

## МАТЕМАТИЧЕСКИЕ МОДЕЛИ В СЕЛЕКЦИИ РАСТЕНИЙ. СООБЩЕНИЕ I. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ИДЕНТИФИКАЦИИ ГЕНОТИПОВ ПО ФЕНОТИПАМ ПРИ ОТБОРЕ В РАСЩЕПЛЯЮЩИХСЯ ПОКОЛЕНИЯХ\*

И.М. МИХАЙЛЕНКО, В.А. ДРАГАВЦЕВ

Предложена формализованная теория идентификации генотипов по фенотипам в современных селекционных технологиях. В основу теории положены математические модели взаимодействия «генотип—среда», для которых решается обратная информационная задача оценивания величин ненаблюдаемых воздействий семи генетико-физиологических систем на выбранные количественные признаки, подлежащие селекционному улучшению.

**Ключевые слова:** генотипы, фенотипы, математические модели, идентификация, селекционные технологии, генетико-физиологические системы, оценивание.

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

Основные принципы моделирования системы взаимодействия «генотип—среда» и возможности предложенных моделей для решения ряда базовых задач современной генетики и селекции подробно обсуждались нами ранее (1). Из всех задач, решаемых с помощью таких моделей, выделим наиболее важные (рис. 1): оценка механизмов трансгрессий и подбор родительских пар для обеспечения заданного результата скрещивания; оценивание вкладов (например, в продуктивность) генетико-физиологических систем родительских пар; прогнозирование трансгрессий селекционных признаков в потомстве; скрещивание и получение популяции F<sub>2</sub>; выделение (идентификация) генотипов по их фенотипам. Эти задачи неразрывно связаны и как единое целое представляют собой аспекты обобщенной проблемы — строгого управления последовательными этапами генетико-селекционного процесса. С разработкой теории, направленной на решение указанной проблемы, связывают одну из ключевых перспектив современной генетики.



Рис. 1. Блок-схема взаимоотношения задач управления генетико-селекционным процессом.

Как видно из приведенной схемы взаимоотношений решаемых задач (см. рис. 1), они образуют замкнутый контур управления селекционным процессом. Здесь начальным этапом служит подбор родительских пар для обеспечения заданного результата скрещивания, и для оптимизации такого подбора используется задача предсказания результатов скрещива-

\* Статью «Математические модели в селекции растений. Сообщение II. Алгоритмы управления генетико-селекционным улучшением хозяйственно ценных свойств самоопылителей» см. в журнале «Сельскохозяйственная биология», 2013, № 1: 35-41.

ния родительской пары, а реальные результаты скрещивания анализируются на этапе идентификации генотипов по фенотипам, они же используются для коррекции моделей предсказания результатов скрещивания. На каждом из приведенных этапов используются математические модели системы взаимодействия «генотип—среда».

**Общая схема задачи идентификации генотипов по фенотипам.** Несмотря на то, что на схеме (см. рис. 1) эта задача — пятая по счету, в настоящей статье рассматривается именно она как затрагивающая ряд основополагающих принципов, которые позволяют более строго формулировать и решать другие задачи.

По своему содержанию выделение генотипов по фенотипам — довольно сложная с точки зрения научной классификации информационная задача. Ее цель заключается в поиске (или создании) генотипа, включающего в себя максимальное число положительных сдвигов заданных селекционных признаков (СП). Поэтому в основу алгоритма этой задачи положены принцип фоновых признаков (2) и принцип разнонаправленности сдвигов количественного признака отдельной особи под влиянием генетических и экологических причин в двумерных признаковых координатах (3).

Идеальный фонный признак (ФП) имеет нулевую генотипическую дисперсию, поэтому он отражает только экологическое варьирование лимфактора среды (2). Следовательно, особь, имеющая плюсовое отклонение от популяционной средней по ФП, — плюс-модификация, попавшая в лучшую микроэкологическую нишу. В то же время в случае, когда СП у этой особи сдвинут в положительную сторону от популяционной средней, это обычная модификация и нет смысла ее отбирать. Если же у другой особи ФП выражен на уровне среднего показателя по популяции, а СП сдвинут в положительную сторону от среднего в популяции, то это рекомбинация (или мутация), и ее необходимо отобрать для продуктивной селекционной работы.

