Доклады Академии Наук СССР
1940. Том XXVII, № 1

ГЕНЕТИКА

Академик А. Н. КОЛМОГОРОВ

ОБ ОДНОМ НОВОМ ПОДТВЕРЖДЕНИИ ЗАКОНОВ МЕНДЕЛЯ

В происходившей осенью 1939 г. дискуссии по вопросам генетики много внимания уделялось вопросу проверки состоятельности законов Менделя. В принципиальной дискуссии о состоятельности всей менделевской концепции было естественно и законно сосредоточиться на простейшем случае, приводящем по Менделю к расщеплению в отношении 3 : 1. Для этого простейшего случая скрещивания $Aa \times Aa$ при условии доминирования признака A над признаком a менделевская концепция приводит, как известно, к выводу, что в достаточно обширном потомстве (безразлично, составляющем одно семейство или объединяющем много отдельных семейств, полученных от различных пар гетерозиготных родителей типа Aa) отношение числа особей с признаком A (т. е. особей типа AA или Aa) к числу особей с признаком a (типа aa) должно быть близким к отношению 3 : 1. На проверке этого простейшего следствия из менделевской концепции и сосредоточили свое внимание Т. К. Енин (1,2), Н. И. Ермолаева (3) и Э. Кольман (4). Между тем менделевская концепция не только приводит к указанному простейшему заключению о приближенном соблюдении отношения 3 : 1, но и дает возможность предсказать, каковы должны быть в среднем размеры уклонений от этого отношения. Благодаря этому как раз статистический анализ уклонений от отношения 3 : 1 дает новый, более тонкий и исчерпывающий способ проверки менделевских представлений о расщеплении признаков. Задачей настоящей заметки является указание наиболее рациональных, по мнению автора, методов такой проверки и их иллюстрация на материале работы Н. И. Ермолаевой (2). Материал этот, вопреки мнению самой Н. И. Ермолаевой, оказывается блестящим новым подтверждением законов Менделя *.

§ 1. Общие замечания о роли случая в явлениях наследственности

Станем сначала на точку зрения, независимую от менделизма. Пусть в результате скрещивания двух особей α и β получается потомство из n особей. Обычно при этом каждая из двух особей α и β производит число гамет значительно большее, чем число потомков n. Пусть α производит гаметы

$$\alpha_1, \alpha_2, \ldots, \alpha_{k_1},$$

а β — гаметы

$$\beta_1, \beta_2, \ldots, \beta_{k_2}.$$

* На работу Н. И. Ермолаевой обратил мое внимание акад. А. С. Серебровский.

Какие именно из весьма большого числа k_1, k_2 возможных пар гамет будут действительно использованы для получения потомства, зависит от многих обстоятельств. Эти обстоятельства законно разделить на внутренние, определяемые биологическими свойствами особей α и β и гамет α_i и β_j, и на внешние, независимые от этих биологических свойств. Например при опылении у растений решение вопроса о том, какие именно из пыльцевых зерен попадут на рыльце, а какие нет, или о том, каково будет расположение на рыльце попавших на него зерен, зависит от бесчисленного множества факторов совершенно внешнего по отношению к биологическим закономерностям характера. При изучении биологических закономерностей следует считать эти внешние для них обстоятельства оплодотворения случайными и применять к ним аппарат теории вероятностей.

Выбор n пар

$$(\alpha_{i_1}, \beta_{j_1}),\ (\alpha_{i_2}, \beta_{j_2}),\ \ldots,\ (\alpha_{i_n}, \beta_{j_n})$$

возможен

$$s = \frac{k_1!\, k_2!}{n!\,(k_1 - n)!\,(k_2 - n)!}$$

различными способами. В соответствии со сказанным выше дальнейшее исследование должно исходить из предположения, что биологическими факторами определяется для каждого из этих возможных способов выбора лишь некоторая вероятность его действительного осуществления.

Вывод менделевских законов основывается (см. далее § 2) на простейшем допущении, что вероятности, соответствующие любому из этих s возможных способов выбора, одинаковы (и, следовательно, все равны $\frac{1}{s}$). Биологически это допущение обозначает равную жизнеспособность гамет, отсутствие селективного оплодотворения и равную жизнеспособность (по крайней мере, до момента подсчета потомства) особей, получающихся от любой парной комбинации α_i, β_j гамет. Назовем эту простейшую гипотезу для краткости гипотезой независимости (вероятности того или иного набора гамет, используемых для получения потомства от их биологических особенностей).

Как и всякая другая гипотеза о независимости одних явлений от других, интересующая нас гипотеза, взятая в виде абсолютной догмы, не допускающей никаких коррективов, не верна: существует целый ряд твердо установленных примеров уклонений от этой гипотезы, иногда количественно незначительных, а иногда и очень крупных.

