ГЕНЕТИКА
Том V, № 10, 1969
Октябрь

УДК 575.519.95

КИБЕРНЕТИЧЕСКИЕ МЕТОДЫ В КОЛИЧЕСТВЕННОЙ ГЕНЕТИКЕ [1]

Х. АРЕНС, К. БЕЛЛМАН

Институт прикладной математики и механики АН ГДР, Берлин;
Институт растениеводства Немецкой академии сельскохозяйственных наук ГДР, Гросс-Люзевитц

ВВЕДЕНИЕ

При отборе на полигенный признак возникает вопрос об ожидаемой эффективности для любого генотипа в каждом поколении. Однако в такой общей постановке вопрос решить невозможно. Лишь при некоторых очень жестких ограничениях применимы методы оценок количественной генетики, пригодные для предсказания эффективности селекции. Существеннейшим генетическим параметром является при этом наследуемость в узком смысле слова: $h^2 = \sigma_A{}^2/\sigma_P{}^2$. Здесь $\sigma_A{}^2$ — аддитивная генотипическая дисперсия и $\sigma_P{}^2$ — фенотипическая дисперсия признака, по которому ведется отбор. Другой важный параметр — средняя степень доминирования

$$\bar{a} = \frac{2\sigma_D{}^2}{\sigma_A{}^2}.$$

Используя ее, можно оценить, будут ли полезны при селекции комбинативные эффекты [1]. Для вычисления $\bar{a}$ необходима величина $\hat{\sigma}_D{}^2$ — оценка доминантной дисперсии. Кроме того, для вычисления коэффициентов селекции существенна оценка генотипических компонент ковариации. Разработка методов этих оценок проведена в количественной генетике в последнее десятилетие. Оценки генотипических компонент дисперсии и ковариации могут быть получены анализом подобия с помощью специальных моделей дисперсионного анализа. Это дает оценку фенотипической компоненты дисперсии, являющейся линейной комбинацией генотипической компоненты.

Переход от фенотипа к генотипу в таких моделях осуществляется, естественно, при введении ряда допущений. Основные допущения можно сформулировать следующим образом: 1) нормальные диплоидные условия в течение мейоза; 2) отсутствие множественных аллелей; 3) равные частоты аллелей в каждом локусе (0,5); 4) отсутствие сцепления и 5) отсутствие эпистаза (взаимодействий между неаллельными генами). Лишь при выполнении этих исходных условий возможны эффективные оценки генетических параметров. К сожалению, для реальных биологических объектов перечисленные условия выполняются далеко не всегда. Следовательно, возникает вопрос: какой смысл имеют оценки, полученные на основе реально не

---

[1] Перевод Н. В. Глотова и Ю. М. Свирежева.

существующих предпосылок, т. е. какое влияние оказывают нарушения ограничений модели на оценку генотипической компоненты дисперсии? Этого нельзя, однако, выяснить непосредственно в биологических экспериментах, так как для полигенных признаков невозможно установить степень сцепления и характер эпистаза. Кроме того, ничего неизвестно о взаимодействии между этими генетическими факторами.

С помощью математико-кибернетических методов, особенно машинных моделей, можно подойти к решению подобных вопросов. Применение машинных моделей для решения различных задач популяционной генетики и теории отбора уже дало некоторые интересные результаты [2—17]. Такой теоретический подход имеет огромные преимущества. Он позволяет с самого начала исходить из самых общих посылок и рассматривать поведение популяции при изменяющихся генетических параметрах. Моделирование процесса отбора требует, однако, очень внимательного отношения к многочисленным биологическим фактам, имеющимся в настоящее время. Любая модель, естественно, дает лишь упрощенное и абстрактное приближение к действительности. Но в этом и ее большое преимущество, так как модель нацелена на самое существенное и поэтому может помочь обнаружению новых взаимосвязей.

Мы проводили работу в два этапа. Первый этап — создание машинных моделей популяций очень простой генетической структуры; изучение влияния различных генетических параметров и их взаимодействий методом статистических испытаний (метод Монте-Карло); статистический анализ результатов. Второй этап — создание машинной модельной популяции, близкой к биологической действительности. Изучение влияния генетических параметров и их взаимодействий в этом случае.

Для простейшей модели первого этапа использовали ЭВМ Урал-1, а для более сложной модели второго этапа — ЭВМ NE-503 со скоростью 20 000 операций в секунду.

ПРОСТЕЙШАЯ МАШИННАЯ МОДЕЛЬ

Структура простейшей модели. В диаллельном случае любую гамету из m-локусов можно изобразить в виде двоичного числа длины m. Зигота, составленная из двух гамет, изобразится парой двоичных чисел той же длины. Цифрой 1 обозначается доминантный аллель, цифрой 0 — рецессивный. Например, гаметы, содержащие пять локусов с доминантными генами в 1, 2 и 5-м локусах и, соответственно, во 2, 4 и 5-м, могут быть представлены двоичными числами (11001) и (01011). Зигота, образованная из этих двух гамет, изобразится тогда парой $\binom{01011}{11001}$. Каждый вертикальный столбец описывает один локус: $\binom{1}{1}$ — гомозиготный по доминантному аллелю, $\binom{1}{0}$ или $\binom{0}{1}$ — гетерозиготный и $\binom{0}{0}$ — гомозиготный по рецессивному аллелю.

Для построения начального распределения гамет используется следующий алгоритм. Задаются m начальных частот доминантных аллелей p_i ($i = 1, 2, ..., m$) и по некоторой подпрограмме выдается случайное число ξ. Случайная величина ξ равномерно распределена в интервале (0, 1). Если $\xi < p_i$, то в i-том локусе ставится 1; если $\xi \geqslant p_i$, то в i-том локусе ставится 0. Реализация m таких операций дает гамету. Всего строится $2N$-гамет, которые объединяются в популяцию N-зигот.

Каждому генотипу можно поставить в соответствие некоторое число G, называемое ценой генотипа и характеризующее генотипическую отборную ценность зиготы в отношении рассматриваемого признака. Очевидно, что эта величина зависит от степени доминирования в каждом локусе и от эпи-

статических взаимоотношений между локусами. Для простоты будем пока считать, что эпистаз отсутствует. Следующее упрощающее предположение состоит в том, что цена генотипа в целом есть аддитивная функция цен его локусов. Для этого необходимо задать цену каждого локуса в зависимости от степени доминирования в нем. Возможны следующие типы доминирования: D_1 — отсутствие доминирования, D_2 — частичное доминирование, D_3 — полное доминирование и D_4 — супердоминантность. Эти четыре класса мы объединим в два следующим образом: (1) полное доминирование D_3 и (2) неполное доминирование $\bar{D}_3$ (D_1, D_2, D_4).