Применительно к отдельной особи явление разнонаправленности сдвигов (3) позволяет количественно оценить, в какой мере отклонение признака индивида от популяционной средней определяется его генотипом, а в какой — факторами среды. ФП может быть чувствителен ко многим причинам, вызывающим сдвиги СП, кроме одной из генетических причин, к которой ФП либо нечувствителен, либо реагирует смещением в направлении, «ортогональном» к таковому у СП. Собственно, описанное явление и лежит в основе алгоритмов идентификации генотипов по фенотипам.

Конкретизируем информационную ситуацию, в условиях которой предстоит решать подобную задачу. С учетом рассмотренных понятий задача идентификации генотипов по фенотипам по сути сводится к выявлению фоновых и селекционных признаков, по которым та или иная особь может или не может быть отобрана для дальнейшей селекционной работы. В нашем распоряжении есть математическая модель «генотип—среда», которая позволяет прогнозировать количественные признаки у особей или в их популяциях (1). Кроме того, мы располагаем данными мониторинга всех влияющих факторов среды (как управляемых, так и неуправляемых) в течение всего периода вегетации растений, а также сведениями о фактическом росте и развитии особей и их популяций от посева до получения конечных количественных результатов, рассматриваемых нами как признаки. Мы должны классифицировать особей именно по конечным результатам. При этом итогом такой классификации станет разделение всего множества фенотипов нового расщепляющегося поколения на подмножества генотипов, обладающих разными наборами количественных признаков, часть которых относится к хозяйственно ценным. Число особей в отдельных под-

множествах может быть очень небольшим или даже составлять несколько единиц. Учитывая, что для формирования таких подмножеств потребуется моделировать состояние каждой особи индивидуально (без чего невозможно правильная классификация), после формирования подмножеств генотипов целесообразно найти разделяющие их границы, что позволит в дальнейшем значительно упростить и ускорить классификацию отдельных особей в других поколениях, не прибегая к их моделированию. При таком подходе первый этап классификации — обучающий, где в качестве «неидеального (реального) учителя» используются математические модели отдельных особей. Результатом обучения будет определение числа возможных классов и выявление границ подмножеств классов, а второй этап решения задачи собственно и представляет собой оперативную классификацию генотипов особей по их фенотипическим признакам.

На рисунке 2 приведена структурная схема алгоритма классификации генотипов по фенотипам.