Ясно, что столь же неосновательной была бы точка зрения, пытающаяся вовсе отрицать роль внешних с биологической точки зрения случайных обстоятельств в подборе гамет, участвующих в оплодотворении.

Серьезный спор может итти между такими двумя точками зрения:

1) Гипотеза независимости в большинстве случаев является хорошим первым приближением к действительному положению вещей (сторонники менделевской и моргановской генетики).

2) Селективное оплодотворение и неравная жизнеспособность играют всюду столь решающую роль, что рассмотрения, опирающиеся на гипотезу независимости, для биологии бесплодны (школа акад. Т. Д. Лысенко).

§ 2. Расщепление в отношении 3 : 1

Вернемся теперь к более специальным менделевским допущениям для случая скрещивания $Aa \times Aa$ и доминирования признака A. В этом случае принимается, что каждый из родителей образует столько же гамет типа A, как и гамет типа a, пары гамет типов AA и Aa дают потомков, обладающих свойством A, а пары гамет типа aa—потомков, обладающих свойством a. Из этих предположений в соединении с допущением, что k_1 и k_2 несравненно больше, чем n, и с гипотезой не зависимости получается вывод:

I. Вероятность того, что в группе из n потомков окажется ровно m особей с признаком A (остальные же с признаком a) равняется

$$P_n(m)=\frac{n!}{m!\,(n-m)!}\left(\frac{3}{4}\right)^m\left(\frac{1}{4}\right)^{n-m}. \tag{1}$$

Пусть теперь произведено большое число r скрещиваний $Aa \times Aa$ и от каждого из них получено по семейству из

$$n_1,\ n_2;\ \ldots\ ,\ n_r$$

особей, из числа которых, соответственно,

$$m_1,\ m_2,\ \ldots\ ,\ m_r$$

обладают признаком A. Спрашивается, как возможно полнее проверить, согласуется ли такой результат опыта с менделевскими допущениями или нет?

Если число особей в каждом семействе очень мало (например меньше 10), то целесообразно непосредственно проверять формулу (1) при помощи χ^2— критерия Пирсона.

Если каждое семейство сравнительно многочисленно, то целесообразнее другой метод. В этом случае из I вытекает:

II. Нормированные уклонения $\Delta=\left(\frac{m}{n}-\frac{3}{4}\right):\sigma_n$, где

$$\sigma_n=\frac{\sqrt{3}}{4\sqrt{n}}$$

подчиняются приближенно закону Гаусса с единичной дисперсией, т. е. вероятность неравенства

$$\Delta \leqslant x$$

приближенно равна

$$P(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{\xi^2}{2}}\,d\xi. \tag{2}$$

Здесь $\sigma_n=\frac{\sqrt{3}}{4\sqrt{n}}$ есть среднее квадратическое уклонение частоты $\frac{m}{n}$ от $\frac{3}{4}$. Мы видим, что это среднее квадратическое уклонение пропорционально $\frac{1}{\sqrt{n}}$. Следовательно, только в случае очень больших семейств менделевская теория предсказывает большую близость частоты $\frac{m}{n}$ к $\frac{3}{4}$. Например чтобы с вероятностью 0,99 можно было бы утверждать, что

$$\left|\frac{m}{n}-\frac{3}{4}\right| < 0{,}01,$$

n должно быть больше 12 000.

Зато рассмотрение большого числа семейств средней величины дает возможность гораздо более тонкой проверки менделевских допущений при помощи рассмотрения распределения уклонений Δ. В частности, из формулы (2) вытекает, что вероятность неравенства $|\Delta| \leqslant 1$ приблизительно равна 0,68. Следовательно, по менделевской теории среди достаточно большого числа семейств должно быть приблизительно 68% с $\Delta \leqslant 1$ и приблизительно 32% с $|\Delta| > 1$. В работе Н. И. Ермолаевой ([3]) для проверки этого следствия менделевской теории пригодны серия из 98 семейств с расщеплением по окраске цветка и пазухи листа (табл. 4 работы Ермолаевой) и серия из 123 семейств с расщеплением по окраске семядолей (табл. 6 работы Ермолаевой). Остальные серии содержат слишком мало семейств (или групп семейств) для надежной проверки. Результаты приведены в следующей таблице*:

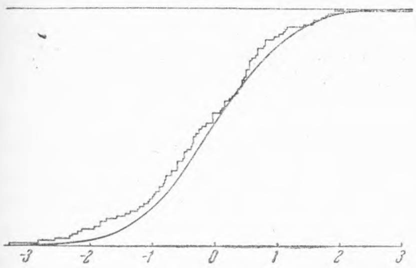
	Расщепление (по окраске) цветка и пазухи листа	в %	Расщепление (по окраске) семядолей	в %	Теоретически в %
Всего семейств	98	100	123	100	100
с $\lvert\Delta\rvert \leqslant 1$ . . .	66	67	85	69	68
с $\lvert\Delta\rvert > 1$ . . .	32	33	38	31	32

При указанном числе семейств в серии совпадение с теорией следует признать очень хорошим. В силу странного недоразумения сама Н. И. Ермолаева утверждает в своей работе, что наличие заметного процента семейств с $|\Delta| > 1$ опровергает менделевскую теорию.