В случае (1) цена генотипа определяется следующим образом: вычисляется логическая сумма для каждого локуса и затем полученные числа суммируются. Например, для $m = 5$: генотип $\left(\begin{smallmatrix}01011\\11001\end{smallmatrix}\right)$, логическое сложение по локусам дает (11011), суммируя, получаем $G = 4$. Таким образом, здесь цена гепотипа совпадает с общим числом гомозиготных по доминантному гену и гетерозиготных локусов.

В случае (2) каждому локусу ставится в соответствие некоторое число, характеризующее степень доминантности признака. Ценой генотипа здесь будет среднее из этих чисел, где осреднение проводится по всем локусам генотипа.

Для учета влияния среды на генотип вводится понятие цены фенотипа. Влияние среды задается в виде нормально распределенной случайной величины $[N(0; \sigma_U^2)]$. Для простоты окружающая дисперсия σ_U^2 выбирается пропорциональной генотипической дисперсии: $\sigma_U^2 = c \cdot \sigma_G^2$. Так как σ_U^2 задается произвольно, то можно изучать признаки с любой степенью наследуемости. Введенное таким образом влияние среды изменяет начальное распределение цен генотипов, давая итоговую фенотипическую цену зиготы.

Далее мы вводим давление отбора. Отбирают $v\%$ «лучших» зигот, имеющих большую фенотипическую ценность, которые и принимают участие в формировании следующего поколения.

Следующее поколение зигот строится путем «разложения» зигот предыдущего поколения на гаметы с их последующим случайным объединением. В нашей модели первому процессу должно соответствовать разложение двойной последовательности нулей и единиц на две одинарные последовательности. При этом нужно учитывать рекомбинацию между гомологичными хромосомами (кроссинговер). Группы сцепления в машинной схеме конструируются путем задания $r = 0{,}5$ (случайная рекомбинация) между «последним локусом» одной хромосомы и «первым локусом» другой. Для создания рекомбинантных гамет применяется следующие два алгоритма.

При небольшом числе локусов (например, $m = 3$) поступаем следующим образом [9]. Пусть для зиготы $\left(\begin{smallmatrix}111\\000\end{smallmatrix}\right)$ частоты рекомбинации между 1-м и 2-м локусами и между 2-м и 3-м равны соответственно r_{12} и r_{23}. Тогда возможно образование $2^3 = 8$ типов разных гамет. Зигота $\left(\begin{smallmatrix}111\\000\end{smallmatrix}\right)$ дает гаметы: (111) с частотой $1/2(1 - r_{12})(1 - r_{23})$; (110) с частотой $1/2(1 - r_{12})r_{23}$; (101) — $1/2r_{12}r_{23}$; (100) — $1/2r_{12}(1 - r_{23})$; (011) — $1/2r_{12}(1 - r_{23})$; (010) — $1/2r_{12}r_{23}$; (001) — $1/2(1 - r_{12})r_{23}$ и (000) — $1/2(1 - r_{12})(1 - r_{23}.)$ Сумма всех частот равна единице. Всем типам гамет ставится в соответствие их частота, гаметы нумеруются и в предположении панмиксии выбираются для образования пар случайно с учетом их частоты. Однако при большом числе локусов, например, при $m = 50$, число разных типов гамет составляет 2^{50}. Случайное образование пар гамет при этом не может быть реализовано ввиду ограниченного количества ячеек памяти.

Здесь, однако, осуществим алгоритм, при котором нет необходимости строить все возможные гаметы. Строгая случайность выборки тем не менее

сохраняется. Пусть ζ — множество родительских зигот и зигота $Z_k (k = 1, 2, ..., N)$ случайно выбрана из ζ. Для m-локусов есть $(m - 1)$ частот рекомбинации между локусами $(r_{12}, r_{23}, \ldots, r_{m-1,m})$. Обе гомологичные хромосомы (А и В) представляют собой последовательности чисел 0 и 1. Случайная величина ξ равномерно распределена в интервале (0, 1). Выбирается случайное число ξ_0 и производится следующая операция. Если $\xi_0 \geqslant 0{,}5$, то начинаем с последовательности А; если $\xi < 0{,}5$, то начинаем с последовательности В. Снова извлекается число ξ_i и принимаем следующее решение. Если $\xi_i < r_{i,i+1}$, то переход в $(i + 1)$-й локус сопровождается переходом к другой гомологичной хромосоме (происходит кроссинговер). Если $\xi_i \geqslant r_{i,i+1}$, то переход в $(i + 1)$-й локус происходит по той же хромосоме, в которой находится i-й локус. Таким образом, из случайно выбранной зиготы Z_k получается одна из 2^m возможных гамет и именно с такой частотой, какую следует ожидать при наличии кроссинговера. Зигота Z_k возвращается обратно в множество ζ родительских зигот, и выбирается новая зигота $Z_{k'}$ (возможно, $k' = k$).

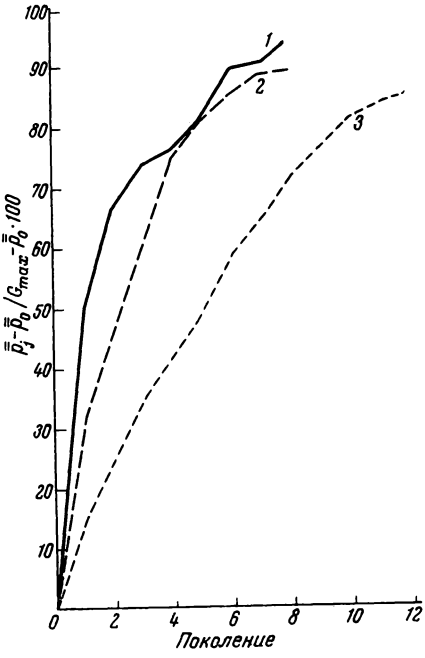

Рис. 1. Зависимость эффективности отбора от числа локусов при отсутствии сцепления (r=0,5):

1 — 5 локусов; *2* — 20 локусов; *3* — 50 локусов

Построенная таким образом первая гамета объединяется с точно так же построенной следующей. Эта процедура повторяется до тех пор, пока не будет восстановлена численность популяции N. Так как каждая гамета формируется случайно, объединение двух друг за другом полученных гамет также является случайным событием, соответствующим панмиксии при перекрестном оплодотворении.

