Sel'skokhozyaistvennaya Biologiya [Agricultural Biology], 2013, № 1, p. 26-34.

UDC 631.522/.524:575.167:51-76

MATHEMATICAL MODELLING IN PLANT BREEDING. I. THEORETICAL BASIS OF GENOTYPES IDENTIFICATION ON THEIR PHENOTYPES DURING SELECTION IN SEGREGATING GENERATIONS

I.M. Mikhailenko, V.A. Dragavtsev
Agrophysical Research Institute, Russian Academy of Agricultural Sciences,
14, Grazhdanskii prosp., St. Petersburg, 196220 Russia,
e-mail: ilya.mihailenko@yandex.ru

Received August 22, 2011

Summary

The authors presented the formalized theory for identification of genotypes on phenotypes in modern breeding technologies. As a base the authors proposed the mathematical models of «genotype—environment» interaction, for which they solves an inverse informational problem during the estimation of sizes of no observed action of seven genetic-physiological system on selected quantitative traits to be improved.

Keywords: genotypes, phenotypes, mathematical models, identification, breeding technologies, genetic-physiological systems, evaluation.

Basic principles of modeling the system of "genotype-environment" interaction and possible ways of using the proposed models in solving basic problems of modern genetics and breeding were discussed in detail previously (1). The most important tasks solved my means of these models are (Figure 1): evaluation of the mechanisms of transgressions and selection of parental pairs for obtaining a desired result of crossing; estimation of the contributions (eg., to productivity) of genetic-physiological systems of the parents; prediction of transgressions of breeding traits in offspring; crossing and obtaining F_2 population; identification of genotypes on the base of phenotypes. These tasks are closely associated as aspects of one generalized task – effective management of successive stages of a genetic-breeding process. Development of the theory for solving this problem may open a prospect for significant advances of modern genetics .

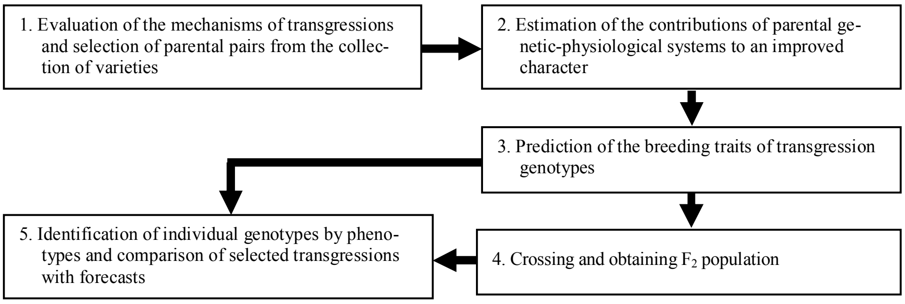
1. Evaluation of the mechanisms of transgressions and selection of parental pairs from the collection of varieties
2. Estimation of the contributions of parental genetic-physiological systems to an improved character
3. Prediction of the breeding traits of transgression genotypes
4. Crossing and obtaining F_2 population
5. Identification of individual genotypes by phenotypes and comparison of selected transgressions with forecasts

Fig 1. The scheme of interactions of the main tasks in management of genetic-breeding process.

As it can be seen from the scheme (Fig. 1), these main tasks are interrelated in a closed cycle of control over the process. The first step is selection of parental pairs capable to provide the desired result of crossing; this process can be optimized using the task of predicting results of such cross. In this case, real results are analyzed at the stage of identification of genotypes by phenotypes and they are used to correct the model for prediction of results of crosses. Each of these stages operates by mathematical models of "genotype-environment" interaction.

General scheme for identification of genotypes on phenotypes. Even though this task is the fifth on the diagram (Fig. 1), it touches upon some fundamental principles that facilitate formulation and solving other tasks.

In the matter, selection of genotypes by phenotypes is the informational task quite difficult for scientific classification. Its target is identification (or creation) of a genotype having the maximum number of positive expressions of specified breeding traits (BRT). That's why the algorithm of this task is based on the principle of background traits (2) and the principle of multi-directional shifts of an individual's quantitative character caused by genetic and environmental causes in two-dimensional coordinates of a trait (3).