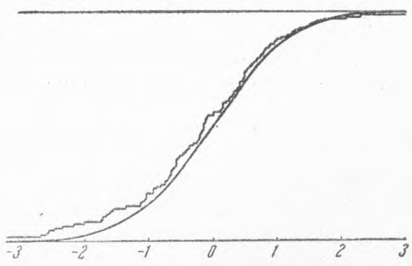
Можно было бы провести аналогичную проверку совпадения с теорией процента семейств, для которых $|\Delta| \leqslant a$ при каком либо $a \neq 1$. Например теория предсказывает, что приблизительно у 50% семейств должно быть $|\Delta| \leqslant 0{,}674$. Однако лучше всего сразу произвести полную проверку близости фактически наблюдаемого распределения уклонений Δ к теоретическому гауссовскому. Для этого вычерчиваем на одном и том же чертеже теоретическую кривую $y=P(x)$ в соответствии с формулой (2) и эмпирическую ступенчатую кривую

$$y = \frac{q(x)}{r},$$

где r обозначает общее число семейств в данной серии, а $q(x)$—число семейств в серии, для которых $\Delta \leqslant x$. Результаты такой проверки для двух интере-



Фиг. 1.



Фиг. 2.

сующих нас серий Н. И. Ермолаевой показаны на фиг. 1 и 2. Как видно, совпадение эмпирической и теоретической кривой в обоих случаях получается достаточно хорошее. Чтобы оценить, является ли наблюдаемое между ними расхождение допустимым при данной численности серий, следует воспользоваться выведенными мною ранее формулами [см. В. Рома-

* Сводная табл. 1 работы Н. И. Ермолаевой содержит несколько отличные цифры, чем приводятся здесь, так как она учитывает некоторые семейства (2 в первой серии и 4 во второй серии), не внесенные по неизвестным нам причинам в табл. 4 и 6. Наша сводная таблица охватывает только семейства, представленные в табл. 4 и 6. Однако следующие ниже выводы остаются без изменений, если исходить из данных сводной табл. 1 Н. И. Ермолаевой.

новский, «Математическая статистика», стр. 226 (1938)]. Для случаев, изображенных на фиг. 1 и 2, находим по этим формулам

$$\lambda_1 = 0{,}82, \quad \lambda_2 = 0{,}75,$$
$$\Phi(\lambda_1) = 0{,}49, \quad \Phi(\lambda_2) = 0{,}37,$$

что вполне удовлетворительно.

Мы видим, таким образом, что большей близости частот $\frac{m}{n}$ по отдельным семействам к их среднему значению $\frac{3}{4}$, чем получилось у Н. И. Ермолаевой, при данной численности семейств и нельзя было бы ожидать по менделевской теории.

Если бы в какой-либо достаточно обширной серии семейств уклонения $\frac{m}{n}$ от $\frac{3}{4}$ были бы систематически меньше, чем требует теория, то это в такой же мере опровергало бы применимость к этой серии семейств сформулированных выше допущений, как и систематическое превышение теоретически предсказываемых размеров этих уклонений. Намек на такую систематическую чрезмерную близость частот $\frac{m}{n}$ к $\frac{3}{4}$ имеется в материалах работы Т. К. Енина (1). Однако материалы этой работы недостаточно обширны (25 семейств по сравнению с двумя сериями в 98 и 123 семейства у Н. И. Ермолаевой) и возбуждают ряд других сомнений (сам автор считает их не вполне однородными). Поэтому в детальное их рассмотрение мы входить не будем.

Из упомянутых в начале заметки работ работа Э. Кольмана, не содержащая нового фактического материала, а посвященная анализу материалов Т. К. Енина, целиком основана на непонимании изложенных в нашей заметке обстоятельств.

Поступило
20 II 1940

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] Т. К. Енин, ДАН, XXIV, стр. 176—178 (1939). [2] Т. К. Енин, Доклады ВАСХНИЛ, № 6 (1939). [3] Н. И. Ермолаева, Яровизация, 2 (23), стр. 79—86 (1939). [4] Э. Кольман, Яровизация, 3 (24), стр. 70—73 (1939).