Результаты исследования простейшей модели. На простейшей модели мы изучали зависимость эффективности отбора от силы сцепления, числа локусов и характера доминирования. Изучена также связь между силой сцепления и численностью популяции. Вычислялась средняя цена зиготы для популяции в j-м поколении ν-й повторности. Поскольку для простейшей модели влияние среды не вводится, итоговая фенотипическая цена совпадает с ценой генотипа:

$$\overline{P}_j^{(\nu)} = \frac{1}{N} \sum_{k=1}^{N} G_{kj}^{(\nu)}, \qquad \begin{array}{l} j = 1, 2, \ldots, 8, (12), (15) \\ \nu = 1, 2, \ldots, 5, (\ldots, 20). \end{array}$$

Численность популяции $N = 50$. Средняя по повторностям равна

$$\overline{\overline{P}}_j = \frac{1}{n} \sum_{\nu=1}^{n} P_j^{(\nu)}.$$

Зависимость эффективности отбора от числа локусов при отсутствии сцепления ($r = 0{,}5$). В этом случае были выбраны следующие генетические параметры: число локусов $m = 5$, 20, 50; начальные частоты аллелей по всем локусам $p_i = q_i = 0{,}5$; отбирается 10% «лучших» зигот; полное доми-

нирование; отсутствие сцепления ($r = 0{,}5$). Предполагается также отсутствие эпистаза и константность среды. Эффективность отбора для j-го поколения характеризовалась отношением приращения $\overline{\overline{P}}_j - \overline{\overline{P}}_0$ к общему приращению $G_{\max} - \overline{P}_0$. На рис. 1 можно видеть, что эффективность отбора уменьшается при увеличении числа локусов. Специальный анализ сходимости результатов отбора показал, что она также уменьшается при заданных параметрах с увеличением числа локусов [6].

Зависимость эффективности отбора от числа локусов при тесном сцеплении ($r = 0{,}005$). При тесном сцеплении ($r = 0{,}005$) тормозящий эффект больше-

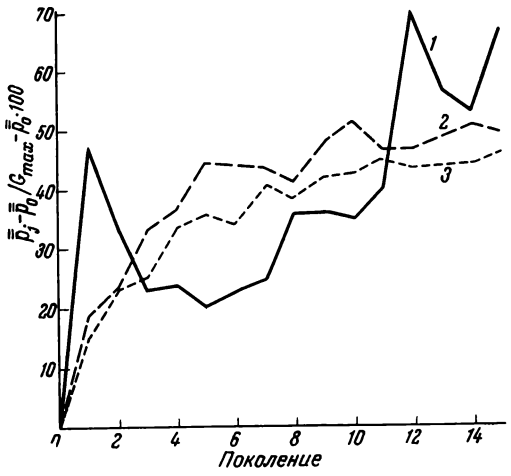

Рис. 2

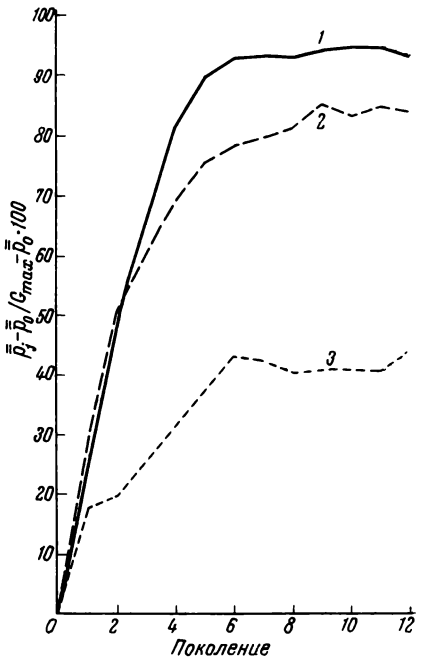

Рис. 3

Рис. 2. Зависимость эффективности отбора от числа локусов при тесном сцеплении (r=0,005):

1 — 5 локусов; *2* — 20 локусов; *3* — 50 локусов

Рис. 3. Зависимость эффективности отбора от средней степени доминирования:

1 — генотип А, частичное доминирование ($\bar{a} = 0{,}5$); *2* — генотип В, полное доминирование ($\bar{a} = 1{,}0$); *3* — генотип С, супердоминантность ($\bar{a} = 1{,}5$)

го числа локусов выражен слабо (рис. 2). Здесь значительно ниже и эффективность отбора: в то время как при отсутствии сцепления отбор в течение восьми поколений всегда давал относительное приращение больше 70% (см. рис. 1), при тесном сцеплении относительное приращение за 14 поколений составляет лишь около 50% (см. рис. 2). Эти различия статистически достоверны.

Зависимость эффективности отбора от степени доминирования. Еще одним параметром, влияющим наряду с числом локусов и силой сцепления на генетический состав популяции, является средняя степень доминирования всех аллелей $\bar{a}$. Этот параметр имеет большое практическое значение. Знание его позволяет принять решение, продолжать ли работу по индивидуальному или семейному отбору внутри штамма или проводить скрещивания между штаммами. Аналитически этот вопрос исследовался нами ранее [1]. Было показано, что любые способы отбора, основанные на использовании аддитивной генотипической дисперсии σ_A^2 (например, массовый отбор),

становятся все менее действенными с увеличением средней степени доминирования. В машинных экспериментах мы приходим к тому же выводу.

В этой модели число локусов $m = 20$, сцепление отсутствует. Испытываются три генотипа с разной средней степенью доминирования. Для генотипа А в 70% всех локусов аллели не проявляют доминирования, в 15% проявляют частичное доминирование, в 10% — полное и в 5% — сверхдоминирование; средняя степень доминирования для генотипа А $\bar{a} = 0{,}5$. Для генотипа В соответствующее распределение локусов по характеру доминирования 5, 15, 70 и 10%, средняя — 1,0. Для генотипа С : 5, 10, 15 и 70% и средняя степень доминирования 1,5. Зависимость эффективности отбора от средней степени доминирования $\bar{a}$ приведена на рис. 3. Для каждого варианта проведены четыре повторности. После отбора в первом поколении разницы между вариантами недостоверны (уровень значимости $\alpha = 0{,}01$). После отбора в течение двух поколений достоверны разницы между вариантами А и С и между В и С. Начиная с третьего поколения достоверны разницы между всеми тремя генотипами. Несмотря на оговорки, касающиеся простоты модели, можно сделать следующее предварительное заключение. Если по данному признаку имеется высокая степень доминирования, то целесообразно проводить скрещивания между штаммами. Это будет более эффективно, чем использование аддитивных эффектов.