An ideal background trait (BGT) has zero genotypic variance, so it only reflects ecological variation of environmental lim-factor (2). Therefore, an individual with above-zero deviation from mean level of BGT over a population is plus-modification in the best microecological niche. At the same time, if this individual shows a positive shift of BRT relative its mean level in population, this is a common modification not significant for selection. However, if another individual has BGT expressed at the average level of population and BRT shifted positively over its mean value in population, this is recombination (or mutation) that must be selected for further use in productive breeding work.

Concerning a particular individual, the phenomenon of multi-directional shifts (3) allows quantifying how much the deviation a trait from a mean level of population is determined by its genotype, and to what extent – by environmental factors. BGT may be sensitive to many factors that cause shifts of BRT except the one of genetic reasons to which BGT is insensitive or it responds "orthogonally" to BRT. In fact, this phenomenon is the base of algorithms for identification of genotypes by phenotypes.

The informational situation for a current task should be itemized. In the view of the abovementioned concepts, the task of identification of genotypes on phenotypes is based on identification of BGTs and BRTs according to which a particular individual can or can't be selected as proper for further breeding. There's the mathematical model "genotype-environment" for prediction of quantitative traits in individuals or populations (1). Along with it, a breeder has the data about all influencing environmental factors (both controlled and uncontrolled) monitored for a whole growing season, as well as real parameters of growth and development of individuals and populations recorded from sowing to final quantitative results assumed here as traits. Individuals must be to classified by final output, and such classification would divide the whole set of phenotypes of a new splitting generation to subsets of genotypes with different groups of quantitative traits some of which are economically important. These selected subsets may include a very small number of animals, or even few ones. Since the formation of such subsets is based on modeling the state of each individual (as essential step of correct classification), after completion of this process it is advisable to determine boundaries of the subsets, which then will simplify classification of individuals in other generations without modeling them. Using this approach, the first stage of classification is training with "imperfect (real) teacher" – mathematical models of particular individuals. The result of this training is determining the number of possible classes and revealing the boundaries of subsets of classes. The second stage of solving the task actually is an operative classification of individuals' genotypes on the base of their phenotypic characteristics.

Figure 2 shows the schematized algorithm of classification of genotypes by phenotypes.

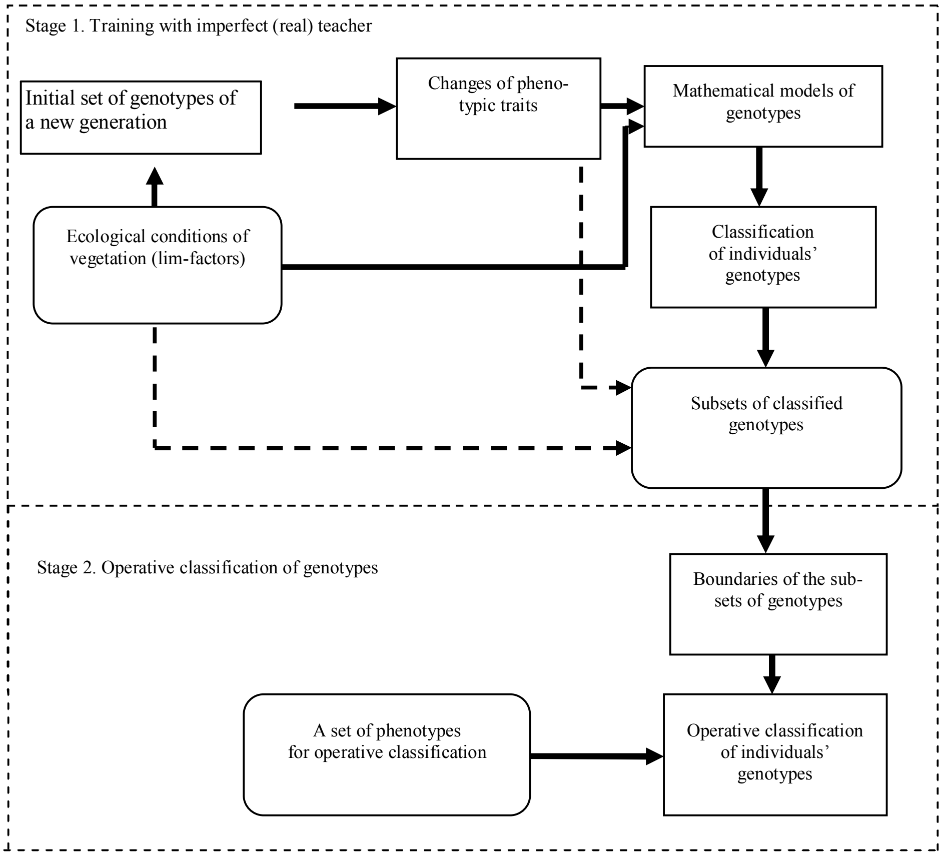
Stage 1. Training with imperfect (real) teacher
Initial set of genotypes of a new generation
Changes of phenotypic traits
Mathematical models of genotypes
Ecological conditions of vegetation (lim-factors)
Classification of individuals' genotypes
Subsets of classified genotypes
Stage 2. Operative classification of genotypes
Boundaries of the subsets of genotypes
A set of phenotypes for operative classification
Operative classification of individuals' genotypes