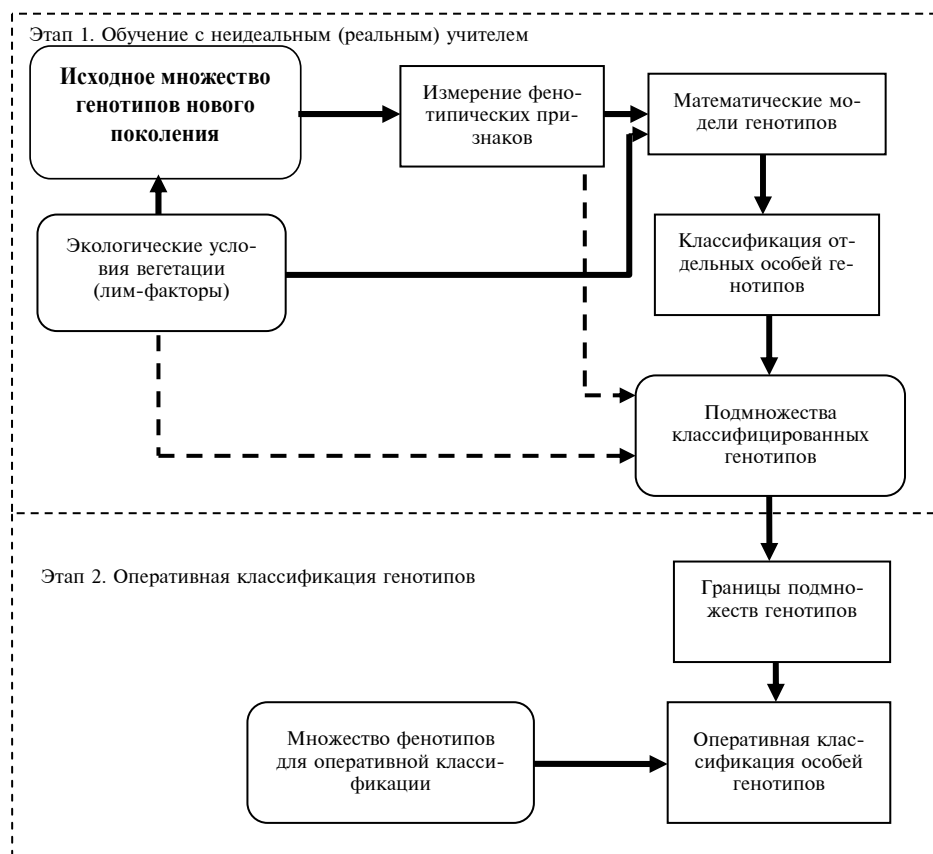


Рис. 2. Блок-схема, описывающая общий алгоритм классификации генотипов по фенотипам.

При разработке алгоритма общей классификации генотипов на основе приведенных выше принципов дадим краткую характеристику эволюции моделей, используемых в современной генетике. В 1984 году к двум уже имеющимся моделям, описывающим соотношения гены—признаки, а именно модели Г. Менделя (4) и модели Р. Фишера, К. Мазера и С. Райта (5), добавилась третья — модель эколого-генетической организации количественных признаков (МЭГОКП) (6). В 1984-2008 годах были теоретически обоснованы и экспериментально подтверждены 23 следствия третьей модели. Важнейшие из них — расшифровка природы и прогнозирование

трансгрессий, экологически зависимого гетерозиса, сдвигов знаков и величин генотипических и генетических (аддитивных) корреляций, эффектов взаимодействия генотип—среда, изменений чисел генов и амплитуды генетической изменчивости свойств продуктивности, генетического гомеостаза и др. (7). В 2008 году работоспособность модели была полностью подтверждена на молекулярном уровне совместно с учеными-генетиками из Германии (8), что позволило перевести модель 1984 года в ранг теории эколого-генетической организации количественных признаков (ТЭГОКП). Эта теория привела к изменению классической модели Р. Фишера:

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

где  $\Psi_i$  — фенотипическое значение количественного признака у  $i$ -й особи,  $\mu$  — средняя величина количественной характеристики признака в популяции,  $\gamma_i$  — генотипическое отклонение признака особи от средней величины,  $\pi_i$  — экологическое отклонение признака особи от средней величины.

Новая предложенная модель (9) описывает интегральное свойство продуктивности  $i$ -й особи:

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

где  $\Psi_i$  — фенотипическое значение признака продуктивности у  $i$ -й особи;  $\mu$  — среднее значение зерновой продуктивности в популяции;  $\gamma_{attr,i}$  — отклонение аттракции продуктов фотосинтеза из стеблей и листьев в колос;  $\gamma_{mic,i}$  — отклонение распределения продуктов аттракции между зернами и мякиной в колосе;  $\gamma_{ad,i}$  — отклонение влияния системы адаптивности на продуктивность, измеряемую общей сухой биомассой растения;  $\gamma_{imm,i}$  — влияние горизонтальной устойчивости на продуктивность;  $\gamma_{ef,i}$  — влияние оплаты биомассой лимитирующих факторов почвенного питания;  $\gamma_{tol,i}$  — отклонение толерантности к загущению;  $\gamma_{ont,i}$  — отклонение генетической вариабельности длительности фаз онтогенеза;  $\gamma_{com,i}$  — отклонение генетической конкуренции растений за влагу, питание, свет и т.д.;  $\pi_{com,i}$  — отклонение негенетической конкуренции, вызванное неодинаковыми начальными условиями роста,  $\pi_{ont,i}$  — отклонение, вызванное сменой лим-факторов в онтогенезе в период закладки и развития признака,  $\pi_i$  — средовое отклонение.