Влияние ограничения численности популяции. Рассмотрим еще один результат машинного эксперимента. По Р. А. Фишеру, оценка генетических параметров популяции, таких как средняя μ, аддитивная генотипическая дисперсия σ_A^2 и других, возможна при следующих условиях: (1) — бесконечно большая популяция; (2) — отсутствие сцепления; (3) — отсутствие эпистаза.

Предположение (1) необходимо для выполнения условий уравнения Харди — Вайнберга. Если (1) выполняется, то средняя в случае m локусов равна

$$\mu = \sum_{i=1}^{m} \mu_i = \sum_{i=1}^{m} \left(p_i^2 y^{(i)}_{\binom{1}{1}} + 2p_i q_i y^{(i)}_{\binom{1}{0}} + q_i^2 y^{(i)}_{\binom{0}{0}}\right).$$

Принимая в нашей модели полное доминирование, получаем

$$\mu_i = p_i^2 + 2p_i q_i = 2p_i - p_i^2 \quad (0 \leqslant p_i \leqslant 1).$$

ЭВМ устанавливала частоту гена p_i и вычисляла отсюда μ_i и, следовательно, μ. Методом Монте-Карло, с другой стороны, для каждого поколения j вычислялось в 5—20 повторностях $\bar{P}_j$ и отсюда было получено $\bar{\bar{P}}_j$. Мы будем сравнивать значения μ_j (вычисленные в предположениях Р. А. Фишера) с точно вычисленными значениями $\bar{\bar{P}}_j$ для двух вариантов. В варианте А ограничена численность популяции ($N = 50$), но сцепление отсутствует ($r = 0{,}5$). В варианте В ограничена численность популяции ($N = 50$) и вводится тесное сцепление ($r = 0{,}005$). Эпистаз отсутствует в обоих вариантах. Результаты для обоих вариантов сравнивали с μ после 3, 6 и 8 поколений отбора.

Для варианта А гипотеза H_0: $\mu_j = \bar{\bar{P}}_j$ выполняется лишь при пяти локусах ($m = 5$), при $m = 20$, 50 справедлива альтернативная гипотеза H_1: $\mu_j \neq \bar{\bar{P}}_j$. Таким образом, нарушение условия бесконечности популяции недопустимо.

Для варианта В во всех случаях выполняется гипотеза H_0, т. е. одновременное несоблюдение условий (1) и (2) приводит к уничтожению систематических ошибок. Другими словами, сцепление может преодолеть влияние ограниченной численности популяции.

Хромосома	Параметры	Характеристика локусов	Длина генетической карты
1	P	13 40 81 24 34 34 30 92 FM_1 [84] 28 90 38 31 18	74,2
	r	8 3 7 3 4 6 0,05 12 2 8 6 9 6	
	D	□ × K_1 × × ○ × ○ × △ ○ × × ○ ×	
2	P	16 43 67 84 36 31 (85) 32	5,2
	r	14 8 0,05 10 13	
	D	△ × ○ ○ ○ ○ × ○	
3	P	32 41 19 45 31 44 FM_2 [31] 14 31 14	44,3
		K_2	
	r	14 7 0,05 4 0,05 0,08 0,08 0,08 19	
	D	× × × × × × × × × ×	
4	P	44 HM_3 [19] 76 76 21 36 26 HM_4 [39] 43 25	50,1
	r	4 14 0,05 12 3 2 0,1 9 6	
	D	× × ○ × □ ○ △ △ × ○	
5	P	30 (24) 39 71 83 22 33 30 80 80 23 36 (34) 78	49,3
		K_3	
	r	10 0,05 4 1 26 3 5 6	
	D	○ ○ ○ □ ○ ○ ○ × ○ × ○ ○ △ ×	
6	P	38 FM_5 [39] 87 21 25 18	40,1
		K_4	
	r	3 0,05 12 19 6	
	D	× × × ○ × □	
7	P	29 78 FM_6 [66] HM_7 [25] 9 38 HM_8 [31] 29	38,3
	r	30 0,05 8	
	D	× ○ △ × × × △ ○	
8	P	36 32 25 33 49 27 32 FM_9 [14] 30 69	37,1
		K_5	
	r	4 0,1 0,05 12 0,1 5 4 11 1	
	D	○ × △ ○ × × × △ △ △	
9	P	26 25 36 HM_{10} [28] 20 24 18 43	36,1
	r	3 6 4 0,05 9 7 6	
	D	△ ○ × △ × ○ × △	
10	P	34 47 81 22 35 12 29 15 72 19 11	27,3
		K_5	
	r	4 3 0,05 17 3 0,1 0,05	
	D	△ × × × × ○ × □ ○ × ○	

Модификаторы	FM_1	FM_2	HM_3	HM_4	FM_5	FM_6	HM_7	HM_8	FM_9	HM_{10}
Эффект модификаторов	·2	·5	·1/2	·1/4	·3	·2	·1/2	·1/4	·4	·1/2

Примечание: × - 1; ○ - 2; △ - 3; □ - 4; □ - 5; ◯ - 6

Рис. 4. Стуктура генотипа в популяции II:

P — частота доминантного аллеля · 10^{-2}; *r* — степень сцепления · 10^{-2}; *D* — характер доминирования: *1* — отсутствие доминирования; *2* — частичное; *3* — полное; *4* — супердоминантность; K_1 — K_6 — группы комплементарных генов; *5* — модификаторы, *FM* — усилители, *HM* — ослабители; *6* — «сильные» гены

СЛОЖНАЯ МАШИННАЯ МОДЕЛЬ

Каким образом нарушение некоторых предположений, лежащих в основе модели (например, введение сцепления, эпистаза и т. п.), влияет на оценку генотипической дисперсии? Этот вопрос изучался на более сложной машинной модели. Использовали численные методы оценки генотипической дисперсии ([18], эксперимент I). Исследование проводили параллельно для популяции I, которая удовлетворяет обычным требованиям,