Fig 2. The algorithm of classification of genotypes on phenotypes.

In modeling of the abovementioned algorithm for general classification of genotypes, it should be briefly noted the evolution of mathematical models used in modern genetics. At first, there were two earliest models describing relations of genes and traits – the model of G. Mendel (4) and the model of R. Fisher, K. Mather, and S. Wright (5). The third one dated 1984 was the model of eco-genetic organization of quantitative traits (MEGOQT) (6) along with its 23 conclusions theoretically predicted and experimentally confirmed in 1984-2008. The most important of them are used to predict and determine the nature of transgressions, ecologically dependent heterosis, change of sign and value of genotypic and genetic (additive) correlations, effects of genotype-environment interaction, changes in number of genes and the amplitude of genetic variation of production traits, genetic homeostasis, etc. (7). In 2008, operability of this model was conclusively demonstrated at the molecular level during the research performed in collaboration with German scientists and geneticists (8), which allowed ranking the model of 1984 as the theory of eco-genetic organization of quantitative traits (TEGOQT). This theory has changed a classical Fisher's model:

$$\Psi_i = \mu + \gamma_i + \pi_i \qquad [1],$$

where Ψ_i — phenotypic value of a quantitative trait in i^{th} individual, μ — mean value of a quantitative trait over a population, γ_i — genotypic variance relative to a mean level, π_i — ecological variance relative to a mean level.

The new proposed model (9) describes integral production trait of i^{th} individual as:

$$\Psi_i = \mu + \gamma_{\text{attr},i} + \gamma_{\text{mic},i} + \gamma_{\text{ad},i} + \gamma_{\text{imm},i} + \gamma_{\text{ef},i} + \gamma_{\text{tol},i} + \gamma_{\text{ont},i} + \gamma_{\text{com},i} + \pi_{\text{com},i} + \pi_{\text{ont},i} + \pi_i \qquad [2],$$

where Ψ_i — phenotypic value of productive trait of i^{th} individual; μ — mean grain productivity in population; $\gamma_{\text{attr},i}$ — variance of the attraction of photosynthesis products by ear from stems and leaves; $\gamma_{\text{mic},i}$ — variance of microdistribution of the attraction products between grains and chaff in the ear; $\gamma_{\text{ad},i}$ — variance of the adaptive effects in productivity determined from total dry plant biomass; $\gamma_{\text{imm},i}$ — the effect of horizontal immobility on productivity; $\gamma_{\text{ef},i}$ — the effect of "payment" by biomass in response to limiting edaphic factors; $\gamma_{\text{tol},i}$ — variance of tolerance to thickening; $\gamma_{\text{ont},i}$ — variance of genetic diversity in duration of ontogeny phases; $\gamma_{\text{com},i}$ — variance of genetic competitiveness of plants for moisture, nutrition, light, etc.; $\pi_{\text{com},i}$ — variance of non-genetic competitiveness of plants caused by initial growing conditions; $\pi_{\text{ont},i}$ — variance caused by ontogenic changes of lim-factors in the period of formation and development of a character; π_i — variance associated with environment.