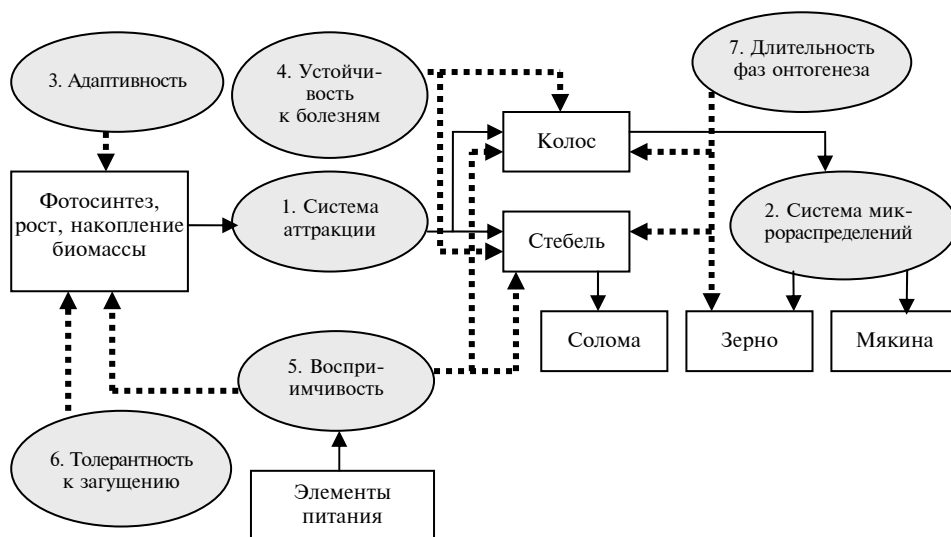


Рис. 3. Схема модели «генотип—среда» для зерновых культур.

Расшифруем каждую составляющую модели в виде конкретных состояний генетико-физиологических систем и модульной структуры модели: 1 — система аттракции (масса стебля  $\varphi_{11}$  и колоса  $\varphi_{12}$ , то есть товарной и нетоварной части растения; 2 — система микрораспределений (масса зерновой  $\varphi_{21}$  и незерновой  $\varphi_{22}$  части колоса — мякины, остей и др.); 3 — система адаптивности, то есть устойчивости к климатическим и химическим стрессорам среды (степень замедления ростовых процессов под воздействием неблагоприятных факторов среды — стрессоров, скорость и время восстановления нормальных ростовых процессов); 4 — система полигенного иммунитета (устойчивости растений к вредителям и возбудителям болезней, выработка растениями защитных веществ и механизмов); 5 — система восприимчивости (ответа) на дозы элементов почвенного питания (параметры чувствительности продуктивных характеристик к дозам элементов питания); 6 — система толерантности к загущению (параметры чувствительности продуктивных показателей к густоте посева); 7 — система варибельности периодов онтогенеза (используется в селекции с целью вывести критическую фазу онтогенеза из-под воздействия стрессового фактора среды).

Отразим приведенные характеристики генетико-физиологических систем сначала на структуре модульной модели «генотип—среда» для зерновых культур (рис. 3), а затем на математической модели этого основного (выходного) модуля (4):

$$\begin{aligned} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_{3j} \end{bmatrix} &= \begin{bmatrix} a_{11}(\varphi_3) & a_{12}(\varphi_2) & a_{13}(\varphi_1) \\ a_{21}(\varphi_2) & a_{33}(\varphi_3) & 0 \\ a_{31}(\varphi_1) & 0 & a_{33} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_{3j}(t) \end{bmatrix} + \begin{bmatrix} b_1(\varphi_5) \\ b_2 \\ b_3 \end{bmatrix} [u(t)] + \\ &+ \begin{bmatrix} 0 & c_{12}(\varphi_3) & c_{13}(\varphi_3) \\ 0 & 0 & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \begin{bmatrix} f_1(t) \\ f_2(t) \\ f_3(t) \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & d_{14} & 0 & d_{16} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & d_{34} & 0 & d_{36} & 0 \end{bmatrix} \begin{bmatrix} \Delta\varphi_1 \\ \Delta\varphi_2 \\ \Delta\varphi_3 \\ \Delta\varphi_4 \\ \Delta\varphi_5 \\ \Delta\varphi_6 \\ \Delta\varphi_7 \end{bmatrix} + \begin{bmatrix} \xi_1(t) \\ \xi_2(t) \\ \xi_3(t) \end{bmatrix}, \end{aligned} \quad [3],$$