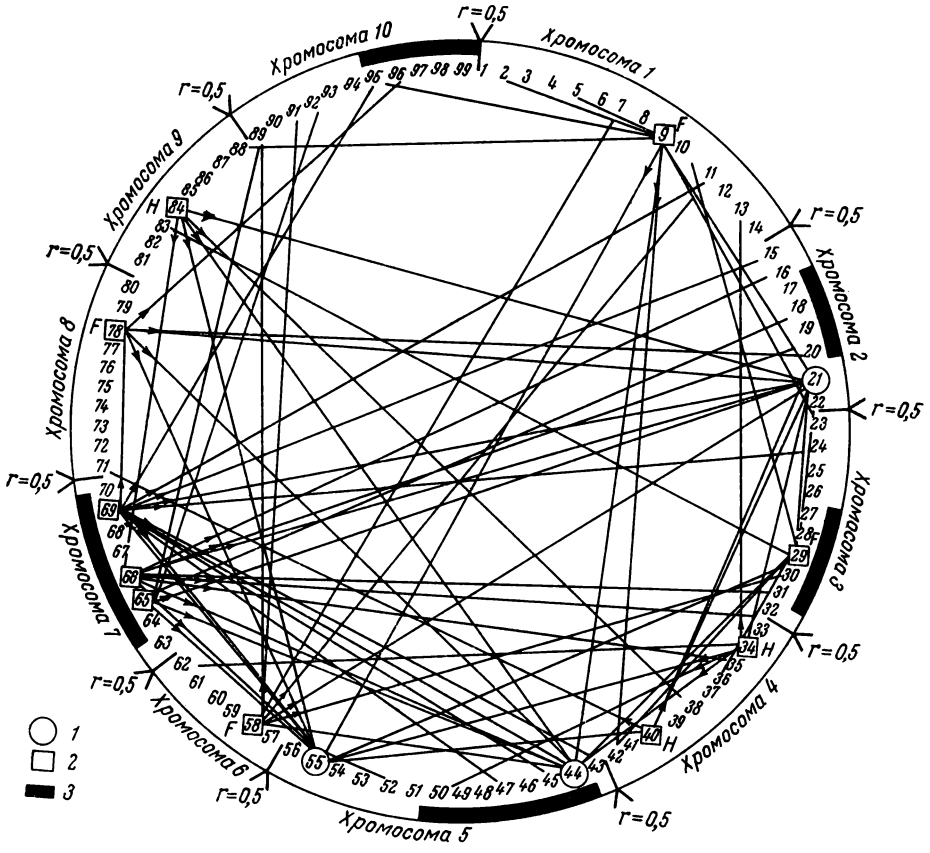

Рис. 5. Характер эпистатических взаимоотношений для генотипа II:

1 — «сильные» гены; *2* — модификаторы; *F* — усилители; *H* — ослабители; *3* — блоки очень тесно сцепленных генов (r = 0,0005); 1, 2,...— порядковый номер локуса

характерным для формальной модели, и для популяции II, не удовлетворяющей этим требованиям и сходной по основным признакам с реальными природными популяциями.

Структура сложной машинной модели. Опишем вначале структуру популяции II (рис. 4, 5), указав затем, чем от нее отличается структура популяции I. В основу модели положены цитологические данные по кукурузе [19]. Взято 10 пар хромосом с учетом их относительных длин в метафазе. По хромосомам случайно распределены 99 локусов. Частоты рекомбинаций между локусами приведены на рис. 4. Для хромосом 2, 3, 5 и 10 мы имеем тесно сцепленные блоки генов (r = 0,0005). Для генов, находящихся в разных хромосомах, естественно, r = 0,5. В 99 локусах гены подразделяются на: 1) «сильные» — 3 локуса; 2) «слабые» — 86 локусов и 3) гены-модификаторы — 10 локусов. Для генов. называемых нами «силь-

ными», фенотипический эффект выбран в 5 раз больше, чем для «слабых» генов. Среди 86 непосредственно действующих на признак «слабых» генов введены 6 групп комплементарных генов K_1 — K_6 (см. рис. 4). Гены-модификаторы не оказывают прямого влияния на признак, действуя лишь опосредованно всегда на «сильные» гены и иногда на «слабые». Усилителями

Рис. 6. Распределение частот доминантного аллеля по локусам в популяции II

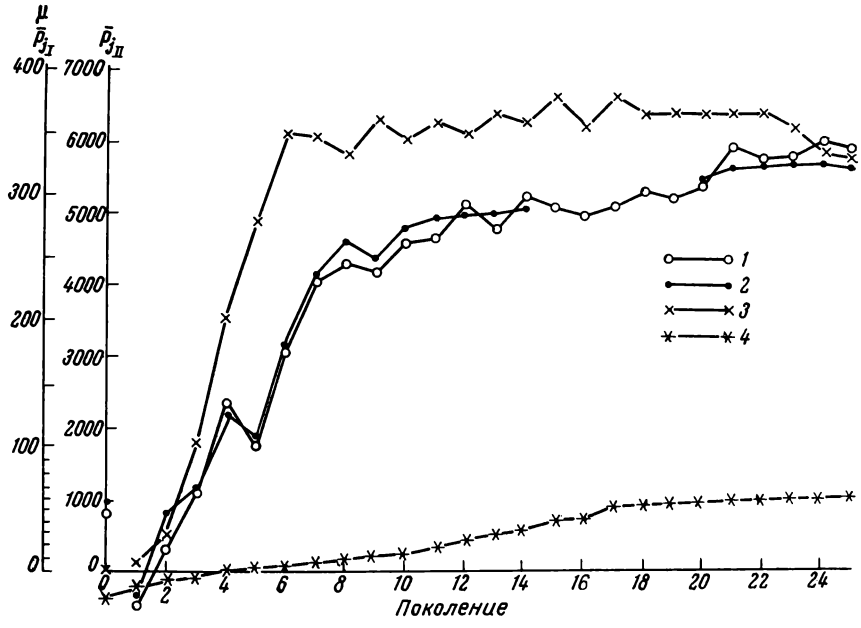

Рис. 7. Изменение средней фенотипической цены зиготы $\bar{P}_j$ и вычисленной средней μ:

1 — $\bar{P}_j$ для генотипа I; *2* — μ для генотипа I; *3* — $\bar{P}_j$ для генотипа II; *4* — μ для генотипа II

являются 5 модификаторов (*FM*) и 5 — ослабителями (*HM*). В последней строке рис. 4 приведены заданные значения эффективности модификаторов. Весь комплекс эпистатических взаимоотношений показан на рис. 5. Отношения доминирования для каждого локуса также приведены на рис. 4: для 48 локусов доминирование отсутствует, для 31 — частичное, для 15 — полное, для 5 — сверхдоминирование. Частоты доминантного гена в каждом локусе (см. рис. 4) выбраны так, что в целом получается дву-

вершинная кривая распределения генных частот p_i по локусам (рис. 6). Влияние среды вводится в виде нормально распределенной случайной величины $u_k[N(0;\ \sigma_u^2)]$, так, что коэффициент наследуемости для данного признака $h = 0{,}3$. Численность популяции $N = 500$, популяция панмиктична. В каждом поколении для формирования популяции особей следующего поколения отбираются 10% «лучших» зигот.