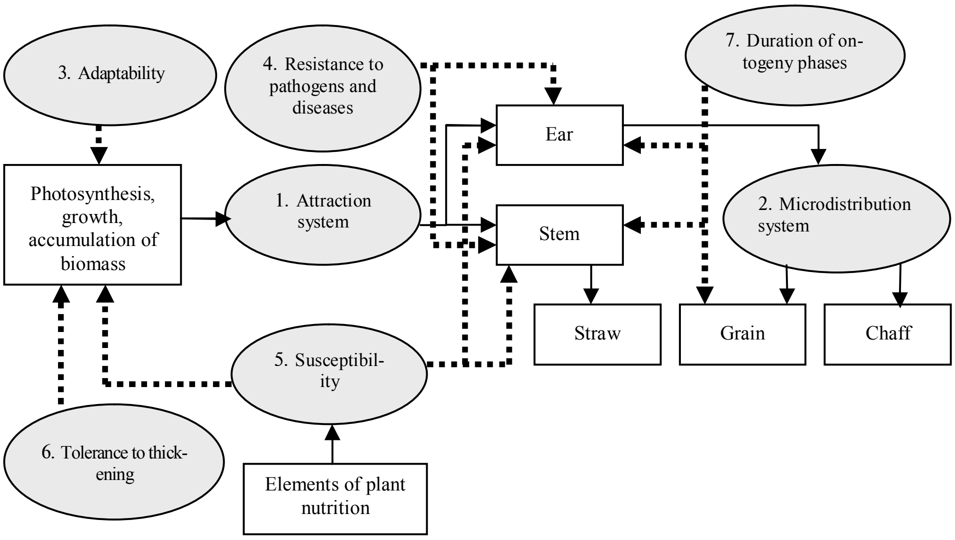
3. Adaptability
4. Resistance to pathogens and diseases
7. Duration of ontogeny phases
Ear
Photosynthesis, growth, accumulation of biomass
1. Attraction system
Stem
2. Microdistribution system
Straw
Grain
Chaff
5. Susceptibility
6. Tolerance to thickening
Elements of plant nutrition

Fig. 3. The scheme of "genotype-environment" model for cereal crops.

Each component of this model presents a specific state of individual's genetic-physiological systems described by a modular structure: 1 – the system of attraction (weight of stem φ_{11} and ear φ_{12}, i.e. commodity and non-commodity parts of a plant); 2 – the system of microdistribution (weight of grain φ_{21} and non-grain φ_{22} parts of the ear – chaff, awns, etc.); 3 - the system of adaptability, i.e. resistance to climatic and chemical environmental stressors (degree of slowing growth processes due to unfavorable factors – stressors, recovery rate and time needed to restore normal growth processes); 4 – the system of polygenic immunity (plant resistance to pests and diseases, synthesis of protective substances and development of defense mechanisms), 5 – the system of susceptibility (response) to doses of soil nutrition elements (sensitivity parameters of production traits to doses of nutrients); 6 – the system of tolerance to thickening (sensitivity parameters of productive traits to high sowing rate); 7 – the system of diversity in duration of ontogeny periods (used in plant breeding to move a critical ontogeny phase away from the onset of environmental stressor).