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

где приняты следующие обозначения:  $x_{1i}$  — масса зерна в колосе  $i$ -й особи,  $x_{2i}$  — масса мякины в колосе,  $x_{3i}$  — масса соломы в колосе,  $u$  — обеспеченность (управление) азотным питанием;  $f_1$  — световой фактор продуктивности,  $f_2$  — температурный фактор продуктивности,  $f_3$  — влага, как фактор продуктивности;  $\varphi_{1...7}$  — влияние генетико-физиологических систем;  $\xi_1, \xi_2, \xi_3$  — случайные возмущения, отражающие информационную неопределенность модели;  $a_{kj}, b_{ki}, c_{kj}, d_{kj}$  — динамические параметры модели.

Представим модель [3] в более компактной векторно-матричной форме:

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

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

в которой все переменные и параметры объединены в соответствующие векторы и матрицы.

Модель [4] определяет состояние  $i$ -й особи, при этом влияние лим-факторов, неодинаковое для всех особей, как и действие генетико-физиологических систем, приводит к возмущениям состояний отдельных особей и возникновению средовых и генетических дисперсий. Такие возмущения можно представить в следующем виде:

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

где  $U_E, U_\varphi$  — векторы функций чувствительности состояния модуля соответственно к экологическим и генетическим возмущениям;  $\Delta E_i, \Delta \varphi_i$  — векторы вариаций наблюдаемых экологических факторов и ненаблюдаемых генетических воздействий.

Выражение [5] отражает моделируемые вклады экологических и генетических факторов. Вместе с тем селекционер чаще всего имеет дело с наблюдаемыми вариациями признаков, которые мы будем обозначать как  $\Delta Y_i$ . В этом случае смысл классификации генотипов заключается в установлении причин наблюдаемых вариаций признаков отдельных особей по сравнению со среднепопуляционными значениями. В том случае, если такими причинами служат экологические факторы, мы имеем дело с модификациями одного и того же генотипа, а при установлении генетических причин — с новым генотипом.

Введем в рассмотрение квадратический функционал качества классификации:

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

форма которого отражает «разбалансировку» моделируемых и наблюдаемых сдвигов СП, возникающих за счет ненаблюдаемых воздействий семи генетико-физиологических систем.

Минимизация критерия [6] по ненаблюдаемым воздействиям генетических факторов  $\Delta \varphi_i$  для известных вариаций экологических факторов  $\Delta E_i(t)$  позволит оценить вклады каждой из генетико-физиологических систем в проявление изменений у анализируемой особи:

$$\Delta \varphi_i = \arg \min_{\Delta \varphi_i} \int_{t_0}^t [(\Delta X_i(t | \Delta E_i) - \Delta Y_i(t))^T (\Delta X_i(t | \Delta E_i) - \Delta Y_i(t))] dt \quad [6a].$$

Теперь, располагая областями допустимых значений воздействий генетико-физиологических систем для отдельных генотипов  $\Omega_k$ , где  $k = 1, 2, 3, \dots, K$  — индексы генотипов (классов), можно определить решающее правило классификации:

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

Располагая исходной информацией об отклонениях экологических факторов  $\Delta E_i(t)$  и наблюдаемых отклонениях количественных признаков  $i$ -й особи  $\Delta Y_i(t)$ , мы посредством процедур [2]-[7] можем определить ее принадлежность к генотипу с заданными признаками. Фиксируя во время проведения этой процедуры для каждого класса множества  $\Delta E_k(t), \Delta Y_k(t)$ , можно построить границы между генотипами в пространстве экологических отклонений и вариаций количественных признаков в виде специальной аппроксимирующей функции:

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

В этом случае решающее правило выглядит следующим образом:

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

здесь  $c$  — пороговое значение правила, являющееся одним из параметров решающего правила.