В популяции I сцепление отсутствует, гены-модификаторы и комплементарные гены не отличаются от «слабых» генов. Отсутствуют какие бы то ни было взаимоотношения между неаллельными генами. Отношения доминирования те же, что и в популяции II. Генные частоты по всем локусам одинаковы ($p_i = q_i = 0{,}5$).

Результаты исследования сложной модели. *Изменение средней цены зиготы.* На рис. 7 показано возрастание в чреде поколений средней $\bar{P}_j$ для генотипов I и II, а также вычисленной средней

$$\mu = \sum_{i=1}^{m} a(p_i - q_i) + 2p_iq_id\,.$$

Для популяции I вычисленная μ хорошо согласуется с наблюдаемой. Вследствие действия модификаторов $\bar{P}_j$ для генотипа II расположена значительно выше, чем $\bar{P}_j$ для генотипа I (обратите внимание на шкалу!), а также много выше, чем μ, вычисленная исходя из a, d, и q. В то время как $\bar{P}_j$ для генотипа I растет на протяжении всего времени отбора, соответствующий подъем для генотипа II прекращается после шестого поколения, а затем устанавливается некоторый постоянный уровень.

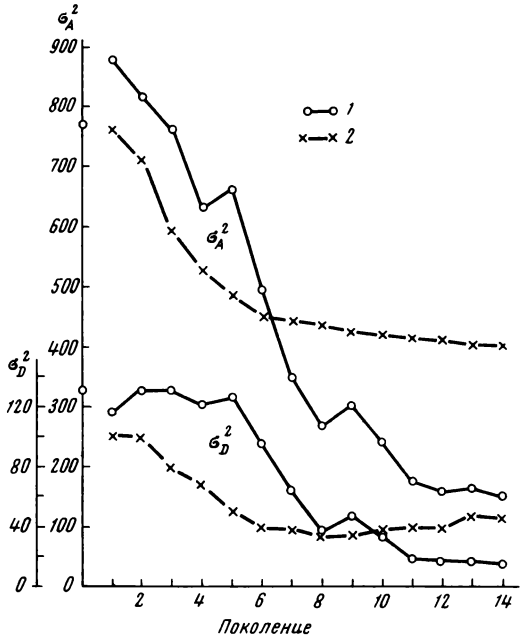

Рис. 8. Изменение аддитивной генотипической дисперсии σ_A^2 и доминантной дисперсии σ_D^2, вычисленных по генным частотам:

1 — генотип I; *2* — генотип II

Изменение генотипической дисперсии. Для модельных популяций, конструируемых на основе простых предположений (популяция I), теория предсказывает уменьшение аддитивной генотипической дисперсии σ_A^2 и доминантной дисперсии σ_D^2 при увеличении средней. Это имеет место и в наших экспериментах (рис. 8). Можно видеть, что обе дисперсии для генотипа I уменьшаются быстрее, чем для генотипа II. Это объясняется тем, что увеличение частоты доминантных аллелей по большинству локусов при наличии сцепления и эпистаза (генотип II) идет медленнее, чем в отсутствие сцепления и неаллельных взаимодействий (генотип I).

Изменение генетической структуры популяции. Отбор действует почти исключительно на модификаторы (рис. 9). После шести поколений все доминантные аллели генов-ослабителей (*HM*) элиминируются, в то время как частота доминантных аллелей генов-усилителей (*FM*) резко возрастает. Оба процесса повышают генотипическую цену зигот и ведут к весьма быстрому увеличению $\bar{P}_j$ в популяции II. С генетической точки зрения харак-

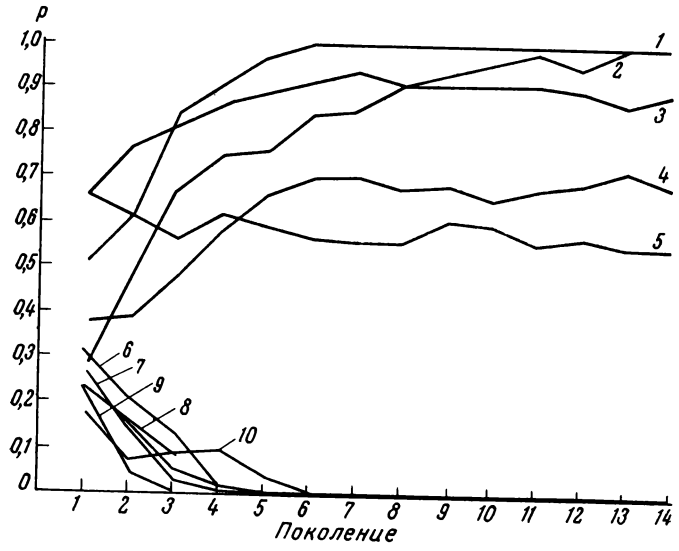

Рис. 9. Изменение частот доминантных аллелей модификаторов:

1 — FM_5; *2* — FM_9; *3* — FM_6; *4* — FM_2; *5* — FM_1; *6* — HM_7; *7* — HM_4; *8* — HM_{10}; *9* — HM_8; *10* — HM_3

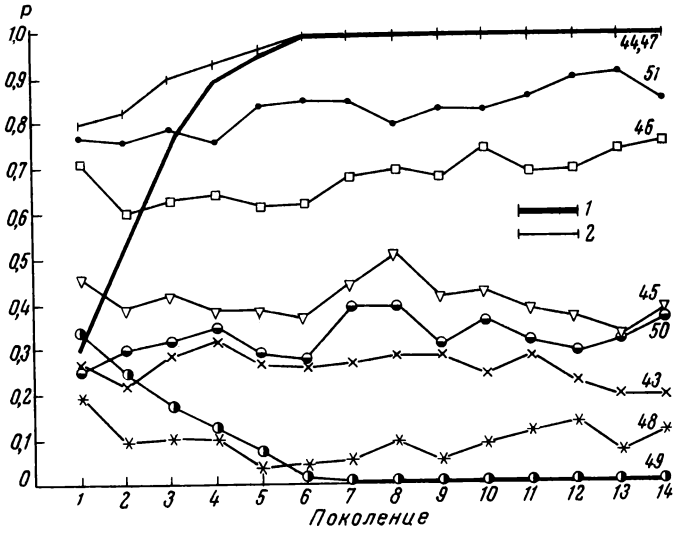

Рис. 10. Изменение частот доминантного аллеля «сильного» гена (№ 44) и восьми очень тесно сцепленных с ним «слабых» генов (r=0,0005):