These parameters of genetic and physiological systems are described firstly in the modular structure of "genotype-environment" model for cereal crops (Fig. 3), and then in the mathematical model of main (output) module (4):

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\end{bmatrix}_i = \begin{bmatrix}a_{11}(\varphi_3) & a_{12}(\varphi_2) & a_{13}(\varphi_1)\\ a_{21}(\varphi_2) & a_{33}(\varphi_3) & 0\\ a_{31}(\varphi_1) & 0 & a_{33}\end{bmatrix}\begin{bmatrix}x_1(t)\\ x_2(t)\\ x_3(t)\end{bmatrix}_i + \begin{bmatrix}b_1(\varphi_5)\\ b_2\\ b_3\end{bmatrix}[u(t)]+$$

$$+\begin{bmatrix}0 & c_{12}(\varphi_3) & c_{13}(\varphi_3)\\ 0 & 0 & c_{23}\\ c_{31} & c_{32} & c_{33}\end{bmatrix}\begin{bmatrix}f_1(t)\\ f_2(t)\\ f_3(t)\end{bmatrix}+\begin{bmatrix}0 & 0 & 0 & d_{14} & 0 & d_{16} & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & d_{34} & 0 & d_{36} & 0\end{bmatrix}\begin{bmatrix}\Delta\varphi_1\\ \Delta\varphi_2\\ \Delta\varphi_3\\ \Delta\varphi_4\\ \Delta\varphi_5\\ \Delta\varphi_6\\ \Delta\varphi_7\end{bmatrix}+\begin{bmatrix}\xi_1(t)\\ \xi_2(t)\\ \xi_3(t)\end{bmatrix}, \quad [3],$$

$$t \in [t_0(\varphi_7); T(\varphi_7)]$$

with the following notations: x_{1i} — grain weight per ear in i^{th} individual; x_{2i} — weight of chaff per ear; x_{3i} — weight of straw per ear; u — supply (control) of nitrogen nutrition; f_1 — luminous efficiency factor; f_2 — temperature factor of productivity; f_3 — moisture as a productive factor; $\varphi_{1\ldots7}$ — influence of genetic and physiological systems; ξ_1, ξ_2, ξ_3 — random disturbances reflecting informational uncertainty of the model; a_{kj}, b_{ki}, c_{kj}, d_{kj} — dynamic parameters of the model.

The model [3] can be represented in compact vector-matrix form:

$$\dot{X}_i = A(\varphi_1,\varphi_2,\varphi_3)X(t) + b(\varphi_5)u(t) + C(\varphi_3)F(t) + D*[\varphi_4(t)\varphi_6(t)] + \xi(t),$$
$$t \in [t_0(\varphi_7); T(\varphi_7)] \quad [4],$$

in which all variables and parameters are combined in corresponding vectors and matrixes.

The model [4] shows the state of i^{th} individual where the effect of lim-factors (specific for each individual, as well as the action of genetic-physiological systems) results in perturbances of states of particular individuals and the occurrence of environmental and genetic variances. Such perturbances can be shown as:

$$\Delta X_i = X - X_i = U_E \Delta E_i + U_\varphi \Delta\varphi_i,$$
$$U_E = \frac{\partial X_i}{\partial E},\ U_\varphi = \frac{\partial X_i}{\partial \varphi}, \quad [5],$$

where U_E, $U_{\varphi u}$— vectors of sensitivity functions of the module's state to, respectively, ecological and genetic perturbances; ΔE_i, $\Delta\varphi_i$ — vectors of the diversity of observed ecological factors and unobserved genetic effects.

The expression [5] mathematically reflects contributions of ecological and genetic factors to the model. Along with it, a breeder most often deals to the observed diversity of traits that here are noted as ΔY_i. In this case, classification of genotypes becomes the task of determining the causes of observed diversity of traits in particular individuals compared with mean population value. Finding such cause in ecological factors means the presence of modifications of one genotype, while finding the cause in genetic factors means the presence of a new genotype.

This task involves the introduction of a quadratic functional for the quality of classification

$$J_i = \int_{t_0}^{t} [(\Delta X_i(t) - \Delta Y_i(t))^{T} (\Delta X_i(t) - \Delta Y_i(t))] dt \qquad [6],$$

whose form reflects "misbalance" of simulated and observed changes of BRT caused by unobserved effects of the seven genetic-physiological systems.