Таким образом, мы рассмотрели всю схему идентификации генотипов по наблюдаемым фенотипам, представленную на рисунке 1. Здесь вся процедура предварительной сепарации генотипов и модификаций особей составляет этап обучения с «учителем» более простого решающего правила [8], [9]. Поскольку указанный алгоритм может допускать ошибки, такой «учитель» — «неидеальный» или, более точно, «реальный» (10).

**А л г о р и т м и д е н т и ф и к а ц и и.** Введем в рассмотрение гамильтониан системы:

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

где  $\lambda$  — вектор сопряженных переменных, являющийся решением в обратном времени системы

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

С учетом ввода новых вспомогательных переменных процедура минимизации критерия идентификации [5] по ненаблюдаемому воздействию генетико-физиологических систем будет иметь вид следующей многошаговой процедуры:

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

где  $j$  — номер рабочей итерации процесса минимизации критерия [5].

При достижении итераций [12] условий полученные оценки воздействий генетико-физиологических систем в дальнейшем будут обозначаться  $\Delta \varphi_i^*$ . При сепарации полученных значений векторов на подмножества классов по правилу [7] их границы удобно задавать системой неравенств:

$$\Omega_k : \varphi_{lk \min} \leq \varphi_{lk} \leq \varphi_{lk \max}, l = \overline{1, 7}, \quad [13],$$

где  $l$  — индексы генетико-физиологических систем.

Систему неравенств [13], посредством которых в пространстве воздействий семи генетико-физиологических систем осуществляется сепарация особей, в дальнейшем мы будем называть «эколого-генетическим портретом» генотипа, имея в виду, что в свете развиваемой теории (ТЭ-ГОКП) он считается единственно возможным представлением различий для генотипов.

Заметим, что векторы воздействий генетико-физиологических систем  $\Delta \varphi_i^*$  служат нам только «метками», или ориентирами для формирования подмножеств причинно-следственных отношений:

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

здесь  $\Delta \tilde{E}_{ki}$  — среднее значение вектора экологических вариаций в заключительный межфазовый период.

Именно для таких множеств мы вводим более простые решающие

правила. Для этого сначала объединим вектор экологических причин  $\Delta\tilde{E}$  и вектор последствий  $\Delta Y$  в единый вектор  $Z^T = [\Delta\tilde{E}, \Delta Y]^T$  и определим основные статистические характеристики классов в множествах [14] — векторы математических ожиданий  $M_{Zk}$  и матрицы ковариаций  $K_{Zk}$ , а также вероятности появления классов, оценки которых представляют собой отношение чисел особей, попавших во множества отдельных классов  $I_k$ , к общему числу исследуемых особей, то есть

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

По указанным характеристикам легко построить разделяющие функции классов (11):

$$\Phi_{k,k+1}(Z) = Z^T (K_{Zk} - K_{Zk+1})Z + 2(M_{Zk} K_{Zk}^{-1} M_{Zk}^T - M_{Zk+1} K_{Zk+1}^{-1} M_{Zk+1}^T)Z \quad [16],$$

а также пороговое число  $c$  правила [9]:

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

где  $|K|$  означает норму матрицы.

Очевидно, что согласно [9] для каждой новой реализации причинно-следственных признаков  $Z^T = [\Delta\tilde{E}, \Delta Y]^T$  необходимо попарно сопоставить между собой все возможные генотипы. Как уже указывалось выше, векторы воздействий генетико-физиологических систем  $\Delta\varphi_i^*$  служат «метками» для формирования подмножеств причинно-следственных отношений

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

Однако для решения последующих задач селекции нам потребуется статический вариант модели «экологическое возмущение—реакция генетико-физиологических систем».

Для этого нам придется сформировать идентификационное множество:

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

на котором мы оцениваем параметры матрицы  $W$  искомой модели:

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

Здесь в качестве иллюстрации для решения задачи мы рассмотрели только один из модулей общей модели системы «генотип—среда». В случае необходимости включения в характеристики генотипов других количественных признаков размерность задачи может быть увеличена без изменения сути предложенного подхода. При этом важная особенность разработанной теории идентификации генотипов заключается том, что она предусматривает решение этой задачи в течение всего периода онтогенеза, начиная первых фенотипов, то есть с модулей самого низкого уровня иерархии, и завершая ее выходными модулями конечной продукции. Это значительно повышает надежность решения задачи и позволяет более продуктивно использовать всю генотипическую вариабельность, которой располагает генетик-селекционер.