1 — «сильный»; *2* — «слабые» гены. Справа — номера локусов (см. рис. 5)

тер изменений генных частот на рис. 9 крайне интересен тем, что подтверждает часто высказываемое предположение о преимущественном отборе на модификаторы, которые не имеют самостоятельного фенотипического проявления, но влияют на гены, дающие фенотипический эффект. Если разделить весь период отбора на два отрезка — 1—6-е и 7—14-е поколения, то для разных групп локусов мы получим очень разные изменения частот доминантных аллелей на поколение (см. таблицу). Можно видеть, что частота

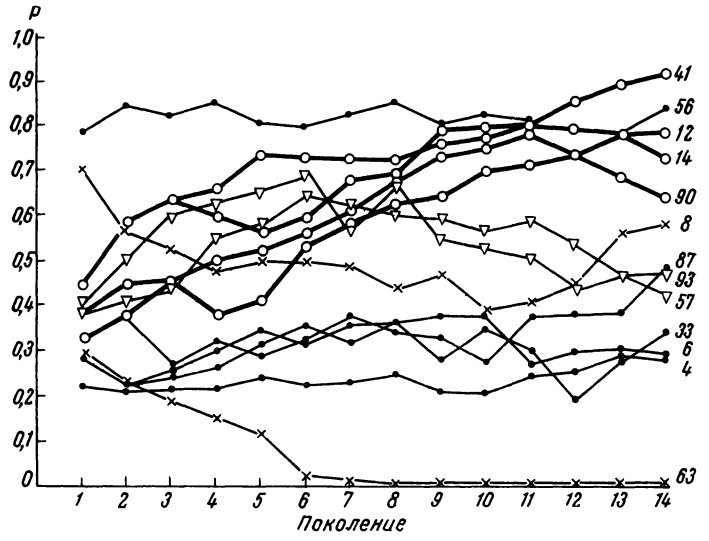

Рис. 11. Изменение частот доминантных аллелей 13 «слабых» генов, не сцепленных или слабо сцепленных

Справа — номера локусов (см. рис. 5)

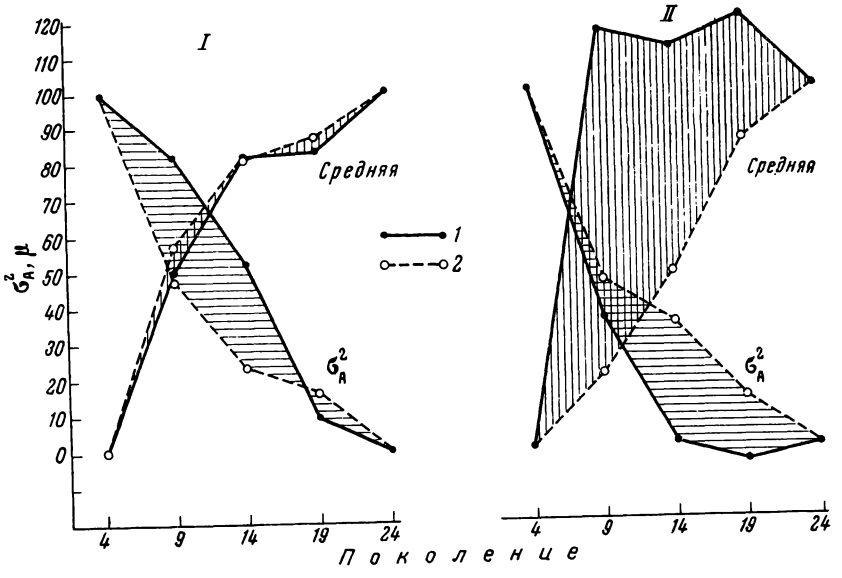

Рис. 12. Изменение средней цены зиготы и аддитивной генотипической дисперсии (100 — разность между 2-м и 4-м поколениями):

1 — наблюдаемая средняя $\overline{P}_j$ и σ_A^2, вычисленная по [18]; *2* — ожидаемая средняя μ и σ_A^2, вычисленная из a, d, q; *I* — генотип I; *II* — генотип II

доминантных ослабителей уменьшается в течение шести поколений на величину $\Delta p = 0{,}06$ (на поколение, на локус). Гены, на которые они влияют, ведут себя так же, но уменьшение частоты для них идет медленнее. Усилители дают возрастание частоты доминантных аллелей, в то время как частоты генов, на которые они влияют, не меняются. «Сильные» гены, на которые одновременно действуют модификаторы, и усилители, и ослабители, поощряются отбором.

Другое интересное обстоятельство связано с одним «сильным» геном, локализованным в блоке очень тесно сцепленных генов (рис. 10). Поразительно, что, несмотря на очень тесное сцепление, частоты генов, расположенных в этом блоке, меняются иначе, чем частота сильного гена. Частота p для «сильного» гена № 44 быстро возрастает, в то время как частоты ряда сцепленных с ним ($r = 0{,}0005$) соседних «слабых» генов (например, № 43 и 45) существенно не меняются. Это указывает на почти случайный дрейф генных частот, сходный с дрейфом «слабых» генов, лежащих вне такого блока сцепления. На рис. 11 жирными линиями выделены лишь четыре систематически идущие вверх кривые. Несущественное повышение $\bar{P}_j$ для популяции II (см. рис. 7) в течение 5—14-го поколений в значительной мере объясняется, по-видимому, этим обстоятельством.