Minimizing the criterion [6] for unobserved effects of genetic factors $\Delta\varphi_i$ with known variation of ecological factors $\Delta E_i(t)$ allows estimation of contributions from each genetic-physiological system to expression of changes in a particular individual:

$$\Delta\hat{\varphi}_i = \arg\min \xrightarrow[\Delta\varphi_i]{} \int_{t_0}^{t} [(\Delta X_i(t|\Delta E_i) - \Delta Y_i(t))^{T} (\Delta X_i(t|\Delta E_i) - \Delta Y_i(t))] dt \qquad [6a].$$

Now, have established the acceptable region of effects of genetic-physiological systems for particular genotypes Ω_k, where $k = 1, 2, 3....K$ — indices of genotypes (classes), it is possible to determine the decision rule of classification:

$$k_i = k^*, \quad \text{if } \Delta\varphi_i \in \Omega_k^* \qquad [7].$$

The initial information about variances of ecological factors $\Delta E_i(t)$ and the observed variances of quantitative traits of i^{th} individual $\Delta Y_i(t)$ can be treated with the procedure [2]-[7] to reveal its relation to a genotype with specific parameters. Tracing time of each procedure for each class of the subset $\Delta E_k(t)$, $\Delta Y_k(t)$ allows distinguishing the boundaries of genotypes in the space of ecological variances and variances of quantitative traits in the form of a special approximating function:

$$L_k = \Phi_{k,\, k+1} (\Delta E, \Delta Y) \qquad [8].$$

In this case, the decision rule is expressed as:

$$\begin{aligned} k_i &= k^*, \quad \text{if} \quad \Phi_{k,k+1}(\Delta E, \Delta Y) - c \le 0, \\ k_i &= k^*, \quad \text{if} \quad \Phi_{k,k+1}(\Delta E, \Delta Y) - c \rhd 0, \end{aligned} \qquad [9];$$

where c — threshold value of the rule, one of parameters of the decision rule.

So, the proposed scheme of identification of genotypes by phenotypes shown in Figure 1 was explained in detail. In this scheme, the whole procedure of preliminary separation of genotypes and individuals' modifications corresponds to the stage of training with a "teacher" of more simple decision rule [8], [9]. Since this algorithm isn't free from errors, this "teacher" is imperfect, or, more exactly, it's "real" (10).

Algorithm of identification.

Introduction of Hamiltonian of the system:

$$\begin{aligned} H_i &= (\Delta X_i(t|\Delta E) - \Delta Y_i(t))^{T} (\Delta X_i(t|\Delta E) - \Delta Y_i(t)) + \\ &+ \lambda^{T}[A(\varphi_1, \varphi_2, \varphi_3)\Delta X_i(t) + b(\varphi_5)\Delta_i u(t) + C(\varphi_3)\Delta F_i(t) + D * [\Delta\varphi_4(t)\Delta\varphi_6(t)] \end{aligned} \qquad [10],$$

where λ — vector of conjugate variables, which is a solution in reversed time of the system

$$\begin{aligned} \dot{\lambda}_i &= -\frac{\partial H}{\partial \Delta X} = -2[(\Delta X_i(t|\Delta E) - \Delta Y_i(t)) + A^{T}(\varphi_1, \varphi_2, \varphi_3)\lambda_i], \\ & t \in (t, t_0),\ \lambda(t) = 0. \end{aligned} \qquad [11].$$

Considering the introduction of new auxiliary variables, minimization of identification criterion [5] for unobserved effect of genetic-physiological systems becomes a multi-step procedure expressed as:

$$\Delta\hat{\varphi}_{i,j+1} = \Delta\hat{\varphi}_{i,j} - \gamma_j \frac{\partial H_i}{\partial \Delta\varphi_{i,j}}, \qquad [12],$$

where j — number of a working iteration in the procedure of minimization of criterion [5].