Итак, предложена формализованная теория идентификации генотипов по их фенотипам, включающая в себя оценивание (с помощью математической модели и специального алгоритма оптимизации) величин ненаблюдаемых вкладов семи генетико-физиологических систем в продуктивность особи; классификацию особей с помощью заданной системы неравенств по полученным уровням вкладов генетико-физиологических систем в продуктивность особи; формирование для каждого из классов осо-



бей подмножеств вариаций экологических факторов и вариаций количественных признаков с одновременным оцениванием многомерных статистических характеристик для этих объединенных подмножеств; определение по статистическим характеристикам вариаций экологических факторов и вариаций количественных признаков границ отдельных классов генотипов, по которым может быть реализован алгоритм упрощенной идентификации генотипов по фенотипам.

#### Л И Т Е Р А Т У Р А

1. Михайленко И.М., Драгавцев В.А. Основные принципы математического моделирования системы «генотип—среда». *Сельскохозяйственная биология*, 2010, 3: 31-34.
2. Драгавцев В.А. Метод оценки роли наследственности и среды в развитии признаков растений, не требующий смены поколений. *Ботанический журнал*, 1966, 7: 939-946.
3. Дьяков А.Б., Драгавцев В.А. Разнонаправленность сдвигов количественного признака индивидуального организма под влиянием генетических и средовых причин в двумерных системах признаковых координат. В кн.: *Алгоритмы эколого-генетической инвентаризации генофонда и методы конструирования сортов сельскохозяйственных растений по урожайности, устойчивости и качеству (методические рекомендации, новые подходы)* /Под ред. В.А. Драгавцева. СПб, 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*. Edinburgh, 1950, London, 1952.
6. Драгавцев В.А., Литун П.П., Шкель Н.М., Нечипоренко Н.Н. Модель эколого-генетического контроля количественных признаков растений. *Доклады АН СССР*, 1984, 274(3): 720-723.
7. Кочерина Н.В., Драгавцев В.А. Введение в теорию эколого-генетической организации количественных признаков растений и теорию селекционных индексов. СПб, 2008.
8. Чесноков Ю.В., Почепня Н.В., Бёрнер А., Ловассер У., Гончарова Э.А., Драгавцев В.А. Эколого-генетическая организация количественных признаков растений и картирование локусов, определяющих агрономически важные признаки у мягкой пшеницы. *Доклады Академии наук (РАН)*, 2008, 418, 5: 1-4.
9. Драгавцев В.А. Эколого-генетический скрининг генофонда и методы конструирования сортов сельскохозяйственных растений по урожайности, устойчивости, качеству, (новые подходы). СПб, 1998: 25.
10. Миленький А.В. *Классификация сигналов в условиях неопределенности*. М., 1975.
11. Пугачев В.С. *Теория вероятностей и математическая статистика*. М., 1979.

*ГНУ Агрофизический институт Россельхозакадемии,*  
196220 г. Санкт-Петербург, Гражданский просп., 14,  
e-mail: [ilya.mihailenko@yandex.ru](mailto:ilya.mihailenko@yandex.ru)

*Поступила в редакцию*  
*22 августа 2011 года*

#### 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*

#### S u m m a r y

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.

---

#### Научные собрания

**ЮБИЛЕЙНАЯ XXV ЗИМНЯЯ МОЛОДЕЖНАЯ НАУЧНАЯ ШКОЛА  
«ПЕРСПЕКТИВНЫЕ НАПРАВЛЕНИЯ ФИЗИКО-ХИМИЧЕСКОЙ  
БИОЛОГИИ И БИОТЕХНОЛОГИИ»,  
посвященная 30-летию Научно-образовательного центра ИБХ РАН  
(г. Москва, 11-15 февраля 2013 года)**

**Контакты и информация:** [press@ibch.ru](mailto:press@ibch.ru), <http://www.ibch.ru>