predicting progeny variance from parental divergence in hard red spring wheat // *Crop Sci.*, 1998. – Vol. 38. №1. – pp. 243-248.
6. Kim H.S., Ward R.W. Patterns of RFLP-based genetic diversity in germplasm pools of common wheat with different geographical or breeding program origins // *Euphytica*. – 2000. V. 115. pp. 197-208.
7. Ngyen Tkhan' Tuan, Smiryaev A. V. Bazhenova S.S. Efficiency assessment of parental pairs selection methods in spring wheat. *Izvestia of Mosk. Timir. Agr. Academy*, issue 6, 2010, pp. 67-74.
8. Smiryaev A. V. Ecological and ontogenetic variability of the quantitative trait in a series of comparative field trials. *Dokl. VASHNIL*, 1985, no. 8, pp. 23-25.
9. Smiryaev, A.V., Analysis of Similarity between Genotypes with Respect to Ecological and Ontogenetic Variability of a Growth Trait, *Russian Journal of Genetics*, 1997, vol. 33, no. 1, pp. 50–55.
10. Smiryaev A. V. , Pyl'nev V. V. Biometric–Genetic Analysis of Genotype Dissimilarity in Diallel Crosses of Spring Wheat, *Russian Journal of Genetics*, 2008, vol. 44, no. 2, pp. 234–239. <http://www.springerlink.com/content/qq46q7020x6707m5/fulltext.pdf>
11. Smiryaev, A.V., Pyl'nev, V.V., Tao Yun-Shen, Comparison of Relative Variability of Growth Traits and Genetic Similarity in Winter Wheat Cultivars, *Russian Journal of Genetics*, 1997, vol. 33, no. 1, pp. 56–64.
12. Шевелуха В.С. Рост растений и его регуляция в онтогенезе. -М: Колос, 1992. -258 с
13. Шмальгаузен, И.И. Факторы эволюции. Теория стабилизирующего отбора.// М.: Наука, 1968. 451 с.
14. Жученко А.А. Экологическая генетика культурных растений. Кишинев: Штиинца, 1980. 587 с.
15. Lloid E., Ledermann W., Handbook of Applicable Mathematics. John Wiley & Sons Ltd., 1984, 508 p.