Изменение частоты доминантного аллеля (Δp) для некоторых групп локусов (в % на поколение)

Тип локуса	Локусы, для которых отсутствует доминирование			Локусы, для которых имеется разная степень доминирования		
	Поколения					
	1—6	7—14	1—14	1—6	7—14	1—14
Модификаторы-ослабители	—6	0	—1,7	—6	0	—1,7
«Слабые» гены, на которые влияют ослабители	—1,4	2,5	0,05	—1,8	1,5	—0,01
Модификаторы-усилители	5,3	0	2,3	5,0	0	2,2
«Слабые» гены, на которые влияют усилители	0,8	1,5	1,3	—0,3	1,1	0,05
«Сильные» гены	2,3	0,7	1,3	9,3	0	4,0
Гены, на которые не влияют модификаторы	0,5	0,3	0,4	—0,5	1,6	0,1

Важнейший вопрос, ради которого было предпринято машинное изучение генотипов I и II, заключался в попытке получить представление о том, как отказ от того или иного предположения, положенного в основу модели, влияет на величину генетических параметров. Представлялось интересным сравнение σ_A^2 и μ, вычисленных, с одной стороны, согласно [18], и, с другой, на основе a, d и q (рис. 12). Для генотипа I (отсутствие сцепления, эпистаза, равные генные частоты) и для генотипа II (наличие сцепления, эпистаза, неравные генные частоты) отмечается уменьшение эффективности отбора, которое сказывается на оценках параметров и может быть обнаружено по изменению генных частот. Нужно заметить также, что в случае генотипа II порядки величин вычисленных, с одной стороны, по a, d и q, а с другой, по [18], различны.

ВЫВОДЫ

Работа посвящена применению кибернетических методов в количественной генетике.

Результаты, полученные в этой работе, показывают целесообразность и перспективность применения математико-кибернетических методов(по-

строение машинных моделей) при изучении некоторых важных проблем теории отбора.

Таблиц — 1, иллюстраций — 12, библиография — 19 назв.

Поступила в редакцию
11 сентября 1967 г.

Литература

1. K. Bellmann, H. Ahrens. Zum gegenwärtigen Stand der Anwendungsmöglichkeiten der biometrischen Genetik in der Pflanzenzüchtung. 1. Teil: Die verschiedenen Formen der genetischen Variabilität und ihre Bedeutung in der Pflanzenzüchtung. Züchter, 35, 156, 1965.
2. H. Ahrens, K. Bellmann. Statistische Analyse des Konvergenzverhaltens von Selektionsprozessen bei der Auslese auf polyfaktoriell bedingte Merkmale. Monatsber. Dtsch. Akad. Wiss., Heft 7/8, 568, 1965.
3. J. S. Allan, A. Robertson. The effect of initial reverse selection upon total selection response. Genet. Res., 5, 68, 1964.
4. R. W. Allard, P. E. Hansche. Some parameters of population variability and their implications in plant breeding. Advances Agron., 16, 281, 1964.
5. J. A. Arthur, H. Abplanalp. Studies using computer simulation of reciprocal reccurent selection. Genetics, 55, 233, 1964.
6. K. Bellmann, H. Ahrens. Modellpopulationen in der Selektionstheorie und einige Ergebnisse aus Simulationsstudien. Züchter, 36, 172, 1966.
7. T. P. Bogyo. Some population parameters as affected by truncation selection. Genetics, 52, 429, 1965.
8. N. R. Bohidar, D. W. Patel. A Monte-Carlo investigation of interaction between linkage and selection under stochastic models (1). Biometrics, 20, 660, 1964.
9. A. S. Fraser. Simulation of genetic systems by automatic digital computers. I. Austral. J. Biol. Sci., 10, 484, 1957.
10. A. S. Fraser. Simulation of genetic systems by automatic digital computers. II. Austral. J. Biol. Sci., 10, 492, 1957.
11. A. S. Fraser. Simulation of genetic systems by automatic digital computers. 5-linkage, dominance and epistasis. In: Biometrical Genetics. Pergamon Press, 1960, p. 70.
12. J. L. Gill, B. A. Clemmer. Selection, linkage and the coefficient of inbreeding. Genetics, 52, 445, 1965.
13. S. K. Jain, R. W. Allard. The simulation of selection models involving linkage, epistasis and inbreeding. Genetics, 50, 259, 1964.
14. F. G. Martin, C. C. Cockerham. High speed selection studies. In: Biometrical Genetics. Pergamon Press, 1960, p. 35.
15. R. F. Nassar. Effect of correlated gene distribution due to sampling on the diallel analysis. Genetics, 52, 9, 1965.
16. A. Robertson. Computer simulation of genetic selection experiments. Biometrics, 20, 393, 1964.
17. K. Bellmann. Über eingie Möglichkeiten zur Nachbildung genetischer Strukturen und Prozesse und Rechenautomaten. Abhandl. Dtsch. Akad. Wiss., Berlin, Kl. Math., Phys. u. Tech., № 4, 21, 1967.
18. R. E. Comstock, H. F. Robinson. Estimation of average dominance of genes. In: Heterosis. ed. by J. Gowen, Iowa State Coll. Press, 1952, p. 494.
19. B. McClintock. Chromosome morphology in *Zea mays*. Science, 69, 629, 1929.

CYBERNETIC METHODS IN QUANTITATIVE GENETICS

H. AHRENS, K. BELLMANN

Institute of Applied Mathematics and Mechanics, German Academy of Sciences, Berlin; Institute of Plant Breeding, German Academy of Agricultural Sciences, Gross-Lüsewitz, GDR

Summary

By means of mathematic-cybernetical methods machine models designed to simulate genetic structures and processes were developed. The mathematical theory of estimation of genetic parameters requires certain assumptions which are not fulfilled in re-

all biological populations. Therefore biassed estimates of genetic parameters are possible. The models described here are intended to investigate the influence of the nonfulfilled assumptions on the estimated genetic parameters. The creation of genotypic populations characterized by certain genetic conditions (linkage, epistasis, gene frequency, degree of dominance) are described in detail. Furthermore. the evaluation of the genotype and phenotype, their dependence on allelic and nonallelic gene interactions, the formation of gametes in relation to given recombination values are described. The results of simulated selection processes suggest, that in the case of constant selection coefficients, close linkage and a high degree of dominance diminish the genetic gain per generation. It is possible to overcome the effect of finiteness on the Hardy — Weinberg structure by means of certain linkage relations in the inheritance of polygenic characters.

After the first investigations on relative simplebuilt genotypes (20, 50 loci, without epistasis) the main problem (influence of non fulfilled model assumptions on the estimates of genetic parameters) was treated on the basis of model genotypes better adapted to the real biological conditions. For this purpose the experiment I of Comstock and Robinson (1952) was simulated in 2 populations: the first population with fulfilled assumptions (original gene frequencies = 0.5, without linkage, without epistasis); the second population with infringed assumptions (original gene frequencies = 0.5, with linkage, with epistasis). Preliminary results of the studies involving these two genotypes are reported.