Have reached by iterations [12] of specified conditions, the obtained estimates of effects of eco-genetic systems then are designated as $\Delta\varphi_i^*$. During separation of the obtained values of vectors into subsets of classes according to the rule [7], it is convenient to express their boundaries as a system of inequalities:

$$\Omega_k : \varphi_{lk\,\min} \le \varphi_{lk} \lhd \varphi_{lk\,\max},\ l = \overline{1,7}, \qquad [13],$$

where l — indices of genetic-physiological systems.

The system of inequalities [13] used for separation of individuals within the space of effects of the seven eco-genetic systems then will be named as "eco-genetic portrait" of a genotype, assuming that in the view of the developed TEGOQT theory this is an only possible representation of differences between genotypes.

It should be noted that vectors of effects of genetic-physiological systems $\Delta\varphi_i^*$ serve here only as "labels", or reference points in formation of the subsets of cause-and-effect relations:

$$\Omega_{kEX} : (\Delta\tilde{E}_{ki}, \Delta X_{ki}),\ i = 1, I_k \qquad [14];$$

where $\Delta\tilde{E}_{ki}$ — mean value of the vector of ecological variations in final interphase period.

Assuming such subsets, there are introduced more simple decision rules. Firstly, the vector of ecological causes $\Delta\tilde{E}$ should be combined with the vector of effects ΔY to form one vector $Z^T = [\Delta\tilde{E}, \Delta Y]^T$ and determine basic statistical characteristics of classes in the subsets [14] — vectors of mathematical expectations M_{Zk} and matrices of covariance K_{Zk}, as well as probabilities of the occurrence of classes whose estimates are reflected by the ratio: (number of individuals in subsets of separate classes I_k / total number of examined individuals), i.e.

$$p_k = \frac{I_k}{\sum_k I_k} \quad [15].$$

These characteristics allow simple description of separating functions of classes (11):

$$\Phi_{k.k+1}(\mathrm{Z}) = \mathrm{Z}^{\mathrm{T}}(\mathrm{K}_{Zk} - \mathrm{K}_{Zk+1})\mathrm{Z} + 2(\mathrm{M}_{Zk}\mathrm{K}_{Zk} - \mathrm{M}_{Zk+1}\mathrm{K}_{Zk+1})\mathrm{Z} \quad [16],$$

and threshold value c of the rule [9]:

$$c = 2\ln\frac{p_{k+1}}{p_k} + \ln\frac{|\mathrm{K}_{Zk}|}{|\mathrm{K}_{Zk+1}|} + \mathrm{M}_{Zk}^{\mathrm{T}}\mathrm{K}_{Zk}^{-1}\mathrm{M}_{Zk} - \mathrm{M}_{Zk+1}^{\mathrm{T}}\mathrm{K}_{Zk+1}^{-1}\mathrm{M}_{Zk+1} \quad [17],$$

where $|\mathrm{K}|$ is norm of the matrix.

Obviously, that according to [9] each new expression of cause-and-effect relations $\mathrm{Z}^{\mathrm{T}} = [\Delta\tilde{\mathrm{E}}, \Delta\mathrm{Y}]^{\mathrm{T}}$ necessitates pairwise comparison of all possible genotypes. As it was already mentioned above, vectors of effects of genetic-physiological systems $\Delta\varphi_i^*$ serve as "labels" in formation of the subsets of cause-and-effect relations:

$$\Omega_{kEX} : (\Delta\tilde{\mathrm{E}}_{ki}, \Delta\mathrm{X}_{ki}),\ i = 1, I_k\ .$$

However, further solving of breeding tasks should use a statistical variant of the model "ecological perturbation – response of genetic-physiological systems". In this regard, a following identification set should be formed:

$$\Omega_{E\Delta\varphi} : (\Delta\varphi_i, \Delta\tilde{\mathrm{E}}_i),\ i = \overline{1, I},$$

which allows to assess the parameters of matrix W in the required model:

$$\Delta\varphi = \mathrm{W}^{\mathrm{T}}\Delta\tilde{\mathrm{E}} \quad [18].$$

The presented way of solving the task operates with one of many modules in the general system "genotype-environment". If required to add other quantitative traits to characteristics of genotypes, dimensionality of the task may be expanded with no changes in the matter of the proposed approach. The important feature of the developed theory of identification of genotypes is its suitability for solving this task through the whole ontogenesis from earliest phenophases (i.e. modules of the lowest hierarchical level) up to terminal output modules of final product. This fact significantly improves reliability of solution and allows more efficient utilization of all genotypic diversity available for a breeder.

Thus, it was proposed the formalized theory of identification of genotypes by their phenotypes that includes evaluation (with mathematical model and special algorithm for optimization) of the values of unobserved contributions of the seven genetic-physiological systems to individual's productivity; classification of individuals using the established system of inequalities for contributions of genetic-physiological systems to individuals' productivity; separating individuals to classes and determining specific subsets of variations of ecological factors and variations of quantitative traits in each class, along with simultaneous evaluation of multidimensional statistical characteristics of these complex subsets; revealing the boundaries of separate classes of genotypes on the base of statistical characteristics of variation of ecological factors and variation of quantitative traits allowing the use of the algorithm for simplified identification of genotypes on phenotypes.

REFERENCES

1. Mikhailenko I.M., Dragavtsev V.A. *Sel'skokhozyaistvennaya Biologiya* [*Agricultural Biology*], 2010, 3: 31-34.
2. Dragavtsev V.A. *Botanicheskii Zhurnal*, 1966, 7: 939-946.
3. D'yakov A.B., Dragavtsev V.A. *Algoritmy ekologo-geneticheskoi inventarizatsii genofonda i metody konstruirovaniya sortov sel'skokhozyaistvennykh rastenii po urozhainosti, ustoichivosti i kachestvu (metodicheskie rekomendatsii, novye podkhody)* /Pod redaktsiei V.A. Dragavtseva [Algorithms of Ecology-Genetic Inventory of Gene Pool and Methods of Development Crop Varieties for Productivity, Resistance and Quality: Guidelines and New Approaches. V.A. Dragavtsev (ed.)]. St. Petersburg, 1994: 22-47.
4. Mendel G. Versuche uber Pflanzen Hybriden. *Verhandlungen des naturforschenden Vereins in Brunn*, 1865, 4: 3-47.
5. Wright S. The genetics of quantitative variability. *Quantitative inheritance. Edinburh*, 1950, London, 1952.
6. Dragavtsev V.A., Litun P.P., Shkel' N.M., Nechiporenko N.N. *Doklady AN SSSR*, 1984, 274(3): 720-723.
7. Kocherina N.V., Dragavtsev V.A. *Vvedenie v teoriyu ekologo-geneticheskoi organizatsii kolichestvennykh priznakov rastenii i teoriyu selektsionnykh indeksov* [Introduction into the Theory of Ecological and Genetic Organization of Quantitative Traits of Plants and the Theory of Selection Indices]. St. Petersburg, 2008.
8. Chesnokov Yu.V., Pochepnya N.V., Berner A., Lovasser U., Goncharova E.A., Dragavtsev V.A. *Doklady Akademii Nauk (RAN)*, 2008, 418, 5: 1-4.
9. Dragavtsev V.A. *Ekologo-geneticheskii skrining genofonda i metody konstruirovaniya sortov sel'skokhozyaistvennykh rastenii po urozhainosti, ustoichivosti, kachestvu, (novye podkhody)* [Ecology-Genetic Screening of Plant Gene Pool and Methods of Creating Varieties of Agricultural Plants on Productivity, Resistance, and Quality: New Approaches]. St. Petersburg, 1998: 25.
10. Milen'kii A.V. *Klassifikatsiya signalov v usloviyakh neopredelennosti* [Classification of Signals in Uncertainty Conditions]. Moskva, 1975.
11. Pugachev V.S. *Teoriya veroyatnostei i matematicheskaya statistika* [Probability Theory and Mathematical Statistics]. Moscow, 1979.